Volume Preface

There are probably not many psychologists who have spent much time thinking about creating a handbook. The prevalent reasons for becoming a psychologist—scientific curiosity, the need for personal expression, or the desire for fame and fortune—would be unlikely to bring to mind the idea of generating a handbook. At the same time, most would agree that a handbook can be remarkably useful when the need arises. The chapters can provide the background for a grant proposal, the organization of a course offering, or a place for graduate students to look for a research problem. If presented at the right time, the clearly worthwhile aspects of this otherwise most unlikely endeavor can make it an attractive opportunity; or, at least in retrospect, one could imagine saying, “Well, it seemed like a good idea at the time.” Even if there are a few simple and sovereign principles underlying all personality processes and social behavior, they were not consciously present when organizing this volume. Instead, what was terribly salient were the needs and goals of potential users of this volume: What would a reader need to know to have a good understanding of the current theoretical and empirical issues that occupy present-day thinkers and researchers? What could the highly sophisticated investigators who were selected to write the chapters tell the reader about the promising directions for future development? The chapters in this volume provide both thorough and illuminating answers to those questions, and, to be sure, some can be grouped into a few sections based on some common, familiar themes. For those readers who want more information about what chapters would be useful or who are open to being intrigued by the promise of some fascinating new ideas, this is a good time to take a brief glimpse at what the chapters are about.

An immediately pressing question for the editors centered on what content to include and whom to invite for the individual chapters. There are probably many ways to arrive systematically at those decisions, but then there is the intuitive method, which is easier, at least in that it can introduce a slight element of self-expression. The first chapter of this volume is a clear manifestation of the self-expressive mode. It comprises the thoughts of one of this volume’s editors and contains a creative series of proposals concerning both the logic and the derivations of employing evolutionary theory as a basis for generating personality attributes, personality being the initial topic of the two major subjects that compose this fifth volume of the 12-volume Handbook of Psychology.

Chapters 1 and 2 of this book are subsumed under the general heading of contexts. The thought here is that both personality and social psychology, broad though they may be in their own right, should be seen as components of even wider fields of study, namely evolution and culture.

Evolution provides a context that relates to the processes of the time dimension, that is, the sequences and progressions of nature over the history of life on earth. Evolutionary theory generates a constellation of phylogenetic principles representing those processes that have endured and continue to undergird the ontogenetic development and character of human functioning. As such, these principles may guide more effective thinking about which functions of personality are likely to have been—and to persist to be—the most relevant in our studies. Similarly, culture provides a context that relates to the structure and processes of the space dimension, that is, the larger configuration of forces that surround, shape, and give meaning to the events that operate in the more immediate social psychological sphere. The study of culture may explicate the wide constellation of influences within which social behaviors are immersed and that ever so subtly exert direction, transform, and control and regulate even the most prosaic events of ordinary social communications and relationships. A few additional words should be said in elaboration of these two contextual chapters.

Admittedly theoretical and speculative, the paper by Theodore Millon outlines several of what he has deduced as the universal polarities of evolution: first, the core aims of existence, in which the polarities of life preservation are contrasted with life enhancement; second, life’s fundamental modes of adaptation, counterposing ecologic accommodation and ecologic modification; third, the major strategies of species replication, setting reproductive nurturance in opposition to reproductive propagation; and fourth, a distinctly human polarity, that of predilections of abstraction, composed of comparative sources of information and their transformational processes. Millon spells out numerous personality implications of these polarities and articulates sources of support from a wide range of psychological
The next set of eight chapters of the volume represent the creative and reflective thinking of many of our most notable theoretical contributors to personology. They range from the genetic and biologic to the interpersonal and factorial. Each contributor is a major player in contemporary personality thought and research.

Before we proceed, a few words should be said concerning the current status of personologic theory. As he wrote in a 1990 book, *Toward a New Personology*, the first editor of this volume commented that the literature of the 1950s and 1960s was characterized by egregious attacks on the personality construct—attacks based on a rather facile and highly selective reading of then-popular research findings. And with the empirical grounding of personality in question and the consequential logic of personologic coherence and behavioral consistency under assault, adherents of the previously valued integrative view of personality lost their avant garde respectability and gradually withdrew from active publication. Personality theory did manage to weather these mettle-some assaults, and it began what proved to be a wide-ranging resurgence in the 1970s. By virtue of time, thoughtful reflection, and, not the least, disenchantment with proposed alternatives such as behavioral dogmatism and psychiatric biochemistry, the place of the personality construct rapidly regained its formal solid footing. The alternatives have justly faded to a status consonant with their trivial character, succumbing under the weight of their clinical inefficacy and scholarly boredom. By contrast, a series of widely acclaimed formulations were articulated by a number of contemporary psychological, psychoanalytic, interpersonal, cognitive, factorial, genetic, social, neurobiologic, and evolutionary theorists. It is to these theorists and their followers that we turn next.

Bringing the primitive and highly speculative genetic thought of the early twentieth century up to date by drawing on the technologies of the recent decade, W. John Livesley, Kerry L. Jang, and Philip Anthony Vernon articulate a convincing rationale for formulating personality concepts and their structure on the basis of trait-heritability studies. In a manner similar to Millon, who grounds his personologic concepts on the basis of a theory of evolutionary functions, Livesley et al. argue that genetic research provides a fundamental grounding for deriving complex trait constellations; these two biologically anchored schemas may ultimately be coordinated through future theoretical and empirical research. The authors contend that most measures of personality reflect heritable components and that the phenotypic structure of personality will ultimately resemble the pattern of an underlying genetic architecture. They assert, further, that etiologic criteria such as are found in genetics can offer a more objective basis for appraising personologic structure than can psychometrically based phenotypic analyses. Moreover, they believe that the interaction of multiple genetic factors will fully account for the complex patterns of trait covariances and trait clusters.

Continuing the thread of logic from evolution to genetics to the neurochemical and physiological, Marvin Zuckerman traces the interplay of these biologically based formulations to their interaction with the environment and the generation of learned behavioral traits. Writing in the spirit of Edward Wilson’s concept of *consilience* and its aim of bringing a measure of unity to ostensibly diverse sciences, Zuckerman spells out in considerable detail the flow or pathways undergirding four major personality trait concepts: extroversion/sociability; neuroticism/anxiety; aggression/agreeableness; and impulsivity/sensation seeking. Recognizing that detailed connections between the biological and the personological are not as yet fully developed, Zuckerman goes to great pains, nevertheless, to detail a wide range of strongly supporting evidence, from genetic twin studies to EEG and brain imaging investigations of cortical and autonomic arousal, to various indexes of brain neurochemistry.

Shifting the focus from the biological grounding of personality attributes, Robert F. Bornstein provides a thoughtful essay on both classical psychoanalytic and contemporary models of psychodynamic theory. He does record, however, that the first incarnation of psychoanalysis was avowedly biological, recognizing that Freud in 1895 set out to link psychological phenomena to then-extant models of neural functioning. Nevertheless, the course of analytic theory has
evolved in distinctly divergent directions over the past century, although recent efforts have been made to bridge them again to the challenge of modern neuroscience, as Bornstein notes. His chapter spells out core assumptions common to all models of psychoanalysis, such as classical analytic theory, neoanalytic models, object-relations theory, and self psychology, as well as contemporary integrative frameworks. Threads that link these disparate analytic perspectives are discussed, as are the key issues facing twenty-first-century analytic schemas.

No more radical a contrast with psychoanalytic models of personality can be found than in theories grounded in the logical positivism and empiricism that are fundamental to behavioral models, such as those articulated in the chapter by one of its primary exponents, Arthur W. Staats. Committed to a formal philosophical approach to theory development, Staats avers that most personality models lack formal rules of theory construction, possessing, at best, a plethora of different and unrelated studies and tests. Staats’s theory, termed psychological behaviorism, is grounded in learning principles generated originally in animal research, but more recently put into practice in human behavioral therapy. Like Clark Hull, a major second-generation behavioral thinker, he believes that all behavior is generated from the same primary laws. In his own formulations, Staats explicates a unified model of behavioral personology that is philosophically well structured and provides a program for developing diverse avenues of systematic personality research.

An innovative and dynamic framework for coordinating the cognitive, experiential, learning, and self-oriented components of personology (termed CEST) is presented in the theoretical chapter by Seymour Epstein. The author proposes that people operate through two interacting information-processing modes, one predominantly conscious, verbal, and rational, the other predominantly preconscious, automatic, and emotionally experiential. Operating according to different rules, it is asserted that the influence of the experiential system on the rational system is akin to what psychoanalysis claims for the role of the unconscious, but it is conceptualized in CEST in a manner more consistent with contemporary evolutionary and cognitive science. Epstein details the application of his CEST model for psychotherapy, notably by pointing out how the rational system can be employed to correct problems generated in the experiential system. Also discussed is the importance of designing research that fully recognizes and encompasses the interplay between these two information-processing systems.

The chapter by Charles S. Carver and Michael F. Scheier represents the current status of their decades-long thought and research on self-regulatory models of personality functioning. Anchored in a sophisticated framework of feedback schemas, the authors emphasize a major facet of personality processing, the system of goals that compose the self, how the patterns of a person’s goals are related, and the means by which persons move toward and away from their goals. As a consequence of their research, the authors have come to see that actions are managed by a different set of feedback processes than are feelings. Aspirations are recalibrated in reasonably predictable ways as a function of experience; for example, successes lead to setting higher goals, whereas failures tend to lower them. Conflicting goals often call for the suppression of once-desired goals, resulting in goal shifts, scaling back, disengagements, and, ultimately, lapses in self-control. Carver and Scheier view their goal as closely related to other contemporary schemas, such as dynamic systems theory and connectionism.

In their richly developed chapter, Aaron L. Pincus and Emily B. Ansell set out to create a new identity for interpersonal theory that recognizes its unique aspects and integrative potential. They suggest that the interpersonal perspective can serve as the basis for integrating diverse theoretical approaches to personality. Given its focus on interpersonal situations, this perspective includes both proximal descriptions of overt behavioral transactions and the covert or intrapsychic processes that mediate those transactions, including the internal mental representations of self and other. In addition to reviewing the work of the major originators (e.g., Sullivan, Leary) and contemporary thinkers in interpersonal theory (e.g., Benjamin, Kiesler), the authors believe that there continues to be a need for a more complete integration of the interpersonal perspective with motivational, developmental, object-relations, and cognitive theories of human behavior. Similarly, they argue for a further identification of those catalysts that stimulate the internalization of relational experiences into influential mental representations.

The current popularity among psychologists of various five-factor formulations of personality in contemporary research is undeniable. Despite the extensive literature in the area, these formulations have not been as thoroughly dissected, critically examined, and explicated as they are in Willem K. B. Hofstee’s chapter on the structure of personality traits. The author asserts that concepts such as personality are shaped and defined largely by the operations employed to construct them. Hence, several procedures applied under the rubric of the number five have been employed to characterize trait adjectives describing the structure and composition of the personality concept. Hofstee differentiates four operational modules that constitute the five component paradigms: The first set of operations reflects standardized self-report questionnaires; the second comprises the lexical approach
based on selections from a corpus of a language; the third relies on a linear methodology employing a principal components analysis of Likert item scales; and the fourth produces rival hierarchical and circumplex models for structuring trait information. Hofstee concludes his chapter by proposing a family of models composed of a hierarchy of generalized semicircumplexes.

 Appropriately placed at the conclusion of the social psychology section, Aubrey Immelman’s chapter comprises a synthesis of personality and social behavior. It not only examines the history of personality inquiry in political psychology but also offers a far-reaching and theoretically coherent framework for studying the subject in a manner consonant with principles in contextually adjacent fields, such as behavioral neuroscience and evolutionary ecology. Immelman provides an explicit framework for a personality-based risk analysis of political outcomes, acknowledging the role of filters that moderate the impact of personality on political performance. Seeking to accommodate a diversity of politically relevant personality characteristics, he bridges conceptual and methodological gaps in contemporary political study and specifically attempts a psychological examination of political leaders, on the basis of which he imposes a set of standards for personality-in-politics modeling.

 By way of confession, the social psychology chapters in this volume were selected for the most part after simply jotting down the first thoughts about what areas to include and who would be good candidates to write the chapters. Fortunately, subsequent scanning of a few well-known introductory texts and prior handbooks did nothing to alter those initial hunches that came so immediately and automatically to mind. For the most part, the vast majority of the chapters cover contemporary perspectives on traditional social psychological issues; however, a few introduce new, highly active areas of inquiry (e.g., justice, close relationships, and peace studies).

 At this point, it would be nice to describe the central theme, the deep structure underlying the organization of the social psychology chapters. But, as most readers know, social psychology and social behavior are too broad and varied for that kind of organization to be valid, much less useful. For the past 50 years or so, social psychology has done remarkably well examining the various aspects of social behavior with what Robert Merton termed theories of the midrange—his theory of relative deprivation being a good example.

 The social psychology chapters easily fall in to a few categories based on the nature of the issues they address. Four chapters focus on the social context of fundamental psychological structures: social cognitions, emotions, the self concept, and attitudes. These, together with the chapter on emotional psychology, provide a natural introduction to the social processes and interpersonal dynamics that follow.

 In the chapter on social cognition, Galen V. Bodenhausen, C. Neil Macrae, and Kurt Hugenberg, point out that the substance of the chapter contains an excellent review of the available literature describing the types of mental representations that make up the content of social cognitions; how various motives and emotions influence those cognitions; and the recent very exciting work on the nature, appearance, and consequences of automatic as well as more thoughtfully controlled processes. This chapter would be an excellent place for someone to get an overview of the best that is now known about the cognitive structures and processes that shape understanding of social situations and mediate behavioral reactions to them.

 No less fundamental are the questions of the sources of people’s emotions and how they influence behavior. The chapter by José-Miguel Fernández-Dols and James A. Russell provides a review of the theories and empirical evidence relevant to the two basic approaches to emotions and affect: as modular products of human evolutionary past and as script-like products of human cultural history. Whether one fully accepts their highly creative and brave integration of these two approaches employing the concept of core affect, their lucid description of the best available evidence together with their astute analytic insights will be well worth the reader’s time and effort. In addition, it would be remarkably easy to take their integrative theoretical model as the inspiration, or at least starting point, for various lines of critically important research.

 Roy F. Baumeister and Jean M. Twenge clearly intend that their readers fully appreciate their observation that the self-concept is intrinsically located in a social processes and interpersonal relations. In fact, as they state, the self is constructed and maintained as a way of connecting the individual organism to other members of the species. It would be easy to view this as a contemporary example of teleological theorizing (i.e., explaining structures and processes referring to a functional purpose); however, the authors go to considerable length to provide evidence explicitly describing the underlying dynamics. This includes issues such as belongingness, social exclusion, and ostracism, as well as the more familiar concerns with conformity and self-esteem. The authors make a good case for their proposition that one of the self’s crucial defining functions is to enable people to live with other people in harmony and mutual belongingness.

 The notion that people walk around with predispositions to think, feel, and act with regard to identifiable aspects of their world has a long and noble tradition in social
between high- and low-effort processes of attitude change. From the target person and those that engage the target’s processes that involve relatively automatic low-effort reactions to the evidence and theories relevant to changing explicit attitudes. After a relatively brief discussion of the current structure of attitudes, the dimensions on which they differ, how they are formed and related to beliefs and values, and their functions in social relations and behavior. Of particular importance is the identification of those issues and questions that should be addressed in future research. For example, the evidence for the distinction between implicit and explicit attitudes opens up several areas worthy of investigation.

Ever since the seminal work of Barker and his colleagues, social psychologists have recognized the importance of considering the built environments as well as sociocultural contexts in arriving at an adequate understanding of human thought, feelings, and actions. In their chapter on environmental psychology, Gabriel Moser and David Uzzell adopt the idea exemplified in Barker’s early field research that psychologists must recognize that the environment is a critical factor if they are to understand how people function in the real world. As Moser and Uzzell demonstrate, much has been discovered about the environment-person relationship that falls nicely within the context created by that early work. The authors note that not only do environmental psychologists work in collaboration with other psychologists to understand the processes mediating these relationships, but they also find themselves in collaborative efforts with other disciplines, such as architects, engineers, landscape architects, urban planners, and so on. The common focus, of course, consists of the cognitions, attitudes, emotions, self-concepts, and actions of the social participants.

The next chapters consider the dynamics involved in interpersonal and social processes that lead to changes in people’s attitudes and social behavior.

Recognizing the important distinction between implicit and explicit attitudes, in their chapter on persuasion and attitude change Richard E. Petty, S. Christian Wheeler, and Zakary L. Tormala report that as yet there is no way to change implicit attitudes. Their main contribution consists of presenting the evidence and theories relevant to changing explicit attitudes. After a relatively brief discussion of the currently influential elaboration likelihood model, their chapter is organized around the important distinction between processes that involve relatively automatic low-effort reactions from the target person and those that engage the target’s thoughts and at times behavioral reactions. The distinction between high- and low-effort processes of attitude change provides a comfortable and rather meaningful framework for organizing processes as seemingly disparate as affective priming, heuristic-based reactions, role playing, dissonance, information integration, and so on.

Andrzej Nowak, Robin R. Vallacher, and Mandy E. Miller’s chapter on social influence and group dynamics has several noteworthy features, one of which is the range of material that they have included. The chapter is so nicely composed and lucidly written that the reader may not easily appreciate the wide range of material, both theory and evidence, that is being covered. For example, the chapter begins with the more traditionally familiar topics such as obedience and reactance, moves on to what is known about more explicit efforts to influence people’s behavior, and then addresses the interpersonal processes associated with group pressure, polarization, and social loafing. All that is pretty familiar to most psychologists. However, the authors finally arrive at the most recent theoretical perspectives involving cellular automata that naturally lend themselves to the use of computer simulations to outline the implicit axiomatic changes in complex systems. What an amazing trip in both theories and method! Is it possible that what the authors identify as the press for higher order coherence provides a coherent integration of the entire social influence literature?

The transition from these initial chapters to those remaining can be roughly equated with the two dominant concerns of social psychologists. Up to this point, the chapters were most concerned with basic social psychological processes: scientific understanding of the interpersonal processes and social behavior. The remaining chapters exemplify social psychologists’ desire to find ways to make the world a better place, where people treat each other decently or at least are less cruel and destructive. Three of these chapters consider the social motives and processes that are involved in people helping and being fair to one another, whereas the last three examine harmful things that can happen between individuals and social groups, ranging from acts of prejudice to open warfare. The last chapter offers an introduction to what is now known about achieving a peaceful world.

In their chapter on altruism and prosocial behavior, C. Daniel Batson and Adam A. Powell offer a most sophisticated analysis of the relevant social psychological literature. On the basis of his research and theoretical writings, Batson is the most cited and respected psychological expert on prosocial behavior. In this chapter he discusses the evidence for four sources of prosocial behavior. After providing an analysis of the sources of these prosocial motives—enlightened self-interest, altruism, principalism, and collectivism—he then takes on the task of discussing the points of possible conflict and cooperation among them. One might
argue with his evidence for the ease with which the principalist motives—justice and fairness—can be corrupted by self-interest, and thus his conclusion is that prosocial behavior can be most reliably based on altruistic (i.e., empathy-based) motives. I suspect, however, that Kurt Lewin would have been very pleased with this highly successful example of the potential societal value of good social psychological theory.

Leo Montada, in the chapter on justice, equity, and fairness in human relations, provides a very content-rich but necessarily selective review of what is known about how justice appears in people’s lives, the various aspects of justice, and their social and individual sources, as well as interpersonal consequences. At the same time that he leads the reader through a general survey of the justice literature, he provides the reader with highly sophisticated insights and critical analyses. It is clear from the outset of this chapter that Montada is a thoroughly well-informed social scientist approaching one of the fundamental issues in human relations: how and why people care about justice in their lives, what forms that concern takes, and how important those are concerns in shaping how they treat one another.

Margaret S. Clark and Nancy K. Grote’s chapter can be viewed as the integration of several literatures associated with close relationships, friendships, and marriages—romantic and familial. They focus on the social-psychological processes associated with “good relationships”: those that they define as fostering members’ well-being. This chapter provides the most recent developments in Clark’s important distinction between communal and exchange relationships and includes the report of an important longitudinal study examining the relationship between conflict and fairness in close relationships. They find that conflict in a relationship leads to increased concern with issues of fairness that then lead the participants even further from the important communal norms based on mutual concern for one another’s welfare.

Kenneth L. Dion’s chapter on prejudice, racism, and discrimination looks at various aspects of the darker side of interpersonal relations. In the first section of the chapter, Dion leads the reader to a very thoughtful and complete review of the various explanations for prejudice, racism, and discrimination. Beginning with the classic and contemporary versions of the authoritarian personality theories, he discusses just-world, belief congruence, and ambivalence literatures. Dion does a masterful job of leading the reader through the more recently developed distinction between automatic and controlled processes, as well as social dominance theory and multicomponent approaches to intergroup attitudes. But that is only the beginning. Reflecting his own earlier research interests, Dion devotes the second section of his chapter to the psychology of the victim of prejudice and discrimination. This section integrates the most recent findings in this highly active and productive area of inquiry. Dion describes the research that has given the familiar self-fulfilling prophecy notion in social psychology new meaning and has provided compelling new insights into the very important ways victims respond to their unfair treatment.

The chapter by John F. Dovidio, Samuel L. Gaertner, Victoria M. Esses and Marilynn B. Brewer examines the social-psychological processes involved in interpersonal and intergroup relations. This includes both the sources of social conflict and those involved in bringing about harmony and integration. The origins of the important work reported in this chapter can be traced to the initial insights of European social psychologists who recognized that when people think in terms of “we” rather than “I,” there is a strong tendency also to react in terms of “us” versus “them” (i.e., in-group vs. out-group). The consequences, of course, include favoring members of the in-group and discriminating against members of the salient out-groups. After describing what is known about the psychological processes involved in these biased reactions, the authors then consider those processes that can preclude or overcome those destructive biases and promote harmony and social integration.

Joseph de Rivera’s chapter takes a similar path, by first focusing on those social-psychological processes involved in aggression and violence, and then with that as background presenting his recommendations concerning how positive peace can be promoted. For de Rivera this does not simply mean an absence of open conflict, but rather a benevolent and supportive environment, as well as societal norms, that promote individual processes involving harmony and well-being. In describing the various means for generating a global culture of peace, he also makes the case for the importance of individual’s personal transformation in creating and maintaining a culture of peace. De Rivera offers the reader a highly sophisticated use of the social-psychological research and theory to arrive at specific recommendations for solving, arguably, the most important issues of our lives: the achievement of a peaceful, caring, nurturing social environment. Ambitious? Yes. But de Rivera generates the framework of his own perspective out of the best of what social science has to offer.

We trust the readers of this volume on personality and social psychology will find the chapters it contains to be both provocative and illuminating. It has been an honor and a joy to edit a book written by so many able, inspiring, and cooperative authors, whom we thank personally for their thoughtful and stimulating contributions.

THEODORE MILLON

MELVIN J. LERNER
Contents

Handbook of Psychology  Preface    vii
Irving B. Weiner

Volume Preface    ix
Theodore Millon and Melvin J. Lerner

Contributors    xvii

PART ONE
CONTEXTS

1  EVOLUTION: A GENERATIVE SOURCE FOR CONCEPTUALIZING
THE ATTRIBUTES OF PERSONALITY    3
Theodore Millon

2  CULTURAL PERSPECTIVES ON PERSONALITY AND SOCIAL PSYCHOLOGY    31
Joan G. Miller and Lynne Schaberg

PART TWO
PERSONALITY

3  GENETIC BASIS OF PERSONALITY STRUCTURE    59
W. John Livesley, Kerry L. Jang, and Philip A. Vernon

4  BIOLOGICAL BASES OF PERSONALITY    85
Marvin Zuckerman

5  PSYCHODYNAMIC MODELS OF PERSONALITY    117
Robert F. Bornstein

6  A PSYCHOLOGICAL BEHAVIORISM THEORY OF PERSONALITY    135
Arthur W. Staats

7  COGNITIVE-EXPERIENTIAL SELF-THEORY OF PERSONALITY    159
Seymour Epstein

8  SELF-REGULATORY PERSPECTIVES ON PERSONALITY    185
Charles S. Carver and Michael F. Scheier

9  INTERPERSONAL THEORY OF PERSONALITY    209
Aaron L. Pincus and Emily B. Ansell
# PART THREE

## SOCIAL PSYCHOLOGY

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>SOCIAL COGNITION</td>
<td>257</td>
</tr>
<tr>
<td></td>
<td>Galen V. Bodenhausen, C. Neil Macrae, and Kurt Hugenberg</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>EMOTION, AFFECT, AND MOOD IN SOCIAL JUDGMENTS</td>
<td>283</td>
</tr>
<tr>
<td></td>
<td>José-Miguel Fernández-Dols and James A. Russell</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>ATTITUDES IN SOCIAL BEHAVIOR</td>
<td>299</td>
</tr>
<tr>
<td></td>
<td>James M. Olson and Gregory R. Maio</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>THE SOCIAL SELF</td>
<td>327</td>
</tr>
<tr>
<td></td>
<td>Roy F. Baumeister and Jean M. Twenge</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>PERSUASION AND ATTITUDE CHANGE</td>
<td>353</td>
</tr>
<tr>
<td></td>
<td>Richard E. Petty, S. Christian Wheeler, and Zakary L. Tormala</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>SOCIAL INFLUENCE AND GROUP DYNAMICS</td>
<td>383</td>
</tr>
<tr>
<td></td>
<td>Andrzej Nowak, Robin R. Vallacher, and Mandy E. Miller</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>ENVIRONMENTAL PSYCHOLOGY</td>
<td>419</td>
</tr>
<tr>
<td></td>
<td>Gabriel Moser and David Uzzell</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>CLOSE RELATIONSHIPS</td>
<td>447</td>
</tr>
<tr>
<td></td>
<td>Margaret S. Clark and Nancy K. Grote</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>ALTRUIISM AND PROSOCIAL BEHAVIOR</td>
<td>463</td>
</tr>
<tr>
<td></td>
<td>C. Daniel Batson and Adam A. Powell</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>SOCIAL CONFLICT, HARMONY, AND INTEGRATION</td>
<td>485</td>
</tr>
<tr>
<td></td>
<td>John F. Dovidio, Samuel L. Gaertner, Victoria M. Esses, and Marilynn B. Brewer</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>PREJUDICE, RACISM, AND DISCRIMINATION</td>
<td>507</td>
</tr>
<tr>
<td></td>
<td>Kenneth L. Dion</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>JUSTICE, EQUITY, AND FAIRNESS IN HUMAN RELATIONS</td>
<td>537</td>
</tr>
<tr>
<td></td>
<td>Leo Montada</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>AGGRESSION, VIOLENCE, EVIL, AND PEACE</td>
<td>569</td>
</tr>
<tr>
<td></td>
<td>Joseph de Rivera</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>PERSONALITY IN POLITICAL PSYCHOLOGY</td>
<td>599</td>
</tr>
<tr>
<td></td>
<td>Aubrey Immelman</td>
<td></td>
</tr>
</tbody>
</table>

### Author Index

<table>
<thead>
<tr>
<th>Author</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Subject Index

<table>
<thead>
<tr>
<th>Subject</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
PART ONE

CONTEXTS
In the last year of the twentieth century, voters elected a group of Kansas school board members who supported the removal of the concept of evolution from the state’s science curriculum, an act that indicated the extent to which evolutionary ideas could incite intense emotional, if not irrational opposition on the part of unenlightened laymen. Retrospectively appalled by their prior action, in the following year Kansan voters rescinded their perverse judgment and chose new board members who intended to restore the concept.

The theory of evolution was reinstated not because the electors of Kansas, a most conservative and religious state, suddenly became agnostic, but because they realized that rejecting the idea would deny their children the necessity of remaining in touch with one of the fundamentals of modern science; they realized that this could, in effect, allow their children to fall behind, to be bereft of a basic science, and to be both a misinformed and misguided generation. Their children could become embarrassingly backward in a time of rapidly changing technology.

Might not the same ambivalence be true of our own field, one composed of ostensibly sophisticated and knowledgeable scientists? Might we not be so deeply mired in our own traditions (scholarly religions?) that we are unable to free ourselves from the habit of seeing our subject from no vantage point other than those to which we have become accustomed? Are we unable to recognize that behavior, cognition, the unconscious, personality—all of our traditional subjects—are merely diverse manifestations of certain common and deeper principles of functioning, processes, and mechanisms that have evolved either randomly or adaptively through history and time? Do we psychologists have a collective phobia about laws that may represent the fundamental origins of our traditional subjects? Does the search for and application of such laws push our emotional buttons, perhaps run hard against our habitual blinders, so much so as to prevent us from recognizing their value as a potential generative source that may more fully illuminate our science?

PERSONOLOGY’S RELATIONSHIP TO OTHER SCIENCES

It is the intent of this chapter to broaden our vistas, to furnish both a context and a set of guiding ideas that may enrich our studies. I believe it may be wise and perhaps even necessary to go beyond our current conceptual boundaries in psychology, more specifically to explore carefully reasoned, as well as intuitive hypotheses that draw their laws and principles if not their substance from contextually adjacent sciences such as evolution. Not only may such steps bear new conceptual fruits, but they may also provide a foundation that can undergird and guide our own discipline’s explorations. Much of personology, no less psychology as a whole, remains adrift, divorced from broader spheres of scientific knowledge; it is isolated from firmly grounded if not universal principles, leading us to continue building the patchwork quilt of
concepts and data domains that characterize our field. Preoccupied with but a small part of the larger puzzle of nature or fearful of accusations of reductionism, we may fail to draw on the rich possibilities to be found in parallel realms of scientific pursuit. With few exceptions, cohering concepts that would connect our subject domain to those of its sister sciences in nature have not been adequately developed.

It appears to me that we have become trapped in (obsessed with?) horizontal refinements. A search for integrative schemas and cohesive constructs that link its seekers closely to relevant observations and laws developed in other scientific fields is needed. The goal—albeit a rather ambitious one—is to refashion our patchwork quilt of concepts into a well-tailored and aesthetically pleasing tapestry that interweaves the diverse forms in which nature expresses itself (E. O. Wilson, 1998).

What sphere is there within the psychological sciences more apt than personology to undertake the synthesis of nature? Persons are the only organically integrated system in the psychological domain, evolved through the millennia and inherently created from birth as natural entities rather than culture-bound and experience-derived gestalts. The intrinsic cohesion of nature’s diverse elements that inheres in persons is not a rhetorical construction, but rather an authentic substantive unity. Personological features may often be dissonant and may be partitioned conceptually for pragmatic or scientific purposes, but they are segments of an inseparable physicochemical, biopsychosocial entity.

To take this view is not to argue that different spheres of scientific inquiry must be collapsed or even equated, but rather that there may be value in seeking a single, overarching conceptual system that interconnects ostensibly diverse subjects such as physics, biology, and psychology (Millon, 1990; E. O. Wilson, 1998). Arguing in favor of establishing explicit links between these domains calls for neither a reductionistic philosophy, nor a belief in substantive identicity, nor efforts to so fashion the links by formal logic. Rather, one should aspire to their substantive concordance, empirical consistency, conceptual interfacing, convergent dialogues, and mutual enlightenment.

A few words should be said concerning the undergirding framework used to structure an evolutionary context for a personology model. Parallel schemas are almost universally present in the literature; the earliest may be traced to mid-nineteenth-century philosophers, most notably Spencer (1855) and Haeckel (1874). More modern but equally speculative systems have been proposed by keen and broadly informed observers such as Edward Wilson (1975), Cosmides and Tooby (1987, 1989) and M. Wilson and Daley (1992), as well as by empirically well-grounded methodologists, such as Symons (1979, 1992) and D. M. Buss (1989, 1994). Each of their proposals fascinates either by virtue of its intriguing portrayals or by the compelling power of its logic or its data. Their arguments not only coordinate with but also are anchored to observations derived specifically from principles of modern physical and biological evolution. It is these underpinnings of knowledge on which the personological model presented in this chapter has been grounded and from which a deeper and clearer understanding may be obtained concerning the nature of both normal and pathological personality functioning.

On the Place of Theory in Personology

The following discussion is conjectural, if not overly extended in its speculative reach. In essence, it seeks to explain the structure and styles of personality with reference to deficient, imbalanced, or conflicted modes of evolutionary survival, ecological adaptation, and reproductive strategy. Whatever one’s appraisal of these conjectures, the model that follows may best be approached in the spirit in which it was formulated—an effort to provide a context for explicating the domains of personological science in the hope that it can lead to a clearer understanding of our subject. All sciences have organizing principles that not only create order but also provide the basis for generating hypotheses and stimulating new knowledge. A contextual theory not only summarizes and incorporates extant knowledge, but is heuristic—that is, it has “systematic import,” as Hempel (1965) has phrased it, in that it may originate and develop new observations and new methods.

It is unfortunate that the number of theories that have been advanced to “explain” personality is proportional to the internecine squabbling found in the literature. However, and ostensibly toward the end of pragmatic sobriety, those of an antitheory bias have sought to persuade the profession of the failings of premature formalization, warning that one cannot arrive at the desired future by lifting science by its own bootstraps. To them, there is no way to traverse the road other sciences have traveled without paying the dues of an arduous program of empirical research. Formalized axiomatics, they say, must await the accumulation of so-called hard evidence that is simply not yet in. Shortcutting the route with ill-timed systematics, they claim, will lead us down primrose paths, preoccupying attentions as we wend fruitlessly through endless detours, each of which could be averted by our holding fast to an empiricist philosophy and methodology.

No one argues against the view that theories that float, so to speak, on their own, un concerned with the empirical domain, should be seen as the fatuous achievements they are and the travesty they make of the virtues of a truly coherent conceptual system. Formal theory should not be pushed far beyond the data, and its derivations should be linked at all points to
established observations. However, a theoretical framework can be a compelling instrument for coordinating and giving consonance to complex and diverse observations—if its concepts are linked to relevant facts in the empirical world. By probing beneath surface impressions to inner structures and processes, previously isolated facts and difficult-to-fathom data may yield new relationships and expose clearer meanings. Scientific progress occurs when observations and concepts elaborate and refine previous work. However, this progression does not advance by brute empiricism alone, by merely piling up more descriptive and more experimental data. What is elaborated and refined in theory is understanding, an ability to see relations more plainly, to conceptualize categories and dimensions more accurately, and to create greater overall coherence in a subject—to integrate its elements in a more logical, consistent, and intelligible fashion.

A problem arises when introducing theory into the study of personality. Given our intuitive ability to “sense” the correctness of a psychological insight or speculation, theoretical efforts that impose structure or formalize these insights into a scientific system will often be perceived as not only cumbersome and intrusive, but alien as well. This discomfiture and resistance does not arise in fields such as particle physics, in which everyday observations are not readily available and in which innovative insights are few and far between. In such subject domains, scientists not only are quite comfortable, but also turn readily to deductive theory as a means of helping them explicate and coordinate knowledge. It is paradoxical but true and unfortunate that personologists learn their subject quite well merely by observing the ordinary events of life. As a consequence of this ease, personologists appear to shy from and hesitate placing trust in the obscure and complicating, yet often fertile and systematizing powers inherent in formal theory, especially when a theory is new or different from those learned in their student days.

Despite the shortcomings in historic and contemporary theoretical schemas, systematizing principles and abstract concepts can “facilitate a deeper seeing, a more penetrating vision that goes beyond superficial appearances to the order underlying them” (Bowers, 1977). For example, pre-Darwinian taxonomists such as Linnaeus limited themselves to apparent similarities and differences among animals as a means of constructing their categories. Darwin was not seduced by appearances. Rather, he sought to understand the principles by which overt features came about. His classifications were based not only on descriptive qualities, but also on explanatory ones.

On the Place of Evolutionary Theory in Personology

It is in both the spirit and substance of Darwin’s explanatory principles that the reader should approach the proposals that follow. The principles employed are essentially the same as those that Darwin developed in seeking to explicate the origins of species. However, they are listed to derive not the origins of species, but rather the structure and style of personalities that have previously been generated on the basis of clinical observation alone. Aspects of these formulations have been published in earlier books (Millon, 1969, 1981, 1986, 1990; Millon & Davis, 1996); they are anchored here, however, explicitly to evolutionary and ecological theory. Identified in earlier writings as a biosocial learning model for personality and psychopathology, the theory we present seeks to generate the principles, mechanisms, and typologies of personality through formal processes of deduction.

To propose that fruitful ideas may be derived by applying evolutionary principles to the development and functions of personological traits has a long (if yet unfulfilled) tradition. Spencer (1870), Huxley (1870), and Haeckel (1874) offered suggestions of this nature shortly after Darwin’s seminal Origins was published. The school of functionalism, popular in psychology in the early part of this century, likewise drew its impetus from evolutionary concepts as it sought to articulate a basis for individual difference typologies (McDougall, 1932).

In recent decades, numerous evolution-oriented psychologists and biologists have begun to explore how the human mind may have been shaped over the past million years to solve the problems of basic survival, ecological adaptation, and species replication and diversification. These well-crafted formulations are distinctly different from other, more traditional models employed to characterize human functioning.

The human mind is assuredly sui generis, but it is only the most recent phase in the long history of organic life. Moreover, there is no reason to assume that the exigencies of life have differed in their essentials among early and current species. It would be reasonable, therefore—perhaps inevitable—that the study of the functions of mind be anchored to the same principles that are universally found in evolution’s progression. Using this anchor should enable us to build a bridge between the human mind and all other facets of natural science; moreover, it should provide a broad blueprint of why the mind engages in the functions it does, as well as what its essential purposes may be, such as pursuing parental affection and protection, exploring the rationale and patterns of sexual mating, and specifying the styles of social communication and abstract language.

In recent times we have also seen the emergence of sociobiology, a new science that has explored the interface between human social functioning and evolutionary biology (E. O. Wilson, 1975, 1978). The common goal among both sociobiological and personological proposals is the desire not only to apply analogous principles across diverse scientific realms, but also to reduce the enormous range of behavioral
and trait concepts that have proliferated through modern history. This goal might be achieved by exploring the power of evolutionary theory to simplify and order previously disparate personological features. For example, all organisms seek to avoid injury, find nourishment, and reproduce their kind if they are to survive and maintain their populations. Each species displays commonalities in its adaptive or survival style. Within each species, however, there are differences in style and differences in the success with which its various members adapt to the diverse and changing environments they face. In these simplest of terms, differences among personality styles would be conceived as representing the more-or-less distinctive ways of adaptive functioning that an organism of a particular species exhibits as it relates to its typical range of environments. Disorders of personality, so formulated, would represent particular styles of maladaptive functioning that can be traced to deficiencies, imbalances, or conflicts in a species’ capacity to relate to the environments it faces.

A few additional words should be said concerning analogies between evolution and ecology on the one hand and personality on the other. During its life history, an organism develops an assemblage of traits that contribute to its individual survival and reproductive success, the two essential components of fitness formulated by Darwin. Such assemblages, termed complex adaptations and strategies in the literature of evolutionary ecology, are close biological equivalents to what psychologists have conceptualized as personality styles and structures. In biology, explanations of a life history strategy of adaptations refer primarily to biogenic variations among constituent traits, their overall covariance structure, and the nature and ratio of favorable to unfavorable ecological resources that have been available for purposes of extending longevity and optimizing reproduction. Such explanations are not appreciably different from those used to account for the development of personality styles or functions.

Bypassing the usual complications of analogies, a relevant and intriguing parallel may be drawn between the phylogenetic evolution of a species’ genetic composition and the ontogenic development of an individual organism’s adaptive strategies (i.e., its personality style, so to speak). At any point in time, a species possesses a limited set of genes that serve as trait potentials. Over succeeding generations, the frequency distribution of these genes will likely change in their relative proportions depending on how well the traits they undergird contribute to the species’ “fittedness” within its varying ecological habitats. In a similar fashion, individual organisms begin life with a limited subset of their species’ genes and the trait potentials they subserve. Over time the salience of these trait potentials—not the proportion of the genes themselves—will become differentially prominent as the organism interacts with its environments. It “learns” from these experiences which of its traits fit best (i.e., most optimally suit its ecosystem). In phylogenesis, then, actual gene frequencies change during the generation-to-generation adaptive process, whereas in ontogenesis it is the salience or prominence of gene-based traits that changes as adaptive learning takes place. Parallel evolutionary processes occur—one within the life of a species, and the other within the life of an organism. What is seen in the individual organism is a shaping of latent potentials into adaptive and manifest styles of perceiving, feeling, thinking, and acting; these distinctive ways of adaptation, engendered by the interaction of biological endowment and social experience, comprise the elements of what is termed personality styles. It is a formative process in a single lifetime that parallels gene redistributions among species during their evolutionary history.

Two factors beyond the intrinsic genetic trait potentials of advanced social organisms have a special significance in affecting their survival and replicability. First, other members of the species play a critical part in providing postnatal nurturing and complex role models. Second, and no less relevant, is the high level of diversity and unpredictability of their ecological habitats. This requires numerous, multifaceted, and flexible response alternatives that are either preprogrammed genetically or acquired subsequently through early learning. Humans are notable for unusual adaptive pliancy, acquiring a wide repertoire of styles or alternate modes of functioning for dealing with both predictable and novel environmental circumstances. Unfortunately, the malleability of early potentials for diverse learnings diminishes as maturation progresses. As a consequence, adaptive styles acquired in childhood and usually suitable for comparable later environments become increasingly immutable, resisting modification and relearning. Problems arise in new ecological settings when these deeply ingrained behavior patterns persist, despite their lessened appropriateness; simply stated, what was learned and was once adaptive may no longer fit. Perhaps more important than environmental diversity, then, is the divergence between the circumstances of original learning and those of later life, a schism that has become more problematic as humans have progressed from stable and traditional to fluid and inconstant modern societies.

From the viewpoint of survival logic, it is both efficient and adaptive either to preprogram or to train the young of a species with traits that fit the ecological habitats of their parents. This wisdom rests on the usually safe assumption that consistency if not identicality will characterize the ecological conditions of both parents and their offspring. Evolution is spurred when this continuity assumption fails to hold—when
formerly stable environments undergo significant change. Radical shifts of this character could result in the extinction of a species. It is more typical, however, for environments to be altered gradually, resulting in modest, yet inexorable redistributions of a species’ gene frequencies. Genes that subserve competencies that proved suited to the new conditions become proportionately more common; ultimately, the features they engender come to typify either a new variant of or a successor to the earlier species.

All animal species intervene in and modify their habitats in routine and repetitive ways. Contemporary humans are unique in evolutionary history, however, in that both the physical and social environment has been altered in precipitous and unpredictable ways. These interventions appear to have set in motion consequences not unlike the “equilibrium punctuations” theorized by modern paleontologists (Eldredge & Gould, 1972). This is best illustrated in the origins of our recent borderline personality epidemic (Millon, 1987):

Central to our recent culture have been the increased pace of social change and the growing pervasiveness of ambiguous and discordant customs to which children are expected to subscribe. Under the cumulative impact of rapid industrialization, immigration, urbanization, mobility, technology, and mass communication, there has been a steady erosion of traditional values and standards. Instead of a simple and coherent body of practices and beliefs, children find themselves confronted with constantly shifting styles and increasingly questioned norms whose durability is uncertain and precarious. Few times in history have so many children faced the tasks of life without the aid of accepted and durable traditions. Not only does the strain of making choices among discordant standards and goals beset them at every turn, but these competing beliefs and divergent demands prevent them from developing either internal stability or external consistency. (p. 363)

Murray has said that “life is a continuous procession of explorations... learnings and relearnings” (1959). Yet, among species such as humans, early adaptive potentials and pliancies may fail to crystallize because of the fluidities and inconsistencies of the environment, leading to the persistence of what some have called immature and unstable styles that fail to achieve coherence and effectiveness.

Lest the reader assume that those seeking to wed the sciences of evolution and ecology find themselves fully welcome in their respective fraternities, there are those who assert that “despite pious hopes and intellectual convictions, these two disciplines have so far been without issue” (Lewontin, 1979). This judgment is now both dated and overly severe, but numerous conceptual and methodological impediments do face those who wish to bring these fields of biological inquiry into fruitful synthesis—no less employing them to construe the styles of personality. Despite such concerns, recent developments bridging ecological and evolutionary theory are well underway, and hence do offer some justification for extending their principles to human styles of adaptation.

To provide a conceptual background from these sciences and to furnish a rough model concerning the styles of personality, four domains or spheres of evolutionary and ecological principles are detailed in this chapter. They are labeled existence, adaptation, replication, and abstraction. The first relates to the serendipitous transformation of random or less organized states into those possessing distinct structures of greater organization; the second refers to homeostatic processes employed to sustain survival in open ecosystems; the third pertains to reproductive styles that maximize the diversification and selection of ecologically effective attributes; and the fourth, a distinctly human phenomenon, concerns the emergence of competencies that foster anticipatory planning and reasoned decision making.

What makes evolutionary theory and ecological theory as meritorious as I propose them to be? Are they truly coextensive with the origins of the universe and the procession of organic life, as well as human modes of adaptation? Is extrapolation to personality a conjectural fantasy? Is there justification for employing them as a basis for understanding normal and pathological behaviors?

Owing to the mathematical and deductive insights of our colleagues in physics, we have a deeper and clearer sense of the early evolution and structural relations among matter and energy. So too has knowledge progressed in our studies of physical chemistry, microbiology, evolutionary theory, population biology, ecology, and ethology. How odd it is (is it not?) that we have only now again begun to investigate—as we did at the turn of the last century—the interface between the basic building blocks of physical nature and the nature of life as we experience and live it personally. How much more is known today, yet how hesitant are people to undertake a serious rapprochement? As Barash (1982) has commented:

Like ships passing in the night, evolutionary biology and the social sciences have rarely even taken serious notice of each other, although admittedly, many introductory psychology texts give an obligatory toot of the Darwinian horn somewhere in the first chapter... before passing on to discuss human behavior as though it were determined only by environmental factors. (p. 7)

Commenting that serious efforts to undergird the behavioral sciences with the constructs and principles of evolutionary
biology are as audacious as they are overdue, Barash (1982) notes further:

As with any modeling effort, we start with the simple, see how far it takes us, and then either complicate or discard it as it gets tested against reality. The data available thus far are certainly suggestive and lead to the hope that more will shortly be forthcoming, so that tests and possible falsification can be carried out. In the meanwhile, as Darwin said when he first read Malthus, at least we have something to work with! (p. 8)

The role of evolution is most clearly grasped when it is paired with the principles of ecology. So conceived, the so-called procession of evolution represents a series of serendipitous transformations in the structure of a phenomenon (e.g., elementary particle, chemical molecule, living organism) that appear to promote survival in both its current and future environments. Such processions usually stem from the consequences of either random fluctuations (such as mutations) or replicative reformations (e.g., recombinant mating) among an infinite number of possibilities—some simpler and others more complex, some more and others less organized, some increasingly specialized and others not. Evolution is defined, then, when these restructurings enable a natural entity (e.g., species) or its subsequent variants to survive within present and succeeding ecological milieus. It is the continuity through time of these fluctuations and reformations that comprises the sequence we characterize as evolutionary progression.

THREE UNIVERSAL POLARITIES OF EVOLUTION

As noted in previous paragraphs, existence relates to the serendipitous transformation of states that are more ephemeral, less organized, or both into those possessing greater stability, greater organization, or both. It pertains to the formation and sustenance of discernible phenomena, to the processes of evolution that enhance and preserve life, and to the psychic polarity of pleasure and pain. Adaptation refers to homeostatic processes employed to foster survival in open ecosystems. It relates to the manner in which extant phenomena adapt to their surrounding ecosystems, to the mechanisms employed in accommodating to or in modifying these environments, and to the psychic polarity of passivity and activity. Replication pertains to reproductive styles that maximize the diversification and selection of ecologically effective attributes. It refers to the strategies utilized to replicate ephemeral organisms, to the methods of maximizing reproductive propagation and progeny nurturance, and to the psychic polarity of self and other. These three polarities have forerunners in psychological theory that may be traced back to the early 1900s.

Some Historical Notes

A number of pre–World War I theorists proposed polarities that were used as the foundation for understanding a variety of psychological processes. Although others formulated parallel schemas earlier than he, I illustrate these conceptions with reference to ideas presented by Sigmund Freud. He wrote in 1915 what many consider to be among his most seminal works, those on metapsychology and in particular, the paper entitled “The Instincts and Their Vicissitudes.” Speculations that foreshadowed several concepts developed more fully later both by himself and by others were presented in preliminary form in these papers. Particularly notable is a framework that Freud (1915/1925) advanced as central to understanding the mind; he framed these polarities as follows:

Our mental life as a whole is governed by three polarities, namely, the following antitheses:

• Subject (ego)-Object (external world)
• Pleasure-Pain
• Active-Passive

The three polarities within the mind are connected with one another in various highly significant ways. We may sum up by saying that the essential feature in the vicissitudes undergone by instincts is their subjection to the influences of the three great polarities that govern mental life. Of these three polarities we might describe that of activity-passivity as the biological, that of the ego-external world as the real, and finally that of pleasure-pain as the economic, respectively. (pp. 76–77, 83)

Preceding Freud, however, aspects of these three polarities were conceptualized and employed by other theorists—in France, Germany, Russia, and other European nations as well as in the United States. Variations of the polarities of active-passive, subject-object, and pleasure-pain were identified by Heymans and Wiersma in Holland, McDougall in the United States, Meumann in Germany, Kollarits in Hungary, and others (Millon, 1981; Millon & Davis, 1996).

Despite the central role Freud assigned these polarities, he failed to capitalize on them as a coordinated system for understanding patterns of human functioning. Although he failed to pursue their potentials, the ingredients he formulated for his tripartite polarity schema were drawn upon by his
disciples for many decades to come, seen prominently in the progressive development from instinct or drive theory, in which pleasure and pain were the major forces, to ego psychology, in which the apparatuses of activity and passivity were central constructs, and, most recently, to self-psychology and object relations theory, in which the self-other polarity is the key issue (Pine, 1990).

Forgotten as a metapsychological speculation by most, the scaffolding comprising these polarities was fashioned anew by this author in the mid-1960s (Millon, 1969). Unacquainted with Freud’s proposals at the time and employing a biosocial-learning model anchored to Skinnerian concepts, I constructed a framework similar to Freud’s “great polarities that govern all of mental life.” Phrased in the terminology of learning concepts, the model comprised three polar dimensions: positive versus negative reinforcement (pleasure-pain); self-other as reinforcement source; and the instrumental styles of active-passive. I (Millon, 1969) stated:

By framing our thinking in terms of what reinforcements the individual is seeking, where he is looking to find them and how he performs we may see more simply and more clearly the essential strategies which guide his coping behaviors.

These reinforcements [relate to] whether he seeks primarily to achieve positive reinforcements (pleasure) or to avoid negative reinforcements (pain).

Some patients turn to others as their source of reinforcement, whereas some turn primarily to themselves. The distinction [is] between others and self as the primary reinforcement source.

On what basis can a useful distinction be made among instrumental behaviors? A review of the literature suggests that the behavioral dimension of activity-passivity may prove useful. . . . Active patients [are] busily intent on controlling the circumstances of their environment. . . . Passive patients . . . wait for the circumstances of their environment to take their course . . . reacting to them only after they occur. (pp. 193–195)

Do we find parallels within the disciplines of psychiatry and psychology that correspond to these broad evolutionary polarities?

In addition to the forerunners noted previously, there is a growing group of contemporary scholars whose work relates to these polar dimensions, albeit indirectly and partially. For example, a modern conception anchored to biological foundations has been developed by the distinguished British psychologist Jeffrey Gray (1964, 1973). A three-part model of temperament, matching the three-part polarity model in most regards, has been formulated by the American psychologist Arnold Buss and his associates (Buss & Plomin 1975, 1984). Circumplex formats based on factor analytic studies of mood and arousal that align well with the polarity schema have been published by Russell (1980) and Tellegen (1985). Deriving inspiration from a sophisticated analysis of neuroanatomical substrates, the highly resourceful American psychiatrist Robert Cloninger (1986, 1987) has deduced a threefold schema that is coextensive with major elements of the model’s three polarities. Less oriented to biological foundations, recent advances in both interpersonal and psychoanalytic theory have likewise exhibited strong parallels to one or more of the three polar dimensions. A detailed review of these and other parallels has been presented in several recent books (e.g., Millon, 1990; Millon & Davis, 1996).

The following pages summarize the rationale and characteristics of the three-part polarity model. A few paragraphs draw upon the model as a basis for establishing attributes for conceptualizing personality patterns.

Aims of Existence

The procession of evolution is not limited just to the evolution of life on earth but extends to prelife, to matter, to the primordial elements of our local cosmos, and, in all likelihood, to the elusive properties of a more encompassing universe within which our cosmos is embedded as an incidental part. The demarcations we conceptualize to differentiate states such as nonmatter and matter, or inorganic and organic, are nominal devices that record transitions in this ongoing procession of transformations, an unbroken sequence of re-formed elements that have existed from the very first.

We may speak of the emergence of our local cosmos from some larger universe, or of life from inanimate matter, but if we were to trace the procession of evolution backward we would have difficulty identifying precise markers for each of these transitions. What we define as life would become progressively less clear as we reversed time until we could no longer discern its presence in the matter we were studying. So, too, does it appear to theoretical physicists that if we trace the evolution of our present cosmos back to its ostensive origins, we would lose its existence in the obscurity of an undifferentiated and unrecoverable past. The so-called Big Bang may in fact be merely an evolutionary transformation, one of an ongoing and never-ending series of transitions.

Life Preservation and Life Enhancement: The Pain-Pleasure Polarity

The notion of open systems is of relatively recent origin (Bertalanffy, 1945; Lotka, 1924; Schrodinger, 1944), brought to bear initially to explain how the inevitable consequences of the second law of thermodynamics appear to be circumvented in the biological realm. By broadening the ecological
field so as to encompass events and properties beyond the local and immediate, it becomes possible to understand how living organisms on earth function and thrive, despite seeming to contradict this immutable physical law (e.g., solar radiation, continuously transmitting its ultimately exhaustible supply of energy, temporarily counters the earth’s inevitable thermodynamic entropy). The open system concept has been borrowed freely and fruitfully to illuminate processes across a wide range of subjects. In recent decades it has been extended, albeit speculatively, to account for the evolution of cosmic events. These hypotheses suggest that the cosmos as known today may represent a four-dimensional “bubble” or set of “strings” stemming either from the random fluctuations of an open meta-universe characterized primarily by entropic chaos or of transpositions from a larger set of dimensions that comprise the properties of an open mega-universe—that is, dimensions beyond those we apprehend (Millon, 1990).

By materializing new matter from fluctuations in a larger and unstable field—that is, by creating existence from non-existence (cold dark matter)—any embedded open system might not only expand, but also form entities displaying anti-entropic structure, the future survival of which is determined by the character of parallel materializations and by the fortuitous consequences of their interactions (including their ecological balance, symbiosis, etc.). Beyond fortuitous levels of reciprocal fitness, some of these anti-entropic structures may possess properties that enable them to facilitate their own self-organization; that is to say, the forms into which they have been rendered randomly may not only survive, but also be able to amplify themselves, to extend their range, or both, sometimes in replicated and sometimes in more comprehensive structures.

Recent mathematical research in both physics and chemistry has begun to elucidate processes that characterize how structures “evolve” from randomness. Whether one evaluates the character of cosmogenesis, the dynamics of open chemical systems, or repetitive patterns exhibited among weather movements, it appears that random fluctuations assume sequences that often become both self-sustaining and recurrent. In chemistry, the theory of dissipative (free energy) structures (Prigogine 1972, 1976) proposes a principle called order through fluctuation that relates to self-organizational dynamics; these fluctuations proceed through sequences that not only maintain the integrity of the system but are also self-renewing. According to the theory, any open system may evolve when fluctuations exceed a critical threshold, setting in motion a qualitative shift in the nature of the system’s structural form. Similar shifts within evolving systems are explained in pure mathematics by what has been termed catastrophic theory (Thom, 1972); here, sudden switches from one dynamic equilibrium state to another occur instantaneously with no intervening bridge. As models portraying how the dynamics of random fluctuation drive prior levels of equilibrium to reconstitute themselves into new structures, both catastrophe and dissipative theories prove fruitful in explicating self-evolving morphogenesis—the emergence of new forms of existence from prior states.

There is another equally necessary step to existence, one that maintains “being” by protecting established structures and processes. Here, the degrading effects of entropy are counteracted by a diversity of safeguarding mechanisms. Among both physical and organic substances, such as atoms and molecules, the elements comprising their nuclear structure are tightly bound, held together by the strong force that is exceptionally resistant to decomposition (hence the power necessary to split the atom). More complicated organic structures, such as plants and animals, also have mechanisms to counter entropic dissolution—that is to say, to maintain the existence of their lives.

Two intertwined strategies are required, therefore: one to achieve existence, the other to preserve it. The aim of one is the enhancement of life—creating and then strengthening ecologically survivable organisms; the aim of the other is the preservation of life—avoiding circumstances that might terminate (entropically decompose) it. Although I disagree with Freud’s concept of a death instinct (Thanatos), I believe he was essentially correct in recognizing that a balanced yet fundamental biological bipolarity exists in nature, a bipolarity that has its parallel in the physical world. As he wrote in one of his last works, “The analogy of our two basic instincts extends from the sphere of living things to the pair of opposing forces—attraction and repulsion—which rule the inorganic world” (Freud, 1940, p. 72). Among humans, the former may be seen in life-enhancing acts that are attracted to what we experientially record as pleasurable events (positive reinforcers), the latter in life-preserving behaviors oriented to repel events experientially characterized as painful (negative reinforcers). More is said of these fundamental if not universal mechanisms of countering entropic disintegration in the next section.

To summarize, the aims of existence reflects a to-be or not-to-be issue. In the inorganic world, to be is essentially a matter of possessing qualities that distinguish a phenomenon from its surrounding field—not being in a state of entropy. Among organic beings, to be is a matter of possessing the properties of life as well as being located in ecosystems that facilitate the enhancement and preservation of that life. In the phenomenological or experiential world of sentient organisms, events that extend life and preserve it correspond largely to metaphorical terms such as pleasure and pain; that
is to say, recognizing and pursuing positive sensations and emotions, on the one hand, and recognizing and eschewing negative sensations and emotions, on the other.

Although there are many philosophical and metapsychological issues associated with the nature of pain and pleasure as constructs, it is neither our intent nor our task to inquire into them here. That they recur as a polar dimension time and again in diverse psychological domains (e.g., learned behaviors, unconscious processes, emotion, and motivation, as well as their biological substrates) has been elaborated in another publication (Millon, 1990). In this next section, I examine their role as constructs for articulating attributes that may usefully define personality.

Before we proceed, let us note that a balance must be struck between the two extremes that comprise each polarity; a measure of integration among the evolutionary polarities is an index of normality. Normal personality functioning, however, does not require equidistance between polar extremes. Balanced but unequal positions emerge as a function of temperamental dispositions, which, in their turn, are modified by the wider ecosystems within which individuals develop and function. In other words, there is no absolute or singular form of normal personality. Various polar positions and the personality attributes they subserve result in diverse styles of normality, just as severe or marked imbalances between the polarities manifest themselves in diverse styles of abnormality (Millon & Davis, 1996).

Moreover, given the diverse and changing ecological milieu that humans face in our complex modern environment, there is reason to expect that most persons will develop multiple adaptive styles, sometimes more active, sometimes less so, occasionally focused on self, occasionally on others, at times oriented to pleasure, at times oriented to the avoidance of pain. Despite the emergence of relatively enduring and characteristic styles over time, a measure of adaptive flexibility typifies most individuals: Persons are able to shift from one position on a bipolar continuum to another as the circumstances of life change.

**Personality Implications**

As noted, an interweaving and shifting balance between the two extremes that comprise the pain-pleasure polarity typifies normal personality functioning. Both of the following personality attributes should be met in varying degrees as life circumstances require. In essence, a synchronous and coordinated personal style would have developed to answer the question of whether the person should focus on experiencing only the enhancement of life versus concentrating his or her efforts on ensuring its preservation.

**Avoiding Danger and Threat: The Life Preservation Attribute.** One might assume that an attribute based on the avoidance of psychic or physical pain would be sufficiently self-evident not to require specification. As is well known, debates have arisen in the literature as to whether normal personality functioning represents the absence of mental disorder—that is, the reverse side of the mental illness or abnormality coin. That there is an inverse relationship between health and disease cannot be questioned; the two are intimately connected both conceptually and physically. On the other hand, to define a healthy personality solely on the basis of an absence of disorder does not suffice. As a single attribute of behavior that signifies both the lack of (e.g., anxiety, depression) and an aversion to (e.g., threats to safety and security) pain in its many and diverse forms does provide a foundation upon which other, more positively composed attributes may rest. Substantively, however, positive personal functioning must comprise elements beyond mere nonnormality or abnormality. And despite the complexities of personality, from a definitional point of view normal functioning does preclude nonnormality.

Turning to the evolutionary aspect of pain avoidance, that pertaining to a distancing from life-threatening circumstances, psychic and otherwise, we find an early historical reference in the writings of Herbert Spencer, a supportive contemporary of Darwin. In 1870 Spencer averred:

> Pains are the correlative of actions injurious to the organism, while pleasures are the correlatives of actions conducive to its welfare.

> Those races of beings only can have survived in which, on the average, agreeable or desired feelings went along with activities conducive to the maintenance of life, while disagreeable and habitually avoided feelings went along with activities directly or indirectly destructive of life.

> Every animal habitually persists in each act which gives pleasure, so long as it does so, and desists from each act which gives pain. . . . It is manifest that in proportion as this guidance approaches completeness, the life will be long; and that the life will be short in proportion as it falls short of completeness.

> We accept the inevitable corollary from the general doctrine of Evolution, that pleasures are the incentives to life-supporting acts and pains the deterrents from life-destroying acts. (pp. 279–284)

More recently, Freedman and Roe (1958) wrote:

> We . . . hypothesize that psychological warning and warding-off mechanisms, if properly studied, might provide a kind of psychological-evolutionary systematics. Exposure to pain, anxiety, or danger is likely to be followed by efforts to avoid a
repetition of the noxious stimulus situation with which the experience is associated. Obviously an animal with a more highly developed system for anticipating and avoiding the threatening circumstance is more efficiently equipped for adaptation and survival. Such unpleasant situations may arise either from within, in its simplest form as tissue deprivation, or from without, by the infliction of pain or injury. Man’s psychological superstructure may be viewed, in part, as a system of highly developed warning mechanisms. (p. 458)

As for the biological substrate of pain signals, Gray (1975) suggests two systems, both of which alert the organism to possible dangers in the environment. Those mediating the behavioral effects of unconditioned (instinctive?) aversive events are termed the fight-flight system (FFS). This system elicits defensive aggression and escape and is subserved, according to Gray’s pharmacological inferences, by the amygdala, the ventromedial hypothalamus, and the central gray of the midbrain; neurochemically, evidence suggests a difficult-to-unravel interaction among aminobutyric acids (for example, gamma-aminobutyric acid), serotonin, and endogenous opiates (for example, endorphins). The second major source of sensitivity and action in response to pain signals is referred to by Gray as the behavioral inhibition system (BIS), consisting of the interplay of the septal-hippocampal system, its cholinergic projections and monoamine transmissions to the hypothalamus, and then on to the cingulate and prefrontal cortex. Activated by signals of punishment or nonreward, the BIS suppresses associated behaviors, refocuses the organism’s attention, and redirects activity toward alternate stimuli.

Harm avoidance is a concept proposed by Cloninger (1986, 1987). As he conceives the construct, it is a heritable tendency to respond intensely to signals of aversive stimuli (pain) and to learn to inhibit behaviors that might lead to punishment and frustrating nonreward. Those high on this dimension are characterized as cautious, apprehensive, and inhibited; those low on this valence would likely be confident, optimistic, and carefree. Cloninger subscribes essentially to Gray’s behavioral inhibition system concept in explicating this polarity, as well as the neuroanatomical and neurochemical hypotheses Gray proposed as the substrates for its pain-avoidant mechanisms.

Shifting from biological-evolutionary concepts, we may turn to proposals of a similar cast offered by thinkers of a distinctly psychological turn of mind. Notable here are the contributions of Maslow (1968), particularly his hierarchical listing of needs. Best known are the five fundamental needs that lead ultimately to self-actualization, the first two of which relate to our evolutionary attribute of life preservation. Included in the first group are the physiological needs such as air, water, food, and sleep, qualities of the ecosystem essential for survival. Next, and equally necessary to avoid danger and threat, are what Maslow terms the safety needs, including the freedom from jeopardy, the security of physical protection and psychic stability, as well as the presence of social order and interpersonal predictability.

That pathological consequences can ensue from the failure to attend to the realities that portend danger is obvious; the lack of air, water, and food are not issues of great concern in civilized societies today, although these are matters of considerable import to environmentalists of the future and to contemporary poverty-stricken nations.

It may be of interest next to record some of the psychic pathologies that can be traced to aberrations in meeting this first attribute of personality. For example, among those termed inhibited and avoidant personalities (Millon, 1969, 1981), we see an excessive preoccupation with threats to one’s psychic security—an expectation of and hyperalertness to the signs of potential rejection—that leads these persons to disengage from everyday relationships and pleasures. At the other extreme of the polarity attribute, we see those of a risk-taking attitude, a proclivity to chance hazards and to endanger one’s life and liberty, a behavioral pattern characteristic of those we contemporaneously label antisocial personalities. Here there is little of the caution and prudence expected in the normal personality attribute of avoiding danger and threat; rather, we observe its opposite, a rash willingness to put one’s safety in jeopardy, to play with fire and throw caution to the wind. Another pathological style illustrative of a failure to fulfill this evolutionary attribute is seen among those variously designated as masochistic and self-defeating personalities. Rather than avoid circumstances that may prove painful and self-endangering, these nonnormal personality styles set in motion situations in which they will come to suffer physically, psychically, or both. Either by virtue of habit or guilt absolution, these individuals induce rather than avoid pain for themselves.

Seeking Rewarding Experiences: The Life Enhancement Attribute. At the other end of the existence polarity are attitudes and behaviors designed to foster and enrich life, to generate joy, pleasure, contentment, fulfillment, and thereby strengthen the capacity of the individual to remain vital and competent physically and psychically. This attribute asserts that existence and survival call for more than life preservation alone—beyond pain avoidance is what we have chosen to term pleasure enhancement.

This attribute asks us to go at least one step further than Freud’s parallel notion that life’s motivation is chiefly that of “reducing tensions” (i.e., avoiding or minimizing pain), maintaining thereby a steady state, if you will, a homeostatic
balance and inner stability. In accord with my view of evolution’s polarities, I would assert that normal humans are also driven by the desire to enrich their lives, to seek invigorating sensations and challenges, to venture and explore, all to the end of magnifying if not escalating the probabilities of both individual viability and species replicability.

Regarding the key instrumental role of “the pleasures,” Spencer (1870) put it well more than a century ago: “Pleasures are the correlates of actions conducive to [organismic] welfare. . . . the incentives to life-supporting acts” (pp. 279, 284). The view that there exists an organismic striving to expand one’s inherent potentialities (as well as those of one’s kin and species) has been implicit in the literature of all times. That the pleasures may be both sign and vehicle for this realization was recorded even in the ancient writings of the Talmud, where it states: “everyone will have to justify himself in the life hereafter for every failure to enjoy a legitimately offered pleasure in this world” (Jahoda, 1958, p. 45).

As far as contemporary psychobiological theorists are concerned, brief mention will be made again of the contributions of Gray (1975, 1981) and Cloninger (1986, 1987). Gray’s neurobiological model centers heavily on activation and inhibition (active-passive polarities) as well as on signals of reward and punishment (pleasure-pain polarity). Basing his deductions primarily on pharmacological investigations of animal behavior, Gray has proposed the existence of several interrelated and neuroanatomically grounded response systems that activate various positive and negative affects. He refers to what he terms the behavioral activation system (BAS) as an approach system that is subserved by the reward center uncovered originally by Olds and Milner (1954). Ostensibly mediated at brain stem and cerebellar levels, it is likely to include dopaminergic projections across various striata and is defined as responding to conditioned rewarding and safety stimuli by facilitating behaviors that maximize their future recurrence (Gray, 1975). There are intricacies in the manner with which the BAS is linked to external stimuli and its anatomic substrates, but Gray currently views it as a system that subserves signals of reward, punishment relief, and pleasure.

Cloninger (1986, 1987) has generated a theoretical model composed of three dimensions, which he terms reward dependence, harm avoidance, to which I referred previously, and novelty seeking. Proposing that each is a heritable personality disposition, he relates them explicitly to specific monoaminergic pathways; for example, high reward dependence is connected to low noradrenergic activity, harm avoidance to high serotonergic activity, and high novelty seeking to low dopaminergic activity. Cloninger’s reward dependence dimension reflects highs and lows on the positive-gratifying-pleasure valence, whereas the harm avoidance dimension represents highs and lows on the negative-pain-displeasure valence. Reward dependence is hypothesized to be a heritable neurobiological tendency to respond to signals of reward (pleasure), particularly verbal signals of social approval, sentiment, and succor, as well as to resist events that might extinguish behaviors previously associated with these rewards. Cloninger portrays those high on reward dependence to be sociable, sympathetic, and pleasant; in contrast, those low on this polarity are characterized as detached, cool, and practical. Describing the undergirding substrate for the reward-pleasure valence as the behavior maintenance system (BMS), Cloninger speculates that its prime neuromodulator is likely to be norepinephrine, with its major ascending pathways arising in the pons, projecting onward to hypothalamic and limbic structures, and then branching upward to the neocortex.

Turning again to pure psychological formulations, both Rogers (1963) and Maslow (1968) have proposed concepts akin to my criterion of enhancing pleasure. In his notion of “openness to experience,” Rogers asserts that the fully functioning person has no aspect of his or her nature closed off. Such individuals are not only receptive to the experiences that life offers, but they are able also to use their experiences in expanding all of life’s emotions, as well as in being open to all forms of personal expression. Along a similar vein, Maslow speaks of the ability to maintain a freshness to experience, to keep up one’s capacity to appreciate relationships and events. No matter how often events or persons are encountered, one is neither sated nor bored but is disposed to view them with an ongoing sense of awe and wonder.

Perhaps less dramatic than the conceptions of either Rogers and Maslow, I believe that this openness and freshness to life’s transactions is an instrumental means for extending life, for strengthening one’s competencies and options, and for maximizing the viability and replicability of one’s species. More mundane and pragmatic in orientation than their views, this conception seems both more substantive theoretically and more consonant a rationale for explicating the role the pleasures play in undergirding reward experience and openness to experience.

As before, a note or two should be recorded on the pathological consequences of a failure to possess an attribute. These are seen most clearly in the personality disorders labeled schizoid and avoidant. In the former there is a marked hedonic deficiency, stemming either from an inherent deficit in affective substrates or the failure of stimulative experience to develop attachment behaviors, affective capacity, or both (Millon, 1981, 1990). Among those designated avoidant personalities, constitutional sensitivities or abusive
life experiences have led to an intense attentional sensitivity to psychic pain and a consequent distrust in either the genuineness or durability of the pleasures, such that these individuals can no longer permit themselves to chance experiencing them, lest they prove again to be fickle and unreliable. Both of these personalities tend to be withdrawn and isolated, joyless and grim, neither seeking nor sharing in the rewards of life.

Modes of Adaptation

To come into existence as an emergent particle, a local cosmos, or a living creature is but an initial phase, the serendipitous presence of a newly formed structure, the chance evolution of a phenomenon distinct from its surroundings. Although extant, such fortuitous transformations may exist only for a fleeting moment. Most emergent phenomena do not survive (i.e., possess properties that enable them to retard entropic decomposition). To maintain their unique structure, differentiated from the larger ecosystem of which they are a part, and to be sustained as a discrete entity among other phenomena that comprise their environmental field requires good fortune and the presence of effective modes of adaptation. These modes of basic survival comprise the second essential component of evolution’s procession.

Ecological Accommodation and Ecological Modification. The Passive-Active Polarity

The second evolutionary stage relates to what is termed the *modes of adaptation*; it is also framed as a two-part polarity. The first may best be characterized as the mode of ecological accommodation, signifying inclinations to passively fit in, to locate and remain securely anchored in a niche, subject to the vagaries and unpredictabilities of the environment, all acceded to with one crucial proviso: that the elements comprising the surroundings will furnish both the nourishment and the protection needed to sustain existence. Although based on a somewhat simplistic bifurcation among adaptive strategies, this passive and accommodating mode is one of the two fundamental methods that living organisms have evolved as a means of survival. It represents the core process employed in the evolution of what has come to be designated as the plant kingdom: a stationary, rooted, yet essentially pliant and dependent survival mode. By contrast, the second of the two major modes of adaptation is seen in the lifestyle of the animal kingdom. Here we observe a primary inclination toward ecological modification, a tendency to change or rearrange the elements comprising the larger milieu, to intrude upon otherwise quiescent settings, a versatility in shifting from one niche to another as unpredictability arises, a mobile and interventional mode that actively stirs, maneuvers, yields, and at the human level substantially transforms the environment to meet its own survival aims.

Both modes—passive and active—have proven impressively capable to both nourishing and preserving life. Whether the polarity sketched is phrased in terms of accommodating versus modifying, passive versus active, or plant versus animal, it represents at the most basic level the two fundamental modes that organisms have evolved to sustain their existence. This second aspect of evolution differs from the first stage, which is concerned with what may be called *existential becoming*, in that it characterizes modes of being: how what has become endures.

Broadening the model to encompass human experience, the active-passive polarity means that the vast range of behaviors engaged in by humans may fundamentally be grouped in terms of whether initiative is taken in altering and shaping life’s events or whether behaviors are reactive to and accommodate those events.

Much can be said for the survival value of fitting a specific niche well, but no less important are flexibilities for adapting to diverse and unpredictable environments. It is here again where a distinction, although not a hard and fast one, may be drawn between the accommodating (plant) and the modifying (animal) mode of adaptation, the former more rigidly fixed and constrained by ecological conditions, the latter more broad-ranging and more facile in its scope of maneuverability. To proceed in evolved complexity to the human species, we cannot help but recognize the almost endless variety of adaptive possibilities that may (and do) arise as secondary derivatives of a large brain possessing an open network of potential interconnections that permit the functions of self-reflection, reasoning, and abstraction. But this takes us beyond the subject of this section of the chapter. The reader is referred elsewhere (Millon 1990) for a fuller discussion of active-passive parallels in wider domains of psychological thought (for example, the “ego apparatuses” formulated by Hartmann (1939) or the distinction between classical and operant conditioning in the writings of Skinner (1938, 1953).

Normal or optimal functioning, at least among humans, appears to call for a flexible balance that interweaves both polar extremes. In the first evolutionary stage, that relating to existence, behaviors encouraging both life enhancement (pleasure) and life preservation (pain avoidance) are likely to be more successful in achieving survival than actions limited to one or the other alone. Similarly, regarding adaptation, modes of functioning that exhibit both ecological accommodation and ecological modification are likely to be more successful...
than is either by itself. Nevertheless, it does appear that the two advanced forms of life on earth—plants and animals—have evolved by giving precedence to one mode rather than both.

**Personality Implications**

As with the pair of criteria representing the aims of existence, a balance should be achieved between the two criteria comprising modes of adaptation, those related to ecological accommodation and ecological modification, or what I have termed the passive-active polarity. Healthy personality functioning calls for a synchronous and coordinated style that weaves a balanced answer to the question of whether one should accept what the fates have brought forth or take the initiative in altering the circumstances of one’s life.

**Abiding Hospitable Realities: The Ecologically Accommodating Attribute.** On first reflection, it would seem to be less than optimal to submit meekly to what life presents, to adjust obligingly to one’s destiny. As described earlier, however, the evolution of plants is essentially grounded (no pun intended) in environmental accommodation, in an adaptive acquiescence to the ecosystem. Crucial to this adaptive course, however, is the capacity of these surroundings to provide the nourishment and protection requisite to the thriving of a species.

Could the same be true for the human species? Are there not circumstances of life that provide significant and assured levels of sustenance and safekeeping (both psychic and physical?) And if that were the case, would not the acquisition of an accommodating attitude and passive lifestyle be a logical consequence? The answer, it would seem, is yes. If one’s upbringing has been substantially secure and nurturant, would it not be normal to flee or overturn it?

We know that circumstances other than those in infancy and early childhood rarely persist throughout life. Autonomy and independence are almost inevitable as a stage of maturation, ultimately requiring the adoption of so-called adult responsibilities that call for a measure of initiative, decision making, and action. Nevertheless, to the extent that the events of life have been and continue to be caring and giving, is it not perhaps wisest, from an evolutionary perspective, to accept this good fortune and let matters be? This accommodating or passive life philosophy has worked extremely well in sustaining and fostering those complex organisms that comprise the plant kingdom. Hence passivity, the yielding to environmental forces, may be in itself not only unproblematic, but where events and circumstances provide the pleasures of life and protect against their pains, positively adaptive and constructive. Accepting rather than overturning a hospitable reality seems a sound course; or as it is said, “If it ain’t broke, don’t fix it.”

Often reflective and deliberate, those who are passively oriented manifest few overt strategies to gain their ends. They display a seeming inertness, a phlegmatic lack of ambition or persistence, a tendency toward acquiescence, a restrained attitude in which they initiate little to modify events, waiting for the circumstances of their environment to take their course before making accommodations. Some persons may be temperamentally ill-equipped to rouse or assert themselves; perhaps past experience has deprived them of opportunities to acquire a range of competencies or confidence in their ability to master the events of their environment; equally possible is a naive confidence that things will come their way with little or no effort on their part. From a variety of diverse sources, then, those at the passive end of the polarity engage in few direct instrumental activities to intercede in events or generate the effects they desire. They seem suspended, quiescent, placid, immobile, restrained, listless, waiting for things to happen and reacting to them only after they occur.

Is passivity a natural part of the repertoire of the human species, does agreeableness serve useful functions, and where and how is it exhibited? A few words in response to these questions may demonstrate that passivity is not mere inactivity but a stance or process that achieves useful gains. For example, universal among mammalian species are two basic modes of learning: the respondent or conditioned type and the operant or instrumental type. The former is essentially a passive process, the simple pairing of an innate or reflexive response to a stimulus that previously did not elicit that response. In like passive fashion, environmental elements that occur either simultaneously or in close temporal order become connected to each other in the organism’s repertoire of learning, such that if one of these elements recurs in the future, the expectation is that the others will follow or be elicited. The organisms do not have to do anything active to achieve this learning; inborn reflexive responses and environmental events are merely associated by contiguity.

Operant or instrumental learning, in contrast, represents the outcome of an active process on the part of the organism, one that requires an effort and execution on its part that has the effect of altering the environment. Whereas respondent conditioning occurs as a result of the passive observation of a conjoining of events, operant conditioning occurs only as a result of an active modification by the organism of its surroundings, a performance usually followed by a positive reinforcer (pleasure) or the successful avoidance of a negative one (pain). Unconditioned reflexes, such as a leg jerk in reaction to a knee tap, will become a passively acquired
conditioned respondent if a bell is regularly sounded prior to
the tap, as will the shrinking reflex of an eye pupil passively
become conditioned to that bell if it regularly preceded expo-
sure to a shining light.

The passive-active polarity is central to formulations of
psychoanalytic theory. Prior to the impressively burgeoning
literature on self and object relations theory of the past two
decades, the passive-active antithesis had a major role in
both classical instinct and post–World War II ego schools of
analytic thought. The contemporary focus on self and object
is considered in discussions of the third polarity, that of self-
other. However, we should not overlook the once key and
now less popular constructs of both instinct theory and ego
theory. It may be worth noting, as well as of special interest
to the evolutionary model presented in this chapter, that the
beginnings of psychoanalytic metapsychology were oriented
initially to instinctual derivatives (in which pleasure and
pain were given prominence), and then progressed subse-
sequently to the apparatuses of the ego (Hartmann, 1939; Ra-
aport, 1953)—where passivity and activity were centrally
involved.

The model of activity, as Rapaport puts it, is a dual one:
First, the ego is strong enough to defend against or control the
intensity of the id’s drive tensions; or second, through the
competence and energy of its apparatuses, the ego is success-
ful in uncovering or creating in reality the object of the id’s
instinctual drives. Rapaport conceives the model of passivity
also to be a dual one: First, either the ego gradually modu-
lates or indirectly discharges the instinctual energies of the
id; or second, lacking an adequately controlling apparatus,
the ego is rendered powerless and subject thereby to instinc-
tual forces. Translating these formulations into evolution-
ary terms, effective actions by the ego will successfully
manage the internal forces of the id, whereas passivity will
result either in accommodations or exposure to the internal
demands of the id.

Turning to contemporary theorists more directly con-
cerned with normal or healthy personality functioning, the
humanistic psychologist Maslow (1968) states that “self-
actualized” individuals accept their nature as it is, despite
personal weaknesses and imperfections; comfortable with
themselves and with the world around them, they do not
seek to change “the water because it is wet, or the rocks be-
cause they are hard” (p. 153). They have learned to accept the
natural order of things. Passively accepting nature, they need
not hide behind false masks or transform others to fit
distorted needs. Accepting themselves without shame or
apology, they are equally at peace with the shortcomings of
those with whom they live and relate.

Where do we find clinical states of personality functioning
that reflect failures to meet the accommodating-agreeable
attribute?

One example of an inability to leave things as they are is
seen in what is classified as the histrionic personality disor-
der. These individuals achieve their goals of maximizing pro-
tection, nurturance, and reproductive success by engaging
busily in a series of manipulative, seductive, gregarious, and
attention-getting maneuvers. Their persistent and unrelenting
manipulation of events is designed to maximize the receipt of
attention and favors, as well as to avoid social disinterest and
disapproval. They show an insatiable if not indiscriminate
search for stimulation and approval. Their clever and often
artful social behaviors may give the appearance of an inner
confidence and self-assurance; beneath this guise, however,
lies a fear that a failure on their part to ensure the receipt of at-
tention will in short order result in indifferent or rejection—
hence their desperate need for reassurance and repeated signs
of approval. Tribute and affection must constantly be reple-
ished and are sought from every interpersonal source. As they
are quickly bored and sated, they keep stirring up things,
becoming enthusiastic about one activity and then another.
There is a restless stimulus-seeking quality in which they can-
not leave well enough alone.

At the other end of the polarity are personality maladapt-
tions that exhibit an excess of passivity, failing thereby to
give direction to their own lives. Several personality disor-
ders demonstrate this passive style, although their passivity
derives from and is expressed in appreciably different ways.
Schizoid personalities, for example, are passive owing to
their relative incapacity to experience pleasure and pain;
without the rewards these emotional valences normally activ-
ate, they are devoid of the drive to acquire rewards, leading
them to become apathetically passive observers of the ongo-
ing scene. Dependent personality styles typically are average
on the pleasure-pain polarity, yet they are usually as passive
as schizoids. Strongly oriented to others, they are notably
weak with regard to self. Passivity for them stems from
deficits in self-confidence and competence, leading to deficits
in initiative and autonomous skills, as well as a tendency to
wait passively while others assume leadership and guide
them. Passivity among so-called obsessive-compulsive per-
sonalities stems from their fear of acting independently,
owing to intrapsychic resolutions they have made to quell
hidden thoughts and emotions generated by their intense self-
other ambivalence. Dreading the possibility of making mis-
takes or engaging in disapproved behaviors, they became
indecisive, immobilized, restrained, and thereby passive.
High on pain and low on both pleasure and self, individuals
with masochistic personality styles operate on the assumption that they dare not expect nor deserve to have life go their way; giving up any efforts to achieve a life that accords with their true desires, they passively submit to others’ wishes, acquiescently accepting their fate. Finally, narcissistic personality styles, especially high on self and low on others, benignly assume that good things will come their way with little or no effort on their part; this passive exploitation of others is a consequence of the unexplored confidence that underlies their self-centered presumptions.

Mastering One’s Environment: The Ecologically Modifying Attribute. The active end of the adaptational polarity signifies the taking of initiative in altering and shaping life’s events. Such persons are best characterized by their alertness, vigilance, liveliness, vigor, forcefulness, stimulus-seeking energy, and drive. Some plan strategies and scan alternatives to circumvent obstacles or avoid the distress of punishment, rejection, and anxiety. Others are impulsive, precipitate, excitable, rash, and hasty, seeking to elicit pleasures and rewards. Although specific goals vary and change from time to time, actively aroused individuals intrude on passing events and energetically and busily modify the circumstances of their environment.

Neurobiological research has proven to be highly supportive of the activity or arousal construct ever since Papez (1937), Moruzzi and Magurno (1949), and MacLean (1949, 1952) assigned what were to be termed the reticular and limbic systems’ both energizing and expressive roles in the central nervous system.

First among historic figures to pursue this theme was Ivan Pavlov. In speaking of the basic properties of the nervous system, Pavlov referred to the strength of the processes of excitation and inhibition, the equilibrium between their respective strengths, and the mobility of these processes. Although Pavlov’s (1927) theoretical formulations dealt with what Donald Hebb (1955) termed a conceptual nervous system, his experiments and those of his students led to innumerable direct investigations of brain activity. Central to Pavlov’s thesis was the distinction between strong and weak types of nervous systems.

Closely aligned to Pavlovian theory, Gray (1964) has asserted that those with weak nervous systems are easily aroused, non-sensation-seeking introverts who prefer to experience low rather than high levels of stimulation. Conversely, those with strong nervous systems would arouse slowly and be likely to be sensation-seeking extroverts who find low stimulation levels to be boring and find high levels to be both exciting and pleasant.

Akin also to the active modality are the more recent views of Cloninger (1986, 1987). To him, novelty-seeking is a heritable tendency toward excitement in response to novel stimuli or cues for reward (pleasure) or punishment relief (pain), both of which leading to exploratory activity. Consonant with its correspondence to the activity polarity, individuals who are assumed to be high in novelty-seeking may be characterized in their personality attributes as impulsive, excitable, and quickly distracted or bored. Conversely, those at the passive polarity or the low end of the novelty-seeking dimension may be portrayed as reflective, stoic, slow-tempered, orderly, and only slowly engaged in new interests.

Turning from ostensive biological substrates to speculative psychological constructs, de Charms (1968) has proposed that “man’s primary motivational propensity is to be effective in producing changes in his environment” (p. 269). A similar view has been conveyed by White (1959) in his concept of effectance, an intrinsic motive, as he views it, that activates persons to impose their desires upon environments. De Charms (1968) elaborates his theme with reference to man as Origin and as Pawn, constructs akin to the active polarity on the one hand and to the passive polarity on the other; he states this distinction as follows:

That man is the origin of his behavior means that he is constantly struggling against being confined and constrained by external forces, against being moved like a pawn into situations not of his own choosing. . . . An Origin is a person who perceives his behavior as determined by his own choosing; a Pawn is a person who perceives his behavior as determined by external forces beyond his control. . . . An Origin has strong feelings of personal causation, a feeling that the locus for causation of effects in his environment lies within himself. The feedback that reinforces this feeling comes from changes in his environment that are attributable to personal behavior. This is the crux of personal causation, and it is a powerful motivational force directing future behavior. (pp. 273–274)

Allport (1955) argued that history records many individuals who were not content with an existence that offered them little variety, a lack of psychic tension, and minimal challenge. Allport considers it normal to be pulled forward by a vision of the future that awakened within persons their drive to alter the course of their lives. He suggests that people possess a need to invent motives and purposes that would consume their inner energies. In a similar vein, Fromm (1955) proposed a need on the part of humans to rise above the roles of passive creatures in an accidental if not random world. To him, humans are driven to transcend the state of merely having been created; instead, humans seek to become the
creators, the active shapers of their own destiny. Rising above the passive and accidental nature of existence, humans generate their own purposes and thereby provide themselves with a true basis of freedom.

**Strategies of Replication**

In their mature stage, organisms possess the requisite competencies to maintain entropic stability. When these competencies can no longer adapt and sustain existence, organisms succumb inexorably to death and decomposition. This fate does not signify finality, however. Prior to their demise, all ephemeral species create duplicates that circumvent their extinction, engaging in acts that enable them to transcend the entropic dissolution of their members’ individual existences.

If an organism merely duplicates itself prior to death, then its replica is doomed to repeat the same fate it suffered. However, if new potentials for extending existence can be fashioned by chance or routine events, then the possibility of achieving a different and conceivably superior outcome may be increased. And it is this co-occurrence of random and recombinant processes that does lead to the prolongation of a species’ existence. This third hallmark of evolution’s procession also undergirds another of nature’s fundamental polarities, that between self and other.

**Reproductive Nurturance and Reproductive Propagation: The Other-Self Polarity**

At its most basic and universal level, the manifold varieties of organisms living today have evolved, as Mayr (1964) has phrased it, to cope with the challenge of continuously changing and immensely diversified environments, the resources of which are not inexhaustible. The means by which organisms cope with environmental change and diversity are well known. Inorganic structures survive for extended periods of time by virtue of the extraordinary strength of their bonding. This contrasts with the very earliest forerunners of organic life. Until they could replicate themselves, their distinctive assemblages existed precariously, subject to events that could put a swift end to the discrete and unique qualities that characterized their composition, leaving them essentially as transient and ephemeral phenomena. After replicative procedures were perfected, the chemical machinery for copying organismic life, the DNA double helix, became so precise that it could produce perfect clones—if nothing interfered with its structure or its mechanisms of execution. But the patterning and processes of complex molecular change are not immune to accident. High temperatures and radiation dislodge and rearrange atomic structures, producing what are termed mutations, alterations in the controlling and directing DNA configuration that undergirds the replication of organismic morphology.

Despite the deleterious impact of most mutations, it is the genetic variations to which they give rise that have served as one of the primary means by which simple organisms acquire traits making them capable of adapting to diverse and changing environments. But isomorphic replication, aided by an occasional beneficent mutation, is a most inefficient if not hazardous means of surmounting ecological crises faced by complex and slowly reproducing organisms. Advantageous mutations do not appear in sufficient numbers and with sufficient dependability to generate the novel capabilities required to adapt to frequent or marked shifts in the ecosystem. How then did the more intricate and intermittently reproducing organisms evolve the means to resolve the diverse hazards of unpredictable environments?

The answer to this daunting task was the evolution of a recombinant mechanism, one in which a pair of organisms exchange their genetic resources: They develop what we term *sexual mating*. Here, the potentials and traits each partner possesses are sorted into new configurations that differ in their composition from those of their origins, generating thereby new variants and capabilities, of which some may prove more adaptive (and others less so) in changing environments than were their antecedents. Great advantages accrue by the occasional favorable combinations that occur through this random shuffling of genes.

Recombinant replication, with its consequential benefits of selective diversification, requires the partnership of two parents, each contributing its genetic resources in a distinctive and species-characteristic manner. Similarly, the attention and care given the offspring of a species’ matings are also distinctive. Worthy of note is the difference between the mating parents in the degree to which they protect and nourish their joint offspring. Although the investment of energy devoted to upbringing is balanced and complementary, rarely is it identical or even comparable in either devotion or determination. This disparity in reproductive investment strategies, especially evident among nonhuman animal species (e.g., insects, reptiles, birds, mammals), underlies the evolution of the male and female genders, the foundation for the third cardinal polarity I propose to account for evolution’s procession.

Somewhat less profound than that of the first polarity, which represents the line separating the enhancement of order (existence-life) from the prevention of disorder (nonexistence-death), or that of the second polarity, differentiating the adaptive modes of accommodation (passive-plant) from those of modification (active-animal), the third polarity,
based on distinctions in replication strategies, is no less fundamental in that it contrasts the maximization of reproductive propagation (self-male) from that of the maximization of reproductive nurturance (other-female).

Evolutionary biologists (Cole, 1954; Trivers, 1974; E. O. Wilson, 1975) have recorded marked differences among species in both the cycle and pattern of their reproductive behaviors. Of special interest is the extreme diversity among and within species in the number of offspring spawned and the consequent nurturing and protective investment the parents make in the survival of their progeny. Designated the \( r \)-strategy and \( K \)-strategy in population biology, the former represents a pattern of propagating a vast number of offspring but exhibiting minimal attention to their survival; the latter is typified by the production of few progeny followed by considerable effort to assure their survival. Exemplifying the \( r \)-strategy are oysters, which generate some 500 million eggs annually; the \( K \)-strategy is found among the great apes, which produce a single offspring every 5 to 6 years.

Not only do species differ in where they fall on the \( r \)- to \( K \)-strategy continuum, but within most animal species an important distinction may be drawn between male and female genders. It is this latter differentiation that undergirds what has been termed the self- versus other-oriented polarity, implications of which are briefly elaborated in the following discussion.

Human females typically produce about four hundred eggs in a lifetime, of which no more than twenty to twenty-five can mature into healthy infants. The energy investment expended in gestation, nurturing, and caring for each child, both before and during the years following birth, is extraordinary. Not only is the female required to devote much of her energies to bring the fetus to full term, but during this period she cannot be fertilized again; in contrast, the male is free to mate with numerous females. And should her child fail to survive, the waste in physical and emotional exertion not only is enormous, but also amounts to a substantial portion of the mother’s lifetime reproductive potential. There appears to be a good reason, therefore, to encourage a protective and caring inclination on the part of the female, as evident in a sensitivity to cues of distress and a willingness to persist in attending to the needs and nurturing of her offspring.

Although the male discharges tens of millions of sperm on mating, this is but a small investment, given the ease and frequency with which he can repeat the act. On fertilization, his physical and emotional commitment can end with minimal consequences. Although the protective and food-gathering efforts of the male may be lost by an early abandonment of a mother and an offspring or two, much more may be gained by investing energies in pursuits that achieve the wide reproductive spread of his genes. Relative to the female of the species, whose best strategy appears to be the care and comfort of child and kin—that is, the \( K \)-strategy—the male is likely to be reproductively more prolific by maximizing self-propagation—that is, adopting the \( r \)-strategy. To focus primarily on self-replication may diminish the survival probabilities of a few of a male’s progeny, but this occasional reproductive loss may be well compensated for by mating with multiple females and thereby producing multiple offspring.

In sum, males lean toward being self-oriented because competitive advantages that inhere within themselves maximize the replication of their genes. Conversely, females lean toward being other-oriented because their competence in nurturing and protecting their limited progeny maximizes the replication of their genes.

The consequences of the male’s \( r \)-strategy are a broad range of what may be seen as self- as opposed to other-oriented behaviors, such as acting in an egotistical, insensitive, uncaring, and minimally communicative manner. In contrast, females are more disposed to be other-oriented, affiliative, intimate, emphatic, protective, communicative, and solicitous (Gilligan, 1982; Rushton, 1985; E. O. Wilson, 1978).

**Personality Implications**

As before, I consider both of the following criteria necessary to the definition and determination of a full personality characterization. I see no necessary antithesis between the two. Humans can be both self-actualizing and other-encouraging, although most persons are likely to lean toward one or the other side. A balance that coordinates the two provides a satisfactory answer to the question of whether one should be devoted to the support and welfare of others (the underlying philosophy of the “Democrats”) or fashion one’s life in accord with one’s own needs and desires (the underlying philosophy of the “Republicans”).

**Constructive Loving: The Other-Nurturing Attribute.**

As described earlier, recombinant replication achieved by sexual mating entails a balanced although asymmetrical parental investment in both the genesis and the nurturance of offspring. By virtue of her small number of eggs and extended pregnancy, the female strategy for replicative success among most mammals is characterized by the intensive care and protection of a limited number of offspring. Oriented to reproductive nurturance rather than reproductive propagation, most adult females, at least until recent decades in Western society, bred close to the limit of their capacity, attaining a reproductive ceiling of approximately 20 viable births.
By contrast, not only are males free of the unproductive pregnancy interlude for mating, but they may substantially increase their reproductive output by engaging in repetitive matings with as many available females as possible.

The other-versus-self antithesis follows from additional aspects of evolution’s asymmetric replication strategy. Not only must the female be oriented to and vigilant in identifying the needs of and dangers that may face each of her few offspring, but it is reproductively advantageous for her to be sensitive to and discriminating in her assessment of potential mates. A bad mating—one that issues a defective or weak offspring—has graver consequences for the female than for the male. Not only will such an event appreciably reduce her limited reproductive possibilities and cause her to forego a better mate for a period of time, but she may exhaust much of her nurturing and protective energies in attempting to revitalize an inviable or infertile offspring. By contrast, if a male indulges in a bad mating, all he has lost are some quickly replaceable sperm, a loss that does little to diminish his future reproductive potentials and activities.

Before we turn to other indexes and views of the self-other polarity, let us be mindful that these conceptually derived extremes do not evince themselves in sharp and distinct gender differences. Such proclivities are matters of degree, not absolutes, owing not only to the consequences of recombinant “shuffling” and gene “crossing over,” but also to the influential effects of cultural values and social learning. Consequently, most normal individuals exhibit intermediate characteristics on this as well as on the other two polarity sets.

The reasoning behind different replication strategies derives from the concept of inclusive fitness, the logic of which we owe to the theoretical biologist W. D. Hamilton (1964). The concept’s rationale is well articulated in the following quote (Daly & Wilson, 1978):

> Suppose a particular gene somehow disposes its bearers to help their siblings. Any child of a parent that has this gene has a one-half of probability of carrying that same gene by virtue of common descent from the same parent bearer. . . . From the gene’s point of view, it is as useful to help a brother or sister as it is to help the child.

> When we assess the fitness of a. . . bit of behavior, we must consider more than the reproductive consequences for the individual animal. We must also consider whether the reproductive prospects of any kin are in any way altered. Inclusive fitness is a sum of the consequences for one’s own reproduction plus the consequences for the reproduction of kin multiplied by the degree of relatedness of those kin [italics added].

> An animal’s behavior can therefore be said to serve a strategy whose goal is the maximization of inclusive fitness.

Mutual support and encouragement represents efforts leading to reciprocal fitness—a behavioral pattern consonant with Darwin’s fundamental notions. Altruism, however, is a form of behavior in which there is denial of self for the benefit of others, a behavioral pattern acknowledged by Darwin himself as seemingly inconsistent with his theory (1871, p. 130). A simple extrapolation from natural selection suggests that those disposed to engage in self-sacrifice would ultimately leave fewer and fewer descendants; as a consequence, organisms motivated by self-benefiting genes would prevail over those motivated by other-benefiting genes, a result leading to the eventual extinction of genes oriented to the welfare of others. The distinguished sociobiologist E. O. Wilson states the problem directly: “How then does altruism persist?” (1978, p. 153). An entomologist of note, Wilson had no hesitation in claiming that altruism not only persists, but also is of paramount significance in the lives of social insects. In accord with his sociobiological thesis, he illustrates the presence of altruism in animals as diverse as birds, deer, porpoises, and chimpanzees, which share food and provide mutual defense—for example, to protect the colony’s hives, bees enact behaviors that lead invariably to their deaths.

Two underlying mechanisms have been proposed to account for cooperative behaviors such as altruism. One derives from the concept of inclusive fitness, briefly described in preceding paragraphs; E. O. Wilson (1978) terms this form of cooperative behavior hard-core altruism, by which he means that the act is “unilaterally directed” for the benefit of others and that the bestower neither expects nor expresses a desire for a comparable return. Following the line of reasoning originally formulated by Hamilton (1964), J. P. Rushton (1984), a controversial Canadian researcher who has carried out illuminating r-K studies of human behavior, explicates this mechanism as follows:

> Individuals behave so as to maximize their inclusive fitness rather than only their individual fitness; they maximize the production of successful offspring by both themselves and their relatives. . . . Social ants, for example, are one of the most altruistic species so far discovered. The self-sacrificing, sterile worker and soldier ants. . . . share 75% of their genes with their sisters and so by devoting their entire existence to the needs of others . . . they help to propagate their own genes. (p. 6)

The second rationale proposed as the mechanism underlying other-oriented and cooperative behaviors Wilson terms soft-core altruism to represent his belief that the bestower’s actions are ultimately self-serving. The original line of reasoning here stems from Trivers’s (1971) notion of reciprocity, a thesis suggesting that genetically based dispositions to
In order to understand this idea more clearly, return with me for a moment to the basic theory of evolution. Imagine a spectrum of self-serving behavior. At one extreme only the individual is meant to benefit, then the nuclear family, next the extended family (including cousins, grandparents, and others who might play a role in kin selection), then the band, the tribe, chiefdoms, and finally, at the other extreme, the highest sociopolitical units. (p. 158)

Intriguing data and ideas have been proposed by several researchers seeking to identify specific substrates that may relate to the other-oriented polarities. In what has been termed the affiliation-attachment drive, Everly (1988), for example, provides evidence favoring an anatomical role for the cingulate gyrus. Referring to the work of Henry and Stephens (1977), MacLean (1985), and Steklis and Kling (1985), Everly concludes that the ablation of the cingulate eliminates both affiliative and grooming behaviors. The proximal physiology of this drive has been hypothesized as including serotonergic, noradrenergic, and opioid neurotransmission systems (Everly, 1988; Redmond, Maas, & Kling, 1971). MacLean (1985) has argued that the affiliative drive may be phylogenetically coded in the limbic system and may undergird the concept of family in primates. The drive toward other-oriented behaviors, such as attachment, nurturing, affection, reliability, and collaborative play, has been referred to as the “cement of society” by Henry and Stevens (1977).

Let us move now to the realm of psychological and social proposals. Dorothy Conrad (1952) specified a straightforward list of constructive behaviors that manifest “reproductive nurturance” in the interpersonal sphere. She records them as follows:

*Has positive affective relationship:* The person who is able to relate affectively to even one person demonstrates that he is potentially able to relate to other persons and to society.

*Promotes another’s welfare:* Affective relationships make it possible for the person to enlarge his world and to act for the benefit of another, even though that person may profit only remotely.
shared emotions, by what clinicians call empathy. His sexual nature may yet lead him to widening ambits of human affection, his acquisitive propensities to an optimum balance of work and leisure, and his aggressive drives to heightened social efficiency through attacks on perils common to all men. (p. 457)

The pathological consequences of a failure to embrace the polarity criterion of others are seen most clearly in the personality maladaptations termed antisocial and narcissistic disorders. Both personalities exhibit an imbalance in their replication strategy; in this case, however, there is a primary reliance on self rather than others. They have learned that reproductive success as well as maximum pleasure and minimum pain is achieved by turning exclusively to themselves. The tendency to focus on self follows two major lines of development.

In the narcissistic personality maladaptive style, development reflects the acquisition of a self-image of superior worth. Providing self-rewards is highly gratifying if one values oneself or possesses either a real or inflated sense of self-worth. Displaying manifest confidence, arrogance, and an exploitive egocentricity in social contexts, this individual believes he or she already has all that is important—him- or herself.

Narcissistic individuals are noted for their egotistical self-involvement, experiencing primary pleasure simply by passively being or attending to themselves. Early experience has taught them to overvalue their self-worth; this confidence and superiority may be founded on false premises, however—it may be unsustainable by real or mature achievements. Nevertheless, they blithely assume that others will recognize their special-ness. Hence they maintain an air of arrogant self-assurance, and without much thought or even conscious intent, benignly exploit others to their own advantage. Although the tributes of others are both welcome and encouraged, their air of snobbish and pretentious superiority requires little confirmation either through genuine accomplishment or social approval. Their sublime confidence that things will work out well provides them with little incentive to engage in the reciprocal give and take of social life.

Those clinically designated as antisocial personalities counter the indifference or the expectation of pain from others; this is done by actively engaging in duplicitous or illegal behaviors in which they seek to exploit others for self-gain. Skeptical regarding the motives of others, they desire autonomy and wish revenge for what are felt as past injustices. Many are irresponsible and impulsive, behaviors they see as justified because they judge others to be unreliable and disloyal. Insensitivity and ruthlessness with others are the primary means they have learned to head off abuse and victimization.

In contrast to the narcissistic form of maladaptation, the antisocial pattern of self-orientation develops as a form of protection and counteraction. These styles turn to themselves first to avoid the depredation they anticipate, and second to compensate by furnishing self-generated rewards in their stead. Learning that they cannot depend on others, individuals with these personality styles counterbalance loss not only by trusting themselves alone, but also by actively seeking retribution for what they see as past humiliations. Turning to self and seeking actively to gain strength, power, and revenge, they act irresponsibly, exploiting and usurping what others possess as just reprisals. Their security is never fully assured, however, even when they have aggrandized themselves beyond their lesser origins.

In both narcissistic and antisocial personality styles, we see maladaptations arising from an inability to experience a constructive love for others. For the one, there is an excessive self-centeredness; for the other, there is the acquisition of a compensatory destructiveness driven by a desire for social retribution and self-aggrandizement.

Realizing One’s Potentials: The Self-Actualizing Attribute. The converse of other-nurturance is not self-propagation, but rather the lack of other-nurturance. Thus, to fail to love others constructively does not assure the actualization of one’s potentials. Both may and should exist in normal, healthy individuals. Although the dimension of self-other is arranged to highlight its polar extremes, it should be evident that many if not most behaviors are employed to achieve the goals of both self- and kin reproduction. Both ends are often simultaneously achieved; at other times one may predominate. The behaviors comprising these strategies are driven, so to speak, by a blend of activation and affect—that is, combinations arising from intermediary positions reflecting both the life enhancement and life preservation polarity of pleasure-pain, interwoven with similar intermediary positions on the ecological accommodation and ecological modification polarity of activity-passivity. Phrasing replication in terms of the abstruse and metaphorical constructs does not obscure it, but rather sets this third polarity on the deeper foundations of existence and adaptation, foundations composed of the first two polarities previously described.

At the self-oriented pole, Everly (1988) proposes an autonomy-aggression biological substrate that manifests itself in a strong need for control and domination as well as in hierarchical status striving. According to MacLean (1986), it appears that the amygdaloid complex may play a key role in driving organisms into self-oriented behaviors. Early studies of animals with ablated amygdalas showed a notable increase in their docility (Kluver & Bucy, 1939), just as nonhuman
primates have exhibited significant decreases in social hierarchy status (Pribram, 1962). Although the evidence remains somewhat equivocal, norepinephrine and dopamine seem to be the prime neurotransmitters of this drive; the testosterone hormone appears similarly implicated (Feldman & Quenzar, 1984).

Regarding psychological constructs that parallel the notion of self-actualization, their earliest equivalent was in the writings of Spinoza (1677/1986), who viewed development as that of becoming what one was intended to be and nothing other than that, no matter how exalted the alternative might appear to be.

Carl Jung’s (1961) concept of individuation shares important features with that of actualization in that any deterrent to becoming the individual one may have become would be detrimental to life. Any imposed “collective standard is a serious check to individuality,” injurious to the vitality of the person, a form of “artificial stunting.”

Perhaps it was my own early mentor, Kurt Goldstein (1939), who first coined the concept under review with the self-actualization designation. As he phrased it, “There is only one motive by which human activity is set going: the tendency to actualize oneself” (1939, p. 196).

The early views of Jung and Goldstein have been enriched by later theorists, notably Fromm, Perls, Rogers, and Maslow.

Focusing on what he terms the sense of identity, Fromm (1955) spoke of the need to establish oneself as a unique individual, a state that places the person apart from others. Further—and it is here where Fromm makes a distinct self-oriented commitment—the extent to which this sense of identity emerges depends on how successful the person is in breaking “incestuous ties” to one’s family or clan. Persons with well-developed feelings of identity experience a feeling of control over their lives rather than a feeling of being controlled by the lives of others.

Perls (1969) enlarged on this theme by contrasting self-regulation versus external regulation. Normal, healthy persons do their own regulating, with no external interference, be it the needs and demands of others or the strictures of a social code. What we must actualize is the true inner self, not an image we have of what our ideal selves should be. That is the “curse of the ideal.” To Perls, each must be what he or she really is.

Following the views of his forerunners, Maslow (1968) stated that self-actualization is the supreme development and use of all our abilities, ultimately becoming what we have the potential to become. Noting that self-actualists often require detachment and solitude, Maslow asserted that such persons are strongly self-centered and self-directed, make up their own minds, and reach their own decisions without the need to gain social approval.

In like manner, Rogers (1963) posited a single, overreaching motive for the normal, healthy person—maintaining, actualizing, and enhancing one’s potential. The goal is not that of maintaining a homeostatic balance or a high degree of ease and comfort, but rather to move forward in becoming what is intrinsic to self and to enhance further that which one has already become. Believing that humans have an innate urge to create, Rogers stated that the most creative product of all is one’s own self.

Where do we see failures in the achievement of self-actualization, a giving up of self to gain the approbation of others? Two maladaptive personality styles can be drawn upon to illustrate forms of self-denial.

Those with dependent personalities have learned that feeling good, secure, confident, and so on—that is, those feelings associated with pleasure or the avoidance of pain—is provided almost exclusively in their relationship with others. Behaviorally, these persons display a strong need for external support and attention; should they be deprived of affection and nurturance, they will experience marked discomfort, if not sadness and anxiety. Any number of early experiences may set the stage for this other-oriented imbalance. Dependent individuals often include those who have been exposed to an overprotective training regimen and who thereby fail to acquire competencies for autonomy and initiative; experiencing peer failures and low self-esteem leads them to forego attempts at self-assertion and self-gratification. They learn early that they themselves do not readily achieve rewarding experiences; these experiences are secured better by leaning on others. They learn not only to turn to others as their source of nurturance and security, but also to wait passively for others to take the initiative in providing safety and sustenance. Clinically, most are characterized as searching for relationships in which others will reliably furnish affection, protection, and leadership. Lacking both initiative and autonomy, they assume a dependent role in interpersonal relations, accepting what kindness and support they may find and willingly submitting to the wishes of others in order to maintain nurturance and security.

A less benign but equally problematic centering on the wishes of others and the denial of self is seen in what is termed clinically as the obsessive-compulsive personality. These persons display a picture of distinct other-directedness—a consistency in social compliance and interpersonal respect. Their histories usually indicate having been subjected to constraint and discipline when they transgressed parental strictures and expectations. Beneath the conforming other-oriented veneer, they exhibit intense desires to rebel and assert their own self-oriented feelings and impulses. They are trapped in an ambivalence; to avoid intimidation and punishment they have
learned to deny the validity of their own wishes and emotions and instead have adopted as true the values and precepts set forth by others. The disparity they sense between their own urges and the behaviors they must display to avoid condemnation often leads to omnipresent physical tensions and rigid psychological controls.

Readers who have reached this final paragraph on the basic three polarities that undergird all physical forms and organic species should have a foundation to move onto our next series of polarities, those which are distinctly human—that is, these polarities relate to personality attributes found almost exclusively in the human species that set us off from all earlier forms of evolution and that pertain to the higher powers and adaptive functions of abstraction and their constituent cognitive modes.

THE DISTINCTLY HUMAN POLARITIES OF EVOLUTION

This group of personality attributes incorporates the sources employed to gather knowledge about the experience of life and the manner in which this information is registered and transformed. Here, we are looking at styles of cognizing—differences (first) in what people attend to in order to learn about life, and (second) how they process information: what they do to record this knowledge and make it useful to themselves.

Predilections of Abstraction

The cognitive features of intelligence are judged by me to be central elements in personological derivations. Comprising the fourth and most recent stage of evolution, they comprise the reflective capacity to transcend the immediate and concrete, they interrelate and synthesize the diversity of experience, they represent events and processes symbolically, they weigh, reason, and anticipate; in essence, they signify a quantum leap in evolution’s potential for change and adaptation.

Cognitive differences among individuals and the manner in which they are expressed have been much overlooked in generating and appraising personality attributes. With an occasional notable exception or two, little of the recent so-called revolution in cognitive science that has profoundly affected contemporary psychology has impacted the study of personology. Historically, the realms of intellect, aptitude, and ability have not been considered to be personality-related spheres of study.

In my view, personology should be broadened to encompass the whole person, an organically unified and unsegmented totality. Consequently, cognitive dimensions and their various styles not only should be included, but also may have a significance equal to that of other functions as a source of personality attributes (Millon, 1990). Unfortunately, the various features comprising cognitive abstraction have only rarely been included as components in personality-oriented concepts and appraisals.

Emancipated from the real and present, unanticipated possibilities and novel constructions may routinely be created cognitively. The capacity to sort, to recompose, to coordinate, and to arrange the symbolic representations of experience into new configurations is in certain ways analogous to the random processes of recombinant replication, but processes enabling manipulation of abstractions are more focused and intentional. To extend this rhetorical liberty, replication is the recombinant mechanism underlying the adaptive progression of phylogeny, whereas abstraction is the recombinant mechanism underlying the adaptive progression of ontogeny. The powers of replication are limited, constrained by the finite potentials inherent in parental genes. In contrast, experiences, abstracted and recomposed, are infinite.

Over one lifetime, innumerable events of a random, logical, or irrational character transpire, are construed, and are reformulated time and again—some of which prove more and others less adaptive than their originating circumstances may have called forth. Whereas the actions of most nonhuman species derive from successfully evolved genetic programs, activating behaviors of a relatively fixed nature suitable for a modest range of environmental settings, the capabilities of both implicit and intentional abstraction that characterize humans give rise to adaptive competencies that are suited to radically divergent ecological circumstances, circumstances that themselves may be the result of far-reaching acts of symbolic and technological creativity.

Although what underlies our self- versus other-oriented attributes stems from differential replication strategies, the conscious state of knowing self as distinct from others is a product of the power of abstraction, the most recent phase of evolution’s procession. The reflective process of turning inward and recognizing self as an object—no less to know oneself, and further, to know that one knows—is a uniqueness found only among humans. Doubling back on oneself, so to speak, creates a new level of reality, consciousness that imbues self and others with properties far richer and more subtle than those that derive from strategies of reproductive propagation and nurturance alone.

The abstracting mind may mirror outer realities but reconstructs them in the process, reflectively transforming them into subjective modes of phenomenological reality, making external events into a plastic mold subject to creative designs. Not only are images of self and others emancipated from
direct sensory realities, becoming entities possessing a life of their own, but contemporaneous time may also lose its immediacy and impact. The abstracting mind brings the past effectively into the present, and its power of anticipation brings the future into the present as well. With past and future embedded in the here and now, humans can encompass at once not only the totality of our cosmos, but also its origins and nature, its evolution, and how they have come to pass. Most impressive of all are the many visions humans have of life’s indeterminate future, where no reality as yet exists.

Four polarities constitute this distinctly human abstraction function. The first two pairs refer to the information sources that provide cognitions. One set of contrasting polarities addresses the orientation either to look outward, or external-to-self, in seeking information, inspiration, and guidance, versus the orientation to turn inward, or internal-to-self. The second set of abstraction polarities contrasts predilections for either direct observational experiences of a tangible, material, and concrete nature with those geared more toward intangible, ambiguous, and inchoate phenomena.

The third and fourth set of abstraction polarities relate to cognitive processing—that is, the ways in which people evaluate and mentally reconstruct information and experiences after they have been apprehended and incorporated. The first of these sets of cognitive polarities differentiates processes based essentially on ideation, logic, reason, and objectivity from those that depend on emotional empathy, personal values, sentiment, and subjective judgments. The second set of these polarities reflects either a tendency to make new information conform to preconceived knowledge, in the form of tradition-bound, standardized, and conventionally structured schemas, versus the opposing inclination to bypass preconceptions by distancing from what is already known and instead to create innovative ideas in an informal, open-minded, spontaneous, individualistic, and often imaginative manner.

Cognitive functions are consonant with our earlier biosocial formulations concerning the architecture of human functioning (Millon, 1990) because we see cognitive processes to be an essential component of our fourfold model regarding how organisms approach their environments. Beyond the driving motivational elements of personality style (as in my formulation of the personality disorders), or the factorial structure of personality (e.g., as explicated in the Big Five model), we seek to conjoin all components of personality style by linking and integrating the various expressions and functions of personality into an overarching and coherent whole.

Several polar dimensions have been proposed through the years as the basis for a schema of cognitive styles. Contrasting terms such as leveling versus sharpening, narrow versus broad, analytic versus synthetic, constricted versus flexible, inductive versus deductive, abstract versus concrete, and convergent versus divergent have been used to illustrate the stylistic differences among cognitive functions. Although each of these pairs contributes to distinctions of importance in describing cognitive processes, few were conceptualized with personality differences in mind, although some may prove productive in that regard.

As noted above, the model formulated by the author separates cognitive activities into two superordinate functions. The first pertains to the contrasting origins from which cognitive data are gathered, or what may be termed information sources; the second pertains to the methods by which these data are reconstructed by the individual, or what we label transformational processes. These two functions—the initial gathering and subsequent reconstruction of information—are further subdivided into two polarities each. As is elaborated later in this chapter, the sources of information are separated into (a) external versus internal and (b) tangible versus intangible. Transformational processes are divided into (a) ideational versus emotional and (b) integrative versus imaginative. The resulting four personality attributes are by no means exhaustive. Rather surprisingly, they turn out to be consonant with a model formulated in the 1920s by Jung (1971a).

Sources of Information

Information may be seen as the opposite of entropy. What energy or nutrients are to physical systems, information is to cognitive systems. A physical system sustains itself by sucking order, so to speak, from its environs, taking in energy or nutrients and transforming them to meet tissue needs; a cognitive system does something similar by sucking information from its environs—that is, taking in data and transforming them to meet its cognitive needs. In much the same way as any other open system, a cognitive structure needs to maintain itself as an integrated and cohesive entity. In the physical world, the integrity of a system is achieved by making adaptations that preserve and enhance the physical structure, thereby precluding the entropic dissipation of its ordered elements. Similarly, a cognitive system achieves its integrity through a variety of preserving and enhancing adaptations that reduce the likelihood of events that may diminish the order and coherence of its knowledge base.

Moreover, an open cognitive system is purposefully focused, as is a physical system. Just as a physical system must be selective about its nutrition sources in order to find those suitable to meet its tissue needs, so, too, must a cognitive system be selective about information sources, choosing and processing particular raw inputs according to specific
cognitive goals. A cognitive system can no more process random input than a physical system can ingest random material. Hence, information (negative entropy) must be acquired selectively rather than randomly or diffusely; some sources of information will be heeded and others ignored or suppressed.

Coherence may be optimized by adopting and maintaining a preferred and regular information source, thereby ensuring a consistent confirmatory bias in favor of a cognitive structure’s world view and organizational architecture. Conversely, a cognitive structure that is exposed to dissonant or contradictory sources or that heeds diverse or multitudinous sources ultimately may be challenged successfully or may be exhausted beyond its ability to maintain coherence. In other words, burdensome processing and discordant sources are likely to result in increasing cognitive entropy. A more structured and coherent focus that strengthens and confirms prior sources of information becomes useful in ensuring optimal cognitive survivability.

**External Versus Internal Orientation Polarity: The Extraceptive and Intraceptive Attributes.** In light of the preceding argument, we see two primary stimulative sources of information, that which originates external to the self and that which originates internally. Whether this polar cognitive orientation is termed external versus internal, extraceptive versus inceptive, or extraverting versus introverting, each polarity provides a replicable reservoir for cognitive information—a selectively narrowed wellspring of knowledge to which the person will continue to be exposed.

A few lines paraphrasing Jung, the originator of the extraverting-introverting dimension, may be of value in highlighting core features of the externally oriented preference. Extraversion, from Jung’s view, was centered in an interest in the external object noted by a ready acceptance of external happenings, a desire to influence and be influenced by external events, a need to join in, and the capacity not only to endure the bustle and noise of every kind, but actually find them enjoyable (Jung, 1971a).

Similarly, Jung clearly states a view paralleling ours in what we have termed the internal orientation. To Jung, the introverted person is “not forthcoming”; he or she “retreats before the external object.” Such an individual is aloof from external happenings and does not join in. Self-communings are a pleasure and the introverted individual experiences his or her own world as a safe harbor, a “carefully tended and walked-in garden, closed to the public and hidden from prying eyes.” The internally oriented person’s own company is best. One who is internally oriented feels at home in one’s own world, a place where changes are made only by oneself. Most significantly, the best work of such individuals is done with their own resources, on their own initiative, and in their own way (Jung, 1971b).

**Tangible Versus Intangible Disposition Polarity: The Realistic and Intuitive Attributes.** Information, whether its source is internal or external to the self, can be classified in numerous ways. A core distinction can be drawn between information that is tangible versus that which is intangible. By tangible we mean identifiable by human sensory capacities, well-defined, distinctive, recognizable, and knowable—referring to phenomena that are concrete, factual, material, realistic, or self-evident. In contrast, information that is termed intangible takes in phenomena that lack an intrinsically distinctive order and structural clarity; they are inherently ambiguous, abstract, insubstantial, vague, mysterious, and obscure. Such phenomena usually can be fathomed only by means that are unknown, unconscious, and percipient, or by glimmerings into their diffuse and elusive nature that are materially tenuous or psychical in form.

The readiness of some individuals to be receptive to information that is well-structured and tangible, and of others to receive information that is obscure and intangible, constitutes, in our view, a fundamental difference in cognitive style that is of appreciable personological significance. Although Jung’s language is only tangentially formulated in cognitive terms, close parallels can be seen between the polarity presented here and that offered by Jung in his distinction between Sensing and Intuiting. As Jung (1933) wrote decades ago:

> Here we should speak of sensation when sense impressions are involved, and of intuition if we are dealing with a kind of perception which cannot be traced back directly to conscious sensory experience. Hence, I define sensation as perception via conscious sensory functions, and intuition as perception via the unconscious. (pp. 538–539)

Favoring tangible, structured, and well-defined sources of information that call upon one’s five senses will no doubt correlate with a wide range of associated behaviors, such as choosing actions of a pragmatic and realistic nature, preferring events in the here and now, and attending to matters calling for facts and quantitative precision.

Jung conceived what we would term the tangible disposition as the fact-minded men in whom intuition is “driven into the background by actual facts.” In contrast, those preferring the intangible, unstructured, and ambiguous world of information are likely to be inspired by possibilities, by challenges, and potentials of an abstract, connotative, and symbolic character, as well as by matters that depend on mystery and speculation. In Jung’s words, “for these persons, actual reality
counts only insofar as it harbors possibilities, regardless of the way things are in the actual present” (Jung, 1971b, p. 539).

Transformational Processes

The first two pairs of cognitive functions were grouped according to attributes that signify choices among the sources and styles of gathering information. These next two pairs of attribute polarities represent amplification preferences and transformational processes, referring to what is done to information after it has been received. Cognitive science has articulated a number of concepts related to the registering, encoding, and organizing of life experiences. These concepts pertain to various questions, such as Through what cognitive mode will information be received and amplified—intellective or affective? and How shall information be organized; will it be assimilated into preformed memory systems or will it be recast through imagination into novel schemas? Although individuals may be positioned on several other continua or polarities—for example, convergent versus divergent, serial versus hierarchical, primary versus secondary, verbal versus visual—it is the author’s view that the most fruitful cognitive distinctions relevant to personality are the pairs selected in this and the following section.

Ideational Versus Emotional Preference Polarity: The Intellective and Affective Attributes. Stated simply, there are essentially two pathways through which experiences pass once recorded by our consciousness or by our senses, if they are of sufficient magnitude to activate an encoded response. The first pathway accentuates information that is conceptual and logical, eliciting a reasoned judgment that signifies in an articulate and organized way that the registered experience makes sense—that is, it is rationally consistent and coherent. The second pathway resonates an emotional response, a subjective feeling reaction, signaling in a somewhat diffuse and global way that the registered event was experienced either as affectively neutral, clearly positive, or distinctly negative.

The ideational pole indicates a preference and elaboration of experience in light of reason and logic. Although life events may derive from internal or external sources and may be of a tangible or intangible nature, the interpretive and evaluative process is inclined toward and augments the objective and impersonal, as events are amplified by means of critical reason and intensified by the application of rational and judicious thought. By diminishing affective engagements—reducing the unruly emotional input of others or the upsetting effects of one’s own affective state—the preference is to sustain and strengthen a high degree of cognitive logic and cohesion. Objective analysis and affective detachment protect against unwanted incursions upon intellectual rationality, but often at the price of promoting processes that tend to be rigid, overcontrolled, and unyielding.

In contrast, experiences processed and amplified emotionally activate subjective states, such as liking versus disliking, feeling good versus feeling bad, comfort versus discomfort, attracted versus repelled, valuing versus devaluing, and so on. Through empathic resonance, the route of enhanced affectivity inclines the individual to record not so much what other people think but rather how they feel. The individual who inclines toward the affective attribute uses feeling vibrations to learn more from the melodic tone that words convey than from their content or logic. The usual modality for those who exhibit an affective bent is that of a subjective reality, a series of more-or-less gut reactions composed of either global or differentiated positive or negative moods. For the most part, the affective amplification style indicates individuals who evince modest introspective analyses, who show an open and direct empathic response to others, and who have a subconscious susceptibility to the emotional facets of experience in as pure a manner as possible.

Integrating Versus Innovating Bias Polarity: The Assimilative and Imaginative Attributes. The second cognitive transformational polarity addresses the question of whether new information is shaped to fit preformed memory schemas (integrated within preexisting cognitive systems), or is organized through the imagination to be cast into innovative and creative forms. Evolutionary theory suggests that the best course may be to reinforce (cognitive) systems that have proved stable and useful. On the other hand, progress will not be made unless promising new possibilities are explored. A beneficial tension in evolution clearly exists between conservation and change, between that of adhering to the habitual and that of unleashing the creative. These two contrasting cognitive biases demonstrate the two options—integrating experiences into already established systems versus exploring innovative ways to structure them.

Assimilators are akin in certain features to persons with well-structured memory systems to which they routinely attach new cognitive experiences. Disposed to operate within established perspectives, assimilators integrate new information to fit previous points of view, exhibiting thereby a high degree of dependability and consistency, if not rigidity, in their functioning. Typically, such people are predictable, conventional, orderly, systematic, decisive, methodical, exacting, formal, disciplined, conscientious, faithful, loyal, and devoted. Hence, in evolutionary terms, the integrating polarity leads to continuity and tradition, or to the maintenance of existing levels of cognitive entropy; this cognitive style
promotes an architectural cohesion that remains unchallenged by variations that could be risky (i.e., potentially diminish established levels of order).

In contrast, those functioning at the innovating pole are characterized by an openness to forming new and imaginative cognitive constructions of a more-or-less impromptu character. They are inclined to search for and enjoy creative ideas and solutions, to find novel ways to order information and to accumulate negative entropy, so to speak, by stepping outside of what is known and given in order to establish a new and potentially higher level of cognitive organization. Innovators stretch beyond confirmed perspectives, seek to broaden interpretations of experience, and are not concerned with demonstrating their reliability. The imaginative attribute is typically associated with being open-minded, spontaneous, and to accumulate negative entropy, so to speak, by stepping outside of what is known and given in order to establish a new and potentially higher level of cognitive organization. Innovators stretch beyond confirmed perspectives, seek to broaden interpretations of experience, and are not concerned with demonstrating their reliability. The imaginative attribute is typically associated with being open-minded, spontaneous, and resourceful.

It is to those who combine these latter two persuasions that the present chapter is heartily addressed.

REFERENCES


During much of its past, psychology represented a culturally grounded enterprise that took into account the constitutive role of cultural meanings and practices in human development. Yet, as recent historical accounts make clear (Jahoda, 1993), this attention to culture was muted during the twentieth century, with psychology dominated by an idealized physical-science model of explanation. This has given rise to the enigma that psychologists find it “difficult to keep culture in mind,” noted by Cole (1996):

On the one hand, it is generally agreed that the need and ability to live in the human medium of culture is one of the central characteristics of human beings. On the other hand, it is difficult for many academic psychologists to assign culture more than a secondary, often superficial role in the constitution of our mental life. (p. 1)

From this type of perspective, which dominates the field, culture is seen as at most affecting the display of individual psychological processes, but not as impacting qualitatively on their form.

However, although culture thus remains in a peripheral role in the contemporary discipline, recent years have seen a reemergence of interest in cultural approaches and an increased recognition of their importance to psychological theory. As reflected in the interdisciplinary perspective of cultural psychology (e.g., Cole, 1990; Greenfield, 1997; J. G. Miller, 1997; Shweder, 1990), culture and psychology are coming to be understood as mutually constitutive processes. It is recognized that human development occurs in historically grounded social environments that are structured by cultural meanings and practices. Cultural meanings and practices are themselves understood to be dependent on the subjectivity of communities of intentional agents. By affecting individuals’ understandings and intentions, cultural meanings and practices, in turn, are recognized to have a qualitative impact on the development of psychological phenomena and to be integral to the formulation of basic psychological theory.

The goal of the present chapter is to highlight some of the insights for understanding personality and social psychology that emerge from a consideration of the cultural grounding of psychological processes. The first section of the chapter considers factors that have contributed to the downplaying of culture in mainstream social psychology and the assumptions that guided some of the earliest research in the traditions of cross-cultural psychology. In the second section, consideration is given to key conceptual developments underlying cultural psychology, recent empirical findings that illustrate the existence of cultural variation in basic social psychological processes, and challenges for future theory and research. In conclusion, consideration is given to the multiple contributions of a cultural perspective in psychology.
disciplinary practices that are affecting these shifts. The first section considers factors that are contributing to the tendency to assign cultural considerations a relatively peripheral role both in social psychology and more generally in the larger discipline. The second section provides an overview of some of the earliest traditions of cultural research in social psychology, highlighting respects in which this research, although groundbreaking in many respects, did not seriously challenge this tendency to downplay the importance of culture in psychology. Finally, attention turns to the core assumptions of cultural psychology, assumptions that highlight the need to accord culture a more integral role in basic psychological theory.

Downplaying of Culture in Mainstream Social Psychology

Signs of the peripheral theoretical role accorded to cultural considerations in social psychology may be seen in its being downplayed in major social psychological publications. Textbooks typically either leave the construct of culture theoretically undefined, treat it as the same as the objective environment or social ecology, or approach it in an eclectic way that lacks conceptual clarity. Likewise, basic theory tends to be presented without any reference to cultural considerations. Culture is treated merely as a factor that influences the universality of certain psychological effects but not as a process that must be taken into account to explain the form of basic psychological phenomena. One example of such a stance can be found in Higgins and Kruglanski’s (1996) recent handbook on basic principles of social psychology: The only citations for culture in the index—with only one exception—refer to pages within the single chapter on cultural psychology by Markus, Kitayama, and Heiman (1996), rather than to any of the other 27 chapters of the volume. In the following discussion, we argue that this downplaying of culture in social psychology reflects to a great degree the tendency to conceptualize situations in culture-free terms, the embrace of an idealized natural-science model of explanation, and the default assumption of cultural homogeneity that dominates the field.

Culture-Free Approach to Situations

A key contribution of social psychology—if not its signature explanatory feature—is its recognition of the power of situations to impact behavior. Such a stance is reflected, for example, in a series of classic studies; salient examples include the Milgram conformity experiment, which demonstrated that to conform with the orders of an experimenter, individuals were willing to inflict a harmful electric shock on a learner (Milgram, 1963), as well as the prison experiment of Zimbardo and his colleagues (Haney, Banks, & Zimbardo, 1973), which demonstrated that individuals who had been thrust into the role relationships of guards and prisoners in a simulated prison behaved in ways that reflected these positions, with the guards behaving abusively and the prisoners becoming passive. It also may be seen in recent lines of inquiry on such topics as individuals’ limited conscious access to their cognitive processes, priming effects, and the mere exposure effect (Bargh, 1996; Bornstein, Kale, & Cornell, 1990; Zimbardo, Banks, Haney, & Jaffe, 1973). Social psychological work of this type has shown that contexts affect behavior in ways that do not depend on conscious mediation and that may even violate individuals’ conscious expectations and motivational inclinations.

Supplementing this focus on the power of situations to affect behavior, it has also been documented that individual differences influence the meaning accorded to situations. This attention to individual differences is evident not only in work on personality processes but also in the attention given to cognitive and motivational schemas as sources of individual variability in behavior. Individual difference dimensions, however, typically are accorded a secondary role to situational influences within social psychological theory. They are believed to affect the display of certain basic psychological dimensions, but they are not often implicated in normative models of psychological phenomena. To give a representative example of such a stance, the theory of communal and exchange relationships has been forwarded to distinguish qualitatively between relationships that are based on need versus those based on exchange considerations (Mills & Clark, 1982). In this model, individual differences are invoked only in a descriptive sense (i.e., to distinguish between persons who are more or less likely to adopt each type of orientation) and not in a theoretical sense (that is, to identify distinctive approaches to relationships beyond those specified in the original conceptual model).

The crucial point is that the approach to situations that dominates social psychological inquiry treats contexts as presenting one most veridical structure that can be known through inductive or deductive information processing. No consideration is given to the possibility that culture is necessarily implicated in the definition of situations or that cultural presuppositions constitute prerequisites of what is considered objective knowledge. It is assumed that variability in judgment arises from differences in the information available to individuals or from differences in their information-processing abilities, resulting in certain judgments’ being more or less cognitively adequate or veridical than others.
(Nisbett & Ross, 1980). Evidence that individuals from different cultural backgrounds maintain contrasting systems of belief, value, or meaning—and that they interpret situations in contrasting ways—tends to be assimilated to an individual-difference dimension. It is viewed as implying that individual differences in attitudes, understandings, or available information may relate to cultural group membership, but not as implying that there is a need to give any independent weight to cultural meanings and practices per se in explanation.

In maintaining the present realist approach to situations and in adopting explanatory frameworks focused on factors in the situation and in the person, cultural considerations are downplayed in theoretical importance. It is assumed that cultural information may substitute for or shortcut individual information processing: The individual comes to learn about the world indirectly through acquiring the knowledge disseminated in the culture. As such, culture is viewed as providing information redundant with that which individuals could obtain by themselves through direct cognitive processing. Wells (1981), for example, maintains that enculturation processes are nonessential to individual knowledge acquisition:

It is difficult for anyone who has raised a child to deny the pervasive influence of socialized processing that surely surfaces as causal schemata originate through secondary sources such as parents... Even though socialized processing may be an important determinant of knowledge about causal forces at one level, it nevertheless begs the question. How is it that the parents knew an answer? The issue is circular. That is precisely the reason that one must consider a more basic factor—namely original processing. (p. 313)

From the present type of perspective, cultural knowledge is seen as necessary neither to account for the nature of individual knowledge nor to evaluate its adequacy.

**Natural Science Ideals of Explanation**

The tendency to downplay the importance of culture in social psychological theory also derives from the field’s embrace of an idealized physical-science model of explanation. Although social psychology makes use of multiple normative models of scientific inquiry, it has typically treated physical science models of scientific inquiry as the ideal approach. This has affected both the goals and methods of inquiry in ways that have tended to marginalize cultural approaches.

In terms of explanatory goals, the foremost aim of psychological explanation has been to identify universal laws of behavior. Adopting the criteria of parsimony and of predictive power as the hallmarks of a successful explanation, psychological inquiry has been conceptualized as involving the identification of deep structural explanatory mechanisms that (it is assumed) underlie overt behavior. Higgins and Kruglanski (1996) outline this vision for social psychological inquiry:

A discovery of lawful principles governing a realm of phenomena is a fundamental objective of scientific research... A useful scientific analysis needs to probe beneath the surface. In other words, it needs to get away from the ‘phenotypic’ manifestations and strive to unearth the ‘genotypes’ that may lurk beneath... We believe in the scientific pursuit of the nonobvious. But less in the sense of uncovering new and surprising phenomena than in the sense of probing beneath surface similarities and differences to discover deep underlying structures. (p. vii)

From this perspective, the assumption is made that fundamental psychological processes are timeless, ahistorical, and culturally invariant, with the principles of explanation in the social sciences no different from those in the natural or physical sciences.

From the present physical-science view of explanation, cultural considerations tend to be regarded as noise; they are consequently held constant in order to focus on identifying underlying processes. Malpass (1988) articulates this type of position:

Cultural differences are trivial because they are at the wrong level of abstraction, and stand as ‘medium’ rather than ‘thing’ in relation to the objects of study. The readily observable differences among cultural groups are probably superficial, and represent little if any differences at the level of psychological processes. (p. 31)

According to this perspective, an explanation that identifies a process as dependent on culturally specific assumptions is regarded as deficient. To discover that a phenomenon is culturally bound is to suggest that the phenomenon has not as yet been fully understood and that it is not yet possible to formulate a universal explanatory theory that achieves the desired goals of being both parsimonious and highly general.

Another consequence of the present physical-science model of explanation is that social psychology has tended to privilege laboratory-based methods of inquiry and to be dismissive of what is perceived to be the inherent lack of methodological control of cultural research. Skepticism surrounds the issue of whether sufficient comparability can be achieved in assessments made in different cultural contexts to permit valid cross-cultural comparisons. Equally serious concerns are raised that methodological weaknesses are inherent in the qualitative methods that are frequently involved in
assessment of cultural meanings and practices. In particular, because such measures are at times based on analyses undertaken by single ethnographers or similar methods, measures used in cultural assessment are seen as characterized by limited reliability and validity, as well as by heavy reliance on interpretive techniques.

It is notable that the adoption within social psychology of a physical-science ideal of explanation also promotes disciplinary insularity. Although there is considerable openness to the integration of biologically based conceptual models and methodologies—a trend seen in the growing interest in neuroscience—there is little or no interest in integrating the theoretical insights and empirical findings from other social science fields, such as anthropology. Rather, the body of knowledge developed within anthropology becomes difficult for social psychologists to assimilate. Thus, for example, psychologists typically treat the findings of anthropological research as merely descriptive or anecdotal, with little attention even given to such findings as a source of hypotheses that might be subject to further testing through controlled social psychological procedures. A situation is then created in which the findings of cultural variability in human behavior (which have been widely documented within anthropology) as well as anthropological tools of interpretive methodological inquiry tend to be given little or no attention in social psychological inquiry.

**Default Assumption of Cultural Heterogeneity**

Finally, the downplaying of the importance of cultural considerations in social psychology also stems from the tendency to assume a universalistic cultural context in recruitment of research participants and in formulation of research questions. This type of stance has led to skewed population sampling in research. As critics (Reid, 1994) have charged, the field has proceeded as though the cultural context for human development is homogeneous; consequently, research has adopted stances that treat middle-class European-American research populations as the default or unmarked subject of research:

Culture . . . has been assumed to be homogenous, that is, based on a standard set of values and expectations primarily held by White and middle-class populations. . . . For example, in developmental psychology, *children* means White children (McLoyd, 1990); in psychology of women, *women* generally refers to White women (Reid, 1988). When we mean other than White, it is specified. (p. 525)

In this regard, slightly over a decade ago, it was observed that fewer than 10% of all hypothesis testing research undertaken in social psychology involved samples drawn from two or more cultures (Pepitone & Triandis, 1987). Likewise, a review conducted of more than 14,000 empirical articles in psychology published between 1970 and 1989 yielded fewer than 4% centering on African Americans (Graham, 1992).

However, it is not only these skewed sampling practices but also the resulting skewed knowledge base brought to bear in inquiry that contributes to the downplaying of the importance of cultural considerations. Commonly, research hypotheses are based on investigators’ translations of observations from their own experiences into testable research hypotheses. In doing this, however, researchers from non-middle-class European-American backgrounds frequently find themselves having to suppress intuitions or concerns that arise from their own cultural experiences. As reflected in the following account by a leading indigenous Chinese psychologist (Yang, 1997), the present type of stance may give rise to a certain sense of alienation among individuals who do not share the so-called mainstream cultural assumptions that presently dominate the field:

I found the reason why doing Westernized psychological research with Chinese subjects was no longer satisfying or rewarding to me. When an American psychologist, for example, was engaged in research, he or she could spontaneously let his or her American cultural and philosophical orientations and ways of thinking be freely and effectively reflected in choosing a research question, defining a concept, constructing a theory and designing a method. On the other hand, when a Chinese psychologist in Taiwan was conducting research, his or her strong training by overlearning the knowledge and methodology of American psychology tended to prevent his or her Chinese values, ideas, concepts and ways of thinking from being adequately reflected in the successive stages of the research process. (p. 65)

It has been suggested, in this regard, that to broaden psychological inquiry to be sensitive to aspects of self emphasized in Chinese culture, greater attention would need to be paid to such presently understudied concerns as filial piety, impression management, relationship harmony, and protection of face (Hsu, 1963, 1985; Yang, 1988; Yang & Ho, 1988). Taking issues of this type into account, researchers of moral development, for example, have challenged the Kohlbergian claim that a concern with human rights fully captures the end point of moral development (Kohlberg, 1969, 1971); such researchers have uncovered evidence to suggest that within Chinese cultural populations, the end point of moral development places greater emphasis on *Ch'ing* (human affection or sentiment) as well as on the Confucian value of *jen* (love, human-heartedness, benevolence, and sympathy; Ma, 1988, 1989).
As a consequence of its tendency to privilege considerations emphasized in European-American cultural contexts, psychology in many cases has focused on research concerns that have a somewhat parochial character, as Moscovici (1972) has argued in appraising the contributions of social psychology:

... The real advance made by American social psychology was... in the fact that it took for its theme of research and for the content of its theories the issues of its own society. Its merit was as much in its techniques as in translating the problems of American society into sociopsychological terms and in making them an object of scientific inquiry. (p. 19)

In proceeding with a set of concepts that are based on a relatively narrow set of cultural experiences, psychological research then has tended to formulate theories and research questions that lack adequate cultural inclusiveness and instead are based on the experiences of highly select populations.

Summary

Despite its concern with social aspects of experience and with units of analysis, such as groups, that are larger than individuals, social psychological inquiry has tended to downplay cultural factors. This downplaying, as we have seen, reflects in part the field’s tendency to give weight both to situational and individual difference considerations, while according no independent explanatory force to cultural factors. Equally, it reflects the field’s embrace of natural-science models of explanation, which emphasize generality as the hallmark of a successful explanation and controlled experimentation as the most adequate approach to scientific inquiry. Finally, in both its sampling practices and in its consideration of research questions, social psychology has privileged a middle-class European-American outlook that gives only limited attention to the perspectives and concerns of diverse cultural and subcultural populations.

Early Research in Cross-Cultural Psychology

Although cultural considerations have tended to be accorded little importance in social psychological theory, there exists a long-standing tradition of research in cross-cultural psychology that has consistently focused attention on them. The scope of work in cross-cultural psychology is reflected in the vast body of empirical research that has been conducted. Empirical work from this perspective is extensive enough to fill the six-volume first edition of the Handbook of Cross-Cultural Psychology (Triandis & Lambert, 1980), as well as numerous textbooks and review chapters (e.g., Berry, Poortinga, Segall, & Dasen, 1992; Brislin, 1983).

Research in cross-cultural psychology shares many of the conceptual presuppositions of mainstream psychology—which explains, at least in part, why it has not fundamentally posed a challenge to the mainstream discipline (see discussion in Shweder, 1990; J. G. Miller, 2001a). These assumptions involve a view of culture as an independent variable affecting psychological processes understood as a dependent variable. From such a perspective, culture is seen as affecting the display or level of development of psychological processes, but not their basic form—a stance similar to the assumption in mainstream social psychology that culture has no impact on fundamental psychological phenomena. Research in cross-cultural psychology also assumes an adaptive approach to culture that is consonant with the view of the environment emphasized in mainstream psychology. Naturally occurring ecological environments are viewed as presenting objective affordances and constraints to which both individual behavior and cultural forms are adapted.

A major thrust of work in cross-cultural psychology has been to test the universality of psychological theories under conditions in which there is greater environmental variation than is present in the cultural context in which the theories were originally formulated. Brief consideration of early cross-cultural research in the traditions of culture and personality, culture and cognition, and individualism-collectivism highlights both the groundbreaking nature of this work as well as the limited extent to which it challenges the core theoretical presuppositions of the mainstream discipline.

Culture and Personality

The research tradition of culture and personality constituted an interdisciplinary perspective that generated great interest and inspired extensive research throughout the middle years of the twentieth century (e.g., LeVine, 1973; Shweder, 1979a, 1979b; Wallace, 1961; J. W. Whiting & Child, 1953; B. B. Whiting & Whiting, 1975). Although many of the classic assumptions of this perspective were subject to challenge, and although interest in this viewpoint diminished after the 1980s, work in culture and personality has served as an important foundation for later work on culture and the development of self.

Some of the earliest work in the tradition of culture and personality adopted a critical case methodology to test the generality of psychological theories. For example, in a classic example of this type of approach, Malinowski tested the universality of the Oedipus complex against case materials from the Trobriand Islands (1959). In contrast to the Freudian
assumption that the father is both the disciplinarian and the mother’s lover, in this society, the mother’s brother, rather than the father, assumed the role of disciplinarian. Based on his analysis, Malinowski concluded that there was no evidence for the occurrence of the Oedipus complex under these societal conditions. Likewise, in another early example of this type of approach, Margaret Mead provided evidence that adolescence does not invariably involve the patterns of psychosocial conflict that are observed in Western populations and that were once assumed in psychological theory to be universal (1928, 1939).

Later work in culture and personality developed models that portrayed culture as an amalgam of parts that conformed to the dominant pattern of individual personality possessed by members of the culture. Such an assumption may be seen reflected, for example, in the stance adopted by Benedict as she portrayed culture and personality as highly integrated entities: “A culture, like an individual, is a more or less consistent pattern of thought and action” (1932, p. 42). Applying this model to an analysis of Japan, Benedict (1946) traced broad consistencies that characterized Japanese values, social institutions, national policy, and interpersonal relations. Similar types of assumptions characterized the national-character studies that were conducted—research that frequently involved studying culture at a distance by relying on sources such as literature, art, and history (Adorno, Frenkel-Brunswick, Levinson, & Sanford, 1950; Gorer, 1955; Gorer & Rickman, 1962). For example, in examining why Nazism was embraced in Germany, researchers identified an assumed “authoritarian” personality that they maintained was characteristic of the German psyche and that they saw as contributing to the emphasis on obedience to authority observed in Germany under Nazi rule (Fromm, 1941).

Still a third thrust of work in culture and personality forwarded a personality—integration-of-culture model (Kardiner, 1945; B. B. Whiting & Whiting, 1975). From this viewpoint, individual personality structure was regarded as adapted to cultural meanings and practices that in turn were regarded as adapted to the demands of particular ecological settings. It was assumed from this perspective that individuals come over time to be socialized to behave in ways that fit what is regarded as the dominant psychological orientation of adults in the culture. As reflected in research that made use of the ethnographic reports compiled in the Human Relations Area Files (HRAF; J. W. Whiting & Child, 1953), studies empirically tested assumed causal relationships between features of the natural ecology, modes of social organization, child socialization, and expressive aspects of culture, such as religious beliefs. In a groundbreaking program of research that stands as one of the most influential contributions of this school of thought, the Six Culture study tested these relations in an investigation that involved conducting behavioral observations of parenting and child behavior in everyday contexts in a worldwide sampling of cultures (J. W. Whiting & Child, 1975). As one example of the many findings from the Six Culture project, it was demonstrated that cultures with rich natural ecologies give rise to societies with complex social structures, which, in turn, lead to the development of egoistic personality dispositions among members of the cultures and to cultural meanings and practices that emphasize competitiveness.

In terms of criticisms, research in the tradition of culture and personality was subject to challenge in terms of the theories of personality and of culture that it embodied (Shweder, 1979a, 1979b). Concerns were raised regarding the determinism of treating culture merely as a concomitant of individual personality, as well as regarding what was viewed as its overly socialized conception of the person—a conception that treated the individual as merely passively conforming to prevailing norms. Additionally, it was argued that work in culture and personality overestimated the thematic nature of cultural forms, as well as failed to take into account the limited longitudinal stability and cross-situational consistency of personality. For example, evidence suggested that what had been interpreted as a difference in personality between cultural populations in fact could be explained in normative terms—as individuals responding to the behavioral expectations of different everyday cultural settings (Shweder, 1975). Thus, the observation was later made that one of the most important influences of culture on individual development is that it provides contrasting socialization experiences rather than affects individual personalities. For example, the degree to which children in different cultures emphasize competitive versus cooperative behavior appears closely linked to whether children spend their days in the competitive atmosphere of formal school settings versus the more prosocial atmosphere of sibling caregiving activities (B. B. Whiting & Edwards, 1988).

In terms of enduring contributions, work on culture and personality succeeded in highlighting the importance of understanding the mutual influence of ecological, psychological, and cultural processes. Methodologically rich, research in this tradition not only demonstrated the importance of integrating both ethnographic and quantitative approaches in psychological investigation, but also called attention to the value of observing behavior in naturalistic contexts and of capturing the dynamics of everyday cultural activities and practices (e.g., Ford, 1967; Honigmann, 1954; LeVine, 1973; Spindler, 1980; Spiro, 1958, 1965, 1982; Wallace, 1961; J. W. Whiting & Child, 1953; B. B. Whiting & Whiting, 1975).
However, although the study of culture and personality left a rich and highly influential legacy with many investigators associated with this tradition at the forefront of contemporary work in cultural psychology, work in culture and personality did not directly move into the issues of culture and basic psychological theory that are being addressed in contemporary research in cultural psychology. Rather, most work in culture and personality assumed psychological universalism or what theorists have characterized as the “postulate of psychic unity” (e.g., Shweder, 1990). Personality theories were treated as having universal validity and thus as applicable in unchanged form in diverse cultural populations. Little consideration was given to respects in which these theories (e.g., psychoanalysis) might themselves be culturally bound.

**Individualism-Collectivism**

Work on individualism-collectivism represents one of the most influential and long-standing traditions of research in cross-cultural psychology. Associated particularly with the early theoretical work of investigators such as Hofstede and Triandis (Hofstede, 1980; Triandis, 1972, 1980, 1988), this perspective has been applied to explain variation in a wide range of behavioral domains on a worldwide scale. Thus, the constructs of individualism-collectivism have been invoked in explaining such diverse phenomena as values (Hofstede, 1980; S. H. Schwartz, 1994), cognitive differentiation (Witkin & Berry, 1975), and modernity (Inkeles, 1974). Embracing the explanatory goals of predictive power and parsimony as well as the quantitative methodological approaches of the mainstream discipline, the primary focus of work on individualism-collectivism has been to forward a universal framework that predicts the nature of both cultural forms and individual psychological experience.

Individualism and collectivism are conceptualized as syndromes of beliefs and attitudes that distinguish different cultural populations. *Collectivism* is seen as encompassing such core ideas as an emphasis on the views, needs, and goals of one’s in-group as having priority over one’s own personal views, needs, and goals, and a readiness to cooperate with in-group members. In contrast, individualism is seen as entailing such core ideas as that of individuals as ends in themselves who should realize their own selves and cultivate their own judgment. In collectivist cultures, in-groups are assumed to influence a broad range of behaviors, with individuals experiencing pressure to conform to in-group norms or leave the groups. In contrast, in individualistic cultures, in-groups are seen as providing only limited norms, with individuals readily able to enter and exit in-groups: The relationship of individuals with their in-groups is of limited intensity.

Further distinctions are made in this broad dichotomy to capture dimensions of variation between different individualistic and collectivist cultures (e.g., Triandis, 1989, 1996). Thus, for example, cultures are seen as differing in terms of which in-groups are important (e.g., family vs. country), the particular collectivist values emphasized (e.g., harmony vs. dignity), and the ease with which individuals can join in-groups and deviate from their norms (e.g., tightness vs. loose-ness of norms; Triandis, 1988). In addition to the global constructs of individualism-collectivism, additional constructs are invoked to explain individual differences. Thus, the constructs of *idiocentrism* and *allocentrism* have been proposed as the psychological manifestations at the level of individual self definitions, beliefs, and attitudes of individualism and collectivism. It is assumed that individuals in all cultures maintain both idiocentric and allocentric aspects of their selves. Cultural differences at the psychological level, then, are seen as reflecting the differential sampling of idiocentric as compared with allocentric features of self in diverse sociocultural contexts (Triandis, 1990, 1996).

In terms of explaining the cultural syndromes of individualism and collectivism, research has shown that factors such as affluence, exposure to mass media, modernization, mobility, movement from rural to urban settings, and industrialization are linked to societal shifts from collectivism toward individualism. In turn, a wide range of psychological consequences are seen as linked to such shifts, with individualism, as compared with collectivism, associated with such outcomes as higher self esteem and subjective well being (e.g., Diener & Diener, 1995; Diener, Diener, & Diener, 1995), values such as being curious and broad-minded as compared with emphasizing family security and respect for tradition (S. H. Schwartz, 1994), as well as direct and frank communication styles, as compared with relatively indirect communication styles that emphasize context and concern for the feelings of the other (Gudykunst, Yoon, & Nishida, 1987; Kim, Sharkey, & Singelis, 1994; Triandis, 1994).

The prototypical research conducted by investigators in the tradition of individualism-collectivism involves multicultural survey or questionnaire research. This work is concerned with developing ecological models of culture that can be invoked to explain the distribution of individualism-collectivism and of related psychological characteristics on a worldwide scale (for review, e.g., see Berry et al., 1992).

In recent years, researchers have shown increased interest in the constructs of individualism and collectivism as a consequence of these constructs being linked to the distinction drawn by Markus and Kitayama (1991) between independent versus interdependent modes of self construal. In introducing the contrast between independent versus interdependent
modes of self-construal, Markus and Kitayama did not adopt all of the assumptions of the individualism-collectivism framework as developed by early cross-cultural psychologists. In contrast to such theorists, for example, they were concerned with the cultural psychological agenda of identifying insights for basic psychological theory of cultural variation (e.g., identifying new culturally based forms of motivation), rather than with the cross-cultural agenda of applying existing psychological theories in diverse cultural contexts (e.g., identifying cultural variation in the emphasis placed on internal vs. external locus of control, as specified by Rotter’s framework). They tended to eschew the use of scale measures of individualism-collectivism; they also did not draw some of the global contrasts made within much work within this framework, such as devaluation of the self in collectivism or of relationships in individualism (see discussion in Kitayama, in press; J. G. Miller, 2002). However, in part as a reflection of the interest in the distinction between independent versus interdependent self-construals introduced by Markus and Kitayama (1991), the number of investigators concerned with individualism and collectivism has grown in recent years, with many investigators drawing on this framework to further the cultural psychological agenda of broadening basic psychological theory (e.g., Greenfield & Cocking, 1994; Greenfield & Suzuki, 1998), and other investigators in social psychology drawing on the framework to further the original agenda of theorists such as Triandis to develop a universal, ecologically based framework to explain psychological variation on a worldwide scale (e.g., Oyserman, Coon, & Kemmelmeier, 2002).

In terms of criticisms, the tradition of cross-cultural research on individualism is limited in its emphasis on testing the generality of existing psychological theories in diverse cultural contexts, and in its inattention to examining the degree to which such theories themselves may be culturally bound and take somewhat contrasting forms in different cultural contexts. This stance represents perhaps the most central reason that mainstream psychologists have tended to view the findings of research on individualism-collectivism as primarily descriptive in nature rather than to view them as contributing to basic psychological theory (e.g., Shweder, 1990). The framework of individualism-collectivism has also been subject to criticism for its global view of culture: Much work in this tradition fails to account for subtleties in cultural meanings and practices, and it has also been criticized for the somewhat stereotypical nature of its portrayal of these two cultural systems (e.g., Dier, 1999). Thus, for example, as numerous theorists have noted (e.g., Markus & Kitayama, 1991; J. G. Miller, 1994, 2002; Rothbaum, Pott, Azuma, Miyake, & Weisz, 2000), much work on individualism-collectivism has failed to recognize that concerns with self have importance in collectivist cultures rather than only in individualistic cultures—although they may take somewhat contrasting forms in the two cultural contexts, just as concerns with relationships have importance but may take different forms in the two cultural contexts. Finally, methodological criticisms have been directed at the widespread use of attitudinal scale measures in work in this tradition (e.g., Kitayama, 2002), with theorists noting the many problems associated with the limited ability of individuals to report on the orientations emphasized in their culture and with the inattention to everyday cultural practices, artifacts, and routines that has characterized much work in this tradition with its reliance on attitudinal indexes of culture.

The individualism-collectivism framework has made major and enduring contributions to understanding culture and society in ecological terms. Work in this tradition has been of great value in providing insight into processes of modernization and cultural change, and it has assisted in modeling how both factors in the physical environment and social structural considerations affect psychological outcomes. The broad framework of individualism-collectivism has also proven useful heuristically as a source of initial research hypotheses, with this distinction embraced—at least in a limited way—not only by investigators concerned with the more universalistic agenda of cross-cultural psychology, but also by some theorists identified more explicitly with cultural psychology (e.g., Greenfield & Suzuki, 1998).

Culture and Cognitive Development

Early work on culture and cognitive development was theoretically diverse and international in character, drawing on Piagetian as well as Vygotskian viewpoints among others. Within Piagetian viewpoints, cross-cultural research was undertaken to test the presumed universality of cognitive developmental theory (Dasen, 1972; Dasen & Heron, 1981). This work involved administering standard Piagetian cognitive tests in different cultures after translating the tests and making minor modifications to ensure their ecological validity. Likewise, in the domain of moral development, Kohlbergian measures of moral judgment were administered in a large number of cultural settings after only minor changes in research protocols were made, such as substituting local names for those originally in the text (e.g., Edwards, 1986; Kohlberg, 1969; Snarey, 1985). The findings on Piagetian tasks suggested that in certain African settings, cognitive development proceeds at a slower rate than that observed in
Geneva, with the highest level of formal operations generally not obtained. Likewise, cross-cultural Kohlberian research indicated that populations not exposed to higher levels of education do not reach the highest (postconventional) stage of moral judgment. Results of this type were generally interpreted as reflecting the cognitive richness of the environment that resulted in more advanced cognitive development in certain cultures over others. They were also interpreted as supporting the universality of cognitive developmental theory. It was concluded that culture is nonessential in development, in that the sequence and end point of developmental change are culturally invariant (e.g., Piaget, 1973).

Inspired by Vygotsky and other Soviet investigators (e.g., Vygotsky, 1929, 1934/1987; 1978; Luria, 1928, 1976), theorists in the early sociocultural tradition of cross-cultural research on cognitive development proceeded by undertaking experiments in diverse cultural settings. However, in contrast to cognitive developmental viewpoints, they assumed that cognitive development has a formative influence on the emergence of basic psychological processes. Rather than viewing development as proceeding independently of cultural learning, cultural learning was assumed to be necessary for development to proceed. Vygotskian theory and related sociocultural approaches emphasized the importance of tool use in extending cognitive capacities. From this perspective, cultural transmission was assumed to be essential, with cognitive development involving the internalization of the tools provided by the culture. Among the key cultural tools assumed to transform minds were literacy and formal schooling, through their assumed effects of providing individuals exposure to abstract symbolic resources and giving rise to modes of reasoning that are relatively decontextualized and not directly tied to practical activity (e.g., Goody, 1968). In viewing cultural processes as a source of patterning of thought, work in the sociocultural tradition shared many assumptions with and may be considered part of cultural psychology. However, at least in its early years, research in this tradition focused on establishing the universality of basic cognitive processes; this linked it closely to other contemporary traditions of cross-cultural cognitive developmental research.

The earliest traditions of cross-cultural experimental research undertaken by sociocultural theorists resembled those of Piagetian researchers in both their methods and their findings. After making only minor modifications, experimental tests were administered to diverse cultural populations. These populations were selected to provide a contrast in the cultural processes thought to influence cognitive development, such as literacy and schooling (e.g., Bruner, Olver, & Greenfield, 1966; Cole, Gay, Glick, & Sharp, 1971). Results revealed that individuals who were illiterate or who lacked formal education scored lower in cognitive development, failing to show such features as abstract conceptual development or propositional reasoning, which appeared as end points of cognitive development in Western industrialized contexts. Such findings supported a “primitive versus modern mind” interpretation of cultural differences, in which it was assumed that the cognitive development of certain populations remains arrested at lower developmental levels. This type of argument may be seen, for example, in the conclusion drawn by Greenfield and Bruner (1969) in drawing links between such observed cross-cultural differences and related differences found in research contrasting cognition among mainstream and minority communities within the United States:

... As Werner (1948) pointed out, ‘development among primitive people is characterized on the one hand by precocity and, on the other, by a relatively early arrest of the process of intellectual growth.’ His remark is telling with respect to the difference we find between school children and those who have not been to school. The latter stabilize earlier and do not go on to new levels of operation. The same ‘early arrest’ characterizes the differences between ‘culturally deprived’ and other American children.

... Some environments ‘push’ cognitive growth better, earlier, and longer than others. . . . Less demanding societies—less demanding intellectually—do not produce so much symbolic embedding and elaboration of first ways of looking and thinking.

(p. 654)

From this perspective, the impact of culture on thought was assumed to be highly general, with individuals fully internalizing the tools provided by their culture and that resulting in generalized cultural differences in modes of thought.

Later experimental research in the sociocultural tradition challenged these early conclusions about global differences in thought and about the transformative impact of cultural tools on minds. Programs of cross-cultural research were undertaken that focused on unpacking the complex cognitive processes that are tapped in standard cognitive tests and in assessing these components under diverse circumstances (Cole & Scribner, 1974). Thus, for example, rather than using the multiple objects that tended to be employed in Piagetian seriation tasks, with their extensive memory demands, researchers employed fewer objects in memory procedures. Also, processes such as memory were assessed in the context of socially meaningful material, such as stories, rather than merely in decontextualized ways, such as through the presentation of words. These and similar modifications showed that cognitive performance varied depending on features of the
Global cultural differences in thought, seemed to be proving something that was already assumed by many anthropologists who held a view of individuals as competent in fulfilling the cognitive demands of their culture. The field had not yet reached the point of articulating a positive agenda of characterizing how culture affected cognition. It was this kind of stance that emerged as sociocultural work, and work on culture and cognition began to turn more explicitly to cultural psychology.

Summary

In sum, early research in cross-cultural psychology laid important groundwork for contemporary research in the newly reemerging framework of cultural psychology. In terms of major empirical findings, this early work challenged the idea that cultural differences map onto personality differences of individual members of a culture, and pointed instead to the role of normative practices in underlying observed differences in individual behavior. It also challenged claims of global differences in cognitive capacity linked to modernizing influences, and instead identified modernizing influences as having localized effects on cognitive capacities. It is important to note, however, that although in many respects it was a precursor to much contemporary work in cultural psychology, early work in these traditions of cross-cultural psychology tended to remain in a relatively peripheral role in the discipline and not to impact fundamentally on psychological theory. Thus, in particular, work on culture and personality never challenged the universality of psychological theories themselves, such as psychoanalysis, but merely applied them in understanding levels of personality development or display of assumed personality traits in different cultures. Equally, work on individualism and collectivism was concerned with developing parameters that affected the level of development of particularly psychological attributes, but not the nature of the attributes themselves. Thus, for example, the prediction was made that self-esteem would be emphasized more in individualistic than in collectivist cultures (e.g., Triandis, 1989), but culture was not assumed to affect qualitatively the nature of self-processes or the relevance of self-esteem as a dimension of self in different cultural contexts (e.g., Heine, Lehman, Markus, & Kitayama, 1999). Finally, early comparative research in the sociocultural tradition approached cognitive processes as culturally dependent, but (at least in these earlier years) tended not to go beyond a contextually based view of cognition and claims of universal cognitive competencies in its implications for psychological theory. In the next section, consideration is given to some of the theoretical insights that underlay the turn from these
earlier traditions of cultural research to a more explicit cultural psychological stance.

**INSIGHTS AND CHALLENGES OF CULTURAL PSYCHOLOGY**

Cultural psychology represents an eclectic interdisciplinary perspective that has many roots. In many (but not all) cases, investigators associated with some of these traditions of research in cross-cultural psychology moved toward a cultural psychological outlook in response to the perceived limitations of some of the conceptual frameworks and goals of their earlier research. Thus, for example, many leading investigators associated with culture and personality, such as individuals who worked on the Six Culture project (B. B. Whiting & Whiting, 1975), as well as those associated with early work in the Vygotskian tradition on culture and thought, are at the forefront of contemporary work in cultural psychology. Equally, however, research in cultural psychology has drawn from disciplinary perspectives outside psychology. Thus, within psychological and cognitive anthropology, many investigators moved in a cultural psychological direction both from a concern that some of the early theories of culture and personality were parochial and needed to be formulated in more culturally grounded terms and from a sense that to understand culture requires attention to psychological and not merely anthropological considerations (e.g., Lutz & White, 1986; T. Schwartz, White, & Lutz, 1992; Shore, 1996; Strauss & Quinn, 1997). Thus, for example, arguments were made that to avoid an oversocialized conception of the person as merely passively conforming to cultural expectations required taking into account the subjectivity of intentional agents (e.g., Strauss, 1992). Equally, in another major research tradition, interest developed in cultural work within sociolinguistics. Thus, in work on language learning, it was recognized that individuals come to acquire not only the code of their language but also the meaning systems of their culture through everyday language use (e.g., Heath, 1983; P. Miller, 1986; Ochs & Schieffelin, 1984). Likewise, it came to be understood that everyday discourse contexts serve as a key context of cultural transmission.

**Key Conceptual Premises**

The perspective of cultural psychology is defined conceptually by its view of culture and psychology as mutually constitutive phenomena. From this perspective, cultural processes are seen as presupposing the existence of communities of intentional agents who contribute meanings and form to cultural beliefs, values, and practices. Equally, psychological functioning is seen as dependent on cultural mediation, as individuals participate in and come to acquire as well as create and transform the shared meaning systems of the cultural communities in which they participate. It is this monistic assumption of psychological and cultural processes as mutually dependent—not the type of methodology adopted—that is central to cultural psychology. Thus, for example, whether an approach employs qualitative versus quantitative methods or comparative versus single cultural analysis does not mark whether the approach may be considered as within the tradition of cultural as compared with cross-cultural psychology.

**Active Contribution of Meanings to Experience**

A core assumption underlying cultural psychology is linked to the insight of the Cognitive Revolution regarding the importance of meanings in mediating behavior (Bruner, 1990). It came to be understood that individuals go beyond the information given as they contribute meanings to experience, with these meanings in turn influencing individuals’ affective, cognitive, and behavioral reactions. The cultural implications of this cognitive shift were not appreciated immediately within psychology. Rather, as Bruner (1990) observes in presenting a brief history of the field, there was a tendency for many years to emphasize the autonomous self-construction of knowledge—indeed, independently of cultural transmission. The cultural implications of the Cognitive Revolution were also not apparent for many years because of the ascendance of information-processing accounts of cognition, which stress the automatic processing of information rather than the more active and creative processes of meaning-making. Nonetheless, although this image of an active constructivist agent for many years was not linked with cultural viewpoints, it formed an important theoretical basis for cultural psychology. The recognition that an act of interpretation mediates between the stimulus and the response established a theoretical basis upon which investigators could draw as they began to appreciate the cultural aspects of meanings and these meanings’ impact on thought and behavior.

**Symbolic Views of Culture**

The development within anthropology of symbolic views of culture (Geertz, 1973; Sahlins, 1976; Shweder & LeVine, 1984) also contributed to the emergence of cultural psychology in that it highlighted the need to go beyond the prevailing tendency to treat culture merely in ecological terms as an aspect of the objective environment. Ecological views of culture are critically important in calling attention to the adaptive
implications of features of the context (Bronfenbrenner, 1979). However, they also are limited in treating the context exclusively in objective terms, as presenting affordances and constraints that are functional in nature. In such frameworks, which have tended to be adopted in both mainstream and cross-cultural psychology, culture is seen as nonessential to interpretation or construction of reality. In contrast, within symbolic approaches, cultural systems are understood as bearing an indeterminate or open relationship to objective constraints rather than being fully determined by objective adaptive contingencies. Within symbolic approaches to culture, it is recognized that cultural meanings serve not merely to represent reality, as in knowledge systems, or to assume a directive function, as in systems of social norms. Rather, they are seen as also assuming constitutive or reality-creating roles. In this latter role, cultural meanings serve to create social realities, whose existence rests partly on these cultural definitions (Shweder, 1984). This includes not only cases in which culturally based social definitions are integral to establishing particular social institutions and practices (e.g., marriage, graduation, etc.) but also cases in which such definitions form a key role in creating psychological realities. Thus, it is increasingly recognized that aspects of psychological functioning (e.g., emotions) depend, in part, for their existence on cultural distinctions embodied in natural language categories, discourse, and everyday practices. For example, the Japanese emotional experience of *amae* (Doi, 1973; Yamaguchi, 2001) presupposes not only the concepts reflected in this label but also norms and practices that support and promote it. As an emotional state, *amae* involves a positive feeling of depending on another’s benevolence. At the level of social practices, *amae* is evident not only in caregiver-child interactions in early infancy (Doi, 1973, 1992), but also in the everyday interactions of adults, who are able to presume that their inappropriate behavior will be accepted by their counterparts in close relationships (Yamaguchi, 2001).

The significance of a symbolic view of culture for the development of cultural psychology was in its complementing the attention to meaning-making heralded by the cognitive revolution. It became clear that not only were meanings in part socially constructed and publicly based, but they also could not be purely derived merely by inductive or deductive processing of objective information. Culture, then, in this way became an additional essential factor in psychological explanation, beyond merely a focus on objective features of the context and subjective features of the person.

**Incompleteness Thesis**

Finally, and most critically, the theoretical grounding of cultural psychology emerged from the realization of the necessary role of culture in completion of the self, an insight that has been termed the *incompleteness thesis* (Geertz, 1973; T. Schwartz, 1992). This stance does not assume the absence of innate capacities or downplay the impact of biological influences as a source of patterning of individual psychological processes. However, without making the assumption that psychological development is totally open in direction, with no biological influences either on its initial patterning or on its subsequent developmental course, this stance calls attention to the essential role of culture in the emergence of higher-order psychological processes. Individuals are viewed not only as developing in culturally specific environments and utilizing culturally specific tools, but also as carrying with them, in their language and meanings systems, culturally based assumptions through which they interpret experience. Although there has been a tendency within psychology to treat this culturally specific input as noise that should be filtered out or controlled in order to uncover basic features of psychological functioning, the present considerations suggest that it is omnipresent and cannot be held constant or eliminated. Rather, it is understood that the culturally specific meanings and practices that are essential for the emergence of higher-order psychological processes invariably introduce a certain cultural-historical specificity to psychological functioning, as Geertz (1973) once noted:

> We are . . . incomplete or unfinished animals who complete or finish ourselves through culture—and not through culture in general but through highly particular forms of it. (p. 49)

From the present perspective, it is assumed that whereas an involuntary response may proceed without cultural mediation, culture is necessary for the emergence of higher-order psychological processes. Wertsch (1995) articulates this point:

> Cultural, institutional, and historical forces are ‘imported’ into individuals’ actions by virtue of using cultural tools, on the one hand, and sociocultural settings are created and recreated through individuals’ use of mediational means, on the other. The resulting picture is one in which, because of the role cultural tools play in mediated action, it is virtually impossible for us to act in a way that is not socioculturally situated. Nearly all human action is mediated action, the only exceptions being found perhaps at very early stages of ontogenesis and in natural responses such as reacting involuntarily to an unexpected loud noise. (p. 160)

Thus, for example, whereas involuntary physiological reactions may be elicited by situational events, whether they become interpreted and experienced in emotional terms depends in part on such input as culturally based theories.
regarding the nature, causes, and consequences of emotions, cultural routines for responding to emotions, natural language categories for defining emotions, and a range of other sociocultural processes.

This assumption of the interdependence of psychological and cultural processes represents the central idea of cultural psychology. Notably, the term cultural psychology was selected by theorists to convey this central insight that psychological processes need to be understood as always grounded in particular socio-cultural-historical contexts that influence their form and patterning, just as cultural communities depend for their existence on particular communities of intentional agents. The present considerations then lead to the expectation that qualitative differences in modes of psychological functioning will be observed among individuals from cultural communities characterized by contrasting self-related sociocultural meanings and practices.

Summary

Among the key conceptual insights giving rise to cultural psychology were the emergence of a view of the individual as actively contributing meanings to experience and an understanding of culture as a symbolic system of meanings and practices that cannot be explained exclusively in functional terms as mapping onto objective adaptive constraints. Crucial to the field’s development was that it also came to be recognized that higher-order psychological processes depend for their emergence on individuals’ participation in particular sociocultural contexts, and thus that culture is fundamental to the development of self.

Select Overview of Empirical Research in Cultural Psychology

The present section examines representative examples of empirical studies that embody this core insight regarding the cultural grounding of psychological processes, an insight that is central to the many traditions of work in cultural psychology (e.g., Cole, 1990, 1996; Markus et al., 1996; J. G. Miller, 1997; Shweder, 1990; Shweder et al., 1998). Although the overview presented here is necessarily highly selective and incomplete, it serves to illustrate ways in which cultural research is offering new process explanations of psychological phenomena as well as identifying fundamental variability in the forms that psychological processes assume.

Sociocultural Traditions of Research

The discussion here makes reference to findings from a diverse range of related viewpoints that have derived from the work of such major cultural theorists as Vygotsky (1978, 1981a, 1981b), Leontiev (1979a, 1979b), Luria (1979, 1981), Bakhtin (1986), and Bourdieu (1977) among others; their work is reflected in the many contemporary traditions of research in sociocultural psychology (e.g., Cole, 1988, 1990; Rogoff, 1990; Valsiner, 1988, 1989; Wertsch, 1979, 1991).

Central to theoretical work within this tradition is an emphasis on the mediated nature of cognition. Human behavior is seen as dependent on cultural tools or on other mediational means, with language recognized as one of the most central of these cultural supports. Embodying a broad lens, sociocultural approaches focus on understanding human activity at phylogenetic, historical, ontogenetic, and microgenetic levels, with cultural practices and activities viewed in terms of their place in larger sociopolitical contexts.

Considerable research in this area focuses on documenting how interaction with cultural tools and participation in everyday cultural activities leads to powerful domain-specific changes in thought. In work on everyday cognition (see review in Schliemann, Carraher, & Ceci, 1997), it has been shown, in fact, that everyday experiences can produce changes that represent an advance on those produced by schooling. For example, Scribner (1984) documented that individuals who work as preloaders in a milk factory and have less formal education than do white-collar workers are able to solve a simulated loading task more rapidly than do white-collar workers through using a more efficient perceptual solution strategy as contrasted with a slower enumerative approach. Likewise, in a growing body of research on expertise, it has been revealed, for example, that compared with novice adult chess players, child chess experts use more complex clustering strategies in organization and retrieval of chess information; they are also more proficient in their memory for chess pieces (Chi, Glaser, & Farr, 1998). Similar effects have equally been shown to occur in the solving of math problems among expert versus novice abacus users (Stigler, 1984).

It is important to note that sociocultural research is also providing new process models of the nature of everyday cognition. For example, recent research on situated cognition has challenged the view of learning as a distinct activity or as an end in itself set off from daily life and has emphasized its embeddedness in everyday activities and social contexts (Lave, 1988, 1993; Lave & Wenger, 1991). Research has revealed, for example, that in contrast to the forms of instruction that occur in formal school settings, learning in everyday situations is more oriented toward practical problems. In part as a result, individuals tend to be more motivationally involved in tasks and spontaneously to search for and generate more flexible task solutions in everyday situations than they do in formal school contexts.
Sociocultural research is also offering new answers to long-standing questions in psychological development. For example, work by Cole and his colleagues (Cole & Engstroem, 1995) has offered a novel process explanation of one of the central theoretical problems of cognitive development and language learning—explaining how individuals can obtain a more powerful conceptual structure if they do not already in some way possess that structure, or how qualitative and not merely quantitative developmental change may occur. In research conducted on the teaching of reading, it has been demonstrated that a range of mediational means, such as simplified reading materials, expert guidance, and so on, are available in everyday socialization contexts that support learning to read. Thus, it is noted that many of the structures entailed in the achievement of competence in reading exist between persons before they appear as individual competencies that may be manifest without this level of cultural support. Equally, in another example, evidence has been obtained to suggest that changes in children’s forms of social participation explain some of the marked advances in cognitive and social functioning that have been linked to the 5- to 7-year-old age shift among the schooled populations that have been subject to most study by cognitive developmental psychologists (Rogoff, 1996).

**Cultural Social Psychological Traditions of Cognitive Research**

Cultural social psychological work on cognition has a more recent history, tracing its origins most directly to early challenges to the universality of certain well-established attributional phenomena. It is giving rise to a rapidly growing experimental literature that points to qualitative cultural variation in basic modes of cognitive processing.

In some of the early groundbreaking work in this tradition, Shweder and Bourne (1984) challenged the completeness of contemporary social psychological theories of social attribution. It was documented that, as compared with European-Americans, Oriyan Hindu Indians place significantly greater emphasis in person description on actions versus abstract traits, with their person descriptions more frequently making reference to the context. Thus, for example, their investigation revealed that whereas European-Americans are more likely to describe a friend by saying she is friendly, Oriyan Indians are more likely to describe the friend by saying she brings cakes to my family on festival days. This type of cultural difference, it was observed, was not explicable in terms of the types of objective ecological or individual psychological factors that had been emphasized in previous studies, such as variation in schooling, literacy, socioeconomic status, linguistic resources, or capacities for abstract thought. Rather, the results appeared explicable only when taking into account cultural factors. In particular, the trends were demonstrated to reflect the contrasting cultural conceptions of the person and related sociocultural practices emphasized in Hindu Indian versus European-American cultural communities.

Subsequent cross-cultural developmental research on social attribution demonstrated that these types of cultural considerations give rise to cultural variation in the paths and endpoints of development (J. G. Miller, 1984, 1987). It was documented that whereas European-American children show an age increase in their reference to traits (e.g., she is aggressive) but no age-related change in their reference to contextual considerations, Hindu Indian children show an age increase in their references to the social context (e.g., there are bad relations between our families) but no age increase in their references to traits. More recently, this type of work has been further extended to understanding the development of individuals’ conceptions of mind, with cultural work calling into question claims that theory of mind understandings develop spontaneously toward an end point of trait psychology—and providing evidence that they proceed in directions that reflect the contrasting epistemological assumptions of local cultural communities (Lillard, 1998).

In other lines of work on social attribution and cognition, culturally based social psychological research is calling into question the universality of various attributional and cognitive tendencies long assumed to be basic to all psychological functioning, such as motives to maintain self-consistency or to emphasize dispositional over situational information. Thus, for example, it has been demonstrated that Japanese college students tend to maintain weaker beliefs in attitude-behavior consistency than do Australian college students (Kashima, Siegal, Tanaka, & Kashima, 1992), while being less prone than are Canadian college students to show cognitive dissonance biases—that is, tendencies to distort social perceptions to make them more congruent with behavior (Heine & Lehman, 1997). Also, relative to European-Americans, various East Asian populations have been documented to display greater sensitivity to situational information in object perception and less vulnerability to the fundamental attributional error (Ji, Peng, & Nisbett, 2000), a tendency to treat behaviors as correspondent with dispositions.

New lines of research in this area are also linking cultural views of the self and related cultural practices to variation in fundamental styles of cognitive processing, such as tendencies to privilege analytic versus dialectical epistemological stances. In one illustration of such a cultural difference, experimental research has demonstrated that American undergraduates tend to treat information in a polarized manner, as
seen in their considering scientific evidence as more plausible when it is presented alone rather than in conjunction with contradictory information (Peng & Nisbett, 1999). In contrast, Chinese undergraduates tend to process information in ways that involve greater acceptance of opposing viewpoints, as seen in their considering scientific evidence as more plausible when it is presented in conjunction with contradictory information rather than alone. Work of this type calls into question the primacy of analytic modes of thought in work in cognitive science, highlighting the salience of fundamentally different styles of cognitive processing in various East Asian cultural populations.

Self-Processes

In the area of the self-concept, psychological research is challenging the long-standing assumption that individuals spontaneously engage in self-maintenance strategies that are oriented toward self-enhancement, and that self-esteem is universally fundamental to psychological well-being. Open-ended attributional research on self-description, for example, has documented that whereas the open-ended self-descriptions of American adults emphasize positive attributes (Herzog, Franks, Markus, & Holmberg, 1998), those of Japanese adults emphasize either weakness or the absence of negative self-characteristics (e.g., I’m poor at math, I’m not selfish). Research has also documented that whereas the scores of Americans on measures of self-esteem tend to be higher than the scale midpoints—an indication of a tendency toward self-enhancement—those of Japanese persons tend to be at or slightly below the scale midpoint, an indication of a tendency to view the self as similar to others (Diener & Diener, 1995).

One of the most far-reaching implications of this type of research is that it calls into question the centrality of self-esteem in psychological functioning in various collectivist cultural communities, and it suggests that other types of self-processes may be more central in everyday adaptation in such contexts. In this regard, cross-national survey research has shown that self-esteem is more closely associated with life satisfaction in individualist than in collectivist cultures (Diener & Diener, 1995). In contrast, it is documented that a concern with maintaining relationship harmony shows a stronger relationship with life satisfaction in collectivist than in individualist cultures (Gabrenya & Hwang, 1996). These contrasting patterns of interrelationship are further documented to distinguish everyday socialization practices and to have important adaptive consequences. Thus, for example, Chinese as well as Japanese mothers tend to be more self-critical of their children’s academic performance than are American mothers (Crystal & Stevenson, 1991), with this stance implicated in the tendencies of Chinese and Japanese versus American mothers to place greater emphasis on their children’s expending effort toward self-improvement and having children who show superior levels of academic achievement (Stevenson & Lee, 1990).

Cultural research on the self is also challenging basic psychological theory in the domain of self-consistency. Social psychological theory has long assumed that individuals are inherently motivated to maintain a consistent view of the self and that such consistency is integral to psychological well-being. This stance is evident not only in classic theories of cognitive dissonance (Festinger, 1957), but also in more recent work on attribution. For example, work on self-verification has shown that individuals tend to prefer information that is consistent rather than inconsistent about themselves (Swann, Wenzlaff, Krull, & Pelham, 1992), as well as that autobiographical memories are structured in ways that preserve a consistent sense of self (Ross, 1989). Equally, work on psychological health has suggested that having an integrated and consistent view of self has adaptive value (Jourard, 1965; Suh, 2000).

A growing body of attributional research in Asian cultures, however, is suggesting that in these cultures the self tends to be experienced as more fluid than is typically observed in U.S. populations, with sensitivity to context valued. Work on self-description has demonstrated, for example, that the self-descriptions of Japanese but not of Americans tend to vary as a function of the presence of others (Kanagawa, Cross, & Markus, 2001). Likewise, experimental research has documented that cognitive dissonance effects tend not to be observed among Japanese as compared with Canadian populations (Heine & Lehman, 1997; Heine & Morikawa, 2000), and that consistency across situations shows a much weaker relationship to psychological well-being among Korean as compared with American populations (Suh, 2000).

Emotions

Emotions provide a particularly challenging area for cultural research because they are phenomena that involve not merely perceptions but also behavioral action tendencies and somatic reactions. They thus entail a biological grounding even as they also involve essential cultural components. Notably, as suggested in the following discussion, culture affects the expression of emotions and their form, as well as their role in mental health outcomes.

One important influence of cultural processes on emotion occurs in the degree of an emotion’s elaboration or suppression. It has been documented that cultural meanings and
practices affect the degree to which particular emotions are hypercognized (in the sense that they are highly differentiated and implicated in many everyday cultural concepts and practices) versus hypocognized (in that there is little cognitive or behavioral elaboration of them; Levy, 1984). Even universal emotions, it has been observed, play contrasting roles in individual experience in different cultural settings. For example, whereas in all cultures both socially engaged feelings (e.g., friendliness, connection) and socially disengaged feelings (e.g., pride, feelings of superiority) may exist, among the Japanese only socially engaged feelings are linked with general positive feelings, whereas among Americans both types of emotions have positive links (Kitayama, Markus, & Kurokawa, 2000).

Cross-cultural differences have also been observed in emotion categories as well as in individuals’ appraisals of emotions. Thus, variation in emotion concepts has been documented not only in the case of culturally specific categories of emotion, such as the concept of amaе among the Japanese (Russell & Yik, 1996; Wierzbicka, 1992), but also among such assumed basic emotions as anger and sadness (Russell, 1991, 1994). It has been shown that Turkish adults make systematically different appraisals of common emotional experiences than do Dutch adults, whose cultural background is more individualist (Mesquita, 2001). Thus, as compared with Dutch adults’ appraisals, Turkish adults tend to categorize emotions as more grounded in assessments of social worth, as more reflective of reality than of the inner subjective states of the individual, and as located more within the self-other relationship than confined within the subjectivity of the individual.

Notably, work on culture and emotions is also providing evidence of the open relationship that exists between physiological and somatic reactions and emotional experiences. For example, research has revealed that although Minangkabu and American men show the same patterns of autonomic nervous system arousal to voluntary posing of prototypical emotion facial expressions, they differ in their emotional experiences (Levenson, Ekman, Heider, & Friesen, 1992). Whereas the Americans tend to interpret their arousal in this type of situation in emotional terms, Minangkabu tend not to experience an emotion in such cases, because it violates their culturally based assumptions that social relations constitute an essential element in emotional experience.

Finally, important cultural influences on the mental health consequences of affective arousal are also being documented. For example, various somatic experiences—such as fatigue, loss of appetite, or agitation—that are given a psychological interpretation as emotions by European-Americans tend not to be interpreted in emotional terms but rather as purely physiological events among individuals from various Asian, South American, and African cultural backgrounds (Shweder, Much, Mahapatra, & Park, 1997). It is notable that such events tend to be explained as originating in problems of interpersonal relationships, thus requiring some form of nonpsychological form of intervention for their amelioration (Rosaldo, 1984; White, 1994).

**Motivation**

Whereas early cross-cultural research was informed exclusively by existing theoretical models, such as Rotter’s framework of internal versus external locus of control (Rotter, 1966), recent work is suggesting that motivation may assume socially shared forms. This kind of focus, for example, is reflected in the construct of secondary control, which has been identified among Japanese populations, in which individuals are seen as demonstrating agency via striving to adjust to situational demands (Morling, 2000; Morling, Kitayama, & Miyamoto, 2000; Weisz et al., 1984). Equally, work in India has also pointed to the existence of joint forms of control, in which the agent and the family or other social group are experienced as together agentic in bringing forth certain outcomes (Sinha, 1990).

In another related area of work on motivation, research is highlighting the positive affective associations linked with fulfillment of role-related responsibilities. This type of documentation notably challenges what has been the assumption informing much psychological theory—that behavior is experienced as most agentic when it is freely chosen rather than socially constrained and that social expectations are invariably experienced as impositions on individual freedom of choice. For example, behavioral research on intrinsic motivation has documented that Asian-American children experience higher intrinsic motivation for an anagrams task that has been selected for them by their mothers than for one that they have freely chosen (Iyengar & Lepper, 1999). In contrast, it is shown that European-American children experience greater intrinsic motivation when they have selected such a task for themselves.

Further support for this view that agency is compatible with meeting role expectations may be seen in attributional research, which has documented that Indian adults indicate that they would want to help as much and derive as much satisfaction in helping when acting to fulfill norms of reciprocity as when acting in the absence of such normative expectations (J. G. Miller & Bersoff, 1994). Such a trend contrasts with that observed among Americans, who assume that greater satisfaction is associated with more freely chosen
helping. These kinds of results challenge prevailing models of communal relationships, which assume that a concern with obligation detracts from a concern with being responsive to the others’ needs (Mills & Clark, 1982). They also challenge models of self-determination, which assume that internalization involves a greater sense of perceived autonomy (Deci & Ryan, 1985). Rather, it appears that in certain collectivist cultures individuals may experience their behavior as demanded by role requirements, while also experiencing themselves as strongly endorsing, choosing to engage in, and deriving satisfaction from the behavior.

In turn, work in the area of morality, relationships, and attachment highlights the need to expand current conceptualizations of motivation. For example, research in the domain of morality with both Hindu Indian populations (Shweder, Mahapatra, & Miller, 1990) as well as with orthodox religious communities within the United States (Jensen, 1997) has documented forms of morality based on concerns with divinity that are not encompassed in existing psychological theories of morality, with their exclusive stress on issues of justice, individual rights, and community (e.g., Kohlberg, 1971; Turiel, 1983). Furthermore, work on moralities of community have documented the highly individualistic cultural assumptions that inform Gilligan’s morality of caring framework (Gilligan, 1982), with its emphasis on the voluntaristic nature of interpersonal commitments. Cross-cultural work conducted on the morality of caring among Hindu Indian populations and cross-cultural work conducted utilizing Kohlbergian methodology have uncovered the existence of forms of duty-based moralities of caring that although fully moral in character, differ qualitatively in key respects from those explained within Gilligan’s framework (J. G. Miller, 1994, 2001b; Snarey & Keljo, 1991).

In terms of relationship research, a growing cross-cultural literature on attachment is suggesting that some of the observed variation in distribution of secure versus nonsecure forms of attachment arises at least in part from contrasting cultural values related to attachment, rather than from certain cultural subgroups’ having less adaptive styles of attachment than others. For example, research conducted among Puerto Rican families suggests that some of the greater tendency of children to show highly dependent forms of attachment reflects the contrasting meanings that they place on interdependent behavior (Harwood, Miller, & Irizarry, 1995). Thus, an analysis of open-ended responses of mothers revealed that compared with European-American mothers, Puerto Rican mothers viewed dependent behavior relatively positively as evidence of the child’s relatedness to the mother. Suggesting that present dimensions of attachment may not be fully capturing salient concerns for Puerto Rican mothers, this work further demonstrated that Puerto Rican mothers spontaneously emphasized other concerns—such as display of respect and of tranquility—that are not tapped by present attachment formulations.

In other research, recent work on attachment among Japanese populations highlights the greater emphasis on indulgence of the infant’s dependency and on affectively based rather than informationally oriented communication in Japanese versus American families (Rothbaum, Weisz, Pott, Miyake, & Morelli, 2000). In contrast to the predictions of attachment theory, however, such forms of parenting are not linked with maladaptive outcomes; rather, these parenting styles have positive adaptive implications, in fitting in with the cultural value placed on amae, an orientation that involves presuming upon another’s dependency and plays an important role in close relationships throughout the life cycle.

Such research has pointed out that the common finding that Japanese attachment more frequently takes what are considered as insecure or overly dependent forms reflects biases in present conceptions of attachment, which fail to take into account the concerns with interdependence in the Japanese context. Furthermore, it is noted that methodologically, the attachment research paradigm presents a separation context that is much rarer and thus much more stressful for Japanese than for American infants. Equally, it is suggested that (rather than treat the individual as the unit of attachment) to fully capture Japanese attachment-related concerns would require treating the individual-caregiver unit rather than the individual alone as the object of attachment assessment, with a focus on how well individuals can anticipate each other’s responses.

**Summary**

Work in cultural psychology is not only documenting cultural variability in psychological outcomes, but is also focused on uncovering respects in which this variation has theoretical implications in pointing to the implicit cultural underpinnings of existing psychological effects, as well as respects in which psychological theory needs to be conceptually expanded to account for culturally diverse modes of psychological functioning. We have seen specifically that cultural work is highlighting the culturally mediated nature of cognition through individuals’ participation in everyday cultural practices and use of culturally specific tools; such work has also uncovered the existence of contrasting culturally based cognitive styles, as well as extensive cultural variation in basic psychological processes involving the self, emotions, and motivation.


Challenges

Whereas there has been a dramatic increase in interest in cultural research in recent years, there nonetheless remains a sense in which cultural perspectives remain in a marginal position in the discipline. This may be seen in the stance adopted for cultural considerations most frequently—to be treated in a diversity sense, as relevant in explaining exceptions from what are assumed to be the general or default patterns—and for psychological theory and psychological generalizations commonly to be formulated without reference to cultural considerations. Concerns have also been raised about the quality of existing cultural research. In this regard, for example, criticisms have been made of the predictive power of recent work in social psychology based on the individualism-collectivism paradigm (e.g., Oyserman et al., 2002; Matsumoto, 1999). Charges have also been made that at least some contemporary cultural research is somewhat simplistic, if not stereotypical, and fails to capture the subtlety of particular cultural outlooks or to forward sophisticated contextually sensitive accounts of psychological functioning (J. G. Miller, 2002). Consideration here is given to ways to overcome some of these limitations and of how to further the promise and potential of cultural psychology to broaden and enrich basic psychological theory.

Process-Oriented Views of Culture

Social psychological traditions of cultural research in particular have been influenced by the views of culture held in the tradition of individualism-collectivism. This link has occurred largely because of the tremendous influence of the distinction introduced by Markus and Kitayama (1991) between interdependent versus independent self-construals. As introduced, this distinction embodied a set of dichotomous contrasts that were presented as characterizing a wide range of cultures, with the independent view of self characteristic of North American as well as many Western European cultural populations and the interdependent view of self characteristic of many Asian and African cultures. Thus, for example, whereas the independent self was defined as “separate from social context, bounded, unitary, stable, and focused on internal private features (abilities, thoughts, feelings)”, the interdependent self was defined in polar opposite ways as “connected with social context, flexible, variable, and focused on external public features (status, roles, relationships)” (Markus & Kitayama, 1991, p. 230).

When presenting this global dichotomy, Markus and Kitayama (1991) cautioned about drawing direct links between the type of general cultural schemas that they were identifying and individual self-representations. In this regard, for example, they noted respects in which individual self-concepts reflect a range of factors, including “gender, race, religion, social class, and one’s particular social and developmental history” (p. 230). They also stressed that both independent and interdependent orientations toward self are found in all societies, although these orientations take somewhat culturally specific forms. However, many social psychological investigators adopted the independent-interdependent self distinction in a nonnuanced manner that has ended up being somewhat stereotypical and simplistic in its characterization of culture and overly global in its views of how culture influences psychological phenomena.

Variation Between and Within Cultural Communities

In future research, it is critical to attend to the variation within different collectivist and individualist cultures and to the frequent overlap between cultural groups. Also, greater attention needs to be given to variation within culture that may be linked to social class, ethnicity, and experiences of discrimination or oppression.

In this regard, recent cultural research that has focused on varieties of individualism and collectivism has been valuable in that it points to psychological consequences linked to such variation. For example, research has suggested that Japanese individuals tend to approach social relations by focusing on the peer group, whereas Chinese individuals tend to adopt more of an authority-directed stance (Dien, 1999). It has also been documented that regional variation occurs in forms of individualism within the United States, such as the concerns with a culture of honor found in southern and western parts of the United States (Nisbett & Cohen, 1996). Notably, sociolinguistic and ethnographic research has also documented that within lower-class and working-class communities within the United States, there tends to be what has been characterized as a “hard defensive” type of individualism, which stresses adoption of abilities to cope in harsh everyday environments, in contrast with the “soft” individualism, which stresses the cultivation of individual uniqueness and gratification within middle-class contexts (Kusserow, 1999).

Attention to Cultural Practices

A limitation of current work on culture has also been the tendency to conceptualize culture purely in idealational terms. This type of stance is reflected in the reliance on scale measures of individualism-collectivism that have tended to portray cultures as systems of value orientations. Current conceptualizations have also been problematic in treating
cultural meanings as individual-difference attitudinal or personality variables—a stance that fails to recognize the multiple motives and personality factors that may be satisfied by given cultural practices, resulting in the lack of a one-to-one relationship between personality and culture.

In future research, it is important to recognize the complexity of cultural meanings. This means acknowledging culture not merely as knowledge about experience or as norms but also as constitutive propositions that serve to define and create social realities. It is equally critical to view cultural meanings as embodied in material artifacts, social institutions, and cultural tools, as well as expressed and communicated in everyday activities and practices. It is important that this type of stance is being recognized in the recent emphasis on the construct of cultural “selfways” or “custom complexes” that treat culture as including ideational and process-oriented elements that are mutually supportive (Greenfield, 1997; Markus, Mullally, & Kitayama, 1997; Shweder et al., 1998). It is important that the present type of concern also expands current understandings of culture in highlighting the frequently implicit and covert nature of cultural meanings, with many cultural commitments experienced by agents as facets of nature rather than of culture per se—a stance that contributes to their motivational force for individuals.

Finally, in future research, there is a need to integrate both symbolic and ecological views of culture. Symbolic views call attention to the arbitrary nature of cultural meanings and the extent to which they rest on nonrational commitments, rather than purely on functional considerations of utility. In turn, ecological approaches call attention to the material aspects of sociocultural systems, pointing to the need to take into account material constraints, resources, and issues of power and control in understanding sociocultural processes. In this regard, it is important to understand respects in which cultural and ecological effects are mutually influential. Thus, for example, research has shown not only that Puerto Rican mothers differ qualitatively in their views of attachment from European-American mothers, but also that both common and culturally specific effects of social class are observed in each case (Harwood et al., 1995).

Culturally Nuanced Models of Cultural Influences

One of the limitations of existing views of cultural influences on psychological processes has been the tendency to treat cultural differences as mapping onto personality differences. Ironically, this was one of the problematic aspects of early work in the tradition of culture and personality. As noted earlier, theorists criticized this work as presenting an overly socialized conception of the person as merely conforming to existing social norms and requirements. It also was criticized for positing an isomorphism between personality and individual motivation, and for failing to recognize the open-ended relationship between them. Notably, another problematic aspect of contemporary treatment of cultural influences has been the tendency to view cultural influences on psychological processes as highly generalized rather than as contextually dependent. This also appears related to a tendency to adopt a dispositional view of cultural effects as giving rise to global orientations that generalize across contexts or as uniform and noncontextually mediated forms of perceptual biases.

To develop more nuanced views of cultural influences on psychological functioning, it is critical, then, to attend both to individual differences and to cultural influences rather than to assume that individual differences map directly onto cultural differences. This involves recognizing the variation in individual attitudinal and personality measures within culture. It also involves taking into account that culture frequently has its impact on psychological processes through affecting individuals’ participation in normative contexts—with their varied normative requirements—rather than through affecting enduring psychological individual-difference variables.

Notably, to develop contextually sensitive views of cultural influences on psychological functioning requires taking into account the variation observed across contexts. Thus, for example, it cannot be assumed that because a concern with social relations and with a more interdependent view of self has been seen in collectivist cultures, individuals from collectivist cultures always give more weight to contextual effects than do individuals from individualist cultures. Rather, it must be recognized that culture influences the meanings given to contexts, and—depending on these meanings—there will be occasions in which individuals from collectivist cultures may show less variation in their judgments across contexts than do individuals from individualist cultures; or in some situations, observed cultural differences may even reverse (e.g., Cousins, 1989).

International and Interdisciplinary Approaches to Scholarship

In order to formulate approaches to culture that are dynamic and nuanced, it is essential for researchers to gain an understanding of the meanings and practices emphasized in the particular cultural communities in which they work. Such an understanding can be promoted through a range of processes, including collaboration with individuals from the culture, spending time in residence in the culture, learning the local language, or any combination of these. It is also likely that...
research that is informed by in-depth understandings of different cultural communities will become more common in psychology in the future. As the field becomes increasingly international and culturally diverse, investigators will be able to bring to their research cultural sensitivities and concerns contrasting with those presently dominating the discipline.

There is equally a need for future research on culture to become increasingly interdisciplinary, with investigators taking into account the conceptual and methodological insights of anthropological and sociolinguistic research traditions and avoiding the present insularity that results from ignoring or dismissing work from different disciplinary viewpoints. This neglect can yield findings considered to have relatively little importance from the perspective of the other traditions. However, given the overlap in concerns across these research traditions and given their contrasting strengths, greater interdisciplinary exchange can only serve to enhance progress in the field.

**Summary**

To enhance the quality of existing cultural research, it is important for investigators to go beyond dichotomous frameworks for understanding cultural differences, such as the global dimensions of individualism-collectivism. These types of frameworks fail to capture the complexity of individual cultural systems, portraying cultures in ways that are overly static, uniform, and isolated. Effort must be made to develop more nuanced views of culture through attending to everyday cultural activities and practices as well as to symbolic culture and ecological dimensions of contexts. Additionally, attention must be given both to individual differences and to cultural influences—the assumption should not be made that individual differences map directly onto cultural variation. Finally, the sensitivity of cultural research stands to be enhanced through researchers’ working to gain a greater understanding of the specific cultural communities which they study.

**CONCLUSION**

In conclusion, the present examination of culture in social psychological theory highlights the importance of recognizing that culture is part of human experience and needs to be an explicit part of psychological theories that purport to predict, explain, and understand that experience. What work in culture aims to achieve, and what it has already accomplished in many respects, is more than to lead investigators to treat psychological findings and processes as limited in generality. Rather than leading to an extreme relativism that precludes comparison, work in this area holds the promise of leading to the formulation of models of human experience that are increasingly culturally inclusive. By calling attention to the cultural meanings and practices that form the implicit context for existing psychological effects, and by broadening present conceptions of the possibilities of human psychological functioning, work in cultural psychology is contributing new constructs, research questions, and theoretical insights to expand and enrich basic psychological theory.

**REFERENCES**


Piaget, J. (1973). Need and significance of cross-cultural studies in genetic psychology. In J. W. Berry & P. R. Dasen (Eds.), *Culture...


Until recently, the study of personality was handicapped by the lack of a systematic taxonomy of constructs to represent individual differences. A confusing array of constructs and measures was available, and different measures of purportedly the same construct often showed little correspondence. This diversity hindered the development of a systematic understanding of individual differences. Recently, the situation began to change with emerging agreement about some of the major dimensions of personality. Broad traits such as neuroticism-stability, extraversion-introversion, and psychoticism-constraint are identified in most analyses of personality traits and part of most descriptive systems. There is also agreement about the way personality is organized. Models based on trait concepts assume that traits differ along a dimension of breadth or generalization and that traits are hierarchically organized, with global traits such as neuroticism subdividing into a set of more specific traits such as anxiousness and dependence (Goldberg, 1993; Hampson, John, & Goldberg, 1986).

Within this framework, attention has focused particularly on the five major factors as a parsimonious taxonomy of personality traits (Goldberg, 1990). Lexical analyses of the natural language of personality description (Digman, 1990; Goldberg, 1990) and subsequent psychometric studies of personality inventories (Costa & McCrae, 1992) have converged in identifying five broad factors typically labeled extraversion or surgency, agreeableness, conscientiousness, emotional stability versus neuroticism, and intellect, culture, or openness. It is widely assumed that this structure is transforming our understanding of personality and that the higher-order structure of personality is becoming more clearly delineated. Enthusiasm for the emergent structure, although understandable because it promises to bring coherence to a field characterized more by conceptual and theoretical debate than by substantive findings, tends to minimize confusions that still exist regarding the number and content of higher-order domains (Zuckerman, 1991, 1995, 1999; Zuckerman, Kulhman, Joireman, Teta, & Kraft, 1993) and nature of the assumed hierarchical arrangement of traits.

These problems remain unresolved despite numerous attempts to explicate personality structure, partly because the methods used incorporate subjective elements regarding choice of analytic strategies and data interpretation, and partly because personality concepts are inherently fuzzy, a factor that contributes to interpretive problems. In this chapter, we examine the contribution that behavioral genetic approaches can make to explicating the structure of personality and resolving issues of the number and content of
domains. The argument we advance is that an approach that contributes to understanding of the causes of trait covariation (as opposed to approaches that simply offer descriptions of trait covariation) offer an important perspective on these intractable taxonomic problems.

**DOMAIN DEFINITION: UNRESOLVED PROBLEMS WITH PHENOTYPIC STRUCTURE**

**Number of Domains**

Despite the dominance of the five-factor approach, disagreement still exists on the number of dimensions required to represent the higher-order structure of personality. Almagor, Tellegen, and Waller (1995), for example, suggested that five factors do not capture all dimensions of the natural language of personality because lexical analyses excluded terms that were evaluative or described temporary states such as mood. When they used an unrestricted set of terms, seven factors were identified. Five factors—Positive Emotionality, Negative Emotionality, Dependability, Agreeableness, Negative Emotionality, and Conscientiousness—corresponded to the five-factor dimensions of Extraversion, Neuroticism, Conscientiousness, Agreeableness, and Openness (negatively), respectively. The remaining factors were evaluative dimensions, Positive Valence and Negative Valence, which are not represented in the five-factor model. They concluded that the seven-factor model provides a better representation of lexical descriptions of personality. McCrae and John (1992) and Widiger (1993) refuted this conclusion, claiming that positive and negative valence factors could be assumed under the five factors.

Whereas Almagor, Tellegen, and Waller (1995) maintained that the five-factor model is too parsimonious, Eysenck (1991) suggested that it is not parsimonious enough. He argued that the five domains differ in abstractness and that the five dimensions could be accommodated within his three-factor model of Psychoticism, Extraversion, and Neuroticism because the Openness and Agreeableness domains are merely facets of Psychoticism. Studies examining the relationship between NEO-PI-R (Neurosis Extraversion Openness-Personality Inventory-Revised) and EPQ-R (Eysenck Personality Questionnaire-Revised), however, suggest that although the two scales overlap they assess unique aspects of personality (Avia et al., 1995; Draycott & Kline, 1995). These problems occur because the five factors, although assumed to be orthogonal, in fact intercorrelate. For example, correlations between NEO-PI-R Neuroticism and Conscientiousness domains and Extraversion and Openness to Experience domains are −.53 and .40 respectively (Costa & McCrae, 1992). These values raise the important issue of what degree of overlap or covariation between domains is tolerable. Whether these values are interpreted as unimportant or substantial depends largely on the investigator’s theoretical perspective.

**Domain Definition**

A related issue is lack of agreement on the lower-order traits that define each domain. Identification of an optimal set of lower-order traits has proved difficult (Costa & McCrae, 1998). Questions about whether a facet belongs to a proposed domain are raised when it consistently correlates with facets comprising another domain. For example, although Costa and McCrae (1992) report a moderate correlation of −.25 between total domain scores for Neuroticism and Agreeableness, the correlations between the Neuroticism facet Angry Hostility and Agreeableness facets Trust, Altruism, and Compliance are −.42, −.34, and −.49, respectively, and the correlation between Angry Hostility and the total Agreeableness domain score is −.47. Similarly, the correlation between the total Neuroticism domain score and the Agreeableness facet Trust is −.37. How this overlap is interpreted often forms the basis of many authors’ claims as to why their model provides the “correct” description of personality. As with the intercorrelations among domains, the interpretations placed on the findings are largely arbitrary.

This problem is also revealed by factor analyses of facet scales. Although factor loadings may conform to simple structure and the hypothesized five-factor pattern, some facets may have an appreciably lower loading than do the other facets defining a domain. This occurs with the NEO-PI-R Neuroticism facet of Impulsiveness. The correlations between Impulsiveness and the other Neuroticism facets range from .31 to .40 (Costa & McCrae, 1992). The median intercorrelation is .35, whereas the median intercorrelation among the other facets is .57. Findings such as these raise questions about the definition of domains and the possibility that additional domains are required to provide a comprehensive taxonomy.

The issue of establishing a coherent set of traits for each domain is related to the interpretation of each domain. Even within the five-factor approach there are differences in the interpretation of some domains, especially the domain that Costa and McCrae label Openness to Experience. They emphasize such defining characteristics as artistic, curious, original, and having wide interests (McCrae & Costa, 1985a, 1985b). In the NEO-PI-R, the factor is defined by ideas
(curious), fantasy (imagination), aesthetics (artistic), actions (wide interests), feelings (excitable), and values (unconventional). Others consider the domain to represent culture or intellect (Digman, 1990; Saucier & Goldberg, 1996). John and Srivastava (1999) maintained that the culture label (Passini & Norman, 1966) is not supported by evidence that traits referring to culture such as civilized, polished, dignified, foresighted, and logical load more highly on the conscientiousness factor. This leaves the alternative interpretation of intellect (Digman & Inouye, 1986; Goldberg, 1990). However, John and Srivastava (1999) concluded that the evidence supports the Costa and McCrae interpretation and that intellect is merely a component of a broader openness factor. This interpretation is supported by studies of the relationship between the domain and measures of cognitive ability. For example, the openness-intellect factor (Understanding, Sentience, Change, and Autonomy) based on the Personality Research Form (Jackson, 1984) correlates highly with measures of crystallized intelligence (e.g., verbal subscales) but less with measures of fluid ability (arithmetic and performance subscales; Ashton, Lee, Vernon, & Jang, 1999).

There are other, less easily resolved confusions about the definition and facet structure of other domains. For example, Conscientiousness according to Costa and McCrae (1995) consists of a single factor defined by competence, order, dutifulness, achievement striving, self-discipline, and deliberateness. Paunonen and Jackson (1996), however, question the unity of conscientiousness: “...The domain is best thought of as three separate, but somewhat overlapping, dimensions related to being (a) methodical and orderly, (b) dependable and reliable, and (c) ambitious and driven to succeed. Moreover, the amount of overlap among these three facets may not be high enough to justify their inclusion in an overall Conscientiousness measure” (p. 55).

The cluster of traits labeled impulsive–sensation seeking poses an even greater problem. Earlier, we discussed problems with the placement of impulsiveness within the five-factor model. The controversy, however, is deeper. For Zuckerman (1991, 1994), impulsivity and sensation seeking define a separate higher-order factor within an alternative five-factor structure. The factor resembles Eysenck’s psychoticism and Tellegen’s (1985) constraint. There appears, therefore, to be strong support for this domain. The five-factor model of Costa and McCrae, however, divides this factor into impulsivity and sensation seeking and assigns them to different domains. Impulsivity is considered part of neuroticism, an interpretation that is not shared by other conceptions of neuroticism, whereas sensation seeking is assigned to extraversion. This leads to similar problems with extraversion. Extraversion is defined by subsets of traits that differ across models. These subsets include such traits as sociability or affiliation, agency, activation, impulsive–sensation seeking, positive emotions, and optimism (Depue & Collins, 1999; Watson & Clark, 1997). Depue and Collins (1999) pointed out that most accounts of extraversion postulate two central features, an interpersonal engagement component consisting of affiliation or sociability and agency, and an impulsivity component that includes sensation seeking. They suggested that impulsive–sensation seeking arises from the interaction of extraversion and a second independent trait represented by Tellegen’s (1985) constraint. This proposal differs, however, from Eysenck’s model that places impulsivity in the psychoticism domain and Costa and McCrae’s proposal that it belongs to neuroticism. It also differs from Gray’s (1973, 1987; Pickering & Gray, 1999) model that considers impulsivity as assessed by questionnaire to be a blend of Eysenck’s higher-order dimensions of extraversion and psychoticism. It appears, therefore, that there are major unresolved definitional problems with most domains that compromise claims that the five-factor model provides a basic assessment framework (McCrae & Costa, 1986).

The existence of such basic uncertainty about the taxonomy of personality traits would seem to suggest that statements that the structure of personality is becoming delineated might be a little premature. Uncertainty about the relationships among traits is a major obstacle to constructing a theory of individual differences and clarification of these issues is essential for the field to advance. The ordering of traits within each domain forms the basis for developing theoretical explanations by defining relationships that require explanation. In effect, a descriptive taxonomy shapes subsequent research and theory development.

Approaches to Domain Definition

In response to these challenges, especially Paunonen and Jackson’s (1996) critique of conscientiousness, Costa and McCrae (1998) outlined six methodological approaches that can be used to demonstrate the unity of any domain: (a) item content analysis, (b) definitions of psychological opposites, (c) examination of empirical correlates, (d) interpreting secondary and tertiary factor loadings, (e) identification of equivalents in specialized languages and (f) case studies. Costa and McCrae (1998) applied these approaches to show that the Conscientiousness domain was unitary in nature. The limitation of these proposals is their reliance on an array of criteria that incorporate a subjective element. The proposal relies on a convergence of evidence across sets of traditional...
The problem with phenotypic analyses is their reliance on constructs that are by their nature fuzzy and imprecise. This is illustrated by the confusion noted about the components of extraversion (Depue & Collins, 1999; Watson & Clark, 1997). Conceptions of extraversion include sociability or affiliation (includes agreeableness, affiliation, social recognition, gregariousness, warmth, and social closeness), agency (surgency, assertion, endurance, persistence, achievement, social dominance, ascendency, ambitiousness), activation (liveliness, talkativeness, energy level, activity level, activity level), impulsive–sensation seeking (impulsivity, sensation seeking, excitement seeking, novelty seeking, boldness, risk taking, unreliability, disorderliness, adventurousness, thrill and adventure seeking, monotony avoidance, boredom susceptibility), positive emotions (positive affect, elatedness, enthusiasm, exuberance, cheerfulness, merriness, joviality), and optimism (Depue & Collins, 1999).

This list reveals the problems faced by attempts to delineate phenotypic structure. Not only does the content of extraversion differ across models, but the definition of each basic or lower-order trait may also differ across models and measures. Moreover, the meaning of putatively distinct traits overlaps so that facet traits defining a given domain shade into each other and into facet traits defining other domains. This fuzziness is probably an inevitable consequence of using natural language concepts that evolved to capture socially significant behaviors that are multidetermined. It adds to concerns that the taxonomies of phenotypic traits may not represent natural cleavages in the way behavior is organized nor reflect underlying etiological structures.

This fuzziness contributes to the considerable variability in personality phenotypes so that minor variations in measures and samples influence the number and contents of factors. The problem is compounded by the fact that many decisions about methodology and analytic strategies have an arbitrary component. More objective criteria are needed to guide decisions on the number of higher-order domains and the location of lower-order or basic traits within domains and to define a systematic set of basic traits. Phenotypic analyses are concerned primarily with describing trait covariation. This evokes the oft-voiced criticism of the five-factor approach—it is descriptive rather than explanatory. The basic problem of why traits are related to each other is not considered. An understanding of etiology of trait covariance, especially genetic etiology, would provide a conceptual foundation for current models. At each level of the trait hierarchy, traits and behaviors, including test items, could be grouped according to a shared etiology. Etiology would provide an additional criterion to supplement the usual psychometric criteria such as proposed by Costa and McCrae (1997) to guide decisions on the number and content of domains. Identification of a robust model of personality structure would be facilitated by evidence that a given phenotypic structure reflects the genetic architecture of personality traits. Unfortunately there are few studies of the genetic architecture underlying multiple personality traits compared to studies of phenotypic structure. Evidence that a given phenotypic structure parallels genotypic structure would support the validity and generalizability of the structure.

**HERITABILITY**

The foundation for an etiological understanding of personality structure and for a behavioral genetic approach is provided by evidence that genetic influences account for approximately 40–60% of the variance for virtually all personality traits, with most of the remaining variance being explained by nonshared environmental effects (Bouchard, 1999; Loehlin & Nicholls, 1976; Plomin, Chipeur, & Loehlin, 1990). The broad traits of extraversion and neuroticism have received most attention. The data from several twin studies yield heritability estimates of approximately 60% for extraversion and 50% for neuroticism. Loehlin (1992) also examined multiple personality scales organized according to the five-factor framework. Estimates of about 40% heritability were obtained for each domain. Subsequent studies using the NEO-PI-R yielded heritability estimates of 41% for neuroticism, 53% for extraversion, 41% for agreeableness, and 40% for conscientiousness (Jang, Livesley, Vernon, & Jackson, 1996; see also Bergeman et al., 1993; Jang, McCrae, Angleitner, Riemann, & Livesley, 1998). Nonadditive genetic effects accounted for 61% the variance in openness to experience.

Although the evidence points to a significant genetic component to personality traits, it has been suggested that traits could be divided into temperament traits that have a substantial heritable component and character traits that are largely environmental in origin. If this is the case and environmental factors give rise to distinct traits, the role of genetic criteria in clarifying trait structure would be limited. The evidence does not, however, support the proposal. Putatively characterological traits such as openness to experience are as heritable as so-called temperament traits. Moreover, molecular genetic studies have found significant allelic associations between so-called character traits such as cooperativeness and
self-directedness as assessed using the Temperament and Character Inventory and the 5-HTTLPR allele (Hamer, Greenberg, Sabol, & Murphy, 1999).

To date, a self-report measure of personality that has no genetic influence has not been identified (Plomin & Caspi, 1998). The qualification should be added that heritability studies have relied largely on self-report measures—alternative methods of assessment may yield different results. However, this was not the case with the few studies using other methods (Heath, Neale, Kessler, Eaves, & Kendler, 1992; Riemann, Angleitner, & Strelau, 1997). Riemann and colleagues (1997), for example, reported a twin study conducted in Germany and Poland that compared assessments of the five factors using self-report questionnaires with peer ratings. Estimates of heritability based on self-report were similar to those reported by other studies. The peer ratings also showed evidence of heritability, although estimates were lower than those obtained from self-reports. Multivariate genetic analyses showed that the same genetic factors contributed to self-report and peer ratings. These results suggest that findings of a heritable component to all self-report measures are likely to generalize to other methods of measurement.

Evidence of heritability alone, however, is not sufficient to justify the use of behavioral genetic criteria to clarify trait structure. It is possible that environmental factors that account for about 50% of the variance have a substantial effect on trait covariation. If this were the case, the finding that traits are genetically related would be of less value in clarifying personality structure. The evidence, however, suggests that the phenotypic structure of traits closely parallels the underlying genetic architecture (Livesley, Jang, & Vernon, 1998; Loehlin, 1987)—a point that is discussed in detail later in this chapter.

It should be noted, however, that information about heritability merely explains the variance in a single trait as opposed to the covariance between traits. Such information has limited value in explicating personality structure. As Turkheimer (1998) argued, all individual differences in behavior are heritable and “...the very ubiquity of these findings make them a poor basis for reformulating scientists’ conceptions of human behavior” (p. 782). Nevertheless, information on heritability forms the foundation for understanding of the etiology of personality. The major contribution of behavior genetics to understanding personality structure, however, comes from multivariate genetic analyses that elucidate the genetic structure underlying multiple traits (Carey & DiLalla, 1994). Multivariate analyses extend univariate analysis of the genetic and environmental influences on a trait to evaluate genetic and environmental components of the covariation between two or more traits (DeFries & Fulker, 1986). It is this extension that promises to contribute to personality theory by explicating the etiological basis for trait covariance by evaluating the degree to which different traits are influenced by the same genetic and environmental factors. This issue is central to resolving some of the problems of personality description and structure.

**THE ETIOLOGICAL BASIS OF COVARIANCE**

The phenotypic covariation between two traits may be due to pleiotropy—that is, the degree to which the traits share a common genetic influence, environmental effects common to both traits, or both. The degree to which two variables have genetic and environmental effects in common is indexed by genetic ($r_G$) and environmental correlation coefficients ($r_E$). These statistics are interpreted as any other correlation coefficient and they may be subjected to other statistical procedures such as factor analysis (Crawford & DeFries, 1978). Genetic and environmental correlation coefficients are readily estimated from data obtained from monozygotic (MZ) and dizygotic (DZ) twin pairs.

The calculation of the genetic correlation is similar to that used to estimate the heritability of a single variable. A higher within-pair correlation for MZ twins than for DZ twins suggests the presence of genetic influences because the greater similarity is directly attributable to the twofold increase in genetic similarity in MZ versus DZ twins. In the multivariate case, a common genetic influence is suggested when the MZ cross-correlation (the correlation between one twin’s score on one of the variables and the other twin’s score on the other variable) exceeds the DZ cross-correlation.

The phenotypic correlation ($r_p$) between two variables (traits), $x$ and $y$, is expressed by the following equation:

$$r_p = (h_x \cdot h_y \cdot r_g) + (e_x \cdot e_y \cdot r_e)$$

where the observed or phenotypic correlation, ($r_p$), is the sum of the extent to which the same genetic ($r_g$) and/or environmental factors ($r_e$) influence each variable, weighted by the overall influence of genetic and environmental causes on each variable ($h_x$, $h_y$, $e_x$, $e_y$, respectively). The terms $h$ and $e$ are the square roots of heritability and environmental effect ($h^2$ and $e^2$) for variables $x$ and $y$, respectively.

It should be noted that a genetic correlation describes *statistical pleiotropism*—that is, the extent to which allelic effects on trait predict allelic effects on the other trait. As Carey (1987) pointed out, statistical pleiotropism is not to be confused with *biological pleiotropism* in which two variables...
share the same loci. Unlike statistical pleiotropism, biological pleiotropism unequivocally links actual genes to behavior.

PHENOTYPIC STRUCTURE AND GENETIC ARCHITECTURE OF PERSONALITY

A critical issue for understanding the etiological structure of personality and for the use of multivariate genetic analyses to clarify personality structure is the degree to which the phenotypic organization of traits reflects an underlying biological structure as opposed to the influence of environmental factors. The evidence indicates that the phenotypic structure of traits closely resembles the underlying genetic architecture and to a lesser degree environmental structure. The evidence also suggests that environmental factors do not appreciably influence trait covariation. These conclusions are based on comparisons of the factors extracted from matrices of phenotypic, genetic, and environmental correlations computed among traits comprising a given model or measure.

In one of the earliest studies of this kind, Loehlin (1987) analyzed the structure of item clusters from the California Psychological Inventory (CPI; Gough, 1989) in samples of MZ and DZ twins. Three matrices were derived that represented the covariance among different traits due to genetic, shared environmental, and nonshared environmental factors. When these matrices were examined with factor analysis, four factors emerged from analyses of genetic covariance that could be interpreted as representing Neuroticism, Extraversion, Openness, and Conscientiousness (few items related to the fifth factor, Agreeableness, are included in the CPI; see McCrae, Costa, & Piedmont, 1993). Analysis of shared environmental effects yielded two factors: family problems and masculinity-femininity. The former is not an aspect of personality per se, and the latter is probably an artifact of the exclusive use of same-sex twins (Loehlin, 1987). It should be noted, however, that shared environmental effects make relatively little contribution to the variance of personality traits. Hence, the important finding is the structure of nonshared environmental effects. Analysis of the nonshared environmental covariance matrix yielded three interpretable factors that resembled Neuroticism, Extraversion, and Conscientiousness. Thus, the structure of nonshared environmental influences largely mirrored genetic influences. This is not an isolated finding: Livesley et al. (1998) found similar structures in genetic and nonshared environmental components of traits related to personality disorder.

Livesley and colleagues (1998) examined the congruence of genetic and phenotypic factor structures and compared phenotypic structure across samples of personality disordered patients and two samples recruited from the general population. The clinical sample consisted of 602 patients with personality disorder. The general population samples consisted of 939 volunteer general population participants and 686 twin pairs. The twin sample allowed the computation of matrices of genetic and environmental correlations that could be compared against the phenotypic structures from all three samples. Personality was assessed with the Dimensional Assessment of Personality Pathology (DAPP; Livesley & Jackson, in press). This measure assesses 18 traits underlying personality disorder diagnoses that were identified in previous studies using a combination of clinical judgments, rational methods, and psychometric procedures (Livesley, 1986; Livesley, Jackson, & Schroeder, 1992).

Phenotypic correlations were computed in all three samples separately, and genetic and environmental correlations were computed on the twin sample. The phenotypic, genetic, and environmental correlation matrices were subjected to separate principal components analyses with rotation to oblimin criteria. Phenotypic structure was similar across all samples. Four factors were extracted from all five matrices (see Tables 3.1 and 3.2).

The first factor, Emotional Dysregulation, represents unstable and reactive affects and interpersonal problems. The factor resembled neuroticism as measured by the NEO-PI-R (Costa & McCrae, 1992; Schroeder, Wormworth, & Livesley, 1992) or the Eysenck Personality Questionnaire (EPQ; Jang & Livesley, 1999) and the DSM-IV diagnosis of borderline personality disorder. The second factor, Dissocial Behavior, was negatively correlated with NEO-PI-R Agreeableness. It described antisocial traits and resembled the DSM-IV Cluster B

<table>
<thead>
<tr>
<th>TABLE 3.1 Rotated Principal Component Factor Loadings: DAPP-BQ Dimensions (clinical sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimension</td>
</tr>
<tr>
<td>Submissiveness</td>
</tr>
<tr>
<td>Cognitive Dysregulation</td>
</tr>
<tr>
<td>Identity Problems</td>
</tr>
<tr>
<td>Affective Instability</td>
</tr>
<tr>
<td>Stimulus Seeking</td>
</tr>
<tr>
<td>Restricted Expression</td>
</tr>
<tr>
<td>Oppositionality</td>
</tr>
<tr>
<td>Intimacy Problems</td>
</tr>
<tr>
<td>Anxiousness</td>
</tr>
<tr>
<td>Conduct Problems</td>
</tr>
<tr>
<td>Social Avoidance</td>
</tr>
<tr>
<td>Insecure Attachment</td>
</tr>
</tbody>
</table>
antisocial personality diagnosis, Eysenck’s Psychoticism, and Zuckerman’s Impulsive–Sensation Seeking. The third factor, labeled Inhibition, was defined by intimacy problems and restricted expression of inner experiences and feelings. The factor correlated negatively with NEO-PI-R and EPQ Extraversion and resembled the DSM-IV avoidant and schizoid personality disorders. The fourth factor, Compulsivity clearly resembled NEO-PI-R Conscientiousness and DSM-IV obsessive-compulsive personality disorder. The loadings derived from the phenotypic correlation matrices were remarkably similar: Congruence coefficients ranged from .94 to .99.

The congruence between factors extracted from the phenotypic and nonshared environmental matrices were also high at .96, .93, .90, .93, and .97 for N, E, O, A, and C, respectively. The congruence of the nonshared environmental factors and normative structure was even higher at .96, .93, .90, .93, and .97 for N, E, O, A, and C, respectively.

The interesting feature of these results is not only that phenotypic structure resembles genetic structure, but also that the structure of environmental effects is similar to the genetic structure. Plomin, DeFries, and McClearn (1990) noted that across a range of studies, “the structure of genetic influences seems to be similar to the structure of [nonshared] environmental influences” (p. 236). They added that this is surprising: “Most of us would probably predict different patterns of genetic and environmental influences” (p. 236). Recently, however, it has been suggested that genetic factors are more important than are environmental influences in shaping trait structure because the resemblance of the structure of nonshared environmental effects to the observed structure of traits may be artifactual (McCrae, Jang, Livesley, Riemann, & Angleitner, in press).

Nonshared environmental effects are usually estimated as a residual term that may include systematic bias such as that introduced by implicit personality theory. Passini and Norman (1966) demonstrated this bias by asking students to rate the personalities of complete strangers. Although each rating was presumably a guess, a clear pattern to the ratings was found. Students who assumed that strangers were talkative also assumed that they were sociable and cheerful. Across a range of targets, these associations defined an Extraversion

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Genetic Factors</th>
<th>Environmental Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Submissiveness</td>
<td>0.91</td>
<td>0.76</td>
</tr>
<tr>
<td>Cognitive Dysregulation</td>
<td>0.66</td>
<td>0.70</td>
</tr>
<tr>
<td>Identity Problems</td>
<td>0.84</td>
<td>0.68</td>
</tr>
<tr>
<td>Affective Lability</td>
<td>0.69</td>
<td>0.70</td>
</tr>
<tr>
<td>Restricted Expression</td>
<td>0.45</td>
<td>0.78</td>
</tr>
<tr>
<td>Oppositionality</td>
<td>0.74</td>
<td>0.54</td>
</tr>
<tr>
<td>Anxiousness</td>
<td>0.96</td>
<td>0.86</td>
</tr>
<tr>
<td>Suspiciousness</td>
<td>0.61</td>
<td>0.45</td>
</tr>
<tr>
<td>Social Avoidance</td>
<td>0.76</td>
<td>0.69</td>
</tr>
<tr>
<td>Narcissism</td>
<td>0.60</td>
<td>0.47</td>
</tr>
<tr>
<td>Insecure Attachment</td>
<td>0.64</td>
<td>0.69</td>
</tr>
<tr>
<td>Stimulus Seeking</td>
<td>0.61</td>
<td>0.81</td>
</tr>
<tr>
<td>Callousness</td>
<td>0.88</td>
<td>0.66</td>
</tr>
<tr>
<td>Rejection</td>
<td>0.82</td>
<td>0.65</td>
</tr>
<tr>
<td>Conduct Problems</td>
<td>0.75</td>
<td>0.69</td>
</tr>
<tr>
<td>Restricted Expression</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>Intimacy Problems</td>
<td>0.93</td>
<td>0.75</td>
</tr>
<tr>
<td>Compulsivity</td>
<td>0.93</td>
<td>0.85</td>
</tr>
<tr>
<td>Suspiciousness</td>
<td>0.45</td>
<td></td>
</tr>
</tbody>
</table>
factor. Factor analysis of the ratings yielded the familiar five factors. Some researchers concluded from such studies that trait structure merely reflects the effects of semantic biases on personality perceptions (Shweder, 1975). Ratings of strangers must contain bias due to implicit personality theory because they cannot be influenced by the true personalities of the targets. It is also likely that self-reports and ratings of well-known targets incorporate a similar bias. For example, two observers may agree that a person is sociable but disagree on the extent of his or her sociability. The observer assigned a higher rating for sociability is also likely to assign a higher rating for cheerfulness and talkativeness. Thus, part of the covariance of these traits may be attributable to systematic biases in person perception that lead to correlated errors in individual judgments. If this is the case, similarities in structure between genetic covariance and nonshared environmental covariance could reflect the biasing effects of implicit personality theory on the latter.

To test for this bias, self-report twin data were supplemented with cross-observer correlations on the NEO-PI-R. This allowed the computation of two matrices of nonshared environmental covariance. The first estimated the covariance due to implicit personality theory bias alone. Factorial analysis of this matrix yielded the familiar five factors. Comparison with normative structure yielded congruence coefficients of .81, .45, .81, .89, and .85 for Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness, respectively. The second matrix of nonshared environmental covariance estimated was free from systematic bias. Factor analysis of this “unbiased” matrix with targeted rotations to the normative NEO-PI-R factors produced low congruence coefficients at .53, .68, .22, .61, and .80 for Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness, respectively. The second matrix yielded two factors. The first resembled a broad form of Conscientiousness with salient loading of the facets Activity, Order, Dutifulness, Achievement Striving, Self-Discipline, and (low) Impulsiveness. The second factor was defined by the facets Warmth, Gregariousness, Positive Emotions, Openness to Feelings, Altruism, and Tender-Mindedness. This combination of Extraversion and Agreeableness facets resembles the Love axis of the Interpersonal Circumplex (Wiggins, 1979). The other interpersonal axis—Dominance—does not appear to be influenced by the nonshared environment. Assertiveness did not load on either factor.

These results suggest that when the conventional estimates of nonshared environmental covariances are decomposed into implicit personality theory bias and true nonshared effects, much of the resemblance to the five-factor structure appears attributable to bias. Overall, these studies point to the conclusion that genetic factors are largely responsible for the observed pattern of trait covariation.

THE HIERARCHICAL STRUCTURE OF PERSONALITY

Beyond problems with the content of personality taxonomies, there are also uncertainties about the nature of the proposed hierarchical structure of traits and the relationship between higher- and lower-order traits. Factor analytic studies provide consistent evidence that specific traits are organized into more global entities. Lexical studies also show that natural language reflects this structure. Substantial agreement exists among individuals in judgments of trait breadth (Hampson et al., 1986). Despite this evidence, the nature and origins of the hierarchy are unclear. This is clearly a problem that requires explanation.

Fundamental differences exist among models on the way the personality hierarchy is conceptualized. The lexical approach seems to consider the higher-order domains to be lexical categories that impose structure on personality descriptors by organizing them into clusters that are not necessarily discrete or equally important (Saucier & Goldberg, 1996). The lexical structure “provides a framework for description, but not necessarily for explanation” (Saucier & Goldberg, 1996, p. 24–25). Saucier and Goldberg also asserted that “as a representation of phenotypes based on natural language, the Big Five structure is indifferent and thus complementary to genotypic representations of causes, motivations, and internal personality dynamics” (p. 42). The higher-order terms do not appear, therefore, to have any significance beyond that of description.

Traits psychologists, including other five-factor theorists, make different assumptions. For Allport (1961), a trait is “a neuropsychic structure” (p. 347) and therefore an explanatory concept. Eysenck also adopted this approach: Traits have heritable biological basis. Similarly, the five-factor model assumes that traits are “endogenous basic tendencies” with a substantial heritable component (McCrae & Costa, 1996, p. 72). For Eysenck and Costa and for McCrae, traits are explanatory as well as descriptive. In contrast to the lexical approach, the five-factor model assumes that domains are equally important and equal in breadth.

Assumptions that trait theories make about the psychobiological basis for the higher-order domains initially created uncertainty about the status of the lower-order traits. Most research effort has been directed toward understanding higher-order factors and little attention has been paid to parsing these domains into more specific components. Until
recently, it was unclear whether the lower-order traits were merely facets of the higher-order traits or distinct entities with their own etiology. The use of the term *facet* to describe the lower-order traits, a convention adopted by Costa and McCrae, implies that they are merely exemplars or components of a more fundamental global trait. In this sense, the facet traits can be understood in terms of the domain sampling approach used in test construction in which facets are merely arbitrary ways to subdivide global traits to ensure adequate domain sampling. Identification of general genetic factors that have a broad influence on personality phenotypes also raises questions about the significance of the lower-order or facet traits—in particular, whether these traits are heritable simply because of their association with the broader domains or whether they are also subject to specific genetic influences. Clarification of this issue is critical to constructing an explanatory account of personality structure.

**Heritability of Lower-Order Traits**

If lower-order traits are only subcomponents of broader traits, all variance in a facet apart from error variance should be explained by the variance in the global trait. Recently, however, behavioral genetic research has suggested that lower-order traits have a distinct heritable component (Jang et al., 1998; Livesley et al., 1998). These studies estimated whether lower-order traits have a unique genetic basis when the heritable component of higher-order traits is removed from them. Jang and colleagues (1998) partialled out all of the common variance due to each of higher-order Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness scales from the 30 facet scales of the NEO-PI-R. When the residual variances on the facets were subjected to heritability analyses, a substantial genetic influence remained. Additive genetic effects accounted for 25 to 65% of the reliable specific variance, with most heritabilities ranging from .20 to .35 (see Table 3.3).

When these values were corrected for unreliability, the values increased to the usual range observed for personality traits. The implication is that these traits are not merely facets of more general traits, but rather distinct heritable entities.

A similar approach was used to study the residual heritability of the 18 traits underlying personality disorder (Livesley et al., 1998). Factor scores were computed for the four factors described previously. A standardized residual score for each scale was computed by regressing the four factor scores on each of the 18 basic traits. Monozygotic twin correlations were higher that the dizygotic twin correlations for all 18 traits. Estimates of the heritability of the residual trait scores showed substantial residual heritability for 11 of the 18 basic traits that ranged from .26 for Intimacy Problems to .48 for Conduct Problems.

These studies, in contrast to studies of phenotypic structure, point to the significance of the lower-order traits. Although these traits have tended to be neglected in personality research, they appear to be important for understanding personality. This suggests that a bottom-up approach to personality structure would provide additional information to complement that provided by the traditional top-down approach of the three- and five-factor models that identify the higher-order domains first and then seek to define an appropriate complement of facet traits. Before considering these issues in greater depth, it is important to recognize a limitation of the methods used. The regression method does not model genetic effects directly, and the results need to be replicated using multivariate genetic analyses. This introduces another feature of behavioral genetic analyses that is pertinent to...
understanding the genetic basis of personality: the use of pathways models to evaluate competing models of personality.

**Independent and Common Pathways Models**

In heritability analyses, components of variance are estimated by fitting models to the observed covariance matrices. In the univariate case, the heritability of a variable is estimated by comparing the similarity (estimated by Pearson’s $r$) of MZ to DZ twins. In the bivariate case, common genetic influences are suggested when the MZ cross-correlation exceeds the DZ cross-correlation used to compute the genetic correlation, $r_G$. The multivariate extension of this idea is found in two general classes of path analytic models that are pertinent to personality research: independent and common pathways models (see Figures 3.1 and 3.2; McArdle & Goldsmith, 1990; Neale & Cardon, 1992). The independent pathway model specifies direct links from one or more genetic and environmental influences common to each variable and unique genetic and environmental effects to each variable. The common pathways model is a more stringent version of the independent pathways model. The primary difference between the two models is that the common pathways model postulates that of the covariation in a set of variables is mediated by a single latent variable that has its own genetic and environmental basis. Both models provide the opportunity to examine variance specific to each variable—that is, each lower-order trait. Factor analytic studies of personality have been concerned with reducing the covariance between lower-order traits to fewer factors. Residual variance specific to each trait is neglected. Biometric path models applied to twin data decompose this variance into etiological components. This makes it possible to evaluate the significance of these specific traits.

These models offer the opportunity to evaluate the hierarchical structure of personality by comparing the fit of the two models to the same data set. The common pathways model is the biometric equivalent to the traditional model of exploratory factor analysis used to delineate the phenotypic structure of traits. As applied to each of the five-factor domains, the model postulates a single latent factor for each domain that mediates the effects of genetic and environmental effects on each lower-order trait. In the case of NEO-PI-R Neuroticism, a latent variable of neuroticism is hypothesized through which genetic and environmental factors influence the six facets of Anxiety, Hostility, Depression, Self-Consciousness, Impulsivity, and Vulnerability. In contrast, the independent pathways model postulates direct genetic and environmental effects on each facet trait. The fits of these models provide an opportunity to evaluate different conceptions of personality structure. If the common pathways model provides the best fit, the implication is that the hierarchical structure of personality arises from the effects of higher-order factors that have a genetic and environmental basis. The task is then to explain how this entity differs from lower-order or facet traits and the role it plays in the formation of the

---

**Independent Pathways Model**

$G =$ additive genetic effects common to all variables  
$E =$ nonshared environmental effects common to all variables  
$g =$ additive genetic effects unique to each variable  
$e =$ nonshared environmental effects unique to each variables

---

**Figure 3.1** Independent pathways model; $G =$ additive genetic effects common to all variables, $E =$ nonshared environmental effects common to all variables, $g =$ additive genetic effects unique to each variable, and $e =$ nonshared environmental effects unique to each variable.
The Hierarchical Structure of Personality

Common Pathways Model

- \( G \) = additive genetic effects common to all variables
- \( E \) = nonshared environmental effects common to all variables
- \( g \) = additive genetic effects unique to each variable
- \( e \) = nonshared environmental effects unique to each variable

Hierarchical Structure of Personality

- Anxiety
- Hostility
- Depression
- Self-conscious
- Impulsivity
- Vulnerability

Figure 3.2 Common pathways model; \( G = \) additive genetic effects common to all variables, \( E = \) nonshared environmental effects common to all variables, \( g = \) additive genetic effects unique to each variable, and \( e = \) nonshared environmental effects unique to each variable.

hierarchy. If the independent pathways model provides the best fit, however, the implication is that the higher-order constructs of phenotypic analyses do not reflect the effects of a phenotypic entity, but rather the pleiotropic action of the genes shared by all lower-order or facet traits that define the domain. Under these circumstances, the task is to elucidate the mechanisms that lead to trait clusters. Regardless of which model provides the best fit to the data, a useful feature of both models is that the magnitude of the path coefficients between each facet scale and the common genetic factor or latent variable along with information on the magnitude of genetic and environmental influences unique to each facet provides the basis for determining which facets should be grouped together within the taxonomy.

Five-Factor Model

Jang and colleagues (in press) fit common and independent pathways models to evaluate the coherence of the five domains assessed with the NEO-PI-R. The models were applied separately to a sample of 253 identical and 207 fraternal twin pairs from Canada and 526 identical and 269 fraternal pairs from Germany. The two samples made it possible to examine the universality of the etiological basis for personality structure by investigating whether the same genetic and environmental factors influenced personality traits in the two samples and whether they had similar effects.

For each sample, a single-factor common pathways model and a series of independent pathway models specifying variable numbers of genetic and nonshared environmental factor were fit to the six facets defining each domain. Shared environmental effects were omitted from the models because their effects were minimal. For each domain, the best fit was obtained with an independent pathways model. Table 3.4 illustrates the findings for the Neuroticism domain. An independent pathways model specifying two genetic factors and two nonshared environmental factors provided the most satisfactory explanation of the covariance between the six Neuroticism facets in the two samples. In both samples, the first genetic factor was marked by the Angry Hostility facet and, to a lesser extent, Anxiety. The second factor influenced all facets except Angry Hostility and Impulsivity. The depression facet had the highest loading in both samples.

In addition to demonstrating that the independent pathways model provided the best fit, these findings also suggest that the broad domains of personality are nonhomogeneous. This raises important questions about the factors that account for the apparent hierarchical structure of personality traits and the nature and conceptual status of the higher-order
dimensions. These conclusions are, however, based on a single study using only a single measure of personality. Replication is clearly needed, given the results’ significance for understanding trait structure. The conclusions are, however, similar to those drawn from a study of personality disorder traits (Livesley & Jang, 2000).

**Personality Disorder Traits**

Livesley and Jang (2000) investigated the etiological structure of personality disorder by fitting independent and common pathways models to the 18 lower-order traits of personality disorder assessed by administering the DAPP to a volunteer sample of 686 twin pairs. Each trait consists of two or more specific traits so that a total of 69 specific traits define the 18 basic traits. The 18 traits in turn define four higher-order factors. Thus the DAPP system incorporates three levels of construct (higher-order factors, lower-order traits, and specific traits) whereas the NEO-PI-R has only two levels (domains and facets). This makes it possible to explore the genetic architecture of personality in more detail. For example, the basic trait of Anxiousness is defined by four specific dimensions: trait anxiety, guilt proneness, rumination, and indecisiveness. Each basic trait represents a single phenotypic factor. If personality is inherited as a few genetic dimensions represented by the four higher-order factors, a single genetic dimension should underlie each basic trait that is shared by other traits constituting the higher-order factor. Evidence of a genetic effect specific to each trait would be provided by evidence that the 18 basic traits are composed of two or more genetic dimensions.

A one-factor common pathways model did not provide a satisfactory fit for any of the 18 basic traits. On the other hand, an independent pathways model postulating a single genetic dimension explained the covariation among specific traits for 12 of the 18 basic trait scales: Anxiousness, Cognitive Dysregulation, Compulsivity, Conduct Problems, Identity Problems, Insecure Attachment, Intimacy Problems, Oppositionality, Rejection, Stimulus Seeking, Submissiveness, and Suspiciousness. The results of model fitting for illustrative scales are provided in Table 3.5. For three of these scales, Intimacy Problems, Rejection, and Stimulus Seeking, the common genetic dimension accounted for little of the variance for one or more of the specific trait scales, indicating that a specific genetic factor influenced these traits. Two genetic dimensions were found to underlie four scales: Affective Lability, Narcissism, Restricted Expression, and Social Avoidance. Three common genetic dimensions contributed to Callousness (see Table 3.5).

Multivariate analyses of normal and disordered personality traits suggest that multiple genetic and environmental factors influence the covariant structure of traits. They also confirm the findings of the regression analyses that many lower-order traits are influenced by one or more genetic dimensions specific to those traits. Finally, in both sets of analyses, the common pathways model did not provide a better fit to the data than did the independent pathways model. This suggests that the general genetic dimensions found by Livesley and colleagues (1998) and others by factor analyzing matrices of genetic correlations do not influence each trait through a latent phenotypic variable, but rather exert a direct influence on each trait.

### IMPLICATIONS FOR PERSONALITY STRUCTURE

The studies described in the previous section reveal a complex genetic basis for personality. Multiple genetic dimensions differing in the breadth of their effects contribute to personality phenotypes (Jang et al., 1998; Livesley et al., 1998; Livesley & Jang, 2000). Some are relatively specific dimensions that influence single phenotypic traits, whereas others have broader effects influencing multiple phenotypically distinct but covarying traits. Consequently, many traits appear to be influenced by multiple genes and gene systems. Similarly, trait covariance seems to arise from multiple genetic effects. Genetic effects on traits appear to be direct rather than mediated by higher-order entities. These findings require replication. Nevertheless they appear to challenge
models of personality positing links between specific genetically based neurotransmitter systems and specific personality traits. They also suggest a different conception of the trait hierarchies from that assumed by many trait taxonomies.

Hierarchical Structure
Factor analyses of genetic correlations and the modeling studies cited in the previous section identified general genetic factors that account for trait covariation. The model-fitting analyses also confirmed conclusions based on regression analyses that lower-order traits are not merely components of higher-order traits, but rather are distinct etiological entities. It appears that each basic or facet trait is influenced by general and specific genetic factors. Genetic dimensions that affect multiple traits appear to influence each trait directly rather than indirectly through a higher-order phenotypic entity. This raises questions about the basis for the hierarchy consistently identified by factor analytic studies and the conceptual status of higher-order constructs like neuroticism and extraversion and their role in theories of individual differences.

Although the facets delineating each of the five-factor domains covary due to shared genetic effects, it is not necessary to invoke a higher-order latent construct to explain this covariation. This raises the possibility that higher-order constructs such as neuroticism merely represent the pleiotropic action of genes. If this is the case, neuroticism and other higher-order domains are not entities that are distinct from the specific traits that delineate them. They are not traits in Allport’s sense of distinct phenotypic entities with an underlying biology, but rather heuristic devices that represent clusters of traits that covary because of a common genetic effect. This is consistent with the conception of domains as lexical categories (Saucier & Goldberg, 1996). Nevertheless, facet traits defining domains such as neuroticism and extraversion overlap sufficiently to justify grouping them into an overall global measure.

The model of trait structure implied by these findings differs from that of traditional trait theories. With traditional models in which lower-order traits are nested within a few higher-order factors, it follows that any statement about the higher-order factor applies to all subordinate traits. This is not the case with the model proposed because each basic trait has its own specific etiology. A second difference is that traditional hierarchical models seem to assume that trait taxonomies are similar to any classification based on set theory principles. At each level in the hierarchy, categories are assumed to be exhaustive and exclusive (Simpson, 1961). Exhaustiveness means that trait categories exist to classify all subordinate traits, whereas exclusiveness refers to the principle that each subordinate feature can be classified into only one superordinate trait. Considerable effort has been expended in attempts to delineate a structure with these properties. Indeed, this is the reason for debate on number and

| TABLE 3.5 Illustrative Scales: Multivariate Genetic Analyses of the DAPP-DQ Facet Scales |
|-----------------------------------|---|---|---|---|---|---|---|
|                                  | A1 | A2 | A3 | E1 | E2 | A  | C  |
| Rejection                        |    |    |    |    |    |    |    |
| Rigid Cognitive Style            |    |    |    | .43| .38| .56|    |
| Judgmental                       | .13|    |    | .46| .34| .57|    |
| Interpersonal Hostility          | .46|    |    | .31| .34| .51|    |
| Dominance                        | .53|    |    |    | .40| .61|    |
| Restricted Expression            |    |    |    | .12| .49| .33| .49|
| Self-Disclosure                  | .55| .30|    | .12| .49| .33| .49|
| Affective Expression             | .31| .58|    | .21| .56|    | .45|
| Angry Affects                    | .24| .38|    | .75|    | .49|    |
| Positive Affects                 | .40| .51|    |    | .46| .33| .52|
| Self-Reliance                    | .55| .15|    |    | .50| .23| .59|
| Callousness                      |    |    |    |    |    |    |    |
| Contemptuousness                 | .36| .27| .42| .28| .15| .44| .57|
| Egocentrism                      | .28| .36| .28| .46| .21| .47| .48|
| Exploitation                     | .26| .54| .43| .35| .18|    | .51|
| Irresponsibility                 | .40| .33| .23| .22| .23| .40| .65|
| Lack of Empathy                  | .53| .20| .16| .33| .23| .26| .65|
| Remorselessness                  | .42| .16| .34|    | .76| .32|    |
| Sadism                           | .36|    | .66| .27| .14| .40| .65|

$\chi^2 = 52.45, \ df = 54, \ p = .53$

$\chi^2 = 90.63, \ df = 84, \ p = .22$

$\chi^2 = 154.48, \ df = 166, \ p = .73$
content of domains. It also explains Costa and McCrae’s insistence that domains are equal in breadth. If they are not, the five-factor model is open to the criticism that the model is not sufficiently parsimonious, as argued by Eysenck. This theoretical structure is understandable if trait taxonomies are conceptualized only as lexical structures. It is possible, however, that traits at the biological level are not organized in the systematic way proposed by the five-factor model.

There are no a priori reasons to assume that all basic traits must be organized into a hierarchy or that each higher-order domain is equally broad and defined by an equal number of facets as hypothesized by the five-factor model. An equally plausible model is that traits are organized into clusters that differ in the number of basic traits that they subsume and that the hierarchy is incomplete, with some specific traits showing minimal degrees of covariation. This structure is illustrated by the findings regarding the structure of the higher-order dimension of compulsivity identified in studies of personality disorder traits (Livesley et al., 1998). Pathways models identified a single genetic dimension underlying the specific traits that define this construct. Factor analyses show that it is consistently not related to other traits—hence, the three phenotypic traits that delineate compulsivity from separate higher-order factors. Compulsivity is, however, a trait narrower than other higher-order domains. It appears to represent a distinct basic or lower-order trait based on a single genetic dimension that does not have a hierarchical relationship with other basic traits.

Basic-Level Traits: Defining the Basic Unit of Personality

The idea that personality is inherited as a few genetic modules with broad effects and a large number of modules with more specific effects focuses attention on the significance of lower-order or basic traits. These findings are similar to evaluations of hierarchical models of cognitive ability that also provide evidence that specific abilities are heritable (Casto, DeFries, & Fulker, 1995; Pedersen, Plomin, Nesselroade, & McClearn, 1992). Basic traits do not appear to be specific exemplars of the higher-order traits that they define or blends of two or more factors (Hofstee, DeRaad, & Goldberg, 1992). Rather, they are discrete genetic entities with their own biological basis. This suggests that personality models that reduce traits to a few global domains do not reflect the genetic architecture of normal or disordered personality. As noted earlier, personality research has tended to neglect these traits in favor of more global dimensions. Yet evidence of specificity of genetic effects suggests that the basic traits are the fundamental building blocks of personality that are more important for understanding personality than are the global constructs that have traditionally been the focus of research and explanation. This approach again raises the question of how basic traits should be conceptualized and defined, as well as which criteria are relevant to defining domains.

Costa and McCrae (1998) noted the challenges of delineating a comprehensive set of basic traits. The specificity of genetic effects also reveals the challenge involved because of the large number of genetic dimensions that are likely to be involved. A genetic perspective does, however, provide a definition of a basic dimension that could facilitate the identification and assessment of these traits. The usual psychometric criteria used to develop homogeneous scales could be supplemented with the genetic criterion that a basic trait scale represents a single specific genetic dimension. With this approach, items assessing a basic trait would form a genetically homogeneous unit as opposed to a factorially homogeneous unit. Items could then be selected according to their correlation with the underlying genetic dimension. Thus items forming a scale would share the same general and specific genetic etiology. With this approach, the goal would be to use behavioral genetic techniques to bring about definitions of the phenotype that correspond to what Farone, Tsuang, and Tsuang (1999) refer to as “genetically crisp categories” (p. 114).

An example of this approach is provided by a study of the genetic structure of the Eysenck Personality Questionnaire (Heath, Eaves, & Martin, 1989). This instrument has three broad scales composed of 21 to 25 items that assess Neuroticism, Extraversion, and Psychoticism. Heath and colleagues extracted a common genetic and environmental factor for Neuroticism and Extraversion, indicating that these items are etiologically homogeneous. In contrast, little evidence was found for a common genetic factor for the Psychoticism items. Subsequent analyses showed that the items formed into two distinct genetic factors: paranoid attitudes and hostile behavior. The results of such a systematic evaluation of item etiology could be used to form etiologically homogeneous scales.

This approach could be used either to develop new scales or modify existing scales so that they resemble the underlying genetic architecture more closely. This could be achieved by applying differential weights that index the influence of specific genetic and environmental influences on different traits. In this way, questions about the phenotypic structure of personality are addressed, and scales could be constructed so that they do not reflect competing genetic and environmental influences.

The estimation of genetic and environmental factor scores is a relatively new and active area of research. Sham et al. (2001) recently described a method that permits these genetic
factor scores to be computed. Their method uses the following equation:

\[ \gamma = \gamma \Sigma^{-1}x \]  

(3.2)

where \( \gamma \) = factor score for the common genetic factor, \( \beta \) = the factor loadings of each variable on the genetic factor of interest (i.e., the column vector of estimated path coefficients that represent the correlations between the common genetic or environmental factor and the observed measures), \( \Sigma^{-1} \) = correlation matrix between all of the variables (i.e., the inverse of the correlation matrix of the observed measures), and \( x \) = each person’s score or response to each of the variables (i.e., column vector of observed values on the measures). Other methods are also available to compute genetic and environmental factor scores (Thomis et al., 2000).

**Domain Content**

As discussed earlier, the facet structure of several five-factor domains is still unclear. The same behavioral genetic approach used to define and measure basic trait scales could also be applied to the delineation of domain content. The unity of a domain is demonstrated by evidence that a single common genetic factor influences all the facets composing the domain. This approach could be used to clarify the location of impulsivity within the higher-order structure. The five-factor model locates impulsivity in Neuroticism, whereas Eysenck places it within Extraversion. As noted earlier, the bivariate correlations of this facet with other Neuroticism facets assessed with the NEO-PI-R are lower than correlations between other facets. Etiological data could be used to relocate impulsivity with other traits with which it shares a common etiology. Alternatively, the item content could be changed based on genetic and environmental etiology so that correlations with the other Neuroticism facets are increased (of the loadings on the common factors are increased). In the case of the DAPP scales, impulsivity is part of the phenotypic trait of stimulus seeking along with sensation seeking and recklessness. Multivariate genetic analyses showed that a single common genetic factor underlies this dimension that is defined by sensation seeking and recklessness (see Table 3.5). Impulsivity has a low loading on the factor and a substantial specific heritable component. It appears that impulsivity as defined within the DAPP structure is a specific heritable entity and not the result of interaction between extraversion and constraint or psychoticism as suggested by Depue and Collins (1999) or extraversion and psychoticism as suggested by Gray (1970, 1973, 1987; Pickering & Gray, 1999), although it is consistent with Gray’s argument that impulsivity is a fundamental dimension of temperament.

The findings of behavioral genetic studies of personality structure also have implications for attempts to identify the putative genes for personality. Most molecular genetic studies of personality use an analytic strategy that correlates a total personality trait score such as Neuroticism with variations in the candidate allele (Lesch et al., 1996). As the studies described show, the total scale score confounds multiple genetic and environmental effects and reduces the power to detect putative loci. The use of etiological factor scores that index the proportions of the personality phenotype directly attributable to specific genetic and environmental effects (Boomsma, 1996; Sham et al., 2001; Thomis et al., 2000) could reduce these confounds.

**UNIVERSALITY OF TRAIT STRUCTURE**

Most models of personality traits including Eysenck’s three-factor model (Eysenck & Eysenck 1992), the five-factor model, and diagnostic categories of personality disorder proposed in the DSM-IV (American Psychiatric Association, 1994) assume that the taxonomies proposed reflect a universal structure. This assumption is also assumed to apply to the measures developed to assess these constructs. The only differences that these models of personality (and their measures) permit between cultures and other groups (e.g., gender) are quantitative in nature; they typically mean differences in trait levels or severity. If these assumptions are correct, we should find that the etiological architecture of personality is also invariant across cultures and other basic groupings. We discuss this idea with respect to cross-cultural comparisons and the effects of gender.

**Cross-Cultural Comparisons**

Multiple studies show that the observed factorial structure of scales such as the NEO-PI-R is stable across cultures. For example, McCrae and Costa (1997) reported that the five-factor structure is consistent across samples from the United States, Western Europe, and Asia (see also Costa & McCrae, 1992; McCrae et al., 2000). The issue of cross-cultural stability also applies to etiological structure. Earlier, we described fitting an independent pathways model to the six facets defining NEO-PI-R domains in independent samples of German and Canadian twins. The universality of genetic effects can be evaluated by testing the equivalence of the genetic and environmental structures across independent samples. It is possible to test whether: (a) the same genetic and environmental factors influenced the Canadian and German samples; and (b) whether these factors influenced each sample to the same
Two tests of equivalency were applied. The first evaluated equivalency of model form by testing the hypothesis that the same kind and number of genetic parameters are required to explain the data across the two samples. Sample differences are hypothesized to be limited to differences in the magnitude of the genetic and environmental influence exerted on a domain’s facet scales. If equivalence of model form was supported across the samples, the next step was to evaluate the magnitude of genetic and environmental influences across samples. This was accomplished by applying a model with the same parameters to both samples. That is, the model specified the same number and type of factors in both samples and identical and constrained the factor loadings to be identical.

The results of tests of model form and magnitude for NEO-PI-R Neuroticism are shown in Table 3.6. The same number and types of genetic and environmental influences (two additive genetic and two nonshared environmental common factors) were identified in both samples, suggesting that the structure of neuroticism was similar across the samples. When the factor loadings on the common factors from the German sample were made to be the same as those on the Canadian sample (and vice versa), the model no longer fit the data. The results suggested that the primary differences between the German and Canadian samples were limited to the magnitude rather than kind of genetic and environmental effects supporting the claim that the factorial structure of the NEO-PI-R facets is universal.

### Table 3.6 Model-Fitting Statistics

<table>
<thead>
<tr>
<th>Model</th>
<th>( \chi^2 )</th>
<th>( p )</th>
<th>RMSEA</th>
<th>AIC</th>
<th>( \chi^2 )</th>
<th>( p )</th>
<th>RMSEA</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuroticism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1a</td>
<td>199.91</td>
<td>.00</td>
<td>.040</td>
<td>−64.09</td>
<td>216.56</td>
<td>.00</td>
<td>.039</td>
<td>−47.44</td>
</tr>
<tr>
<td>2s</td>
<td>172.11</td>
<td>.00</td>
<td>.036</td>
<td>−79.89</td>
<td>149.82</td>
<td>.07</td>
<td>.019</td>
<td>−102.18</td>
</tr>
<tr>
<td>3c</td>
<td>151.12</td>
<td>.00</td>
<td>.029</td>
<td>−88.88</td>
<td>135.86</td>
<td>.15</td>
<td>.015</td>
<td>−104.14</td>
</tr>
<tr>
<td>4d</td>
<td>144.88</td>
<td>.03</td>
<td>.029</td>
<td>−83.12</td>
<td>131.14</td>
<td>.13</td>
<td>.016</td>
<td>−96.86</td>
</tr>
<tr>
<td>5e</td>
<td>145.12</td>
<td>.03</td>
<td>.030</td>
<td>−82.88</td>
<td>130.40</td>
<td>.14</td>
<td>.014</td>
<td>−97.60</td>
</tr>
<tr>
<td>6f</td>
<td>210.86</td>
<td>.00</td>
<td>.043</td>
<td>−61.14</td>
<td>220.57</td>
<td>.00</td>
<td>.038</td>
<td>−51.43</td>
</tr>
</tbody>
</table>

Note. All models specified additive genetic and nonshared environmental factors unique to each facet. \( df = 132 \), one common additive and one common nonshared environmental factor. \( df = 126 \), one common additive and two common nonshared environmental factors. \( df = 120 \), two common additive and two common nonshared environmental factors. \( df = 114 \), three additive and three common nonshared environmental factors. \( df = 136 \), common pathways model.

### Gender Differences

Personality tests are usually constructed to minimize gender-based differences by eliminating items whose intercorrelations with the other items can be attributable to gender and eliminating items evoking marked gender differences in endorsement. The approach yields scales that are applicable to both females and males but it overlooks the possibility of gender differences in the etiology. Behavioral genetic methods may be used to determine whether the same genetic and environmental factors influence personality measure scores in males and females and whether the etiological architecture underlying the factorial structure of a personality measure is the same in males and females.

The first question can be answered by fitting sex-limitation models to personality data (Neale & Cardon, 1992). This is accomplished by fitting a simple extension of the usual heritability model that uses data from same- and opposite-sex twin pairs to test whether the same genetic factors operate in males and females. In this case, gender differences are limited to differences in the magnitude of genetic and environmental influences. Another form of sex-limited gene expression occurs when different genes control the expression of a trait that is measured in the same way in males and females. With this form of sex-limitation, it is also possible to determine whether the same genes are present in both sexes but only expressed in one sex. This is evaluated by comparing the similarities of opposite-sex DZ twin pairs with same-sex DZ pairs. Sex-specific genetic influences are suggested when the similarity of opposite-sex pairs is significantly less than the similarities of male or female DZ pairs. The difference in the correlation is attributable to the gender composition of each zygosity group. When the same and opposite-sex DZ correlations are similar, gender differences are not indicated.

Only a few studies have investigated sex-limited gene expression in normal personality. The most notable is Finkiel and McGue’s (1997) study that showed that the same genetic loci influence 11 out of the 14 scales of Multidimensional Personality Questionnaire (MPQ; Tellegen, 1982) in males and females. The heritable influences on the remaining three
traits—Alienation, Control, and Absorption—indicated that the genetic influences were gender-specific. Jang, Livesley, and Vernon (1998) reported some evidence for sex-limited gene expression in 18 traits delineating personality disorder measured by the DAPP. All dimensions except Submissiveness in males, and Cognitive Dysfunction, Compulsivity, Conduct Problems, Suspiciousness, and Self-Harm in females were significantly heritable. Sex-by-genotype analyses suggested that the genetic influences underlying all but four DAPP dimensions (Stimulus Seeking, Callousness, Rejection, Insecure Attachment) were specific to each gender, whereas environmental influences were the same in both genders across all dimensions. Furthermore, the four higher-order dimensions derived from the 18 basic traits (Livesley et al., 1998) were also heritable across sex, and genetic effects were in common to both genders; the exception was Dissocial Behavior, which was not heritable in females.

Such evidence of sex-limited effects challenges the assumed universality of trait taxonomies. However, it could be argued that the results based on the DAPP and MPQ are atypical. The DAPP is a specialized scale designed primarily to assess personality dysfunction. The scale does not cover such areas of normal personality as Openness to Experience (Jang & Livesley, 1999; Schroeder et al., 1992) because abnormal variants of Openness are not included in clinical descriptions of personality disorder. The MPQ, unlike other scales, routinely reveals nonadditive genetic effects due to dominance (Waller & Shaver, 1994). This suggests that it may assess content different from that tested by scales such as the NEO-PI-R, which reveals genetic effects that are additive (e.g., Jang et al., 1998).

A more appropriate evaluation of the assumption of universality would be to examine sex-limited gene expression on a major model of personality such as the five-factor model. Evaluation of whether the same genes are present across different samples is similar to the evaluation of cross-cultural effects. Jang, Livesley, Riemann, and Angleitner (in press) applied sex-limitation models to NEO-FFI data obtained from the Canada and German twin samples described earlier. Two general models were fit to the data. The first specified additive genetic and nonshared environmental influences for females and males and a male-specific genetic factor. The second tested whether heritable influences common to males and females were the same across the two samples. Table 3.7 reports the intrapair twin correlations for each zygosity group in each sample. The MZ male and MZ female correlations exceed their respective DZ correlations, suggesting the presence of heritable influences on each NEO-FFI domain in each sample. Of particular interest is the comparison between the DZ opposite-sex correlations and the same-sex DZ correlations. In both samples, the DZ opposite-sex correlation for Conscientiousness was near zero, suggesting the presence of differential gender effects. The final form of the best-fitting model is presented in Table 3.8. The results suggest that genetic and environmental influences common to males and females influence four of the five FFM domains. The exception was Conscientiousness, for which gender-specific additive genetic influences operate. However, the external events and experiences specific to each twin—nonshared environmental influences—are common to males and females. The results also suggest that the type and magnitude of genetic and environmental influence were the same across the two groups, supporting the notion that the five-factor model as assessed by the NEO-FFI is applicable to different cultures and genders.

This study has several limitations. The first is that the sample sizes are rather small in both samples, especially male DZ twin pairs and opposite-sex pairs. The twin covariances associated with these two zygosity types, especially the opposite-sex pairs, are crucial for the validity of the analyses. The availability of relatively few twin pairs calls into question the stability of the correlations and thus the detection of sex-limited genes—as was obtained for Conscientiousness. Second, the study used the NEO-FFI, the short form of the NEO-PI-R. The full scale might produce different results because long versions of these scales sample domains more

<table>
<thead>
<tr>
<th>NEO-FFI Domain</th>
<th>Canadian Sample</th>
<th>German Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MZ F M F M-F</td>
<td>MZ F M F M-F</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>.41 .53 .22 .35 .13</td>
<td>.49 .52 .36 .20 .15</td>
</tr>
<tr>
<td>Extraversion</td>
<td>.50 .49 .34 .30 .23</td>
<td>.57 .57 .34 .25 .17</td>
</tr>
<tr>
<td>Openness</td>
<td>.63 .51 .28 .36 .20</td>
<td>.57 .50 .44 .26 .10</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>.50 .46 .14 .33 .26</td>
<td>.43 .42 .37 .10 .10</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>.47 .50 .28 .38 .01</td>
<td>.57 .46 .40 .23 .05</td>
</tr>
</tbody>
</table>

Sample sizes (pairs): 102 165 61 129 73 104 425 38 163 68
thoroughly. As such, the present results should be considered tentative until replicated on a larger sample using full-scale versions, as well as other measures of personality.

These analyses suggest that although most personality traits are influenced by the same genes in both genders (the implication being, e.g., DRD4 influences novelty seeking in both men and women), this is not true for all traits. The previous section suggested several explanations, but it is also possible that at the molecular level, different genes (or yet-to-be-discovered polymorphisms) differentially influence personality across genders. If this is the case, the genetic and environmental architecture of some scales may differ by gender. This could be evaluated by fitting independent and common pathways models to data from sister pairs and brother pairs separately and constraining the models (in form and magnitude) to be equal across gender groups. The sex-limitation model described previously that uses data from brother-sister pairs to test for gender-specific effects can be expanded to the multivariate case to further explore gender differences in personality. As far as we are aware, few multivariate genetic analyses of gender differences have been conducted, probably because many studies have limited data collection to sister pairs or have difficulty obtaining data from brother pairs (Lykken, McGue, & Tellegen, 1987).

ENVIRONMENTAL EFFECTS

Although our primary concern is with the genetic basis for personality structure, any discussion of genetic influence would be incomplete without reference to environmental factors. Twin studies consistently show that about 50% of the variance in personality traits is explained by environmental factors and that most of this is accounted for by nonspecific influences; common environmental influences do not appear to contribute to personality variation (Plomin & Daniels, 1987). This etiological model derived from twin studies is confirmed by a large-scale study of Neuroticism by Lake, Eaves, Maes, Heath, and Martin (2000) that showed that individual differences in neuroticism were not transmitted from parent to offspring via the environment but rather by genetic factors. The size and unique features of their data set (45,880 twin pairs and their relatives on two continents) allowed them to test models of genetic transmission as well as gene-environment correlations. The results suggest that the environment exerts a contemporaneous influence on individual differences in neuroticism. That is, its effects are located in the current environment as opposed to being preset like genetic factors that are passed to individuals from their parents.

Although nonshared environmental factors are important, the nature of these variables and the way they affect personality remain unclear. Despite considerable research effort (e.g., Hetherington, Reiss, & Plomin, 1994; Turkheimer & Waldron, 2000) using a variety of methods (Baker & Daniels, 1990; Hetherington et al., 1994; Reiss et al., 1994; Vernon, Lee, Harris, & Jang, 1996) the results have been uniformly disappointing: Few nonshared influences on personality have been identified (Turkheimer & Waldron, 2000). Most studies have, however, investigated the effects of the nonshared environment on the single variables; few studies have examined the effects of the nonshared environment on trait covariance. The study by McCrae et al. (in press) and the illustrative multivariate genetic analyses of the NEO-PI-R and DAPP presented earlier suggest that the nonshared environmental factors have an influence on personality structure different from that of genetic factors. They do not appear to contribute to trait substantially to the trait covariation described by trait taxonomies.

### TABLE 3.8 Parameter and Standard Error Estimates Produced by the Best-Fitting Sex-Limitation Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>N</th>
<th>E</th>
<th>O</th>
<th>A</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>h&lt;sub&gt;i&lt;/sub&gt;</td>
<td>.86 ± .03</td>
<td>.84 ± .03</td>
<td>.80 ± .02</td>
<td>.85 ± .03</td>
<td>.86 ± .03</td>
</tr>
<tr>
<td>e&lt;sub&gt;i&lt;/sub&gt;</td>
<td>.82 ± .02</td>
<td>.84 ± .02</td>
<td>.88 ± .01</td>
<td>.83 ± .02</td>
<td>.82 ± .02</td>
</tr>
<tr>
<td>h&lt;sub&gt;m&lt;/sub&gt;</td>
<td>.80 ± .04</td>
<td>.84 ± .04</td>
<td>.80 ± .02</td>
<td>.89 ± .03</td>
<td>—</td>
</tr>
<tr>
<td>e&lt;sub&gt;m&lt;/sub&gt;</td>
<td>.88 ± .03</td>
<td>.84 ± .03</td>
<td>.88 ± .01</td>
<td>.79 ± .03</td>
<td>.85 ± .03</td>
</tr>
<tr>
<td>h′&lt;sub&gt;m&lt;/sub&gt;</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.83 ± .04</td>
</tr>
</tbody>
</table>

Note. N = Neuroticism; E = Extraversion; O = Openness to Experience; A = Agreeableness; C = Conscientiousness; h<sub>i</sub>, e<sub>i</sub>, h<sub>m</sub>, e<sub>m</sub> = additive genetic and nonshared environmental effects common to males and females; h′<sub>m</sub> = male-specific additive genetic effects.
It also appears that the environment does not have an effect that is independent of preexisting genetic factors. Reviews by Reiss, Neiderhiser, Hetherington, and Plomin (2000) and Caspi and Bem (1990) document personality-environment interaction and the way the individuals select and create their own environment. Genetic factors influence the environmental variables that are the focus of attention and the situations that the individual selects. For example, some kinds of life events are not independent of the individual; rather, their occurrence is influenced by such traits as Neuroticism and Extraversion (e.g., Poulton & Andrews, 1992; Magnus, Diener, Fujita, & Pavot, 1993). Saudino, Pedersen, Lichtenstein, and McClearn (1997) showed that all genetic variance on controllable, desirable, and undesirable life events in women was common to the genetic influences underlying EPQ Neuroticism and Extraversion, NEO-FFI Openness to Experience (Costa & McCrae, 1985). Genetic influences underlying personality scales had little influence on uncontrollable life events because this variable was not heritable. Kendler and Karkowski-Shuman (1997) showed that the genetic risk factors for major depression increased the probability of experiencing significant life events in the interpersonal and occupational-financial domains, probably because individuals play an active role in creating their own environments. Heritable factors, such as personality and depression, influence the types of environments sought or encountered. Jang, Vernon, and Livesley (2000) report significant genetic correlations between the Family Environment Scale (FES: Moos & Moos, 1974) subscale of Cohesiveness and DAPP higher-order factors of Emotional Dysregulation (−.45), Inhibition (−.39), FES Achievement Orientation and DAPP Dissocial Behavior (.38), Inhibition (−.58), and FES Intellectual Cultural Orientation and DAPP Emotional Dysregulation (−.34). These results help to explain why measures of the environment often have a heritable component: They often reflect genetically influenced traits (Saudino et al., 1997).

Using the factor score approach described earlier, Thomis et al. (2000) computed genetic factor scores for measures of muscle strength obtained from a sample of MZ and DZ twins. The twins were then subjected to a 10-week muscle strength training regimen. The muscle strength genetic factor scores explained the greatest proportion of the variance pre- and posttraining, indicating that genes are switched on, so to speak, in response to stress due to training, thus demonstrating the existence of gene-environment interaction. Findings such as these suggest that the environmental factors that influence traits are partially dependent on preexisting genetically based personality traits. For example, a person scoring highly on a genetically based trait like sensation seeking will seek out environments conducive to the expression of this personality genotype, such as engaging exciting sports. For this reason, molecular genetic studies designed to identify the genes for personality need to incorporate measures of personality that separate the effects of genes and environment on the phenotype.

MOLECULAR GENETICS

From a genetic perspective, dimensions of individual differences in personality are complex traits. That is, multiple genes and gene systems and multiple environmental factors influence each trait (Plomin, DeFries, McClearn, & McGuffin, 2000). The emergence of molecular genetics prompted considerable optimism about the possibility of identifying the genetic component or quantitative trait loci (QTLs) of these traits. Such a development would radically change the nature of personality research by enabling investigators to link behavioral dimensions to underlying molecular genetic structures. This would provide a more powerful way to resolve trait taxonomic issues that the behavioral genetic approaches discussed. The results of such studies have, however, been inconsistent, replications have often failed, and progress has been slower than expected.

One of the earliest studies investigated the relationship between Novelty Seeking and dopamine D4 or DRD4 receptor (Cloninger, Adolfsson, & Svrakic, 1996). Earlier Cloninger (1987; Cloninger, Svrakic, & Przybeck, 1993) proposed a model of personality that postulated that the expression of each personality trait is modulated by a specific genetically controlled neurotransmitter system. Specifically, Novelty Seeking is controlled by the dopaminergic system, Harm Avoidance by the serotonin system, and Reward Dependence by norepinephrine. Cloninger and colleagues (1996) reported a polymorphism of the D4 receptor that accounted for about 10% of the variance. Several replications have been reported (Benjamin, Greenberg, & Murphy, 1996; Ebstein, Novick, & Umansky, 1996; Ebstein, Segman, & Benjamin, 1997) along with many failed replications (Ebstein, Gritsenko, & Nemanov, 1997; Malhotra, Goldman, Ozaki, & Breier, 1996; Ono et al., 1997; Pogue-Geile, Ferrell, Deka, Debski, & Manuck, 1998; Vandenbergh, Zonderman, Wang, Uhl, & Costa, 1997).

Similarly, several studies have demonstrated a relationship between the serotonergic system and Harm Avoidance, Neuroticism, or related constructs (Hansenne & Ansseau, 1999; Rinne, Westenberg, den Boer, & van den Brink, 2000), and significant associations were reported with the serotonin transporter gene, 5-HTTLPR (Katussagi et al., 1999). However, several studies have failed to replicate
these findings (Flory et al., 1999; Gelernter, Kranzler, Coccaro, Siever, & New, 1998; Hamer et al., 1999; Herbst, Zonderman, McCrae, & Costa, 2000). Gustavsson et al. (1999) also failed to replicate these findings using the Karolinska Scales of Personality.

These inconsistencies can be attributed to conceptual and measurement issues. The early studies in particular were often based on a conceptual model that assumed that personality is influenced by relatively few genes, each accounting for substantial variance. As noted, the evidence does not support this approach. There has also been a tendency to assume that each trait was linked to a specific neurotransmitter system. More recently, however, attention has focused on pleiotropic effects by investigating the possibility that a given polymorphism influences several traits. Work on the serotonin transporter gene, for example, suggests that it is not associated with a single trait but rather has a pleiotropic relationship with Neuroticism and Agreeableness. Studies on humans and primates suggest that altered serotonin activity is related to negative emotional states such as depression, anxiety, and hostility, and to social behaviors such as dominance, aggression, and affiliation with peers (Graeff, Guimarães, De Andrade, & Deakin, 1996; Knutson et al., 1998; Murphy et al., 1998). Knutson and colleagues (1998) found that administration of the specific serotonin reuptake inhibitor, paroxetine, decreased negative affect and increased social affiliation in normal human subjects. Lesch and colleagues (1996) reported that individuals carrying the 5-HTTLPR-S allele had increased total scores on NEO-PI-R Neuroticism and the facets of Anxiety, Angry Hostility, Depression, and Impulsiveness. The allele accounted for 3 to 4% of the total variance in these scales. Unexpectedly, the allele was also associated with a decreased NEO-PI-R Agreeableness score. Greenberg et al. (1999) recently replicated these findings. Hamer et al. (1999) showed that 5-HTTLPR-S genotypes were significantly associated with increased Harm Avoidance (which correlates .66 with NEO-PI-R Neuroticism) and decreased Self-Directedness (correlated -.64 with NEO-PI-R Neuroticism), Reward Dependence, and Cooperativeness (shown to correlate .43 and .66 with NEO-PI-R Agreeableness). These effects accounted for .80%, 1.98%, .97%, and 2.60% of the total variance in these scores, respectively. Mazzanti et al. (1998), Peirson et al. (2000), and Benjamin et al. (2000) have reported replications.

Measurement problems contributing to inconsistent findings include the use of measures with less-than-optimal psychometric properties and the use of relatively broad personality constructs. Comparison of the dopamine–novelty seeking and serotonin-neuroticism studies suggests that the serotonin-neuroticism literature is less ambiguous than the dopamine–novelty seeking literature. These differences appear to be related to scale properties. Inconsistent findings may also be due to the confounding of genetic and environmental influences on the phenotypes. As we have tried to show, many constructs and scales are etiologically heterogeneous.

Twin studies estimating statistical pleiotropy could contribute to molecular genetic studies by identifying traits that are etiologically homogeneous units and etiologically related. Molecular genetic work could then be used to confirm these associations by identifying the actual genes that account for trait covariance. This would provide the strongest basis for revising personality models and allocating traits to etiologically related domains.

CONCLUSIONS

The thesis of this chapter is that behavioral genetic approaches promise to provide an additional perspective that may help to resolve some of the more intractable problems in delineating and conceptualizing personality structure. The evidence reviewed suggests an alternative perspective on the trait structure of personality that complements traditional conceptions. Although trait theory has largely concentrated on mapping personality in terms of broad global traits, the evidence suggests that personality is inherited as a large number of genetic dimensions that have relatively specific effects on personality phenotypes and a smaller number of genetic dimensions that have broader effects, perhaps through a modulating influence on related dispositions. These dimensions with broader effects appear to account for some of the observed covariation among traits. They do not appear, however, to exert these effects through higher-order phenotypic structures, but rather through a direct influence on each basic trait. We assume that these common features are more likely to involve modulating functions or common mechanisms that regulate each trait in a given cluster.

These tentative conclusions suggest the need to reconsider traditional models of the hierarchical structure of personality in which traits are organized into broad domains due to the effects on broad dispositions. Instead, the organization of traits into clusters is assumed to arise from the pleiotropic effects of genetic dimensions that affect multiple traits. Under these circumstances, it is conceivable that not all traits are organized into clusters of covarying features, but rather remain relatively distinct characteristics. Nor is it inevitable the traits are hierarchically organized in similar ways across domains. That is, it is possible that the symmetrical hierarchical structure avidly sought by trait theorists and students of psychopathology does
not reflect the way personality is organized at a genetic level. An equally feasible structure would involve considerable differences in complexity across domains. Some domains may consist of a relatively large number of traits, whereas others may consist of only one or two genetically homogeneous traits. These assumptions are consistent with the lexical view of Saucier and Goldberg (1996), who argued that the five domains are merely a convenient way of organizing lower-order traits and that there is no inherent reason to assume that domains are equal in breadth or in pervasiveness.

Although behavioral genetic analyses show that environmental factors exert a considerable influence on personality, they do not appear to influence the structural relationships among traits to any appreciable extent. Instead, environmental factors appear to exert a more contemporaneous effect on trait expression. The nature of these factors and the way that they function remain important topics of research.

REFERENCES


Herbst, J. H., Zondervan, A. B., McCrae, R. R., & Costa, P. T. (2000). Do the dimensions of the Temperament and Character Inventory map a simple genetic architecture? Evidence from...


Whether we speak of mice or men, every member of a species is the same as other members in many respects but different in others. One task of personality psychology is to describe the basic behavioral differences and discover their origins. Description of personality is usually in terms of observable traits, and various models have been proposed to classify them. Biology has confronted a similar task in the classification of species (taxonomy). Taxonomy has been based on phenomenal and functional similarities and differences but more recently has been moving in the direction of using evolutionary analyses to define species in terms of their ancestries. Psychology still depends on phenomenal similarities and differences. As the genome reveals its secrets, both fields will eventually turn to DNA for the classification task.

There are two basic pathways for the second task, the search for the sources of individual differences. These are shown in Figure 4.1. One pathway is the biological beginning in behavioral genetics. Genes make proteins into neurons, and neurons are organized into brain and nervous systems. Neurons operate through chemical neurotransmitters and the enzymes that govern their production and catabolism, as well as through hormones produced in other loci. This is the biochemical level. Differences in biochemical makeup result in differences in neural activity and reactivity or physiology. Physiological differences affect conditionability, both of the classical and operant types. Individuals differ in both their conditionability and their sensitivities to conditioned stimuli associated with reward and punishment.

The second pathway begins with the largest social unit, culture. Cultures are subdivided into specific societies defined by geography or class groupings defined by wealth, occupation, and education. Neighborhood provides the more proximal influences on behavior. The family of origin and peers transmit the influences of society, albeit with individual variations on modal mores, values, and behavior patterns. Observational learning combined with social reinforcement is the mechanism of influence at the next level. At this point there is a convergence of the pathways because the different
mechanisms of learning combine to produce behavioral traits. These traits are usually specific to certain types of situations. Depending on their generality and strength they combine to form what we call personality traits.

Both of these pathways have a historical origin in the evolutionary history of the species. Genetic changes account for the origin and changes (over long periods of time) in the species. Cultures represent the collective solutions of the human species to the basic demands of evolution: survival and reproduction. Cultural evolution is more rapid than biological evolution. Significant changes can occur within a generation, as with the sudden impact of computer technology on the current generation.

This chapter describes the biological pathway up to, but not including, conditioning. For each of four dimensions of personality I describe theory and research at each level of analysis along this pathway starting at the top (physiology). At the genetic level I describe primarily the studies of molecular genetics that link specific genes to traits. The biometric genetic studies are covered in the chapter by Livesly, Jang, and Vernon in this volume. The molecular studies link genes more directly to the neurological and biochemical levels on the way up to personality traits. An analysis of this type was conducted a decade ago (Zuckerman, 1991). Advances occur rapidly in the neurosciences. Ten years is equivalent to at least several decades in the social sciences. I have made an attempt to survey the changes since my last attempt. In a chapter I can hope only to highlight some of these advances and will reserve a more thorough review for a revision of my 1991 book. My approach draws heavily on comparative studies of other species as any psychobiological model must do (Gosling, 2001; Zuckerman, 1984, 1991), but I cannot do so within the constraints of a single chapter. I will limit comparative studies to those in which there are clear biological markers in common between animal and human models.

**TEMPERAMENT AND PERSONALITY TRAITS**

Researchers of temperament in children and behavioral traits in other species have typically included certain dimensions like emotionality, fearfulness, aggressiveness, approach versus withdrawal (in reactions to novel stimuli), general activity, playfulness, curiosity, sociability versus solitariness, and inhibition versus impulsivity (Strelau, 1998). From the 1950s through the 1970s personality trait classification was dominated by two models: Eysenck’s (1947) three-factor theory (extraversion, neuroticism, and psychoticism) and Cattell’s (1950) 16-factor model. Eysenck’s (1967) model was biologically based with an emphasis on genetics, physiology, and conditioning. Gray’s (1982, 1987) model is a bottom-up model that starts with behavioral traits in animals and extrapolates to human personality. He places his three behavioral dimensions (anxiety, impulsivity, fight-flight) within the axes of Eysenck’s dimensions, but not lying on the axes of those dimensions or being precise equivalents of them.

The first five-factor model originated in lexical studies of trait-descriptive adjectives in language done in the 1960s (Norman, 1963; Tupes & Christal, 1961) with its roots in a much earlier study by Fiske (1949). Interest in this model reawakened in the 1980s (Digman & Inouye, 1986; Goldberg, 1990; Hogan, 1982; McCrae & Costa, 1985). Most of these studies used adjective rating scales. The translation of the model into a questionnaire form (NEO-PI-R; Costa & McCrae, 1992a) increased the use of the scales by personality investigators. The five factors incorporated in this test are labeled extraversion, neuroticism, agreeableness, conscientiousness, and openness to experience. The five factors have been replicated in studies in many countries although with some differences—particularly on the last factor, openness. The enthusiasts for the Big Five insist it is the definitive and final word on the structure of personality (Costa & McCrae, 1992b), although critics regard this claim as premature (Block, 1995; Eysenck, 1992; Zuckerman, 1992). One of the criticisms of the model is its atheoretical basis in contrast to Eysenck’s development of his factors from theory as well as empirical factor
analytic studies of questionnaire content. However, recent studies in behavior genetics have used the model, and some of the data from earlier studies has been translated into the form of these five factors (Loehlin, 1992).

Two recent models have been derived from biosocial theories. Based on factor analyses of scales used in psychobiological studies of temperament and personality, Zuckerman and Kuhlman developed a five-factor model dubbed the alternative five (Zuckerman, Kuhlman, & Camac, 1988; Zuckerman, Kuhlman, Thornquist, & Kiers, 1991). This model was translated into a five-factor questionnaire (Zuckerman-Kuhlman Personality Questionnaire, or ZKPQ) on the basis of item and factor analyses (Zuckerman, Kuhlman, Joireman, Teta, & Kraft, 1993). The five factors are sociability, neuroticism-anxiety, impulsive sensation seeking, aggression-hostility, and activity. This model was used as the framework for a volume on the psychobiology of personality (Zuckerman, 1991).

Cloninger (1987) developed a personality model for both clinical description and classification of personality. The theory is biologically based and, like Zuckerman’s, uses the monoamine neurotransmitters as fundamental determinants of personality differences. The factors included in the most recent version of his questionnaire include novelty seeking, harm avoidance, reward dependence, persistence, cooperativeness, persistence, self-directedness, and self-transcendence (Cloninger, Przybeck, Svrakic, & Wetzel, 1994). Much of the recent psychobiological research in personality and psychopathology has used Cloninger’s system and questionnaires.

Builders of personality trait models often give different names to what are essentially the same traits. But even if one goes by the trait labels alone there are obvious similarities in what are considered the basic personality traits. Extraversion and neuroticism appear in nearly every system. Of course, one cannot take their equivalence for granted until empirical studies are done of their correlational relatedness.

Zuckerman et al. (1993) compared Eysenck’s Big Three, Costa and McCrae’s Big Five, and Zuckerman and Kuhlman’s Alternative Five in a factor-analytic study. A four-factor solution accounted for two thirds of the variance. The first factor was clearly extraversion, and the second was neuroticism with representative scales from all three questionnaires highly loading on their respective factors. The third factor consisted of Eysenck’s psychoticism and Zuckerman and Kuhlman’s impulsive sensation seeking at one pole and the NEO conscientiousness at the other. The fourth factor was defined by NEO agreeableness at one pole and ZKPQ aggression-hostility at the other. The analysis did not yield a fifth factor, possibly because of a lack of representative markers in the three tests. Activity loaded on the extraversion factor, and openness loaded on the agreeableness factor.

Zuckerman and Cloninger (1996) compared the scales of the ZKPQ with those of Cloninger’s Temperament and Character Inventory (TCI). ZKPQ impulsive sensation seeking was highly correlated with TCI novelty seeking ($r = .68$), ZKPQ neuroticism-anxiety with TCI harm-avoidance ($r = .66$), ZKPQ aggression-hostility with TCI cooperativeness ($r = -.60$), and ZKPQ activity with TCI persistence ($r = .46$). These scales showed convergent and discriminant cross validity, but the other scales in both tests had weaker correlations and correlated equally with several measures on the other scales. In Cloninger’s model there is no specific scale for extraversion or sociability.

The personality systems described thus far have been developed using factor analyses of trait dimensions. Many personologists have developed typologies on a rational-theoretical basis. Freud (1914/1957), Erikson (1963), and Maslow (1954) described personality types based on their developmental theories, each stressing the adult expressions of types derived from earlier stages of development. No valid methods of assessment were developed to operationalize these theories, although many clinicians continue to use them to describe personality differences among patients or others.

More recently, Millon and Everly (1985) defined eight types based on the interactions of four primary sources of reinforcement and two kinds of instrumental behavior patterns (active and passive). Some of the resultant types resemble different poles of the standard dimensions of personality. Sociable and introversive personality types resemble the two poles of the extraversion dimension; the inhibited type resembles neuroticism; and the cooperative types sound like agreeableness. The model was developed as a way of integrating personality development of psychopathology, particularly the personality disorders. It has been described as a biosocial theory but has not as yet been widely used in psychobiological research.

The examination of the biosocial bases of personality in this chapter will be organized around four basic personality factors, derived mostly from factor analytic studies, which are the same or quite similar across these studies, have some similarity to traits described in studies of temperament and animal behavior, and have been used in correlative studies of traits and psychobiology in humans. The four traits are extraversion/sociability, neuroticism/anxiety, aggression/agreeableness, and impulsivity/sensation seeking/psychoticism. Although activity is a widely used trait in studies of children and animals, it has not been widely used in studies of humans except for the pathological extreme of hyperactivity.
disorder and is recognized as a primary personality trait only in the Zuckerman-Kuhlman model.

EXTRAVERSION/SOCIABILITY

All models of basic personality, with the exception of Cloninger’s, recognize extraversion (E) as a primary and basic personality factor, but different models have defined it differently. In his earlier model Eysenck regarded E as a combination of two narrower traits: sociability and impulsivity. This amalgam was questioned by Carrigan (1960) and Guilford (1975), who claimed that sociability and impulsivity were independent traits. Sybil Eysenck and Hans Eysenck (1963) initially defended the dual nature of extraversion. However, the introduction of psychoticism (P) into a new version of their questionnaire resulted in a drift of impulsivity-type items to the P dimension, leaving E defined primarily by sociability and activity types of items. Hans and Michael Eysenck (1985) finally defined E in terms of the subtraits: sociable, lively, active, assertive, sensation seeking, carefree, dominant, surgent, and venturesome.

Costa and McCrae (1992a) defined their E superfactor in terms of subscale facets: warmth, gregariousness (sociability), activity, excitement seeking (sensation seeking), and positive emotions. Neither Eysenck nor Costa and McCrae now include impulsivity in the E factor; Eysenck now includes it in the N superfactor, and Costa and McCrae place it in their neuroticism factor. Both Eysenck and Costa and McCrae include activity and sensation seeking as components of their E factors.

Zuckerman et al. (1993) include only sociability and isolation intolerance in their sociability superfactor. In the alternative five, impulsivity and sensation seeking form another primary factor instead of being subsumed under E, and activity comprises another major factor. In spite of these differences in the content of the E factor in the three models, the questionnaire measures of the factors intercorrelate highly and have high loadings on a common factor (Zuckerman et al., 1993).

Cortical Arousal

Eysenck’s (1967) theory of extraversion has shaped much of the psychobiological research on this trait even to the end of the century (Strelau & Eysenck, 1987). The model suggests that introversion-extraversion is based on arousal characteristics of the cerebral cortex as regulated by the reticulocortical activating system. The extravert’s cortex in waking, nonstimulating conditions is underaroused relative to his or her optimal level of arousal. In these conditions the extravert is prone to seek out exciting stimulation in order to increase the level of arousal to a level that makes him or her feel and function better. The introvert is usually closer to an optimal level of arousal in low stimulation conditions and has less need to seek additional stimulation to feel better. The introvert may be overstimulated at a level of stimulation that is positive for the extravert.

The theory was initially tested with measures of brain activity from the electroencephalogram (EEG). Spectrum analyses break the raw EEG into bands characteristic of different degrees of arousal: sleep (delta), drowsiness (theta), relaxed wakefulness (alpha), and alert excitement (beta). Alpha has often been regarded as inversely related to arousal on the assumption that any interruption of this regular wave means an increase in arousal. However, some have used the frequency of alpha within the usual band (8–13 Hz) as a measure of relative arousal or alpha amplitude as an inverse measure of arousal. EEG spectrum characteristics are highly if not completely heritable (Lykken, 1982).

The findings relating extraversion to EEG criteria of arousal in various conditions from nonstimulating to mentally engaged have been summarized by Gale (1983), O’Gorman (1984), and Zuckerman (1991). Gale tried to reconcile the wide variety of results with the hypothesis that differences between introverts and extraverts appear only in moderately active conditions and not in either low stimulation (eyes closed, no stimulation) or activating conditions. Both O’Gorman and Zuckerman concluded that neither Eysenck’s broad hypothesis nor Gale’s narrow hypothesis, limiting the prediction to specific experimental conditions, were consistently supported by studies. Zuckerman noted that among the best studies, those confirming Eysenck’s hypothesis used samples with either all female or equal male and female participants, whereas those with all male or a preponderance of male participants did not support the hypothesis.

A large study utilizing the full spectrum range of EEG, three levels of activating conditions, measures of impulsivity as well as E, and a test of the interaction of personality, arousal level, and performance, found only weak evidence supporting Eysenck’s hypothesis (Matthews & Amelang, 1993). Correlations of .16 (about 3% of the variance) were found between activation in the low arousal bands (delta and theta) and E and one of its components, impulsivity. These correlations controlled for the influence of the other two Eysenck factors, neuroticism and psychoticism. The sociability component of E was not related to any index of cortical arousal. The significant results linking E to low arousal bands were found only in the least stimulating condition (reclining, eyes closed). The fact that the differences were not found in
alpha or beta bands but were found only in the most relaxed condition suggests that the weak correlation may have been due to impulsive extraverts’ getting drowsy or actually falling asleep. Regardless of interpretation, the low level of relationship between personality and arousal in this study could explain the inconsistency of previous studies testing the hypothesis: They simply did not have enough power to detect the relationship with any reliability.

Consistent with Eysenck’s model was the finding that while performing six tasks extraverts tended to perform worse than introverts at higher levels of alpha (indicating lower levels of arousal). Only the alpha band, however, supported the hypothesis of better performance of introverts at lower levels of arousal. Brain imaging using positron-emission tomography (PET) and cerebral blood flow (CBF) have an advantage over EEG because they assess subcortical as well as cortical activation and analyze activity in particular structures or brain loci. The problem with studies using these new techniques is that because of the expense, low numbers of subjects are used and many brain areas are analyzed, increasing the possibilities of both Type I and Type II errors. Replication across studies is one solution to the problem.

Mathew, Weinman, and Barr (1984) found negative correlations between E and CBF indices of activation in all cortical areas in both hemispheres, supporting Eysenck’s hypothesis of higher cortical arousal in introverts than in extraverts. All of their participants were female. Stenberg, Wendt, and Risberg (1993) also found an overall negative correlation \( r = -0.37 \), but this was a function of the high correlation among the female participants; the correlation among the males was close to zero. As with the EEG data, confirmation of the hypothesis was more common in female than in male samples.

Some studies have found hemispheric differences in the relationships between E and activation, but these have not been consistent (Johnson et al., 1999; Stenberg et al., 1993). Studies of subcortical areas of brain have also yielded little in the way of consistent findings except for one: E is associated positively with activation of the anterior cingulate area (Ebmeier et al., 1994; Haier, Sokolski, Katz, & Buchsbaum, 1987; Johnson et al., 1999). The cingulum is the major pathway between the frontal cortex and the limbic system and has been theoretically associated with neuroticism and anxiety rather than E (Zuckerman, 1991).

The results in the two brain imaging studies described, unlike the EEG studies, tend to support Eysenck’s hypothesis of a relationship (albeit a weak one) between E and cortical arousal. There is no clue in his theory, however, why the finding is supported more in females than in males or why subcortical differences in the cingulum, the executive structure of the limbic brain, should be associated with extraversion. In Eysenck’s model limbic arousability is associated with neuroticism, and any association with E would be with introversion rather than extraversion.

General arousal may be too broad a construct to be associated with personality. Arousal is highly dependent on diurnal variation and general stimulation levels. Arousal as a trait would represent the state of the nervous system at a given time under a given set of conditions. In contrast, arousability is the typical immediate reaction of some part of the nervous system to a stimulus with specified characteristics. Eysenck’s (1967) optimal level of stimulation model says that introverts are more arousable at low to moderate intensities of stimulation, but at higher intensities extraverts are more responsive. Introverts have strong reactive inhibition mechanisms that dampen response to high intensities. Strelau (1987), in a model based on neo-Pavlovian theories, states that persons with strong nervous systems are relatively insensitive to stimuli at lower intensities but can process and react to stimuli at higher intensities. For weak nervous system types the opposite is true: They are highly sensitive to low intensities but show inhibition of response at high intensities.

**Cortical Arousal**

Cortical arousability is usually assessed with the cortical evoked potential (EP). A brief stimulus, such as a tone or flash of light, is presented a number of times, and the EEG is digitized at a fixed rate, that is time locked to stimulus delivery time and averaged across trials for a given participant. This process averages out the “noise” and produces a clear waveform representing the typical reaction of that subject to the specific stimulus over a 500-ms period. Although latencies of response vary somewhat for individuals, for most one can identify particular peaks of positivity and negativity. For instance, a peak of positive potential at about 100 ms after the stimulus (P1) represents the first impact of the intensity characteristics of stimuli on the cortical centers. Earlier peaks represent stimulus processing at subcortical centers. The peak at 300 ms after the stimulus (P3) is influenced by novelty, surprise, or unexpectedness of the stimulus and thus represents a higher level of cortical processing in that the stimulus must be compared with previous stimuli.

Stelmack (1990) reviewed the relationship between E and cortical EPs. As might be expected, the results depend on the characteristics of the stimuli used to evoke the EPs as well as the reactor’s age and personality characteristics. For instance, Stelmack said that introverts have greater amplitude EPs in response to low-frequency tones, but there are no differences between introverts and extraverts for high-frequency tones.
If the stimulus attribute had been intensity, these kinds of results might be compatible with Eysenck’s theory of increased sensitivity of introverts to low-intensity stimuli. But the evolutionary type of explanation offered by Stelmack for the greater survival significance of low-frequency sounds is not convincing.

Recent studies have focused on the P300 EP component, many using the “odd-ball” paradigm in which the participant listens with eyes closed to a sequence of tones in which one tone is presented frequently and another one (the oddball) rarely. The rare tone is the signal for some task. These are usually vigilance tasks on which extraverts’ performances and EP reactions are expected to decline more rapidly than those for introverts. However, when the task is made less monotonous or response requirements are high, the differences may disappear or even be reversed with larger EP amplitudes in extraverts (Stenberg, 1994).

The intensity of the stimulus is another factor in the I-E difference. Brocke, Täsche, and Beauducel (1997) found that introverts showed larger P3 reactions to a 40-db stimulus, whereas extraverts showed a larger amplitude of P3 in response to a 60-db stimulus. Introverts’ EP amplitudes decreased going from 40 db to 60 db, whereas extraverts increased going from the less intense to the more intense stimulus. These effects were a function of the impulsivity component rather than the sociability component of the E scale used in the study. The results of studies that vary the experimental conditions suggest that attention and inhibition may be the basic mechanisms governing the nature of the relationship between E and cortical EPs. Responses at the brain-stem level are probably less susceptible to these mechanisms, and Eysenck’s theory does involve the brain stem and other points along the reticulocortical arousal system in I and E.

Stelmack and Wilson (1982) found that extraverts had longer latencies for the EP subcortical wave V (inferior colliculus) for stimulus intensity levels up to but not including 90 db. The direction of the finding was confirmed in a second experiment (Stelmack, Campbell, & Bell, 1993) and in a study by Bullock and Gilliland (1993). Different doses of caffeine and levels of task demand were used in the latter study, but the differences between extraverts and introverts held across all levels of caffeine and task demand. The results support Eysenck’s theory more strongly than those using cortical EPs, which seem more susceptible to stimulus, task, and background arousal factors. A study by Pivik, Stelmack, and Bylsma (1988), however, suggested that Eysenck’s arousal-inhibition hypothesis may not be broad enough. These researchers measured the excitability of a spinal motoneuronal reflex in the leg and found that extraverts showed reduced motoneuronal excitability as measured by reflex recovery functions. These results show that the inhibitory properties of the nervous system related to E may extend well below the reticulocortical level.

Another line of EP research is based on Gray’s (1982, 1987) model of personality. Gray proposed that impulsivity, a dimension close to extraversion, is related to sensitivity to signals (conditioned stimuli) of reward whereas anxiety, close to neuroticism, is related to sensitivity to signals of punishment. This model suggests that the learned biological significance of stimuli, in addition to the intensity of stimulation, governs the strength of reaction to them.

Bartussek, Diedrich, Naumann, and Collet’s (1993) results supported the theory by showing a stronger EP response (P2, N2) of extraverts than introverts to tones associated with reward (winning money) but no differences in tones associated with punishment (losing money). In a later experiment, however, extraverts showed larger P3 EP amplitudes to stimuli associated with both reward and punishment compared to neutral stimuli (Bartussek, Becker, Diedrich, Naumann, & Maier, 1996).

DePascalis and his colleagues also presented findings supporting Gray’s theory. In one study they used a questionnaire scale developed more directly from Gray’s theory measuring the approach tendency (DePascalis, Fiore, & Sparita, 1996). Although they found no effect for E itself, the participants scoring high on the approach scale had higher EP (P6) amplitudes in response to stimuli (words) associated with winning than to those associated with losing, and the reverse was true for low-approach motive subjects.

Eysenck’s and Gray’s theories have also been tested using peripheral autonomic measures of activity like the electrodermal activity (EDA), or skin conductance (SC), heart rate (HR), and blood pressure (BP). These are only indirect measures of cortical activity and reactivity because they occur in the autonomic nervous system (ANS) and are controlled by limbic system centers, which in Eysenck’s model are associated more closely with neuroticism than with E. The results in relation to E are similar to those obtained with more direct cortical measures. Reviews by Smith (1983) and Stelmack (1990) showed mixed and inconclusive findings relating tonic EDA arousal to E, but some evidence of stronger SC responses of introverts than extraverts in response to low-to moderate-intensity stimuli and stronger responses of extraverts in response to high-intensity stimulation. Tonic (base-level) measures of HR (Myrtek, 1984) and BP (Koehler, Scherbaum, Richter, & Boettcher, 1993) are unrelated to E. Young children rated as shy and inhibited had higher and less variable HRs, and a high HR at 21 months is the same behavior pattern at 48 months (Kagan, Reznick, &
Snidman, 1988). Shyness and inhibition, however, are traits that are a mixture of introversion and neuroticism or anxiety; therefore, the correlation with HR could be due to the anxiety component rather than to E.

Eysenck’s model for the trait of extraversion produced a great deal of research in the area of psychophysiology. But psychophysiology has its problems as a branch of neuroscience. Both tonic and phasic psychophysiological measures are highly reactive to environmental conditions. Tonic levels can vary as a function of reactions to the testing situation itself, and phasic reactions depend on the specific qualities of stimulation such as intensity and novelty. It is not surprising that the relationships of physiological measures with personality traits often interact with these stimuli characteristics in complex ways. Eysenck’s theory based on optimal levels of stimulation has received some support. Those based on differences in basal arousal levels are beginning to receive some support from PET studies, although the earlier results with EEG measures remain problematic.

Monoamines

The monoamine neurotransmitter systems in the brain have been the focus of most biosocial theories of personality. The reasons are the evidence of their involvement in human emotional and cognitive disorders and basic emotional and motivational systems in other species. Much of the work with humans has been correlational, comparing basal levels of the neurotransmitters, as estimated from levels of their metabolites in cerebrospinal fluid (CSF), blood, or urine, to personality traits as measured by questionnaires. Of these sources CSF is probably the best because the CSF is in direct contact with the brain. But the indirect relationship of these indicators with brain levels of activity (which can differ in different brain loci) and the fact that some of the metabolites in plasma and urine are produced in the peripheral nervous system make the putative measures of brain amine activity problematic. New imaging methods may eventually overcome these problems by directly viewing the monoamine activities in the brain itself. Added to these problems of validity of measurement is the use of small numbers of subjects in most studies, as well as the use of subjects with certain types of disorders rather than normal subjects. The ethical constraints of giving drugs that affect activity in the brain systems is another barrier, although some of the more recent studies have used such drugs in normals.

The freedom of investigators to experiment directly with the brain in other species has given us a fairly coherent picture of the emotional and motivational functions of the monoamine systems in the brain, and bottom-up theorists have used these findings to extend animal models to human motivations and personality (Gray, 1982, 1987; Mason, 1984; Panksepp, 1982; Soubrí, 1986; Stein, 1978). Top-down theorists have drawn on these findings from the comparative research but have attempted to reconcile them with the relevant research on humans, including clinical and personality studies (Cloninger, Svrakic, & Prszybeck, 1993; Depue & Collins, 1999;Netter, Hennig, & Roed, 1996; Rammsayer, 1998; Zuckerman, 1991, 1995). The problem with building a bridge from two banks is to make it meet in the middle. With these caveats let us first examine the case for extraversion.

The primary monoamines in the brain are norepinephrine, dopamine, and serotonin. The first two are labeled catecholamines because of the similarities in their structures. Serotonin is an indoleamine. These are not independent neurotransmitter systems because activity in one may affect activity in another. Serotonin, for example, may have antagonistic effects on the catecholamines. These kinds of interaction must be kept in mind because most studies relate one neurotransmitter to one personality trait. Some models suggest that this kind of isomorphism of trait and transmitter is the rule. This is a new kind of phrenology based on biochemistry rather than bumps on the head.

To understand the human research one needs to know the pathways of biosynthesis and catabolism (breakdown) of the monoamines because some experiments block the precursors of the transmitter to see its effect on behavior and most use metabolite products of the catabolism to gauge activity in the systems. Figure 4.2 is a simplified diagram showing the

![Figure 4.2](image_url)
stages of production of the monoamines and some of the enzymes (DBH, COMT, MAO) involved in the conversions from one stage to another. The metabolite for dopamine is homovanillic acid (HVA), for norepinephrine it is 3-methoxy-4-hydroxyphenylglycol (MHPG), and for serotonin it is 5-hydroxyindoleacetic acid (5-HIAA).

Theorists are in fair agreement on the role of dopaminergic systems in motivation based on studies of other species: approach and sensitivity to stimuli associated with reward (Crow, 1977; Gray, 1982, 1987; Stein, 1978); foraging and exploration and positive emotions like hope, desire, and joy in humans (Panksepp, 1982; Zuckerman, 1991); and novelty or sensation seeking in animals and humans (Bardo, Donohew, & Harrington, 1996; Cloninger et al., 1993; Le Moal, 1995; Zuckerman, 1984, 1991). I have proposed that the activity of the mesolimbic dopamine system is related to a broad approach trait that includes extraversion, sensation seeking, and impulsivity (Zuckerman, 1991). Considering that dopaminergic reactivity is also related to aggression and sexuality in many species, it is also possible that the third dimension of personality, low socialization, or psychoticism, may also be involved. Gray’s (1987) model linked dopamine and reward sensitivity with impulsivity, a dimension related to high E, P, and N, although his more recent remarks (Gray, 1999) suggest that he is linking dopamine more closely with the P dimension because of this transmitter’s involvement in schizophrenia.

Depue and Collins (1999) defined a broad view of extraversion with two main factors: interpersonal engagement, or affiliation and warmth, and agency, which includes social dominance, exhibitionism, and achievement motivation. Positive affect and positive incentive motivation are more strongly associated with the agentic extraversion factor. Impulsivity and sensation seeking are regarded as constituting an emergent factor representing a combination of extraversion and constraint (a dimension related to Eysenck’s P and Costa and McCrae’s conscientiousness). The “lines of causal neurobiological influence” are suggested to lie along the orthogonal dimensions of extraversion and constraint rather than along the dimension of impulsive sensation seeking. Although Depue and Collins say that this structural system does not mean that positive incentive motivation and its dopaminergic basis are related only to extraversion, the expectation is that they will be more strongly related to agentic extraversion than to impulsive sensation seeking or constraint.

Only a few correlational studies of monoamine CSF metabolites and personality traits were done prior to 1991 (Zuckerman, 1991), and they generally showed few significant relationships between the dopamine metabolite HVA and either extraversion or sensation seeking. This is still the case with studies that simply correlate CSF levels of HVA with questionnaire measures of extraversion, even when there is sufficient power to detect weak relationships (Limson et al., 1991). In fact, the Limson et al. study failed to find any correlations between CSF metabolites of serotonin (5-HIAA), norepinephrine (MHPG), norepinephrine itself, and Dopac and any of the personality measures assessed by the Minnesota Multiphasic Personality Inventory (MMPI), Eysenck Personality Questionnaire (EPQ), or Cloninger’s Temperament Character Inventory (TCI). As with psychophysiological measures, levels of neurotransmitter activity in a resting basal state are not sensitive to variations in personality, at least as the latter is measured in self-report questionnaires. However, studies that attempt to potentiate or attenuate activity in neurotransmitters with agonists or antagonists have yielded some significant findings in regard to personality, even though they typically use very small sample sizes.

Depue, Luciana, Arbisi, Collins, and Leon (1994) challenged the dopamine system with bromocriptine, a potent agonist at D2 receptor sites, and measured the effects using inhibition of prolactin secretion and activation of eye-blink rate, two measures of dopamine activation. The correlations between Positive Emotionality (PE) and baseline measures of the dopamine activity indicators were small and insignificant, but they found significant correlations between the putative measures of dopamine response to the agonist and the PE (an extraversion type measure) factor from Tellegen’s MPQ. Rammsayer’s interpretation of their findings is indicative of higher dopamine reactivity in high-PE persons (extraverts) than in lows, suggesting that the prolactin response would indicate just the reverse (i.e., higher reactivity in the low-PE persons). The disagreements on the meaning of the data are too complicated to elucidate here.

Rammsayer’s interpretation of the findings is supported by PET measures of higher cerebral blood flow to the dopaminergic basal ganglia areas in introverts than in extraverts (Fischer, Wik, & Fredrikson, 1997); but another PET study found no relationship between E and dopamine binding in the basal ganglia (N. S. Gray, Pickering, & Gray, 1994), and still another found a positive relationship with E (Haier et al., 1987). The first two of these studies used normal controls as subjects whereas the Haier et al. study used patients with Generalized Anxiety Disorder, a possible confounding factor.

Rammsayer, Netter, and Vogel (1993), using an inhibitor of tyrosine hydroxylase, thereby blockading dopamine synthesis, found no difference between introverts and extraverts in
either baseline dopamine or reactivity to the blockading agent. Despite the lack of difference in dopaminergic activity or reactivity, they found that reaction time performance was markedly impaired in introverts but not in extraverts by the dopamine blocking agent. In another study, using a chemical that selectively blocks D2 receptors and inhibits dopamine neurons in the limbic and cortical regions of the brain, Rammsayer (1998) again found a detrimental effect on reaction (liftoff) time in introverts but not in extraverts. The agent that was used caused a marked decrease in alertness and cortical arousal, but this effect was equivalent in introverts and extraverts. Both this finding and the performance findings would seem to contradict Eysenck’s arousal explanation for the differences between introverts and extraverts. That theory would predict a more detrimental effect in extraverts because they supposedly start with a lower level of cortical arousal. But the results also raise the question, What is the source of the performance differences between introverts and extraverts if they do not differ in dopamine activity or reactivity?

The answer might lie in the interactions of dopaminergic and other neurotransmitters or hormones or, at another level, in the genetics of the dopaminergic receptors. Considerable interest has developed in a gene associated with the dopamine receptor 4 (DRD4). Allelic variations in this gene have been associated with novelty or sensation seeking, but not with extraversion (Ebstein, Nemeroff, Klotz, Gitsenko, & Belmaker, 1997; Ebstein et al., 1996).

Simple correlative studies have found no relationship between serotonin or norepinephrine and E or other personality variables measured by questionnaires given to adult subjects. A study using CSF from newborns in predicting temperament traits found that infants born with low levels of the serotonin metabolite 5-HIAA showed low sociability at 9 months of age (Constantino & Murphy, 1996). Retest reliability for 5-HIAA in neurologically normal infants was very high (r = .94).

A study of adults with depressive disorder treated with either a noradrenergic or a serotonergic reuptake inhibitor, which increase activity in those systems, showed that there were significant increases in measures of E and gregariousness (sociability) in those treated with these drugs (Bagby, Levitan, Kennedy, Levitt, & Joffe, 1999). The change in E was correlated with the change in depression severity, but the change in sociability was not. Although the result with sociability probably represents a change of state rather than the preillness trait, serotonin and norepinephrine might play some role in the trait as well. Studies of serotonin transporter genes have not shown any relationship to E, although they have to other personality traits (Hamer, Greenberg, Sabol, & Murphy, 1999; Jorm, Henderson, Jacomb, Croft, & Easteal, 1997).

**Monoamine Oxidase**

Monoamine oxidase (MAO) is an enzyme involved in the catabolic deamination of monoamines. Evidence using selective monoamine inhibitors suggests that MAO-Type B, assayed from blood platelets in humans, is preferentially involved in the catabolic breakdown of dopamine more than the other two brain monoamines, norepinephrine and dopamine (Murphy, Aulakh, Garrick, & Sunderland, 1987). Although no direct correlation of platelet and brain MAO has been found, indirect assessments and the effects of MAO inhibitors on depression, as well as a large body of behavioral data, suggest that there must be a connection, if only one limited to certain brain areas. Platelet MAO is normally distributed in the human population, is highly reliable although it increases in brain and platelets with age, and is lower in men than in woman at all ages, and variations are nearly all genetic in origin. Unlike other biochemical variables it does not vary much with changes in state arousal. Thus, MAO has all of the characteristics of a biological trait.

Low levels of MAO-B taken from umbilical cord blood samples in newborn infants were related to arousal, activity, and good motor development (Sostek, Sostek, Murphy, Martin, & Born, 1981). High levels of the enzyme were related to sleep time and general passivity. The relationship with motor development is particularly suggestive of development of the dopamine-influenced basal ganglionic areas of the brain involved in motor coordination. In a study of monkeys living in a colony in a natural environment, low-platelet MAO was related to high sociability, activity, dominance, and sexual and aggressive activity, a broad array of E-type traits described by Depue and Collins (1999) as agentic extraversion. However, in human correlative studies the results relating MAO-B to questionnaire-measured extraversion have been inconsistent (Zuckerman, 1991). The enzyme has more consistently correlated (inversely) with the trait of sensation seeking. But using reported behavioral indices of sociability in college students, low MAO was related to sociability and high MAO to social insolation (Coursey, Buchsbaum, & Murphy, 1979).

**Hormones**

The hormone testosterone (T) is produced by both men and women but is 8 to 10 times as high in men as in women. Plasma T is highly heritable (66%) in young adult males and...
moderately heritable (41%) in females (Harris, Vernon, & Boomsa, 1998). In rats T has reward effects in the nucleus accumbens, the major site of dopaminergic reward. Administration of a dopamine receptor blocker eliminates the rewarding effects of T in rats, suggesting that its rewarding effects are mediated by an interaction with dopamine in the mesolimbic system (Packard, Schroeder, & Gerianne, 1998).

The hormone T affects personality traits and may account in part for many of the personality trait differences between men and women. Men and women do not differ on the pure sociability or affiliative type of extraversion, but they do on the agentic type, which includes dominance, assertiveness, surgency, and self-confidence. To the extent that sensation seeking is associated with extraversion, it is with the agentic type.

Daitzman and Zuckerman (1980) found that T in young males was positively correlated with sociability and extraversion, as well as with dominance and activity and inversely with responsibility and socialization, indicating an association with the agentic type of extraversion. Windle (1994) also found that testosterone was associated with a scale measuring behavioral activation, characterized by boldness, sociability, pleasure seeking, and rebelliousness. Dabbs (2000) also found that T is associated with a type of extraversion characterized by high energy and activity levels and lower responsibility.

Summary

Eysenck’s theory relating cortical arousal to extraversion has been extensively tested using the EEG and, in more recent times, the brain scanning methods. The EEG studies yielded mixed results in which the sources of differences between studies were not clearly apparent. Two cerebral blood flow studies did confirm that extraverts were cortically underaroused related to introverts in female subjects but not in males. Studies measuring cortical arousability have also not clarified the picture. Apparently, experimental conditions affecting attention or inhibition may confound the relationship with E. Some more consistent results have been obtained from EP studies of responses at subcortical levels in which conscious attention is less of a factor. Although Eysenck’s theory is confined to cortical arousal and reactivity, differences between introverts and extraverts have been found at lower levels of the central nervous system, even in a spinal motoneuronal reflex.

Theories of the biochemical basis of extraversion have focused on the monoamine neurotransmitters, particularly dopamine. Simple correlational studies between the monoamine metabolites and trait measures of E have not yielded significant findings, although there is some evidence that drugs that increase noradrenergic or serotonergic activity in depressed patients also increase their extraversion and sociability. This may be an indirect effect of the reduction in depression rather than a direct effect on E. The enzyme MAO-B is involved in regulation of the monoamines, particularly dopamine. Low levels of MAO have been related to arousal and activity in newborn human infants and to sociable behavior in adult humans and monkeys. These results suggest that a dysregulation of the dopamine system may be a factor in extraversion even in its earliest expression in the behavior of newborns. The hormone testosterone is related to E, but more so to E of the agentic type, which is the type characterized by dominance, assertiveness, surgent affect, high energy levels, activity, and irresponsibility, rather than simple sociability and interest in social relationships. This distinction between the two types of E has been hypothesized to be crucial for the relationship between dopamine and E as well (Depue & Collins, 1999).

NEUROTICISM/ANXIETY/HARM AVOIDANCE

Although the broad trait of neuroticism/anxiety includes other negative emotions, such as depression, guilt, and hostility, and character traits such as low self-esteem, neuroticism and anxiety are virtually indistinguishable as traits. Neuroticism is highly correlated with measures of negative affect, but when the negative affect was broken down into anxiety, depression, and hostility components, anxiety had the highest correlation, and hostility the lowest, with the N factor while depression was intermediate (Zuckerman, Joireman, Kraft, & Kuhlman, 1999). Hostility had a higher relationship to a factor defined by aggression.

Eysenck (1967) assumed a continuity between N as a personality trait and anxiety disorders. Indeed, N is elevated in all of the anxiety and depressive mood disorders, and longitudinal studies show that the trait was evident in most persons before they developed the symptoms of the clinical disorder (Zuckerman, 1999). In the first half of the twentieth century, when little was known about the role of the limbic system in emotions, the biological basis of neuroticism and anxiety trait was related to overarousal or arousability of the sympathetic branch of the autonomic nervous system. Such arousal is apparent in state anxiety elicited by anticipation of some kind of aversive stimulus or conditioned stimuli associated with aversive consequences.

Autonomic overarousal is apparent in the primary symptoms of many anxiety disorders. On the assumption of continuity between the N trait and these disorders, it was expected that autonomic arousal, as assessed by peripheral measures such as heart rate (HR), breathing rate (BR), blood pressure (BP), and electrodermal activity (EDA), would be correlated with N. In Eysenck’s (1967) theory, N was ultimately based.
on reactivity of the limbic system, which regulates the ANS, but he did not distinguish particular pathways, structures, or neurotransmitters within that system that were involved in N. Some theories did not even make a distinction between cortical and autonomic arousal in emotions. Eysenck felt that there was some correlation between the two kinds of arousal because of collaterals between the limbic and ascending reticulocortical system. Gray (1982) and others, extrapolating from experimental studies of animals, delineated specific limbic systems involved in anxiety and the neurotransmitters involved in these systems. Neuroimaging studies have attempted to extend these brain models to humans.

**Autonomic Arousal**

Large-scale studies of the relationship between cardiovascular measures, either in resting levels of activity or reactivity to stressful experimental situations, and Measures of N failed to reveal any significant relationships (Fahrenberg, 1987; Myrtek, 1984). On the assumption that high cardiovascular activity put high-N subjects at risk for cardiovascular disease, Almada et al. (1991) investigated the relation between measures of N and subsequent health history in nearly 2,000 men. N was not associated with systolic BP or serum cholesterol but was associated with cigarette smoking and alcohol consumption. When tobacco and alcohol consumption were held constant there was no relationship between N and cardiovascular disease. Similar studies have failed to find any relationships between electrodermal activity and N or trait anxiety (Fahrenberg, 1987; Hodges, 1976; Naveteur & Baque, 1987).

Given the fact that many anxiety disorders do show elevated heart rate and electrodermal reactivity, how can we explain the lack of correlation with N? The answer may lie in the difference between generalized anxiety disorder (GAD) and panic disorder (PD), agoraphobia (Ag), and obsessive-compulsive disorder (OCD). Whereas the latter (PD, Ag, OCD) show elevated basal HRs and frequent spontaneous SCRs, GAD patients show little evidence of this kind of autonomic arousal (Zuckerman, 1991). Their anxiety is expressed cognitively (worry) and in symptoms of muscle tension such as fatigue. In contrast, PD, Ag, and OCD patients complain of autonomic symptoms, such as accelerated heart rate, even when they are not experiencing an actual panic attack (Zuckerman, 1999). Most persons who are high on N probably represent subclinical GAD disorder rather than the other types of anxiety disorders.

**Brain Arousal**

Studies of general cortical arousal using the EEG have historically focused on E, but some of these studies found interactions with N. These effects were inconsistent; some found higher and some reported lower arousal for high-N persons. Application of PET methods has not shown any association of general cortical or limbic arousal with N in situations that were not emotionally provoking (Fischer et al., 1997; Haier et al., 1987). Similar results are seen in anxiety patients; but when anxiety is provoked in patients by presenting them with feared stimuli, increased activity is seen in areas like the orbitofrontal cortex, insular cortex, temporal cortex, and anterior cingulate (Breier et al., 1992; Rauch et al., 1995). These studies identify an anxiety pathway in humans (orbitofrontal-frontal to cingulate to temporal lobe and amygdala) already established in animals, but they do not show a preexisting sensitivity of this pathway in normals scoring high in N. Another study of anxiety patients in non-stimulated conditions, which did use normal controls, found that whole brain blood flow did not distinguish anxiety patients from normals but did find a negative correlation between a depression scale and caudate activation. The previously mentioned study by Canli et al. (2001) found that in a small sample of normal women N correlated with increased brain activation to negative pictures (relative to activation by positive pictures) in left-middle frontal and temporal gyri and reduced activation in the right-middle frontal gyrus. Taken together, the clinical studies and this last study of normals suggests that whole brain activation does not vary with N-Anx, but given negative emotional provocation there may be a reactive disposition in frontal cortex of high-N persons that activates a pathway through the orbitofrontal cortex around the cingulum to the temporal lobe and amygdala.

Davis (1986) argued that the central nucleus of the amygdala is a major center where the input of fear-provoking stimuli is organized and where output to various intermediate nuclei organizes the entire range of behavioral, autonomic, and neurotransmitter reactions involved in panic or fear. A recent MRI study (van Elst, Woermann, Lemieux, & Trimble, 1999) found an enlargement of left and right amygdala volumes in epileptic patients with dysthymia (a chronic kind of neurotic depression). Amygdala volume within the group did not correlate with trait or state anxiety but did correlate positively with a depression inventory. Because anxiety and depression are usually highly correlated and both correlate highly with N, it is not clear why depression alone was related to amygdala volume.

**Monoamines**

Much of the recent exploration of the role of the monoamines in N-Anx have been based on Cloninger’s (1987) biosocial model of personality and therefore used his scale of Harm
Avoidance (HA) instead of the N or anxiety trait scales used by other investigators. HA, however, is not a pure scale of the N factor but lies between the E and N dimensions, constituting a measure of introverted neuroticism. It is defined in the same way that Gray defines trait anxiety: a sensitivity to cues associated with punishment and nonreward (frustration) and a tendency to avoid them.

Gray’s (1982) model suggests that norepinephrine in the dorsal ascending noradrenergic system (DANA) originating in the locus coeruleus is the major neurotransmitter involved in anxiety, although high levels of serotonin may mediate the behavioral inhibition that is associated with high levels of anxiety. Redmond (1977), from a psychiatric viewpoint, sees the DANA as an alarm system at lower levels and a panic provoker at high levels of activity. In contrast to these two theorists, Cloninger, Svrakic, and Przybeck (1993) proposed that high levels of serotonin activity underlie the trait of HA whereas norepinephrine activity is related to another trait called Reward Dependence.

In patients there has been little evidence of higher levels of the norepinephrine metabolite MHPG in anxiety patients compared to normals, although a more recent study by Spivak et al. (1999) showed higher levels of MHPG in plasma of patients with combat-related posttraumatic stress disorder than in controls.

The alpha-2 receptor functions as a homeostatic regulator of the norepinephrine systems, tuning them down when excessive neurotransmitter levels are detected in the synapse. Yohimbine is an antagonist to this receptor and therefore potentiates the activity of the norepinephrine system, just as a broken thermostat results in an overheated room. Yohimbine increases MHPG levels and provokes panic attacks in patients with panic disorders, although it does not have these effects in normal controls (Charney & Heninger, 1986). Cameron et al. (1996) replicated a previous result finding a decreased number of alpha-2 receptors in panic disorder. One might extrapolate that MHPG should correlate with N or anxiety over the range in normals and other patient groups. However, as noted earlier, high N in normals may resemble GAD more than panic disorder. Heinz, Weingarten, Hommer, Wolkowitz, and Linnoila (1999) reported a high correlation between CSF MHPG and an anxiety scale in a combined group of abstinent alcoholics and normals. A stress resistant group, defined by N and similar measures, had lower plasma MHPG after a mild stressor than did a nonresistant (high-N) group (de Leeuwe, Hentschel, Tavenier, & Edelbroek, 1992). Norepinephrine may be one of the factors underlying N, but it may be the dysregulation of norepinephrine by a lack of the receptors needed for this and a consequent tendency to be unable to cope with stress, rather than the basal level of activity in the norepinephrine system, which is related to N.

Cloninger’s biosocial theory of personality proposes that the trait of harm avoidance is related to behavioral inhibition mediated by serotonergic activity in the brain. Earlier studies showed no correlation between between CSF levels of the serotonin metabolite, 5-HIAA, and N. A more recent study has found a positive correlation between CSF 5-HIAA and N but in a sample of depressed patients (Roy, 1999). Constantin and Murphy’s (1996) study of the prediction of infant temperament from CSF levels of 5-HIAA showed no relationship between this metabolite and emotionality, soothability, or activity in infants.

Studies of normals using serotonin challenges, drugs that stimulate serotonergic activity, and indirect measures of serotonin response in normals have yielded mixed results including both positive (Gerra et al., 2000; Hansenne & Ansseau, 1999), nonsignificant (Ruegg et al., 1997), and a negative relationship (Mannuck et al., 1998) with N. The first three of these studies used the HA scale, whereas the last used the N scale, but with a much larger number of normal subjects than in the other studies. Serotonin seems to be implicated in harm avoidance, but the nature of that relationship is open to question. As with other neurotransmitters, the personality-relevant aspects of serotonin may have more to do with receptor number and sensitivity than with basal levels of transmitter activity.

**Hormones**

Daitzman and Zuckerman (1980) found that testosterone (T) in males correlated negatively with various MMPI indexes of anxiety, depression, and neuroticism; that is, subjects with neurotic tendencies were low on T. Dabbs, Hopper, and Jurkovic (1990) reported a significant negative correlation between T and N in one study, but this was not replicated in another larger study of males; and in an even larger study of over 5,000 veterans T was not correlated with any MMPI indexes of neurotic tendencies were low on T. Dabbs, Hopper, and Jurkovic (1990) reported a significant negative correlation between T and N in one study, but this was not replicated in another larger study of males; and in an even larger study of over 5,000 veterans T was not correlated with any MMPI indexes of trait anxiety or N. In still another study Dabbs et al. report significant negative correlations between T and a measure of pessimism in both males and females. T reflects both trait and state characteristics; that is, it is affected by immediate stressful experiences, particularly those involving success or defeat in competitive activities (Dabbs, 2000). The relationship with pessimism may reflect a history of defeat and consequent expectations for future failures. This depressive attitude may underlie negative relationships with N if any such relationships do exist.

Cortisol is one of the end products of activation of the hypothalamic-pituitary adrenocortical (HYPAC) system, a stress-reactive hormonal system. Like T, cortisol reactivity has both trait and state characteristics. Elevated cortisol is associated with major depressive disorder as a trait but is
found in anxiety disorders only when activated by an immediate stressor.

**Molecular Genetics**

Lesch, Bengal, Hells, and Sabol (1996) found an association between a serotonin transporter gene (5-HTTLPR) and the trait of neuroticism, as assessed by three different scales including the NEO N scale and Cloninger’s TCI harm avoidance scale. Individuals with either one or two copies of the short form had higher N scores than individuals homozygous for the long variant of the gene. The association was limited to the N factor of the NEO and the harm avoidant factor of the TCI; none of the other factors in these tests was associated with the genetic variant. However, in a second study by this group (Hamer, Greenberg, Sabol, & Murphy, 1999) the association of the gene with harm avoidance was weaker, and associations were found with TCI traits of cooperativeness and self-directiveness.

Several other studies have not been able to replicate the relationship between the gene variants and N or harm avoidance. This is a common outcome in the hunt for specific genes associated with personality traits or types of psychopathology, even when studies have adequate power and use good methodology. Population differences may account for some of these failures. Even in the studies that are significant the particular gene accounts only for a small portion of the genetic variance. In the Lesch et al. study the 5-HTT polymorphism accounted for 3% to 4% of the total variance for the trait and 7% to 9% of the genetic variance, and 10 to 15 more genes were estimated to be involved. If there is any replication of a gene-trait association, that finding should not be immediately dismissed by subsequent failures of replication, particularly if the finding has a theoretical basis. In this case Cloninger’s theory has suggested the involvement of serotonin in harm avoidance.

The short form of the gene, which is associated with high neuroticism, reduces serotonin uptake and therefore increases serotonergic transmission. Reduced uptake has been associated with anxiety in animal and human models, but paradoxically the serotonin uptake inhibitors are therapeutic agents in depressive disorders and several forms of anxiety disorders. These drugs could achieve their results through the inhibitory effects of serotonin on other systems such as the noradrenergic ones.

**Summary**

A sudden intense surge in anxiety is characterized by arousal of the sympathetic branch of the autonomic nervous system as expressed in elevated heart and breathing rates, blood pressure, sweating, and other signs of activation of this system. This led to the expectation that N or trait anxiety would be related to measures of these indicators either in the basal state or in reaction to stress. Research has generally failed to support this correlational hypothesis. EEG and brain scan studies also fail to reveal a difference in arousal levels as a trait distinguishing high- and low-N individuals. However, PET scan studies, done primarily on patients with anxiety disorders in reaction to fearful stimuli, show heightened reactivity of frontal, insular, and temporal cortex and anterior cingulate to such stimuli. Evidence from studies of animals has implicated the amygdala as a center for organization of the fear response, but brain imaging studies in humans have not yet supplied evidence for this localization.

Much of the research on other species identifies activation of the dorsal ascending noradrenergic system originating in the locus coeruleus as an alarm system activating the entire cortex in states of fear or anxiety. Reactivity of this system is a characteristic of panic disorders during panic attacks compared to the reactions of other types of anxiety disorders and normal controls. Correlational studies of norepinephrine metabolites and N-type trait measures in the basal state have not found a relationship, but at least one study has found a relationship between N and reactivity of a norepinephrine metabolite and response to stress. A hypothesized relationship with the monoamine serotonin has also shown no relationship with N in the basal state and no consistent findings relating N to reactions to drugs that stimulate serotonergic activity. Initial findings of a relationship between a serotonin transporter gene and N-type scales have not been replicated. Hormones like testosterone and cortisol show similar negative findings in the basal state and few findings relating N to reactivity to stress.

The research attempting to find a biological basis for N has had a disappointing outcome, particularly in view of the positive results in experimental research with animals and with humans that suffer from anxiety and mood disorders. Longitudinal research has shown that N is a personality precursor of these disorders, so why does N not show relationships with some of the same biological indicators that characterize the disorders? There may be a kind of threshold effect so that the dysregulation of neurotransmitter systems characteristic of the disorders only emerges at some critical level of persistent stress that is not reproducible in controlled laboratory studies.
Eysenck’s psychoticism scale is one of the best markers for the dimension that consists of scales for impulsivity and sensation seeking at one pole and scales for socialization, responsibility, and restraint at the other pole (Zuckerman et al., 1988, 1991, 1993). In a three-factor solution this factor also includes aggression and capacity to inhibit aggression, but in a four- or five-factor solution aggression and hostility versus agreeableness form a separate factor (Zuckerman et al., 1993). This chapter is organized by the four-factor model.

Cortical Arousal and Arousability

At the time the original studies were done relating conditioning to arousal and the construct “strength of the nervous system” to extraversion, E was measured by scales with two components: E and Impulsivity (Imp). In a theoretical shift, not receiving much attention, Eysenck and Eysenck (1985) reassigned Imp to the P rather than the E dimension. Although nearly all the earlier arousal and conditioning studies focused on E, it was shown that the relationship of E to conditionability (introverts more conditionable than extraverts) depended on the Imp component of E rather than the sociability component (Barratt, 1971; Eysenck & Levey, 1972). A later study showed that classical eyelid conditioning was related most closely to a specific type of Imp, the tendency to act quickly on impulse without thinking or planning. This is the type of Imp, called narrow impulsivity (IMPn), that constitutes a subscale of the older E scale. It is also the type of Imp that has been combined with sensation seeking in the latest ImpSS scale. Conditionability is thought to be a function of arousal; the more aroused a person is, the more conditionable he or she is thought to be. Could this mean that cortical arousal is related to the third dimension (P), including sensation seeking and IMPn, rather than the first (E) dimension of personality? A PET study found negative correlations between P and glucose use in cortex and in thalamic and cingulate areas of the limbic system (Haier et al., 1987). Low cortical and autonomic arousal is a characteristic of the psychopathic (antisocial) personality, which may represent an extreme manifestation of the P dimension of personality (Zuckerman, 1989).

Evidence for a relationship between cortical arousal (EEG) and P and IMPn was found by some investigators (Goldring & Richards, 1985; O’Gorman & Lloyd, 1987); high P and impulsive subjects were underaroused. Sensation seeking, however, was not related to tonic arousal. Instead, sensation seeking—particularly that of the disinhibitory type—has been consistently related to a particular measure of cortical arousability called augmenting-reducing (A-R, Buchsbaum, 1971).

Figure 4.3 Mean visual evoked potential amplitudes (P1-N1) at five levels of light intensity for low and high scorers on the disinhibition subscale of the Sensation Seeking Scale.

A-R assesses the relationship of cortical reactivity, measured as a function of the relationship between the cortical EP and stimulus intensity for any given individual. A strong positive relationship between the amplitude of the EP and the intensity of stimuli is called augmenting, and a negative or zero relationship is called reducing. A-R differences are most often observed at the highest intensities of stimulation, where the reducers show a marked EP reduction and the augmenters continue to show increased EP amplitude. There is an obvious relevance of this measure to Pavlov’s (1927/1960) construct of “strength of the nervous system,” based on the nervous system’s capacity to respond to high intensities of stimulation without showing transmarginal inhibition.

Figure 4.3 shows the first study of the relationship between the Disinhibition (Dis) subscale of the SSS and amplitude of the visual EP. Those scoring high on Dis displayed an augmenting pattern, and those scoring low on this scale showed a strong reducing pattern, particularly at the highest intensity of stimulation. This study was followed by many others, some using visual and others using auditory stimuli. Replications were frequent, particularly for the auditory EP (Zuckerman, 1990, 1991). Replications continue to appear (Brocke, Beauducel, John, Debener, & Heilemann, 2000; Stenberg, Rosen, & Risberg, 1990). A-R has also been found to be related to Imp, particularly cognitive impulsiveness (Brocke, Pritchard, Faulk, & Brandt, 1987).

The A-R model has been extended to other species and used as a biological marker for behavioral traits in animals resembling those in high and low human sensation seekers and impulsive and constrained persons. Cats who showed the
augmenting pattern were active, exploratory, and approached rather than withdrew from novel stimuli. Augmenting cats adapted easily to novel situations, were responsive to a simple reward task, but were poor at learning to inhibit responses where they were only reinforced for low rates of response (Hall, Rapaport, Hopkins, Griffin, & Silverman, 1970; Lukas & Siegel, 1977; Saxton, Siegel, & Lukas, 1987).

Siegel extended this paradigm to a study of two genetically selected strains of rats, one actively avoidant or more aggressive and the other passive and frozen in reaction to shock (Siegel, Sisson, & Driscoll, 1993). The first strain consistently showed the augmenting EP pattern, and the second showed the reducing. Other behavioral characteristics of these strains were consistent with the human model of impulsive sensation seeking: The augmenting strain was aggressive, more willing to ingest alcohol, had high tolerance for barbiturates, and self-administered higher intensities of electrical stimulation in reward areas of the limbic brain than the reducing strain.

Biochemical reactions suggested the basis for behavioral differences in characteristics of stress-reactive neurotransmitter and hormonal responses. Under stress, the augmenting strain showed more dopaminergic activity in the prefrontal cortex of brain, whereas the reducers had a stronger reaction in the hypothalamic-pituitary-adrenal cortex (HYPAC) stress pathway including increased serotonergic activity and corticotropin releasing factor in the hypothalamus and adrenocorticotropic hormone in the pituitary gland. Dopamine is a neurotransmitter implicated in action tendencies and theorized to be the basis of novelty and sensation seeking. Dopamine release would explain the active avoidance patterns that were the basis for selecting the two strains. Conversely, serotonin activity is associated with behavioral inhibition.

**Monoamines**

The animal model described earlier suggests that sensation seeking and related traits in humans may be associated positively with dopaminergic and negatively with serotonergic reactivity. Indirect evidence of this association comes from patients with Parkinson’s disease (PD), in which dopamine is depleted 75% in ventral tegmental neurons. A study of personality of PD patients showed that the PD patients were significantly lower on novelty seeking than controls but did not differ from them on harm avoidance or reward dependence (Menza, Golbe, Cody, & Forman, 1993). The PD patients were more depressed than controls, but depression did not correlate with novelty seeking scores.

Simple correlations between sensation seeking and dopamine and serotonin metabolites (HVA and 5-HIAA) assayed from CSF reveal no correlations between these metabolites and sensation seeking or the P scale or impulsivity scales (Ballenger et al., 1983; Limson et al., 1991). However, the correlational study by Ballenger et al. found a significant negative correlation between norepinephrine in the CSF and sensation seeking. A significant correlation was found between P and dopamine D2 binding in left and right basal ganglia in a PET study of a small group of normal subjects (Gray, Pickering, & Gray, 1994).

An experimental study by Netter, Hennig, and Roed (1996) used drugs that stimulate (agonist) or inhibit (antagonist) activity in the serotonergic and dopaminergic systems and measured their effects on hormonal, emotional-state, and behavioral reactions. Their findings suggested a low responsiveness of the serotonergic system in high sensation seekers, but no association of dopaminergic response to an agonist and sensation seeking. However, craving for nicotine was increased by a dopamine agonist in high sensation seekers, suggesting that dopamine stimulation may induce more approach behavior in high than in low sensation seekers. Experiments in which nicotine or amphetamine is given to participants high or low in sensation seeking or novelty seeking showed that the high sensation/novelty seekers had more intense “highs” or subjective effects in response to these drugs than did low sensation seekers (Hutchison, Wood, & Swift, 1999; Perkins, Wilson, Gerlach, Broge, & Grobe, 2000). The effect for nicotine was most intense for nonsmokers, and the study on amphetamine did not use persons with a drug history. These special reactions of high sensation/novelty seekers to the novel drugs suggests some sensitivity to these dopamine agonists, perhaps in the receptors.

Another study by the German group found that the disinhibition type of sensation seeking and impulsivity, as well as aggression, were correlated with a response to a serotonin antagonist indicating low serotonergic responsivity in impulsive sensation seekers (Hennig et al., 1998).

**Monoamine Oxidase**

Fairly consistent negative relationships have been found between sensation seeking and MAO. A survey of results in 1994 showed low but significant negative correlations between platelet MAO and sensation seeking trait in 9 of 13 groups, and in 11 of 13 groups the correlations were negative in sign. The gender and age differences in sensation seeking are consistent with the gender and age differences in MAO described previously. Low MAO levels are characteristic of disorders characterized by impulsive, antisocial behavior including antisocial and borderline personality disorders, alcoholism and heavy drug abuse, pathological gambling disorder, bipolar
disorder, and attention deficit and hyperactivity disorder in children. MAO is low even in children of alcoholics and bipolar disorders who have not yet manifested the disorders, suggesting that it is a genetic risk marker for these disorders.

In a general normal population, low MAO was associated with use of tobacco, alcohol, and illegal drugs, convictions for crimes other than traffic offenses, and sociability in terms of hours spent with friends (Coursey et al., 1979). A study of low-MAO monkeys living in a natural environment showed that they were more aggressive, dominant, sexually active, and sociable than were high-MAO monkeys (Redmond, Murphy, & Baulu, 1979). Monkeys with high MAO levels were social isolates and passive. This study of another species suggests the evolutionary advantage of sensation seeking as mediated by MAO and possibly dopaminergic systems in the brain. Low MAO, however, is also associated with impulsivity in laboratory tests (Klinteberg et al., 1991), as is sensation seeking (Breen & Zuckerman, 1999; Thornquist & Zuckerman, 1995), and impulsivity in risky situations could be a disadvantaged trait that may lead to premature death. However, the advantage in securing and dominating mates by intimidation of rivals may have outweighed the evolutionary disadvantages of reckless behavior.

In the public mind testosterone is identified with sexual drive and aggressiveness. However research shows that testosterone (T) is associated with a broader range of traits than these two. Androgens and T assayed from blood are correlated with sensation seeking (Daitzman & Zuckerman, 1980; Daitzman, Zuckerman, Sammelwitz, & Ganjam, 1978). Dabbs (2000) and Bogaert and Fisher (1995), using T from saliva, found only nonsignificant tendencies in that direction. A comparison of hypogonadal (low-T) and normal-T men, all referred for complaints of erectile dysfunction, showed that the low T-men were lower on sensation seeking than were the normal-T men (O’Carroll, 1984).

Hormones

Testosterone and sensation seeking in young males are both correlated with their sexual experience, in terms of the number of sex partners they have had (Bogaert & Fisher, 1995; Dabbs, 2000; Daitzman & Zuckerman 1980). Testosterone levels affect sexuality in women as well as men. Androgen levels of married women were related to sexual responsivity, frequency of intercourse, and sexual gratification (Persky et al., 1982). As with MAO, we can see the evolutionary advantage of the behavioral trait based on its biochemical substrate. However, other correlates of T include sociability, dominance, and activity, as well as inverse relationships to socialization and self-control.

The high-T male tends to be assertive, impulsive, and low in self-control, as well as high in sensation seeking. There is much less work on T in women, but what data there are suggest the same kind of personality correlates as in men. Apart from aggression, high-T men were more likely than others to misbehave in school as children, get into legal difficulties as adults, use drugs and alcohol, and go AWOL (absent without leave) while in the army (Dabbs, 2000). Fraternities with high average T levels were generally disorderly and chaotic, and their members were described by an observer as “crude and rude.” The high-T fraternities had more parties, worse grades, and fewer community service activities. Dabbs (2000) suggested that the total effect among members is an outcome of an interaction between T levels of its members and reinforcement of each other for antisocial behavior. In this case, high T is clearly a predisposing factor for low socialization, which these authors describe as “rambunctiousness.”

Testosterone levels reflect both trait and state moods. Although reliability can be found in T levels taken at the same time of day in the same setting, T levels can also be affected by experiences in competition (Dabbs, 2000). Competitors show increases in T when victorious and decreases when defeated. Even sports spectators show increases in T when their team wins and decrease when their team loses.

High levels of cortisol are associated with prolonged stress and depression. Ballenger et al. (1983) found that low levels of CSF cortisol were associated with a P dimension factor that included the P scale, the disinhibition subscale of the Sensation Seeking Scale (SSS), the MMPI hypomania scale, and lifetime number of sexual partners. Low levels of cortisol have been found in prisoners who have a history of psychopathic and violent behavior (Virkkunen, 1985). Low cortisol was also associated with novelty seeking in veterans with posttraumatic stress disorder (Wang, Mason, Charney, & Yehuda, 1997). Low cortisol may indicate a low reactivity to stress, which can be an advantage in some situations but carries the dangers inherent in lack of control and impulsivity. Traits that may have been adaptive in the warrior societies of the past may now confer a disadvantage in more socialized civilizations.

Genetics

Twin studies have found relatively high heritabilities (58%) for sensation seeking whether based on twins raised together (Fulker, Eysenck, & Zuckerman, 1980) or on twins separated shortly after birth and raised in different families (Hur & Bouchard, 1997). Heritability for Cloninger’s NS scale is somewhat lower (40%) but typical of that found for other
personality traits (Heath, Cloninger, & Martin, 1994), but that for impulsivity is lower (15–40%) albeit significant (Eysenck, 1983).

Ebstein et al. (1996) were the first to report an association between the trait of novelty seeking and the gene for the D4 dopamine receptor (D4DR). The longer, usually the 7 repeat form of the 48 base pair sequence, was associated with high scores on Cloninger’s NS scale in an Israeli population. An immediate replication was reported by Benjamin et al. (1996) in an American population using scales from the NEO that approximate the NS factor such as Excitement Seeking and Deliberation (vs. Impulsiveness). Within a year Ebstein and Belmaker (1997) summarized the rapidly growing literature reporting two more replications and three failures to replicate. Since then two more failures to replicate have been reported, one in a Swedish population (Jönsson et al., 1998) and the other in a New Zealand sample (Sullivan et al., 1998). One partial replication was reported in Finland (Ekelund, Lichtermann, Jaervelin, & Peltonen, 1999). The variations in populations among the studies may have something to do with the inconsistent results. The distribution of alleles differs among populations. For instance, in a Japanese population the 7 repeat allele was not found but a comparison of the longer (5 and 6 repeats) with the shorter (2 to 4 repeats) still showed the former to be more characteristic of high novelty seekers (Ono et al., 1997).

As with MAO, the association between sensation or novelty seeking and this genetic marker is given some credence by its association with behavioral traits or disorders characterized by impulsivity and sensation seeking. The longer form of the D4DR has been found in high proportions of opiate abusers (Kotler et al., 1997), persons with pathological gambling disorder (Castro, Ibanez, Torres, Sáiz-Ruiz, & Fernández-Piqueras, 1997), those with attention-deficit-hyperactivity disorder (Swanson et al., 1998), and infants showing less distress in reaction to novel stimuli (Auerbach et al., 1998).

A comparative study was done on the effects of knocking out the D4R gene in mice on tests of approach-avoidance in reaction to novel objects or situations (Dulawa, Grandy, Low, Paulus, & Geyer, 1999). The D4 knockout mice showed reductions in behavioral response to novelty or a decrease in novelty related exploration in comparison to D4 intact mice.

Despite some failures of direct replication the association between novelty seeking and the D4DR receptor gene is given credence by these extensions to psychopathology and behavior in humans and mice. The D4DR association accounts for only about 10% of the genetic variation in the human trait, so other genes are certainly involved. The search is on for such genes. A crucial question is the functional significance of the difference between the alleles associated with high or low sensation seeking. The D4DR gene is expressed mainly in the limbic brain regions associated with emotional and motivational characteristics of sensation seeking. Dopaminergic activity is certainly involved, as has been postulated. But the significance of the D4DR gene in this activity is far from certain. An interesting finding is that the density of D4 receptors is elevated in brains of schizophrenics and that this receptor is the primary target for the antipsychotic drug clozapine (Seeman, 1995).

Summary

The underarousal hypothesis related to E has been more successfully applied to this third dimension of personality. Both EEG and brain imaging studies have found some preliminary evidence of cortical underarousal related to the P dimension and impulsivity. Sensation seeking and impulsivity have been related to the characteristic cortical response to a range of intensities of stimulation. Disinhibited and impulsive persons show an augmentation of cortical response at high intensities of stimulation relative to low intensities, whereas inhibited and constrained individuals show a reducing pattern, particularly at high intensities. This augmenting-reducing paradigm of cortical reactivity has been extended to cats and rats, where it is associated with similar kinds of behavioral reactions and with other kinds of biological reactivity postulated to be the basis of the behavioral traits in humans.

The clinical model for this dimension of personality lies in the psychopathic or antisocial personality disorder. One of the characteristics of this disorder is a lack of emotional reactivity to stimuli associated with punishment and therefore a deficit in learning to avoid reacting to such stimuli. This leads to seeking of high-intensity rewarding stimuli regardless of the risk involved. It is not surprising that psychopaths are all high impulsive sensation seekers and share some of the same biological traits with nonpsychopathic sensation seekers such as low levels of the monoamine oxidase enzyme and high levels of testosterone.

One psychopharmacological theory of the P dimension is that it is based on high dopaminergic reactivity and low serotonergic and noradrenergic reactivity to highly stimulating situations. The low serotonergic reactivity is particularly related to the lack of restraint or behavioral inhibition and the low noradrenergic reactivity to the lack of arousal characteristic of high P, impulsive, and sensation seeking individuals. There is some evidence from studies of humans of a weaker response to serotonin stimulants in high sensation seekers than in low sensation seekers. There is no demonstrated...
A new form of the scale has reduced the number of subscales to four, using factor analyses: physical aggression, verbal aggression, anger, and hostility (Buss & Perry, 1992). Although the subscales are moderately intercorrelated, quite different results have been found for the different subscales of the test in biological studies. Another important distinction in the literature is whether aggression is impulsive. The impulsive type of aggression seems more biologically rooted than instrumental types of aggression, but this confounds two different dimensions of personality.

Although aggression and hostility are correlated in tests and life, they are separated in two of the major trait classification systems. Eysenck’s system includes negative feelings like anger (moodiness) in neuroticism, but aggression and hostility are at the core of the psychoticism dimension. Costa and McCrae (1992a) have angry-hostility as a facet of neuroticism but regard aggression as the obverse of agreeableness. My colleagues and I found that hostility and anger load more highly on N and aggression on P in a three-factor model, but all three correlate with a common factor in a five-factor analysis (Zuckerman et al., 1991).

Aggression has been defined by several methods, including self-report ratings or questionnaires, observer or ratings by others, and life-history variables like membership in groups characterized by violent acts or crimes. Aggression is not a socially desirable trait and this may limit the usefulness of self-report methods in some settings. Laboratory observations may be too specific to the experimental conditions. Persons who committed a violent crime, like murder, may differ depending on how characteristic their violent behavior was before they committed the crime. All methods have methodological problems, but in spite of this there are certain consistencies in results across methods in the literature.

Cortical Arousal and Arousability

Early studies of the EEG in abnormal populations, like violent criminals, used crude qualitative judgments of the EEG records as “abnormal” or “normal” (Volavka, 1995). EEG abnormalities included diffused or focal slowing, spiking or sharp waves in certain areas, and generalized paroxysmal features resembling minor epileptic seizures. The incidence of abnormal records found in samples of prisoners convicted of homicides and habitually violent prisoners was quite high (50–65%) compared to nonviolent prisoners or normal controls (about 5–10%; Volavka, 1995). However, some other studies found no differences between violent and nonviolent offenders.

Studies using quantitative methods showed EEG slowing in offenders, including slowing of alpha activity and an
excess of slow wave (theta) activity. Volavka (1995) pointed out that these results could be due to a variety of factors including developmental retardation, brain injuries, decreased arousal level, cortical disinhibition, or genetic factors. Actually, twin research suggests that most of the activity in spectrum parameters of the EEG is genetically determined (Lykken, 1982).

One limitation of most of these earlier studies was that only prisoners referred for neuropsychiatric evaluation were used. A study by Wong, Lumsden, Fenton, and Fenwick (1994) selected subjects from a population of prisoners who had all been rated for violent behavior, and 70% had received EEG assessment. The prisoners were divided into three groups based on their history of violence. Going from the lowest to the highest violent groups, the percentages of abnormal EEG’s were, respectively, 26%, 24%, and 43%. The most frequent EEG characteristics differentiating the most violent from the least violent groups was focalized EEG abnormalities, particularly in the temporal lobes. Twenty percent of the most violent patients showed abnormal temporal lobe readings compared to 2% to 3% in the other two groups. Computerized tomography (CT) scans confirmed the high incidence of temporal lobe abnormalities in the most violent group.

The cortical EP has also been used to study cortical arousability. A study comparing detoxified alcoholic patients with and without histories of aggression found lower amplitudes of the P300 in the aggressive group (Branchey, Buydens-Branchey, & Lieber, 1988). Aggressive alcoholics often have other characteristics, such as impulsivity and alcoholism, which might have produced the weaker P300 signal. Another study found that impulsive aggressive subjects screened from a college student population also showed lower P300 amplitudes at frontal electrode sites (Gerstle, Mathias, & Stanford, 1998). Still another study showed that a drug that reduced frequency of aggressive acts among prisoners with a history of impulsive aggression also increased the amplitude of the P300 in this group (Barratt, Stanford, Felthous, & Kent, 1997). This effect of the drug was not found in a group of prisoners who had committed premeditated murders. A reduced P300, particularly in the frontal lobes, may be symptomatic of a weakened inhibition from the frontal lobes and may account for the impulsive aspect of the aggression.

Visual imaging methods have been used in the study of violent behavior. Two structural methods are computed tomography (CT) and magnetic resonance imaging (MRI). MRI yields better images for precise assessment of brain structure. PET is used to assess brain activity in specific areas of brain including regions not accessible by ordinary EEG methods. Mills and Raine (1994) reviewed 15 studies of structural brain imaging (MRI, CT) and 5 studies using PET and regional CBF. Subject groups were violent prisoners, convicted murderers, pedophiles, incest offenders, property offenders, and, in some studies, normal controls. Property offenders were regarded as controls for violent offenders. Sexual offenders were not necessarily violent. Nine of the 15 studies using CT or MRI showed some type of structural abnormality, about evenly divided between frontal and temporal or frontotemporal deficits. Frontal abnormalities characterized the violent offenders and frontotemporal the sexual offenders, according to the authors of the review. However, most studies used small samples. The two studies of violent offenders using large samples (Ns of 128 and 148) found no particular localization of abnormalities (Elliot, 1982; Merikangus, 1981). The only study using MRI with any kind of N (another had only 2 cases) found evidence of temporal lobe lesions in 5 of 14 violent patients (Tonkonogy, 1990). The large study by Wong et al. (1994), not included in the review, found that 55% of the most violent group had abnormal CT findings, and 75% of these were temporal lobe findings. Contrary to the hypothesis of Mills and Raine, temporal lobe lesions alone seem to be characteristic of violent patients. More MRI studies are needed to clarify the issue of localization.

The temporal lobe overlays the amygdala and has connections with it. Animal lesion and stimulation studies have found sites in the amygdala that inhibit and others that trigger aggression. Total amygdalectomies in monkeys produce a drop in the dominance hierarchy and an inability to defend against aggression from other monkeys. The comparative data suggest loci for aggression in the amygdala.

Mills and Raine reviewed five PET studies, but of these only one had a near-adequate number of subjects (3 had less than 10) and another compared child molesters with controls. The one study remaining compared 22 murderers with 22 normals and found selective prefrontal dysfunction in the group of murderers (Raine et al., 1994). Temporal lobe damage and functional hypofrontality are not unique to violent offenders but are also found in patients with schizophrenia.

**Cardiovascular Arousal and Arousability**

Numerous studies show that persons who score high on hostility scales show greater anger and cardiovascular arousal, especially blood pressure, in response to stress or perceived attack than do low hostile persons. As an example, a recent study found that among participants who were deliberately harassed in an experiment, the high hostile group who was harassed showed enhanced and prolonged blood pressure, heart rate, forearm blood flow and vascular resistance, and increased norepinephrine, testosterone, and cortisol.
responses than did low hostile subjects who were harassed (Suarez, Kuhn, Schanberg, Williams, & Zimmermann, 1998). This kind of cardiovascular reactivity may occur in frequent situations like stressful marital interactions (Smith & Gallo, 1999), and general day-to-day living (Rääkkönen, Matthews, Flory, & Owens, 1999), and thus put a strain on the cardiovascular system that can result in cardiovascular disease, including hypertension (Lawler et al., 1998; Miller, Dolgoy, Friese, & Sita, 1998) and ischemic heart disease (IHD; Gallagher, Yarnell, Sweetnam, Elwood, & Stansfeld, 1999). Persons with a family history of hypertension exhibit the same pattern of hostility and anger arousal with elevated blood pressure as do those who have developed the disorder suggesting that there may be a genetically influenced source to the cardiovascular overreactivity associated with anger arousability. However, how the anger is dealt with is a factor in vulnerability to heart disease. In a prospective study of nearly 3,000 men in their 50s and 60s Gallager et al. (1999) found that suppressed anger was most predictive of the incidence of IHD even when other risk factors were statistically controlled.

**Monoamines**

Åsberg’s (1994) review of the role of the monoamine neurotransmitters in human aggressiveness and violence attributes a primary importance to the role of serotonin. In animals serotonin is associated with inhibition of aggressive behavior and lowered serotonin with disinhibition of such behavior (Soubrié, 1986). In humans low levels of the serotonin metabolite, 5-HIAA, have been consistently found in those who attempt or complete suicide using violent means and in violent criminal offenders, particularly those characterized by impulsive violence or murder (Åsberg, 1994). Personality disorders like antisocial and borderline disorder have a high incidence of aggressive behaviors and suicide attempts. Homicide and suicide are not antithetical; homicide offenders have increased suicide rates. Within a group with personality disorders a negative correlation was found between CSF 5-HIAA and lifetime aggressive behavior (Brown et al., 1982).

Hormonal responses to serotonergic agonists and antagonists have also been used to assess the reactivity of the system in relation to aggression. They generally support the hypothesis of an inverse relationship between serotonin function and aggression/hostility (Cleare & Bond, 1997; Coccaro, Kavoussi, Sheline, Berman, & Csernansky, 1997; Moss, Yao, & Panzak, 1990).

Tryptophan is a precursor of serotonin in the brain (see Figure 4.2). Tryptophan depletion provides an experimental approach to the serotonin-aggression hypothesis, and unlike correlational studies it can provide evidence of a causal link with aggression. Studies have found that tryptophan depletion increases aggressive responses in laboratory behavioral tests (Cleare & Bond, 1995; Dougherty, Bjork, Marsh, & Moeller, 1999) as well as subjective feelings of anger, aggression, and hostile mood (Cleare & Bond, 1995; Finn, Young, Pihl, & Ervin, 1998), but the effect is limited to persons who are high in trait measures of hostility. The inference is that hostile persons, who are already low in serotonergic activity, tend to react aggressively with even more lowering of serotonin stores. There is the further suggestion that serotonin agonists or selective serotonin reuptake inhibitors (SSRIs) may reduce aggression in aggression-prone persons. A study by Knutson et al. (1998) showed that an SSRI reduced focal indices of hostility through a general decrease in negative affect without altering positive affect. In addition, the SSRI increased agreeableness on a behavioral index and cooperativeness in a puzzle-solving task.

SSRIs are used to treat depression, but can they change other emotions like anger-hostility? A study of SSRI therapy for depressed outpatients showed a significant decrease in anger-hostility as well as neuroticism (Bagby et al., 1999). The decrease in neuroticism, however, was correlated with the decrease in clinical depression severity, whereas the decrease in anger-hostility was independent of the reduction of depression.

NE mediates a primary arousal system in the brain beginning in the locus-coeruleus and extending through limbic structures to innervate all areas of cerebral cortex. As such it has been implicated in the arousal of anxiety as previously discussed. But anger is also associated with an arousal effect as shown by the cardiovascular reactivity in highly hostile and angry persons as previously discussed. A study of aggression in free-ranging monkeys found a negative correlation between aggressivity and CSF 5-HIAA, congruent with the serotonin-aggression hypothesis, but it also found an equally strong positive correlation between aggressivity and CSF MHPG, the NE metabolite. The ballenger et al. (1983) study of humans (normals) found a very high positive correlation (.64) between plasma MHPG and the Assault scale from the BDHS. Use of a noradrenergic challenge revealed a correlation of noradrenergic reactivity and irritability and assault scales (Coccaro et al., 1991).

On the other hand, low levels of the catecholamines epinephrine and NE, obtained from urine, are inversely related to aggressiveness (Magnusson, 1987). Psychopathic youths have low reactivity in these peripheral catecholamine systems in stressful situations (Lidberg, Levander, Schalling, & Lidberg, 1978). The difference may lie between the central
noradrenergic and the peripheral autonomic stress system. Another possibility is that there is a difference between the psychopathic type of aggression, which is often not accompanied by high arousal, and the impulsive-angry type of aggression in which emotional disinhibition is typical. Nettter, Hennig, and Rohrmann (1999) believed that they can distinguish the two types of aggressiveness on the basis of selective types of challenges to the monoamine systems. The serotonergic challenge was primarily correlated with Eysenck’s P scale, assessing the psychopathic type of aggression, whereas another type of challenge more closely related to dopamine reactivity was related to the nonpsychopathic type of aggression.

Increasing levels of brain dopamine in rats increases impulsive aggressive responding, but it takes a great deal of dopamine depletion to reduce aggressive behavior (Volavka, 1995). Little research has been done on dopamine specifically although the aggression producing effects of amphetamine may be a function of stimulation of dopaminergic as well as noradrenergic systems. A study of the neuroendocrine responses to glucose challenge in a group of substance abusers showed that those participants characterized by antisocial hostility had responses suggestive of increased dopaminergic activity (Fishbein, Dax, Lozovsky, & Jaffe, 1992).

**Hormones**

The hypothesis of an influence of T on aggressive behavior has a prescientific origin in that the pacifying effects of castration in animals were known and used for that purpose. Sexual competition among males is one form of aggression influenced by T, but other forms are also affected. Castration reduces aggression in males in most species, and T replacement reverses this effect.

Studies of the relationships between T and hostility or aggression in humans have produced mixed results, but a meta-analysis of such studies found a moderate effect size of .40 over all studies (Archer, Birring, & Wu, 1988). An earlier review by Archer (1991) suggested that results were more positive in studies where behavioral assessments (usually in prisoners) were used and less powerful in studies of trait (self-report) hostility or aggression (usually in college student samples). The newer meta-analysis failed to support this hypothesis. Better results were obtained in studies using salivary T as opposed to T derived from blood. A study using salivary T in 306 students found T positively correlated with aggression and negatively with prosocial scales in both men and women (Harris, Rushton, Hampson, & Jackson, 1996), but in other studies using either blood (Archer et al., 1998) or salivary T (Campbell, Muncer, & Odber, 1997) in large samples of male students no relationship was found. In still another study of blood T in students, both T and estradiol were positively correlated with self-reported aggression in men, but the correlations were negative in women (Gladue, 1991).

More consistent results have been obtained with behavioral (non-self-report) assessments. A study of nearly 700 male prison inmates found that salivary T was related to a history of violent crimes, particularly rape, homicide, and child molestation, as well as violations of prison rules, particularly those involving assault (Dabbs, Carr, Frady, & Riad, 1995). A study of female inmates showed a relationship of T with aggressive dominance in prison but not with the history of criminal violence. A group of alcoholics with a history of violence had elevated levels of serum T relative to other alcoholics (Bergman & Brismar, 1994).

Even among nonclinical samples there is correlational evidence of a relationship between T and aggression. T correlated with more aggressive fighting in men during judo contests (Salvador, Suay, Martinez, Simon, & Brain, 1999) and in amount of shock given to an opponent in a contrived laboratory situation (Berman, Gladue, & Taylor, 1993).

Whether self-report or behavioral, correlational studies cannot establish cause and effect. There is ample evidence in both animal and human studies that aggressive experience in competition may raise T levels in victors or lower them in those who are defeated. Experimental studies in which the effects of raised T levels on aggression are observed might be helpful. Clinical studies of steroid users have shown increased aggressiveness in some of them (Pope & Katz, 1994). Archer (1991) reviewed studies in which T or T-stimulating hormones were given and effects on aggression assessed by self-report. Although there is some evidence that T can affect hostility, there are also some negative findings from other studies. In all likelihood there is a continuous interaction between endogenous levels of T and life experiences (affecting current levels) during life. T makes one more likely to aggress, and aggression or its anticipation raises T levels.

Longitudinal studies may also elucidate the complex causal pattern. In one study a group of boys was followed from 6 to 13 years of age (Tremblay et al., 1998). T at age 12 and body mass predicted social dominance in adolescence but only body mass independently predicted physical aggression. The authors suggest that the relation between aggression and T in adolescents may be mediated by the effect of T in the change in physique in the context of dominance. A similar study followed males from pre- or early adolescence (12–13 years) and found little relationship between early or concurrent measure of T and aggression; the few that were found did not persist over time (Halpern, Udry, Campbell, & Suchindran, 1993).
Short time periods of prediction may confound environmental-developmental interactions that could mask the influence of endogenous levels of T. Windle and Windle (1995), in a retrospective longitudinal study, examined the adult levels of plasma T in four groups: (a) those who were aggressive only in childhood; (b) those who became aggressive as adults; (c) those who were aggressive in both childhood and adulthood (continuity); and (d) those who were low in aggression in childhood and adulthood. Adult onset and continuity (in aggressiveness) groups had higher T levels as adults than the other two groups. Other than aggressiveness, the high-T adult groups had higher rates of antisocial personality and a history of various signs of antisocial behavior. Was the high level of T in these groups a product of their history or a sign of an earlier level of T that affected the development of these behaviors? The authors admit that it is impossible to answer this question.

High levels of cortisol are associated with stress and inhibition and low levels with impulsivity and sensation seeking, as noted previously. In baboons dominant and aggressive males usually have low levels of cortisol and subordinate and nonaggressive primates have higher levels of antisocial behavior. With testosterone, cortisol varies considerably with recent and long-term patterns of experience such as winning or losing in fights. Low levels of cortisol have been found in psychopathic, violent offenders (Virkkunen, 1985), but high levels of cortisol are positively associated with hostility as measured by hostility questionnaires (Keltikangas, Räikkönen, & Adlercreutz, 1997; Pope & Smith, 1991). Chronic feelings of hostility are often associated with anxiety and depression, but the type of impulsive aggression seen in antisocial personality represents a brief state of anger in a generally unemotional personality.

Genetics

Behavior genetic studies of general hostility scales or aggression in children have shown significant heritabilities. However, it is possible that some aspects of hostility or aggression may be more heritable than others. A twin study of adult males using the BDHS revealed heritabilities ranging from 28% for verbal hostility to 47% for assault (Coccaro, Bergeman, Kavoussi, & Seroczynski, 1997). Verbal hostility is the most common form and yet it had the least heritability and the strongest environmental influence. An analysis of the genetic influence on the correlations among the scales that the assault scale had different underlying influences than the other scales which shared a common genetic influence. With the exception of the assault scale the genetic influence underlying the scales is of a nonadditive type suggesting Mendelian dominant or recessive or epistatic mechanisms. If it is the former, there is the likelihood of finding a gene of major effect in the general trait of aggression, apart from physical assault type.

The MAO type-A gene has become a likely candidate for this trait. Aggression in male mice is heightened by deletion of the MAO-A gene (Cases et al., 1995), and a mutation in the gene in a large Dutch family has been linked to mild retardation and impulsive aggressive behavior (Bruner, Nelsen, Breakfield, Ropers, & van Oost, 1993). The mutation is rare, but the gene has a wide range of alleles varying in repeat length. Subjects with one form, in contrast to those with another form, had lower scores on an index of aggression/impulsivity and the Barratt impulsiveness scale (Manuck, Flory, Ferrell, Mann, & Muldoon, 2000). The life history of aggression only approached significance and the BDHS did not show significant differences between allele groups. Apparently, the impulsivity was more salient than the aggressiveness in the combination. Consistent with the association between low serotonin and aggression in the finding that the allele group with the higher impulsive aggression score also showed less response to a serotonergic challenge test.

Just as the findings on the MAO-A gene suggest one source of the link between serotonin and aggression, another gene has been found that suggests a genetic mechanism for the association of norepinephrine with aggression. The adrenergic-2A receptor gene (ADRA2A) plays a role in modulating norepinephrine release in the locus coeruleus. Alleles of this gene were associated with scales for hostility and impulsivity in a younger student sample and impulsivity alone in an older sample (Comings et al., 2000).

Summary

Extreme violence has been associated with EEG evidence of cortical abnormality usually in the form of an excess of slow wave activity (underarousal) or focalized EEG abnormalities in the temporal lobes. Brain scans have confirmed the temporal lobe abnormalities and also found an equal incidence of frontal lobe abnormalities. A reduced P300 cortical EP response has also been found in prisoners with a history of extremely violent behavior. The reduced activity and reactivity in the frontal lobes may reflect a deficit in inhibitory capacity, which is part of the executive function of these lobes. The abnormal activity of the temporal lobe may be symptomatic of abnormal amygdala function because this lobe is in close proximity to the underlying amygdala. An MRI study has revealed temporal lobe lesions in about one third of violent patients. Hostility or anger proneness is related to a high level of cardiovascular, noradrenergic, and testosterone and
cortisol response to stress or perceived attack. Suppressed hostility can lead to cardiovascular disease.

Among the monoamines, serotonin deficit is most highly associated with impulsive aggression. However, low serotonin is associated with depression and suicide as well as aggression and homicide, another example of the multiple trait associations of biological markers. Lack of emotional and behavioral control is the likely consequence of serotonin deficit. Depletion of tryptophan, the precursor of serotonin in the production chain, increases aggressive responses and angry and hostile feelings in laboratory experiments. Augmentation of serotonin, through reuptake inhibitors, can reduce aggression in aggression-prone persons.

Unlike depression, in which both serotonin and norepinephrine depletions are seen, brain norepinephrine (from CSF) tends to be positively correlated with aggressive tendencies in monkeys and humans. However, low levels of peripheral levels of the catecholamines norepinephrine and epinephrine are also related to aggressiveness. We need to distinguish between the type of aggression that occurs in states of high emotional arousal and the cold type of aggression more characteristic of the psychopath. The latter type may be reflected in the low levels of peripheral catecholamine reactivity.

Testosterone is associated with aggression based on behavioral records, but results using self-report measures of hostility or aggression are less conclusive. Prisoners with either histories of extremely violent crimes or characterized by aggression in prison show high levels of testosterone. Testosterone is increased by victory in competitive contests and sexual stimulation and decreased by defeat, raising the old “chicken or egg” problem of causation. The influence of testosterone during development may be mediated by its influence on physique in male adolescents where it is associated with a more muscular mesomorphic body build. Low cortisol levels are found in aggressive types and are also influenced by the outcomes of fights.

Aggression trait is moderately heritable, but its heritability depends on the form it takes. Assaultive aggression is moderately heritable but verbal aggression is only weakly heritable. The gene for MAO of the A type has been linked to aggression in a human family study. Deletion of the MAO-A gene in mice increases their aggressivity, suggesting that the gene is involved in the inhibition or regulation of aggression.

**CONCLUSIONS**

Wilson (1998) described *consilience* as a quality of science that links knowledge across disciplines to create a common background of explanation. Personality psychology, extending from social psychology at the higher level to biopsychology at the more fundamental level, provides a daunting challenge to consilience. The introduction to this chapter presented a model of levels along the biological and social pathways leading up to a merger in personality traits.

Such a levels approach suggests a goal of reductionism, a pejorative term for critics of science and many scientists as well. The artist is contemptuous of the critic’s attempts to reduce his or her art to a textual formula, and the social scientist may resent the presumptuous intrusion of the biological scientist into his or her own complex type of explanation. Wilson, however, views reductionism as a natural mode of science:

> The cutting edge of science is reductionism, the breaking apart of nature into its natural constituents. . . . It is the search strategy employed to find points of entry into otherwise impenetrably complex systems. Complexity is what interests scientists in the end, not simplicity. Reductionism is the way to understand it. The love of complexity without reductionism makes art; the love of complexity with reductionism makes science. (pp. 58–59)

Later, Wilson (1998) admits that reductionism is an oversimplification that may sometimes be impossible. At each level of organization the phenomena may require new laws and principles that cannot be predicted from those at more general levels. My view is that this is always true for levels that involve an interaction between biological traits or genes and experience in the social environment. A learned association cannot be reduced to a specific set of neural events, at least not in the complex brain of a higher organism. It is not inconceivable, however, that the difference in general neural events that make an association more likely in one individual than another is not only explicable but also essential for a complete understanding of the event. Consilience is more possible at the borders of two levels, and this is where the breakthroughs are most likely to take place. As Wilson puts it, “The challenge and the cracking of thin ice are what gives science its metaphysical excitement” (p. 60).

This chapter was organized around a top-down approach, starting with four broad classes of personality traits that are empirically identifiable across several systems of trait description: extraversion/sociability, neuroticism/anxiety, impulsiveness/conscientiousness, and aggression/agreeableness. One way to bypass the complex social determinants of these traits in human societies is to look for appropriate animal models and biological links between behavior in these species and our own. This approach has identified certain biological markers for analogous behavioral traits such as the monoamine
neurotransmitters and enzymes like MAO that regulate them; hormones like testosterone and cortisol; psychophysiological characteristics such as augmenting/reducing of the cortical evoked potential; brain structure and physiology as assessed by brain imaging methods in humans and lesion and stimulation studies in other species; and molecular genetic studies that link genes, biological mechanisms, and behavioral and personality traits.

Simple-minded reductionism would expect one personality or behavioral trait to be associated with one brain structure, one neurotransmitter, one hormone, one physiological pattern of reactivity, and one gene in both humans and other animals. The chapter is organized by personality traits, but if one reads across the traits it is clear that this neat kind of phenomenological isomorphism is not the rule. Evolution may have shaped the nervous system around behavioral mechanisms necessary for adaptation, but evolution did not select for personality traits. The tendency to explore, forage, and approach novel but nonthreatening objects or creatures is part of that adaptation and is important in survival, as is competitive and defensive aggression, cooperation, and even altruism.

If we reverse direction and work up from the biological mechanisms to the personality trait and behavioral levels the fourfold classification at the top becomes blurred. Monoamine reactivities, MAO, testosterone, cortisol, and reactivity of cortical EPs to stimulus variation are related to sociability and sensation seeking, impulsivity and aggression, asocialization, neuroticism, anxiety, and inhibition, but in no simple one-to-one manner. Low levels of serotonergic activity are related to both depression and impulsive aggression producing both violent and impulsive homicides and suicides, sometimes in the same person. Is it the impulsivity, the aggression, or the neuroticism that is related to a serotonin deficiency? High levels of testosterone are related to sociability and social dominance, disinhibitory sensation seeking, aggressivity, asocialization, and low levels to neuroticism and agreeableness. Low levels of MAO are related to sensation seeking, impulsivity, asocial tendencies, and sociability.

Personality traits may be orthogonal, but biological traits do not respect these boundaries. It is almost as if the functional biology of the organism is organized around two basic traits: approach (including sociability, impulsivity, sensation seeking, and aggression) and inhibition/avoidance (or neuroticism/anxiety at the personality trait level). The comparative psychologist Schneirla (1959) put this idea into a postulate: “For all organisms in early ontogenetic stages, low intensities of stimulation tend to evoke approach reactions, high intensities withdrawal reactions” (p. 3). In evolved or more mature organisms Schneirla used the terms “seeking” and “avoidance” in place of “approach” and “withdrawal.” The latter terms convey the idea of reflexive or tropistic mechanisms, whereas the former imply learned behavior. Approach-withdrawal describes a basic dimension of temperament and inhibition/shyness another in infant scales of temperament. These individual differences in infants may represent two biologically based dimensions found in other species, and they may develop into more differentiated characteristics in adult humans.

Genetic dissection is one method of defining the boundaries of biological influence in traits. If both biological and behavioral traits are included in biometric or molecular genetic studies, the genetic covariance between the genetic and the other two can be determined. Rarely are genetic, biological, and behavioral traits all included in one study.

A biosocial approach cannot ignore the complex interactions between biological traits and environmental experiences. In both animals and humans the levels of the hormones testosterone and cortisol influence behavioral interactions with the environment but are in turn influenced by the outcomes of these interactions. There is no reason to think that similar interactions do not occur for the monoamine neurotransmitters. All of these systems are regulated by internal mechanisms. For instance, if there is overactivity in a system, regulators like MAO may catabolize the excess neurotransmitter. There may be more trait stability in the regulator than in the transmitter itself. After repeated experiences, however, there may be changes in the activity of a biological system that are relatively enduring if not irreversible. Environment may even influence the effect of genes by affecting their release. Given the constant interaction between the biological and environmental pathways (Figure 4.1), reductionism of one to the other is impossible. It would be like describing the biological activity of the lungs in the absence of oxygen, the digestive organs in the absence of food, or, using a more relevant analogy, the brain in the absence of stimulation.

Psychology emerged from the biological sciences more than a century ago, although its origins were forgotten by those who wanted a science that would emulate physics and those who wanted to cut all connections with the biological sciences. Fifty years ago, when I entered the field, the founder of behaviorism, Watson, had declared that the outcome of personality was entirely a matter of life experience (conditioning) and had nothing to do with genetics, and Skinner had declared the irrelevancy of the brain in behavior. Despite Freud’s own view that the mysteries of the psyche would one day be understood in terms of biology, his followers advocated an environmental determinism that put the entire weight of explanation on society, the family, and early experience. These early prophets of our science are now historical footnotes, and the science is more cognitive and biosocial with new cross disciplines like cognitive neuroscience emerging. The changes are in large measure due to
Behavior genetics has challenged the radical environmentalist position by showing that nearly all personality traits and even some broad attitudinal traits have a significant degree of genetic determination. It is becoming a truism that genes interact with environment throughout life. But the precise nature of this complex interaction remains obscure. Genes do not make personality traits; they make proteins. The development of molecular behavioral genetics will help solve some of these problems. When we know some of the major genes involved with a personality trait and what these genes make and influence in the nervous system, we will be in a better position to define the biological mechanisms that lie between gene and behavior. Knowing the gene-biological trait link is not sufficient until we can understand the way the biological mechanism interacts with the environment, or more specifically the brain-behavior relationship.

Until recent decades the study of the brain was limited to peripheral measures like the EEG. The brain-imaging methods are only in their infancy but are already influencing the course of our science. The ones like PET or the more effective fMRI can tell us exactly what is happening in the brain after the presentation of a stimulus or condition, as well as where it is happening. The expensiveness of these methods has limited their use to medical settings and to clinical populations. Studies of personality in normals are rare and incidental to the objectives of clinical studies. They usually involve small numbers of subjects with a consequent unreliability of findings. Sooner or later the application of these methods to the study of personality dimensions in nonclinical populations will help to understand exactly what a personality predisposition is in the brain. Longitudinal studies starting with genetic and neurochemical markers and tracking the fate of individuals with these markers through life will enable us to predict both normal variant outcomes and psychopathology.

REFERENCES


References 111


Wong, M. T. H., Lumsden, J., Fenton, G. W., & Fenwick, P. B. C. (1994). Electroencephalography, computed tomography and


CHAPTER 5
Psychodynamic Models of Personality

ROBERT F. BORNESTEIN

THE CORE ASSUMPTIONS OF PSYCHOANALYSIS 117
Primacy of the Unconscious 118
Psychic Causality 118
Critical Importance of Early Experiences 118
THE EVOLUTION OF PSYCHOANALYSIS: GAZING ACROSS THREE CENTURIES 119
Classical Psychoanalytic Theory 119
Neo-Analytic Models 122
Object Relations Theory and Self Psychology 123
Contemporary Integrative Models 124
PSYCHOANALYTIC PERSONALITY THEORIES: BRINGING ORDER TO CHAOS 124
Personality Processes and Dynamics 124

Freud’s psychoanalysis is like Picasso’s cubism. Controversial from the outset, Picasso’s work enchanted some and alienated others, but every twentieth-century painter has responded to it in some way. So it is with Freud’s psychoanalytic theory: Some psychologists love it, others hate it, but almost every psychologist has reacted to it—deliberately or inadvertently, consciously or unconsciously—in his or her own work.

Psychoanalysis and cubism are alike in at least one other respect as well. Both paradigms changed in fundamental ways our view of the world by pointing out limitations in our habitual manner of thinking and perceiving. Cubism compelled us to view a given object or situation from multiple perspectives simultaneously—no single viewpoint can capture the complexity of the scene. Psychoanalysis taught us much the same thing, but instead of looking outward toward the external world, psychoanalysis turned our attention inward. In the process, it altered forever the way we see ourselves.

Evaluating the validity and utility of a theory of personality is never easy, but it is particularly challenging for a theory as complex and far-reaching as psychoanalysis. Psychoanalytic theory touches upon virtually every aspect of human mental life, from motivation and emotion to memory and information processing. Embedded within this larger model is a theory of personality, but it is not always obvious where the personality portion of psychoanalysis begins and other aspects of the model leave off. Because of this, one cannot assess the psychoanalytic theory of personality without examining psychoanalytic theory in toto, with all its complexity, intricacy, and controversy.

This chapter reviews psychodynamic models of personality and their place in contemporary psychology. The chapter begins with a brief discussion of the core assumptions of psychoanalytic theory, followed by an overview of the evolution of the theory from Freud’s classical model to today’s integrative psychodynamic frameworks. I then discuss the common elements in different psychodynamic models and the ways that these models have grappled with key questions regarding personality development and dynamics. Finally, I discuss the place of psychoanalysis within contemporary psychology and the relationship of psychoanalytic theory to other areas of the discipline.

THE CORE ASSUMPTIONS OF PSYCHOANALYSIS

Given the complexity of psychoanalytic theory and the myriad incarnations that the model has assumed over the years, the core assumptions of the psychodynamic framework are surprisingly simple. Moreover, the three core assumptions of
In psychoanalytic terms, the activities of the mind (or psyche) take place outside conscious awareness. Psychodynamic theorists contend that the majority of psychological processes are presumed to be largely unconscious, and unconscious processes are thought to be particularly revealing of personality dynamics (Brenner, 1973; Fancher, 1973). Although aspects of the primacy of the unconscious assumption remain controversial (see Kihlstrom, 1987; McAdams, 1997), research on implicit learning, memory, motivation, and cognition has converged to confirm this basic premise of psychoanalysis (albeit in a slightly modified form). Many mental activities are only imperfectly accessible to conscious awareness—including those associated with emotional responding, as well as more mundane, affectively neutral activities such as the processing of linguistic material (see Bornstein & Pittman, 1992; Greenwald & Banaji, 1995; Schacter, 1987; Stadler & French, 1998). Whether unconscious processes are uniquely revealing of personality dynamics is a different matter entirely, and psychologists remain divided on this issue.

It is ironic that the existence of mental processing outside awareness—so controversial for so long—has become a cornerstone of contemporary experimental psychology. In fact, in summarizing the results of cognitive and social research on automaticity, Bargh and Chartrand (1999) recently concluded that evidence for mental processing outside of awareness is so pervasive and compelling that the burden of proof has actually reversed: Rather than demonstrate unconscious influences, researchers must now go to considerable lengths to demonstrate that a given psychological process is at least in part under conscious control. This conclusion represents a rather striking (and counterintuitive) reversal of prevailing attitudes regarding the conscious-unconscious relationship throughout much of the twentieth century.

**Psychic Causality**

The second core assumption of psychodynamic theory is that nothing in mental life happens by chance—that there is no such thing as a random thought, feeling, motive, or behavior (Brenner, 1973). This has come to be known as the principle of psychic causality, and it too has become less controversial over the years. Although few psychologists accept the principle of psychic causality precisely as psychoanalysts conceive it, most theorists and researchers agree that cognitions, motives, emotional responses, and expressed behaviors do not arise randomly, but always stem from some combination of identifiable biological and/or, psychological processes (Rychlak, 1988).

Although few psychologists would argue for the existence of random psychological events, researchers do disagree regarding the underlying processes that account for such events, and it is here that the psychodynamic view diverges from those of other perspectives. Whereas psychoanalysts contend that unconscious motives and affective states are key determinants of ostensibly random psychological events, psychologists with other theoretical orientations attribute such events to latent learning, cognitive bias, motivational conflict, chemical imbalances, or variations in neural activity (e.g., see Buss, 1991; Danziger, 1997). The notion that a seemingly random event (e.g., a slip of the tongue) reveals something important about an individual’s personality is in its purest form unique to psychoanalysis.

**Critical Importance of Early Experiences**

Psychoanalytic theory is not alone in positing that early developmental experiences play a role in shaping personality, but the theory is unique in the degree to which it emphasizes childhood experiences as determinants of personality development and dynamics. In its strongest form, psychoanalytic theory hypothesizes that early experiences—even those occurring during the first weeks or months of life—set in motion personality processes that are to a great extent immutable (see Emde, 1983, 1992). In other words, the events of early childhood are thought to create a trajectory that almost invariably culminates in a predictable set of adult character traits (Eagle, 1984; Stern, 1985). This is especially of events that are outside the normal range of experience (i.e., very positive or very negative).

The psychodynamic hypothesis that the first weeks or months of life represent a critical period in personality development contrasts with those of alternative theories (e.g., cognitive), which contend that key events in personality development occur somewhat later, after the child has acquired a broad repertoire of verbal and locomotive skills. Freud’s notion of a critical early period in personality development—coupled with his corollary hypothesis that many of the most important early experiences involve sexual frustration or gratification—was (and is) highly controversial. It helped create a decades-long divergence of psychoanalysis from mainstream developmental psychology, which has only recently begun to narrow (Emde, 1992).
THE EVOLUTION OF PSYCHOANALYSIS: GAZING ACROSS THREE CENTURIES

Many psychodynamic ideas—including the core assumptions just discussed—predated Freud’s work and were anticipated by eighteenth and nineteenth century philosophers (Ellenberger, 1970; Hilgard, 1987). Nonetheless, psychoanalytic theory as an independent school of thought was conceived just over 100 years ago, with the publication of Breuer and Freud’s (1895/1955) Studies on Hysteria. Since that time, the history of psychoanalysis can be divided into four overlapping phases: classical psychoanalytic theory, neo-analytic models, object relations theory and self psychology, and contemporary integrative models. Each phase introduced a novel approach to human development and personality.

Classical Psychoanalytic Theory

Given Freud’s background in neurology, it is not surprising that the first incarnation of psychoanalytic theory was avowedly biological. In his early writings, Freud (1895/1966, 1900/1958a) set out to explain psychological phenomena in terms that could be linked to extant models of neural functioning (an ironic goal to say the least, given that psychoanalysis developed in part to explain “neurological” symptoms that had no identifiable neurological basis, such as hysterical blindness and hysterical paralysis).

Because the core principles of classical psychoanalytic theory developed over more than 40 years, there were numerous revisions along the way. Thus, it is most accurate to think of classical psychoanalytic theory as a set of interrelated models, which were often (but not always) consistent with and supportive of each other: the drive model, the topographic model, the psychosexual stage model, and the structural model.

The Drive Model

One consequence of Freud’s determination to frame his theory in quasi-biological terms is that the earliest version of psychoanalytic drive theory was for all intents and purposes a theory of energy transformation and tension reduction (Breuer & Freud, 1895; Freud, 1896/1955c). Inborn (presumably inherited) instincts were central to the drive model, and most prominent among these was the sex drive, or libido. Freud’s interest in (some might say obsession with) sexual impulses as key determinants of personality development and dynamics was controversial during his lifetime, and remains so today (e.g., see Torrey, 1992). At any rate, during the earliest phase of psychoanalytic theory, personality was seen as a by-product of the particular way in which sexual impulses were expressed in an individual.

Freud never fully renounced the drive concept, even after he shifted the emphasis of psychoanalytic theory from inborn instincts to dynamic mental structures with no obvious biological basis (Greenberg & Mitchell, 1983). The concept of cathexis—investment of libidinal (or psychic) energy in an object or act—remained central to psychoanalytic theory even as the drive model waned in influence. As his career drew to a close during the 1930s, Freud (1933/1964a, 1940/1964b) continued to use the concept of cathexis to account for a wide range of psychological processes, from infant-caregiver bonding and infantile sexuality to group behavior and parapraxes (i.e., “Freudian slips”).

As the concept of cathexis became reified in classical psychoanalytic theory, so did the companion concepts of fixation (i.e., lingering investment of psychic energy in objects and activities from an earlier developmental period), and regression (i.e., reinvestment of psychic energy in an earlier stage of development, usually under stress). As should become apparent, the concept of cathexis gradually faded from view, but the concepts of fixation and regression continue to be widely discussed and used to explain a wide range of issues related to personality development and dynamics.

The Topographic Model

At the same time as Freud was refining the drive theory, he was elaborating his now-famous topographic model of the mind, which contended that the mind could usefully be divided into three regions: the conscious, preconscious, and unconscious (Freud, 1900/1958a, 1911/1958b). Whereas the conscious part of the mind was thought to hold only information that demanded attention and action at the moment, the preconscious contained material that was capable of becoming conscious but was not because attention (in the form of psychic energy) was not invested in it at that time. The unconscious contained anxiety-producing material (e.g., sexual impulses, aggressive wishes) that were deliberately repressed (i.e., held outside of awareness as a form of self-protection). Because of the affect-laden nature of unconscious material, the unconscious was (and is) thought to play a more central role in personality than are the other two elements of Freud’s topographic model. In fact, numerous theories of personality ascribe to the notion that emotion-laden material outside of awareness plays a role in determining an individual’s personality traits and coping style (see Hogan, Johnson, & Briggs, 1997; Loevinger, 1987).

The terms conscious, preconscious, and unconscious continue to be used today in mainstream psychology, and research
has provided a surprising degree of support for this tripartite approach in the areas of memory and information processing (Bucci, 1997; Stein, 1997; Westen, 1998). Consciousness is indeed linked with attentional capacity, and studies show that a great deal of mental processing (including perceptual processing) occurs preconsciously (Bornstein, 1999b; Erdelyi, 1985). As noted earlier, the existence of a dynamic unconscious remains controversial, with some researchers arguing that evidence favoring this construct is compelling (Westen, 1998), and others contending that “unconscious” processing can be accounted for without positing the existence of a Freudian repository of repressed wishes and troubling urges and impulses (Kihlstrom, 1987, 1999).

Perhaps the most troubling aspect of the topographic model—for Freud and for contemporary experimentalists as well—concerns the dynamics of information flow (i.e., the mechanisms through which information passes among different parts of the mind). Freud (1900/1958a, 1915/1957, 1933/1964a) used a variety of analogies to describe information movement among the conscious, preconscious, and unconscious, the most well-known of these being his gatekeeper (who helped prevent unconscious information from reaching conscious awareness), and anteroom (where preconscious information was held temporarily before being stored in the unconscious). Contemporary researchers (e.g., Baddeley, 1990) have coined terms more scientific than those Freud used (e.g., central executive, visuospatial scratch pad), but in fact they have not been much more successful than Freud was at specifying the psychological and neurological mechanisms that mediate intrapsychic information flow.

### The Psychosexual Stage Model

Freud clung to the drive model (and its associated topographic framework) for several decades, in part because of his neurological background, but also because the drive model helped him bridge the gap between biological instincts and his hypothesized stages of development. By 1905, Freud had outlined the key elements of his psychosexual stage model, which argued that early in life humans progress through an invariant sequence of developmental stages, each with its own unique challenge and its own mode of drive (i.e., sexual) gratification (Freud, 1905/1953, 1918/1955a). Freud’s psychosexual stages—oral, anal, Oedipal, latency, and genital—are well known even to nonpsychoanalytic psychologists. So are the oral, anal, and Oedipal (or phallic) character types associated with fixation at these stages (Fisher & Greenberg, 1996). From a personality perspective, the psychosexual stage model marks a turning point in the history of psychoanalysis because it was only with the articulation of this model that personality moved from the periphery to the center of psychoanalytic theory.

Table 5.1 illustrates the basic organization of Freud’s (1905/1953) psychosexual stage model. Frustration or overgratification during the infantile, oral stage was hypothesized to result in oral fixation, and an inability to resolve the developmental issues that characterize this period (e.g., conflicts regarding dependency and autonomy). The psychosexual stage model further postulated that the orally fixated (or oral dependent) person would (a) remain dependent on others for nurturance, protection, and support; and (b) continue to exhibit behaviors in adulthood that reflect the oral stage (i.e., preoccupation with activities of the mouth, reliance on food and eating as a means of coping with anxiety). Research supports the former hypothesis, but has generally failed to confirm the latter (Bornstein, 1996).

A parallel set of dynamics (i.e., frustration or overgratification during toilet training) were assumed to produce anal fixation and the development of an anal character type. Because toilet training was viewed by Freud as a struggle for control over one’s body and impulses, the analy fixed individual was thought to be preoccupied with issues of control, and his or her behavior would thus be characterized by a constellation of three traits, sometimes termed the anal triad: obstinacy, orderliness, and parsimony (Masling & Schwartz, 1979). Fixation during the Oedipal stage was presumed to result in a personality style marked by aggressiveness,
The Evolution of Psychoanalysis: Gazing Across Three Centuries 121

**Id:** Present at birth.

**Ego:** Age 2+; develops as a result of imperfect parenting and the child’s need to develop independent coping strategies.

**Superego:** Age 5+; develops when the child becomes capable of internalizing abstract rules and principles as communicated by parents and others.

**The Structural Model**

Ultimately, Freud recognized certain explanatory limitations in the topographic model (e.g., the model’s inability to account for certain forms of psychopathology), and as a result he developed an alternative, complementary framework to explain normal and abnormal personality development. Although the structural model evolved over a number of years, the theoretical shift from topography to structure is most clearly demarcated by Freud’s (1923/1961) publication of The Ego and the Id, wherein he described in detail the central hypothesis underlying the structural model: the notion that intrapsychic dynamics could be understood with reference to three interacting mental structures called the id, ego, and superego. The id was defined as the seat of drives and instincts (a throwback to the original drive model), whereas the ego represented the logical, reality-oriented part of the mind, and the superego was akin to a conscience, or set of moral guidelines and prohibitions (Brenner, 1973). Figure 5.1 illustrates the sequence of development of the id, ego, and superego in Freud’s structural model.

According to the structural model, personality is derived from the interplay of these three psychic structures, which differ in terms of power and influence (Freud, 1933/1964a, 1940/1964b). When the id predominates, an impulsive, stimulation-seeking personality style results. When the superego is strongest, moral prohibitions inhibit impulses, and a restrained, overcontrolled personality ensues. When the ego (which serves in part to mediate id impulses and superego prohibitions) is dominant, a more balanced set of personality traits develop. Table 5.2 summarizes the psychodynamic conceptualization of personality in Freud’s structural model, as well as within the drive, topographic, and psychosexual stage models.

From 1923 until his death in 1939, Freud spent much of his time elaborating the key principles and corollaries of the structural model, and he extended the model to various areas of individual and social life (e.g., humor, mental errors, cultural dynamics, religious belief). He also made numerous efforts to link the structural model to his earlier work in order to form a more cohesive psychodynamic framework. For example, Freud (and other psychoanalysts) hypothesized that oral fixation was characterized in part by a prominent, powerful id, whereas Oedipal fixation was characterized by strong investment in superego activities. At the time of his death, Freud was actively revising aspects of the structural model (Fancher, 1973; Gay, 1988), and it is impossible to know how the model would have developed had Freud continued his work. This much is certain, however: During the decades wherein Freud explicated details of the structural model of the mind, he altered it in myriad ways, and in doing so he laid the foundation for several concepts that—many years later—became key elements of modern psychoanalytic theory.

**TABLE 5.2 Conceptions of Personality Within Classical Psychoanalytic Theory**

<table>
<thead>
<tr>
<th>Model</th>
<th>Conception of Personality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drive</td>
<td>Personality traits as drive (instinct) derivatives.</td>
</tr>
<tr>
<td>Topographic</td>
<td>Unconscious (repressed) material is a primary determinant of personality.</td>
</tr>
<tr>
<td>Psychosexual</td>
<td>Fixation at a particular psychosexual stage leads to an associated character type.</td>
</tr>
<tr>
<td>Structural</td>
<td>Id-ego-superego dynamics determine personality traits and coping strategies.</td>
</tr>
</tbody>
</table>
Neuro-Analytic Models

Following Freud’s 1909 Clark University lectures, psychoanalysis attracted large numbers of adherents from within the medical and lay communities. At first, these adherents followed Freud’s ideas with little questioning and minimal resistance. By the early 1920s, however, competing schools of psychoanalytic thought were beginning to emerge both in Europe and in America. At first, the growth of these alternative psychodynamic frameworks was inhibited by Freud’s strong personality and by the immense international popularity of psychoanalytic theory (Hilgard, 1987; Torrey, 1992). It was only upon Freud’s death in 1939 that competing psychoanalytic perspectives blossomed into full-fledged theories in their own right.

By the mid-1940s, the discipline had splintered into an array of divergent theoretical perspectives. This splintering process, which has continued (albeit in a somewhat abated form) to the present day, is summarized graphically in Figure 5.2. As Figure 5.2 shows, each post-Freudian psychodynamic model was rooted in classical psychoanalytic theory, but each drew upon ideas and findings from other areas of psychology as well.

Several neo-analytic theories became particularly influential in the decades following Freud’s death. Among the most important of these were Jung’s (1933, 1961) analytical psychology, Erikson’s (1963, 1968) psychosocial theory, Sullivan’s (1947, 1953) interpersonal theory, and the quasi-dynamic models of Adler (1921, 1923), Fromm (1941, 1947), and Horney (1937, 1945). These theories shared a Freudian emphasis on intrapsychic dynamics, childhood experiences, and unconscious processes as determinants of personality and psychopathology. However, each neo-analytic theorist rejected the classical psychoanalytic emphasis on sexuality as a key component of personality, and each theory sought to supplant sexuality with its own unique elements. Key features of the most prominent neo-analytic models are summarized in Table 5.3.

Each neo-analytic model in Table 5.3 attained a loyal following during its heyday, but for the most part these neo-analytic models are no longer influential in mainstream psychology. To be sure, aspects of these neo-analytic theories continue to be discussed (and on occasion isomorphically rediscovered by other personality theorists). However, with the exceptions of Erikson and Sullivan, the neo-analytic theories summarized in Table 5.3 have comparatively few adherents today, and they do not receive much attention within the clinical and research communities.

Erikson’s (1963, 1968) psychosocial approach continues to have a strong impact on personality and developmental research (Franz & White, 1985). Sullivan’s (1953, 1956) interpersonal theory not only helped lay the groundwork for

---

**TABLE 5.3 Neo-Analytic Models of Personality**

<table>
<thead>
<tr>
<th>Theorist</th>
<th>Key Assumption</th>
<th>Key Terms/Concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adler</td>
<td>Family dynamics (especially birth order) are primary determinants of personality.</td>
<td>Striving for superiority, inferiority complex</td>
</tr>
<tr>
<td>Erikson</td>
<td>Social interactions between individual and significant others are key in personality development.</td>
<td>Psychosocial stages, developmental crises</td>
</tr>
<tr>
<td>Fromm</td>
<td>Personality is best understood with reference to prevailing social and political (as well as intrapsychic) forces.</td>
<td>Authoritarianism</td>
</tr>
<tr>
<td>Horney</td>
<td>Infantile dependency-powerlessness is key to personality.</td>
<td>Basic anxiety</td>
</tr>
<tr>
<td>Jung</td>
<td>Personality is shaped by spiritual forces as well as by biological and social unconscious variables.</td>
<td>Archetypes, collective unconscious</td>
</tr>
<tr>
<td>Sullivan</td>
<td>Personality can only be conceptualized within the context of an individual’s core relationships.</td>
<td>Personifications, developmental epochs</td>
</tr>
</tbody>
</table>

---

**Figure 5.2** Evolution of psychodynamic models of personality; arrows indicate the influence of earlier theories/perspectives on later ones.

19th-century philosophy, neurology, psychiatry, and academic psychology

Classical Psychoanalytic Theory

Neo-Analytic Models

Self Psychology

Object Relations Theory

Contemporary Integrative Theories

Cognitive, social, and developmental psychology

Behavioral, cognitive, and humanistic treatment models

Behavioral, cognitive, and humanistic treatment models

Behavioral, cognitive, and humanistic treatment models
Object Relations Theory and Self Psychology

Although the influence of most neo-analytic models has waned, two other psychodynamic frameworks that evolved out of Freud’s work—object relations theory and self psychology—remain very much a part of mainstream psychoanalytic theory and practice. Both frameworks developed out of early work in ego psychology, an offshoot of the classical model; this model updated Freud’s thinking on the role of the ego in personality development. Where Freud had conceptualized the ego primarily in terms of its reality-testing and defensive functions, ego psychologists posited that the ego plays another equally important role in intrapsychic life—setting goals, seeking challenges, striving for mastery, and actualizing potential (Hartmann, 1964). Within this line of thinking, the ego was seen as an autonomous, conflict-free structure, rather than an entity that simply responded to the demands of id, superego, and the external world. Ego psychologists’ reconceptualization of the ego set the stage for object relations theory and self psychology.

Object Relations Theory

Although there are several distinct variants of object relations theory (see Greenberg & Mitchell, 1983), they share a core belief that personality can be analyzed most usefully by examining mental representations of significant figures (especially the parents) that are formed early in life in response to interactions taking place within the family (Gill, 1995; Winnicott, 1971). These mental representations (sometimes called introjects) are hypothesized to serve as templates for later interpersonal relationships, allowing the individual to anticipate the responses of other people and draw reasonably accurate inferences regarding others’ thoughts, feelings, goals, and motivations (Sandler & Rosenblatt, 1962). Mental representations of the parents—parental introjects—also allow the individual to carry on an inner dialogue with absent figures. This inner dialogue helps modulate anxiety and enables the person to make decisions consistent with values and beliefs acquired early in life (Fairbairn, 1952; Jacobson, 1964).

One of the most prominent object relations models of personality today is Blatt’s (1974, 1991) anaclitic-introjective framework. Blending psychoanalytic theory with research in cognitive development, Blatt postulated that the structure of an individual’s parental introjects play a key role in personality development and dynamics. When introjects are weak (or even absent), an anaclitic personality configuration results, characterized by dependency, insecurity, and feelings of helplessness and emptiness. When introjects are harsh and demanding, an introjective personality configuration is produced, characterized by feelings of guilt, failure, worthlessness, and self-loathing. A plethora of studies have shown that Blatt’s anaclitic-introjective distinction helps predict risk for psychopathology and physical illness, the form that psychopathology and illness will take, the kinds of stressful events that are likely to be most upsetting to the individual, and the types of interventions that will effect therapeutic change most readily (Blatt & Homann, 1992; Blatt & Zuroff, 1992).

Self Psychology

Self psychologists share object relations theorists’ emphasis on mental representations as the building blocks of personality. However, self psychologists contend that the key introjects are those associated with the self, including selfobjects (i.e., representations of self and others that are to varying degrees merged, undifferentiated, and imperfectly articulated). Self psychology developed in part in response to analysts’ interest in treating severe personality disorders and other treatment-resistant forms of psychopathology (Goldberg, 1980; Kohut, 1971). The development of self psychology was also aided by a recognition that the knowledge base of analytic theory and practice could be enriched if greater attention were paid to the ontogenesis of the self in the context of early child-caregiver relationships (see Mahler, Pine, & Bergman, 1975).

The most widely known self psychology framework was first described by Kohut (1971, 1977). Kohut postulated that empathic and supportive early interactions resulted in the construction of a secure, cohesive autonomous self, with sufficient resources to deal with the stresses and challenges of intimacy. In contrast, disturbances in infant-caregiver interactions were hypothesized to result in damage to the self along with impairments in evocative constancy (i.e., the ability to generate stable mental images of self and absent others) and an inability to tolerate true intimacy with others. A variety of narcissistic disorders result from damage to the self—and although these narcissistic disorders range in severity from moderate to severe, all reflect the individual’s inability to maintain a cohesive sense of self, except when recapitulating specific (often destructive) interaction patterns. Empirical data testing Kohut’s model are less plentiful than those assessing various object relations frameworks, but studies offer indirect support for Kohut’s contention that early difficulties
within the infant-caregiver unit result in subsequent character pathology and may predict the form that character pathology will take (Galatzer-Levy & Cohler, 1993; Masling & Bornstein, 1993).

Contemporary Integrative Models

Object relations theory and self psychology have revived academic psychologists’ interest in psychodynamic ideas during the past several decades, in part because they represent natural bridges between psychoanalytic theory and research in other areas of psychology (e.g., cognitive, social, developmental; see Barron, Eagle, & Wolitzky, 1992; Masling & Bornstein, 1994; Shapiro & Emde, 1995). While object relations theory and self psychology continue to flourish, a parallel stream of theoretical work has developed that focuses on integrating psychodynamic models of personality with ideas and findings from competing clinical frameworks.

As Figure 5.1 shows, contemporary integrative psychodynamic models draw from both object relations theory and self psychology (and to some extent, from classical psychoanalytic theory as well). Unlike most earlier psychodynamic theories, however, these integrative frameworks utilize concepts and findings from other schools of clinical practice (e.g., cognitive, behavioral, humanistic) to refine and expand their ideas. Some integrative models have gone a step further, drawing upon ideas from neuropsychology and psychopharmacology in addition to other, more traditional areas.

There are almost as many integrative psychodynamic models as there are alternative schools of psychotherapeutic thought. Among the most influential models are those that link psychodynamic thinking with concepts from cognitive therapy (Horowitz, 1988; Luborsky & Crits-Christoph, 1990), behavioral therapy (Wachtel, 1977), and humanistic-existential psychology (Schneider & May, 1995). Other integrative models combine aspects of psychoanalysis with strategies and principles from family and marital therapy (Slipp, 1984). Needless to say, not all analytically oriented psychologists agree that these integrative efforts are productive or desirable. Moreover, the question of whether these integrative frameworks are truly psychoanalytic or have incorporated so many nonanalytic principles as to be something else entirely is a matter of considerable debate within the psychoanalytic community.

PSYCHOANALYTIC PERSONALITY THEORIES: BRINGING ORDER TO CHAOS

Given the burgeoning array of disparate theoretical perspectives, a key challenge confronting psychodynamic theorists involves finding common ground among contrasting viewpoints. Although there are dozens of psychodynamically oriented models of personality in existence today, all these models have had to grapple with similar theoretical and conceptual problems. In the following sections, I discuss how contemporary psychodynamic models have dealt with three key questions common to all personality theories.

Personality Processes and Dynamics

Three fertile areas of common ground among psychodynamic models of personality involve motivation, mental structure and process, and personality stability and change.

Motivation

With the possible exception of the radical behavioral approach, every personality theory has addressed in detail the nature of human motivation—that set of unseen internal forces that impel the organism to action (see Emmons, 1997; Loevinger, 1987; McAdams, 1997). Although classical psychoanalytic theory initially conceptualized motivation in purely biological terms, the history of psychoanalysis has been characterized by an increasing emphasis on psychological motives that are only loosely based in identifiable physiological needs (Dollard & Miller, 1950; Eagle, 1984).

During the 1940s and 1950s, evidence from laboratory studies of contact-deprived monkeys (Harlow & Harlow, 1962) and observational studies of orphaned infants from World War II (Spitz, 1945, 1946) converged to confirm that human and infrahumans alike have a fundamental need for contact comfort and sustained closeness with a consistent caregiver. Around this time, developmental researchers were independently formulating theories of infant-caregiver attachment that posited a separate need to relate to the primary caregiver of infancy and specified the adverse consequences of disrupted early attachment relationships (Ainsworth, 1969, 1989; Bowlby, 1969, 1973).

Object relations theorists and self psychologists integrated these developmental concepts and empirical findings into their emerging theoretical models, so that by the late 1960s most psychodynamic psychologists assumed the existence of one or more psychological drives related to contact comfort (e.g., Kernberg, 1975; Kohut, 1971; Winnicott, 1971). Theorists emphasized the critical importance of interactions that take place within the early infant-caregiver relationship, not only because these interactions determined the quality of contact comfort available to the infant, but also because positive interactions with a nurturing caregiver were necessary for the construction of a cohesive sense of self (Kohut, 1971; Mahler et al., 1975); stable, benevolent introjects (Blatt, 1974, 1991);
and useful mental models of self-other interactions (Main, Kaplan, & Cassidy, 1985).

**Mental Structure and Process**

Along with psychoanalysts’ recognition that mental images of self and others were key building blocks of personality came a change in the way the structures and processes of personality were conceptualized. Terms like *introject*, *schema*, and *object representation* gradually took their place alongside those of Freud’s structural model as cornerstones of psychoanalytic theory and therapy (Bornstein, 1996; Greenberg & Mitchell, 1983). Analysts recognized that in addition to mental images of self and others, a key derivative of early relationships was the formation of *internal working models* of self-other interactions (sometimes identified as *scripts*). This alternative conceptualization of the nature of mental structure not only enabled psychodynamic theorists to derive new treatment approaches (especially for working with character-disordered patients), but also helped connect psychodynamic models with research in attachment theory and social cognition (Galatzer-Levy & Cohler, 1993; Masling & Bornstein, 1994).

This language shift not only was due to theoretical changes, but also reflected a need to develop a psychoanalytic terminology that was less abstract and closer to the day-to-day experience of psychoanalytic patients. In fact, close analysis of psychoanalytic discourse during the early days of object relations theory indicated that this terminological evolution was already underway, regardless of the fact that some newfound language was only gradually becoming formalized within the extant psychoanalytic literature.

In this context, Mayman (1976) noted that at any given time, a psychoanalytic theorist or practitioner may use several different levels of discourse to communicate theoretical concepts. At the top of this framework is psychoanalytic metapsychology—the complex network of theoretical concepts and propositions that form the infrastructure of psychoanalysis. Metapsychological terms are often abstract, rarely operationalizable, and typically used in dialogue with other theorists and practitioners. The concepts of libido and self-object are examples of language most closely associated with psychoanalytic metapsychology.

The middle-level language of psychoanalysis incorporates the constructs used by theorists and practitioners in their own day-to-day work. It is the language in which psychoanalysts conceptualize problems and communicate informally—the kind of language likely to turn up in the heart of a case study or in a set of clinical notes. The terms *oral dependent* and *sublimation* are examples of the middle-level language of psychoanalysis.

The bottom level of psychoanalytic language centers on the experience-near discourse that characterizes therapist-patient exchanges within an analytic session. Less formal than Mayman’s (1976) middle-level language, this experience-near discourse is intended to frame psychoanalytic concepts in a way that resonates with a patient’s personal experience without requiring that he or she have any understanding of psychoanalytic metapsychology. When an analyst discusses a patient’s “aggressive impulses” or “sibling rivalry,” that analyst has translated an abstract concept into experience-near terms.

Thus, like most personality theorists, psychoanalysts today conceptualize mental structures and processes on several levels simultaneously. Unfortunately, it has taken psychoanalytic psychologists a long time to develop an experience-near language for day-to-day work—longer perhaps than it has taken psychologists in other areas. On the positive side, however, in recent years psychoanalytic theorists have addressed this issue more openly and systematically than have theorists from other theoretical backgrounds (e.g., see Horowitz, 1991; Kahn & Rachman, 2000).

**Personality Stability and Change**

The parallel conceptualization of psychoanalytic concepts in relational terms introduced a fundamentally new paradigm for thinking about continuity and change in personality development and dynamics. In addition to being understood in terms of a dynamic balance among id, ego, and superego, stability in personality was now seen as stemming from continuity in the core features of key object representations (including the self-representation; see Blatt, 1991; Bornstein, 1996). In this context, personality change was presumed to occur in part because internalized representations of self and other people changed as a result of ongoing inter- and intrapersonal experiences (Schafer, 1999).

This alternative framework influenced psychoanalytic theories of normal personality development and led to a plethora of studies examining the intrapsychic processes involved in therapeutic resistance, transference, and cure (Blatt & Ford, 1994; Luborsky & Crites-Christoph, 1990). It also called theorists’ attention to the critical importance of present-day experiences in moderating long-term psychodynamic processes. One important consequence of newfound concepts of personality stability and change was a continuing shift from past to present in the study of psychodynamics (Spence, 1982).

**Insight, Awareness, and Coping**

As noted earlier, a key tenet of all psychodynamic models is that unconscious processes are primary determinants of thought, emotion, motivation, and behavior. To the degree that
people have only limited introspective access to these underlying causes, they have only limited control over these processes as well. In part as a consequence of their emphasis on unconscious processes, psychodynamic theorists are unanimous in positing that a certain degree of self-deception is characteristic of both normal and abnormal functioning: Not knowing why we are driven to behave in a certain way, but needing to explain our behavior to ourselves, we generate explanations that may or may not have anything to do with the real causes of behavior (e.g., see Bornstein, 1999b). Moreover, when feelings, thoughts, and motivations produce anxiety (including guilt), we invoke coping strategies called *ego defenses* to minimize these negative reactions and to hide them from ourselves (Cramer, 2000).

The once-radical notion of defensive self-deception is now widely accepted among psychoanalytic and nonpsychoanalytic psychologists alike. Research in social cognition (attribution theory in particular) confirms that systematic, predictable distortions in our perceptions of self and others are a normal part of everyday life (Kihlstrom, 1987; Robins & John, 1997). Although the language of attribution theory differs substantially from that of psychoanalysis, scrutiny reveals a remarkable degree of convergence between these two frameworks. Moreover, researchers have begun to bridge the gap between these ostensibly divergent theoretical perspectives, uncovering a surprising degree of overlap in the process.

One area in which psychodynamic models of defensive self-deception diverge from social psychological models of this phenomenon is in the explanations of why these distortions occur. Although both models agree that these distortions stem largely (but not entirely) from self-protective processes, only psychoanalytic theories explicitly link these distortions to an identifiable set of unconsciously determined strategies termed ego defenses. Social cognitive researchers have tended to favor explanatory models that emphasize limitations in the human information-processing apparatus and mental shortcuts that arise from the need to process multiple sources of information simultaneously as key factors in our cognitive biases and distortions of self and others (Robins & John, 1997). Recent work in terror management theory represents a potential bridge between psychodynamic and social-cognitive work in this area, insofar as the terror management theory model specifies how distortions in inter- and intrapersonal perception simultaneously reflect defensive processes and information-processing limitations (Pyszczynski, Greenberg, & Solomon, 1999).

Ironically, the concept of the ego defense—now central to psychodynamic models of personality—did not receive much attention during the theory’s formative years. In fact, Janet paid greater attention to the defense concept than Freud did (Perry & Laurence, 1984), and in certain respects Janet’s position regarding this issue has turned out to be more accurate than Freud’s has (see Bowers & Meichenbaum, 1984). Evidence suggests that a conceptualization of defensive activity as narrowing of consciousness may be more valid and heuristic than is the classic psychoanalytic conceptualization of defense in terms of exclusion (or *barring*) of material from consciousness (Cramer, 2000; cf. Erdelyi, 1985).

Although Freud discussed certain ego defenses (e.g., repression, projection, sublimation) in his theoretical and clinical writings, it was not until Anna Freud’s (1936) publication of *The Ego and the Mechanisms of Defense* that any effort was made to create a systematic, comprehensive listing of these defensive strategies. Most of the ego defenses discussed by A. Freud continue to be discussed today, although some have fallen out of favor, and new ones have been added as empirical research on defenses began to appear following A. Freud’s (1936) seminal work.

In the decades following A. Freud’s (1936) publication, several alternative methods for conceptualizing ego defenses were offered. The most influential of these are summarized in Table 5.4. As Table 5.4 shows, differences among the individual defense, defense style, and defense cluster models have less to do with the way that specific defensive processes are conceptualized and more to do with how these processes are organized and relate to one another. Each approach to conceptualizing and organizing ego defenses has its own associated measurement strategy (technique), its own research base, and its own adherents within the discipline.

The combined influences of unconscious processes and ego defenses raise the unavoidable question of whether within the

<table>
<thead>
<tr>
<th>TABLE 5.4 Perspectives on Ego Defenses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perspective</td>
</tr>
<tr>
<td>-------------------------------------</td>
</tr>
<tr>
<td>Individual defenses</td>
</tr>
<tr>
<td>Defense style approach</td>
</tr>
<tr>
<td>Defense levels-clusters</td>
</tr>
</tbody>
</table>

*Note. Detailed discussions of these three perspectives are provided by Cramer (2000), Ihilevich & Gleser (1986, 1991), and Vaillant (1986).*
Normal and Pathological Functioning

As any psychologist knows, all humans may be irrational, but some are more irrational than others. Like most personality theorists, psychoanalysts see psychopathology as reflected in a greater-than-expected degree of self-destructive, self-defeating (i.e., irrational) behavior (Millon, 1996). In most psychodynamic frameworks, psychopathology is also linked with increased self-deception, decreased insight into the underlying causes of one’s behavior, and concomitant limitations in one’s ability to modify dysfunctional interaction patterns and alter self-defeating responses (Eagle, 1984).

Psychodynamic models conceptualize psychopathology in terms of three general processes: (a) low ego strength, (b) maladaptive ego defenses, and (c) dysfunctional introjects. Low ego strength contributes to psychopathology because the ego cannot execute reality testing functions adequately; intra- and interpersonal distortions increase. Maladaptive defenses prevent the individual from managing stress and anxiety adequately leading to higher levels of self-deception, increased perceptual bias, and decreased insight. Dysfunctional introjects (including a distorted or deficient self-representation) similarly lead to inaccurate perceptions of self and others, but they also foster dysfunctional interaction patterns and propagate problematic interpersonal relationships.

A key premise of the psychoanalytic model of psychopathology is that psychological disorders can be divided into three broad levels of severity (Kernberg, 1970, 1975). The classic conceptualization of this three-level framework invokes the well-known terms neurosis, character disorder, and psychosis. In most instances, neuroses are comparatively mild disorders which affect only a few areas of functioning (e.g., phobias). Character disorders are more pervasive, long-standing disorders associated with problematic social relationships, distorted self-perception, and difficulties with impulse control (e.g., borderline personality disorder). Psychoses are characterized by severely impaired reality testing and low levels of functioning in many areas of life (e.g., schizophrenia).

Although this tripartite model is both theoretically heuristic and clinically useful, it is important not to overgeneralize regarding differences among different levels of functioning. There are great variations in both severity and chronicity within a given level (e.g., certain neuroses may be more debilitating than an ostensibly more severe personality disorder). In addition, there is substantial comorbidity—both within and between levels—so that a disordered individual is likely to show multiple forms of psychopathology (Bornstein, 1998; Costello, 1995).

As Table 5.5 shows, all three dimensions of intrapsychic dysfunction—low ego strength, maladaptive defenses, and dysfunctional introjects—can be mapped onto the tripartite psychopathology model. In this respect, the model represents an integrative framework that links different psychodynamic processes and connects the psychoanalytic model with contemporary diagnostic research. Although the term neurosis is rarely used today in mainstream psychopathology research, perusal of contemporary diagnostic frameworks (including the DSM-IV; APA, 1994) confirms that the tripartite model has had a profound influence on the way practitioners conceptualize and organize psychological disorders (see also Masling & Bornstein, 1994, and Millon, 1996, for discussions of this issue).

<table>
<thead>
<tr>
<th>TABLE 5.5 Levels of Psychopathology in Psychodynamic Theory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>Neurosis</td>
</tr>
<tr>
<td>Character disorder</td>
</tr>
<tr>
<td>Psychosis</td>
</tr>
</tbody>
</table>

PSYCHOANALYSIS AND CONTEMPORARY PSYCHOLOGY: RETROSPECT AND PROSPECT

Psychodynamic models of personality occupy a unique place in contemporary psychology. On the one hand, they continue to be roundly criticized—perceived by those within and outside the discipline as untested and untestable and denigrated by skeptics as a quasi-phenomenological pseudoscience that has hindered the progress of both scientific and clinical psychology. On the other hand, Freud’s theory continues to fascinate many, occupying a central place in undergraduate and graduate psychology texts and influencing in myriad ways our
understanding of ourselves and our culture. In these final sections, I discuss the place of psychoanalysis in contemporary psychology and speculate about its future.

Testing Psychoanalytic Theories

Within the psychoanalytic community, few issues are as controversial as the nature of evidence in psychoanalysis (see Grunbaum, 1984, for a detailed discussion of this issue). Because psychoanalysis focuses on the in-depth understanding of individuals, many of the theory’s adherents argue that research aimed at confirming general principles of human functioning is of little value (e.g., see Gedo, 1999). Others maintain that without a strong nomothetic research base, psychodynamic theory can never be refined and updated based on our evolving understanding of brain, mind, and behavior (Bornstein, 2001).

The controversy regarding the nature of psychoanalytic evidence dates almost to the inception of the theory itself. Although Freud started his career as a researcher, his attitude toward traditional scientific methods became increasingly dismissive as time went on (Fisher & Greenberg, 1996; Masling & Schwartz, 1979). By the 1920s, psychoanalytic theory had become quite distant from its roots in the natural sciences. With this distancing came an increasing discomfort with traditional nomothetic research methods and a shift toward idiographic data, which most theorists and practitioners saw as being ideally suited to both testing and refining psychoanalytic hypotheses via close analysis of clinical material.

Psychoanalytic theories of personality continue to be strongly influenced by data obtained in the treatment setting. The case reports of psychoanalytic practitioners are still used to formulate general principles of psychopathology, after which these case-derived general principles are reapplied to new cases. Although for many years psychoanalytic psychologists accepted the heuristic value of case studies with little outward resistance, this situation is changing, and contemporary theorists and researchers have begun to question the near-exclusive emphasis on case material in psychoanalytic theory-building (Bornstein, 2001; Bornstein & Masling, 1998).

Although psychodynamic theorists have tended to place the greatest value on material derived from the psychoanalytic treatment session, other forms of idiographic evidence (e.g., anthropological findings, literary records) have also been used to assess psychoanalytic ideas. Needless to say, psychodynamic theorists’ devotion to idiographic methods has led to widespread criticism from within and outside psychology. Proponents of the nomothetic approach maintain that idiographic data—especially those obtained behind closed doors—are neither objective nor replicable, and provide little compelling evidence for the validity of psychoanalytic concepts or the efficacy of psychoanalytic treatment (Crews, 1998; Macmillan, 1996).

The Researcher-Practitioner Split

A noteworthy difference between psychoanalysis and other models of personality becomes apparent when one contrasts the theoretical orientations of practitioners with those of academics. Although there are few practicing psychoanalysts outside large metropolitan centers, a sizable minority of clinical psychologists acknowledge the impact of psychodynamic principles on their day-to-day clinical work (Norcross, Karg, & Prochaska, 1997). In contrast, few personality researchers are openly psychodynamic despite the fact that many concepts in contemporary nonanalytic models of personality are rooted to varying degrees in psychodynamic ideas (Bornstein, 2001).

This researcher-practitioner divide is in part political. During the 1960s and 1970s, behavioral, cognitive, and humanistic personality theorists deliberately distanced themselves from psychoanalytic theory. For behaviorists, this distancing was a product of their core assumptions and beliefs, which clearly conflict with those of psychoanalysis. For cognitivists and humanists, however, the split with psychoanalysis was aimed at enhancing the status of their theories. During this era, it was important for these burgeoning models to distinguish themselves from long-standing psychoanalytic principles in order to assert the uniqueness of their perspectives. Even when parallel concepts arose in these models, theorists emphasized the differences from psychoanalysis rather than focusing on their commonality.

The situation has changed somewhat in recent years: Now that the cognitive and humanistic perspectives are well-established, there has been a slow and subtle reconciliation with Freudian ideas. In the case of humanistic psychology, there has even some explicit acknowledgment of the discipline’s Freudian roots. Even contemporary trait approaches—which have historically been strongly bound to the biological and psychometric traditions—have begun to integrate psychodynamic principles into their models and methods (e.g., see Pincus & Wilson, 2001).

Freud’s Cognitive Revolution

The theory that upended mainstream neuroscience a century ago has had a significant impact on cognitive psychology within the past two decades. Although the synergistic interchange between these two fields dates back at least to the
1960s, the impact of Freud’s cognitive revolution only became widely accepted with the publication of Erdelyi’s (1985) landmark analysis of the interface between cognitive psychology and psychoanalysis. Erdelyi’s work demonstrated that many psychoanalytic concepts dovetailed well with prevailing models of perception, memory, and information processing, and set the stage for an increasingly productive interchange between psychodynamic researchers and cognitive psychologists (e.g., see Bucci, 1997; Horowitz, 1988; Stein, 1997).

The language of the topographic model—conscious, unconscious, and preconscious—continues to be used to a surprising degree, even by researchers unaffiliated with (and often unsympathetic to) Freudian ideas. Moreover, recent research in perception without awareness, implicit learning, and implicit memory draws heavily from psychodynamic concepts (Bornstein & Masling, 1998; Bornstein & Pittman, 1992). Despite psychoanalysts’ long-standing resistance to nomothetic research methods, psychoanalytic principles have undeniably been affected by laboratory research in these other related areas.

Although it was largely unacknowledged at the time, the integration of psychoanalysis and cognitive psychology was central to the development of object relations theory and resulted in substantive reconceptualization of such traditional psychoanalytic concepts as transference, repression, and screen (or false) memories (Bornstein, 1993; Bowers, 1984; Eagle, 2000; Epstein, 1998). As cognitive psychology continues to integrate findings from research on attitudes and emotion (resulting in the study of hot, or affect-laden cognitions), the psychodynamics of perception, memory, and information processing are increasingly apparent.

A likely consequence of this ongoing integration will be the absorption of at least some psychodynamic principles into models of problem solving, concept formation, and heuristic use. Studies confirm that systematic distortions and biases in these mental processes are due in part to constraints within the human information-processing system (Gilovich, 1991), but this does not preclude the possibility that motivational factors (including unconscious motives and their associated implicit memories) may also influence psychological processes that were once considered largely independent of personality and psychopathology factors (McClelland, Koestner, & Weinberger, 1989).

**Developmental Issues**

A second domain of contemporary psychology that has been strongly influenced by psychodynamic models is the study of human development. There is a natural affiliation between developmental psychology and the psychodynamic emphasis on stages of growth, familial influences, and the formation of internal mental structures that structure and guide behavior (Eagle, 1996; Emde, 1992; Stern, 1985). Theorists in both areas have built upon and deepened this natural affiliation.

In contrast to cognitive psychology, the exchange between psychoanalysis and developmental psychology has been openly acknowledged from the outset (see Ainsworth, 1969, 1989). Moreover, the psychoanalysis–developmental psychology interface is synergistic: Just as models of child and adolescent development have been affected by psychodynamic concepts, psychoanalytic models of personality formation and intrapsychic dynamics have been affected by developmental research on attachment, emotions, and cognitive development (Emde, 1992). At this point in the history of psychology, the proportion of developmental psychologists receptive to psychoanalytic ideas is probably higher than that found in any other subdiscipline of psychology (with the possible exception of clinical psychology).

Ironically, although Freud denied the existence of personality development past adolescence, there has been a surprising amount of empirical research on the psychodynamics of aging. Beginning with Goldfarb’s (1963) work, theoreticians and researchers have explored myriad aspects of the psychodynamics of late-life development (e.g., see Ainsworth, 1989; Galatzer-Levy & Cohler, 1993). With the advent of more sophisticated multistore models of memory, the links between psychodynamic processes and injury- and illness-based dementia have also been delineated.

**Psychoanalytic Health Psychology**

Over the years, psychoanalysis has had an ambivalent relationship with health psychology (Duberstein & Masling, 2000). In part, this situation reflects Freud’s own ambivalence regarding the mind-body relationship. After all, the great insight that led Freud to develop his topographic and structural models of the mind—in many ways, the raison d’être of psychoanalysis itself—was the idea that many physical symptoms are the product of psychological conflicts rather than of organic disease processes (Bowers & Meichenbaum, 1984; Erdelyi, 1985). Freud’s early interest in conversion disorders and hysteria set the stage for a psychoanalytic psychology that emphasized mental—not physical—explanations for changes in health and illness states.

Beginning in the 1920s, however, Deutsch (1922, 1924) and others argued that underlying psychodynamic processes could have direct effects on the body’s organ systems. The notion that unconscious dynamics could influence bodily functioning directly was extended and elaborated by Alexander
(1950, 1954), who developed a detailed theoretical framework linking specific psychodynamic processes with predictable physiological sequelae and illness states. When Sifneos (1972) articulated his empirically grounded, psychoanalytically informed model of alexithymia (i.e., an inability to verbalize emotions), the stage was set for the development of a truly psychoanalytic health psychology. The key hypotheses of Sifneos’s approach—that unverbalized emotions can have myriad destructive effects on the body’s organ systems—helped lay the groundwork for several ongoing health psychology research programs that are to varying degrees rooted in psychodynamic concepts. Research on health and hardiness (Kobasa, 1979), stress and coping (Pennebaker & O’Heeron, 1984), emotional disclosure and recovery from illness (Spiegel, Bloom, Kraemer, & Gottheil, 1989), and the “Type C” (cancer-prone) personality (Temoshok, 1987) are all based in part in psychodynamic models of health and illness.

The Opportunities and Challenges of Neuroscience

Some of the first contemporary efforts to integrate psychoanalytic principles with findings from neuroscience involved sleep and dreams (Hobson, 1988; Winson, 1985). Although the language of Freudian dream theory is far removed from that of most neuropsychological models, work in this area has revealed a number of heretofore unrecognized convergences between the psychodynamics and neurology of dreaming. In fact, contemporary integrative models of dream formation now incorporate principles from both domains, setting the stage for extension of this integrative effort to other aspects of mental life.

Neuroimaging techniques such as the computerized axial tomography (CAT) scan, the positron-emission tomography (PET) scan, and magnetic resonance imaging (MRI) have begun to play a leading role in this ongoing psychoanalysis-neuroscience integration. Just as neuroimaging techniques have allowed memory researchers to uncover the neural underpinnings of previously unseen encoding and retrieval processes, functional magnetic resonance imaging (fMRI) have enabled dream researchers to record on-line visual representations of cortical activity associated with different sleep stages and experiences.

Two psychodynamically relevant issues now being studied via fMRI (functional MRI) and other neuroimaging techniques are unconscious processes (e.g., implicit perception and learning) and psychological defenses (Schiff, 1999; Walla, Endl, Lindinger, & Lang, 1999). In general, evidence suggests that implicit processes are centered in mid- and hindbrain regions to a greater degree than are explicit processes—a finding that dovetails with Freud’s own hypotheses as well as with recent evolutionary interpretations of psychodynamic principles (Slavin & Kriegman, 1992). Neuroimaging studies of defensive mental operations are still in their infancy, but preliminary findings suggest that the process of biasing and distorting previously-encoded information involves predictable patterns of cortical (and possibly subcortical) activation.

CONCLUSION: THE PSYCHOLOGY OF PSYCHODYNAMICS AND THE PSYCHODYNAMICS OF PSYCHOLOGY

Despite their limitations, psychodynamic models of personality have survived for more than a century, reinventing themselves periodically in response to new empirical findings, theoretical shifts in other areas of psychology, and changing social and economic forces. Stereotypes notwithstanding, psychodynamic models have evolved considerably during the twentieth century and will continue to evolve during the first decades of the twenty-first century as well.

For better or worse, psychoanalytic theory may be the closest thing to an overarching field theory in all of psychology. It deals with a broad range of issues—normal and pathological functioning, motivation and emotion, childhood and adulthood, individual and culture—and although certain features of the model have not held up well to empirical testing, the model does have tremendous heuristic value and great potential for integrating ideas and findings in disparate areas of social and neurological science.

More than a century ago, Freud (1895b) speculated that scientists would be resistant to psychoanalytic ideas because of the uncomfortable implications of these ideas for their own functioning. Whether or not he was correct in this regard, it is true that psychodynamic models of personality provide a useful framework for examining ourselves and our beliefs. Clinical psychologists have long used psychoanalytic principles to evaluate and refine their psychotherapeutic efforts. Scientists have not been as open to this sort of self-scrutiny. There is, however, a burgeoning literature on the biases and hidden motivations of the scientist (Bornstein, 1999a; Mahoney, 1985), and psychodynamic models of personality may well prove to contribute a great deal to this literature.

REFERENCES


CHAPTER 6

A Psychological Behaviorism
Theory of Personality

ARTHUR W. STAATS

This chapter has several aims. One is that of considering the
role of behaviorism and behavioral approaches in the fields of
personality theory and measurement. A second and central
aim is that of describing a particular and different behavioral
approach to the fields of personality theory and personality
measurement. A third concern is that of presenting some of
the philosophy- and methodology-of-science characteristics
of this behavioral approach relevant to the field of personal-
ity theory. A fourth aim is to characterize the field of person-
ality theory from the perspective of this philosophy and
methodology. And a fifth aim is to project some
developments for the future that derive from this theory per-
spective. Addressing these aims constitutes a pretty full
agenda that will require economical treatment.

BEHAVIORAL APPROACHES AND PERSONALITY

Behavioral approaches to personality might seem of central
importance to personology because behaviorism deals with
learning and it is pretty generally acknowledged that learning
affects personality. Moreover, behaviorist theories were
once the models of what theory could be in psychology. But
certain features militate against behaviorism’s significance
for the field of personality. Those features spring from the tradi-
tional behaviorist mission.

Traditional Behaviorism and Personality

One feature is behaviorism’s search for general laws. That is
ingrained in the approach, as we can see from its strategy of
discovering learning-behavior principles with rats, pigeons,
dogs, and cats—for the major behaviorists in the first and sec-
ond generation were animal psychologists who assumed that
those learning-behavior principles would constitute a com-
plete theory for dealing with any and all types of human
behavior. John Watson, in behaviorism’s first generation,
showed this, as B. F. Skinner did later. Clark Hull (1943) was
quite succinct in stating unequivocally about his theory that
“all behavior, individual and social, moral and immoral, nor-
mal and psychopathic, is generated from the same primary
laws” (p. v). Even Edward Tolman’s goal, which he later
admitted was unreachable, was to constitute through animal study a general theory of human behavior. The field of personality, in contrast, is concerned with individual differences, with humans, and this represents a schism of interests.

A second, even more important, feature of behaviorism arises in the fact that personality as conceived in personology lies within the individual, where it cannot be observed. That has always raised problems for an approach that placed scientific methodology at its center and modeled itself after logical positivism and operationism. Watson had decried as mentalistic the inference of concepts of internal, unobservable causal processes. For him personality could only be considered as the sum total of behavior, that is, as an observable effect, not as a cause. Skinner’s operationism followed suit. This, of course, produced another, even wider, schism with personology because personality is generally considered an internal process that determines external behavior. That is the raison d’être for the study of personality.

Tolman, who along with Hull and Skinner was one of the most prominent second-generation behaviorists, sought to resolve the schism in his general theory. As a behaviorist he was concerned with how conditioning experiences, the independent variable, acted on the organism’s responding, the dependent variable. But he posited that there was something in between: the intervening variable, which also helped determine the organism’s behavior. Cognitions were intervening variables. Intelligence could be an intervening variable. This methodology legitimated a concept like personality.

However, the methodology was anathema to Skinner. Later, Hull and Kenneth Spence (1944) took the in-between position that intervening variables should be considered just logical devices, not to be interpreted as standing for any real psychological events within the individual. These differences were played out in literature disputes for some time. That was not much of a platform for constructing psychology theory such as personology. The closest was Tolman’s consideration of personality as an intervening variable. But he never developed this concept, never stipulated what personality is, never derived a program of study from the theory, and never employed it to understand any kind of human behavior. Julian Rotter (1954) picked up Tolman’s general approach, however, and elaborated an axiomatic theory that also drew from Hull’s approach to theory construction. As was true for Hull, the axiomatic construction style of the theory takes precedence over the goal of producing a theory that is useful in confronting the empirical events to which the theory is addressed.

To exemplify this characteristic of theory, Rotter’s social learning has no program to analyze the psychometric instruments that stipulate aspects of personality, such as intelligence, depression, interests, values, moods, anxiety, stress, schizophrenia, or sociopathy. His social learning theory, moreover, does not provide a theory of what personality tests are and do. Nor does the theory call for the study of the learning and functions of normal behaviors such as language, reading, problem-solving ability, or sensorimotor skills. The same is true with respect to addressing the phenomena of abnormal behavior. For example, Rotter (1954) described the Minnesota Multiphasic Personality Inventory (MMPI) but in a very conventional way. There are no analyses of the different personality traits measured on the test in terms of their behavioral composition or of the independent variables (e.g., learning history) that result in individual differences in these and other traits. Nor are there analyses of how individual differences in traits affect other people’s responses to the individuals or of how individual differences in the trait in turn act on the individual’s behavior. For example, a person with a trait of paranoia is more suspicious than others are. What in behavioral terms does being suspicious consist of, how is that trait learned, and how does it have its effects on the person’s behavior and the behavior of others? The approach taken here is that a behavioral theory of personality must analyze the phenomena of the field of personality in this manner. Rotter’s social learning theory does not do these things, nor do the other social learning theories.

Rather, his theory inspired academic studies to test his formal concepts such as expectancy, need potential, need value, freedom of movement, and the psychological situation. This applied even to the personality-trait concept he introduced, the locus of control—whether people believe that they themselves, others, or chance determines the outcome of the situations in which the individuals find themselves. Although it has been said that this trait is affected in childhood by parental reward for desired behaviors, studies to show that differential training of the child produces different locus-of-control characteristics remain to be undertaken. Tyler, Dhawan, and Sinha (1989) have shown that there is a class difference in locus of control (measured by self-report inventory). But this does not represent a program for studying learning effects even on that trait, let alone on the various aspects of personality.

The social learning theories of Albert Bandura and Walter Mischel are not considered here. However, each still carries the theory-oriented approach of second-generation behaviorism in contrast to the phenomena-oriented theory construction of the present approach. For example, there are many laboratory studies of social learning theory that aim to show that children learn through imitation. But there are not programs to study individual differences in imitation, the cause of such differences, and how those differences affect individual differences in important behaviors (e.g., the ability to copy letters, learn new words, or accomplish other actual learning tasks of the child). Bandura’s approach actually began in a loose social learning framework. Then it moved toward a
behavioral approach several years later, drawing on the approach to be described here as well as the approach of Skinner, and later it moved toward including a more cognitive terminology. Mischel (1968) first took a Watsonian-Skinnerian approach to personality and assessment, as did other radical behaviorists. He later abandoned that position (Mischel, 1973) but, like the other social learning theorists, offered no program for study stipulating what personality is, how it is learned, how it functions, and how personality study relates to psychological measurement.

When all is said and done, then, standard behaviorism has not contributed a general and systematic program for the study of personality or personality measurement. It has features that interfere with doing so. Until they are overcome in a fundamental way (which Tolmanian social learning approaches did not provide), these features represent an impassable barrier.

Behavior Therapy and Personality

The major behaviorists such as Hull, Skinner, and Tolman were animal learning researchers. None of them analyzed the learning of functional human behaviors or traits of behavior. Skinner’s empirical approach to human behavior centered on the use of his technology, that is, his operant conditioning apparatus. His approach was to use this “experimental analysis of behavior” methodology in studying a simple, repetitive response of a subject that was automatically reinforced (and recorded). That program was implemented by his students in studies reinforcing psychotic patients, individuals with mental retardation, and children with autism with edibles and such for pulling a knob. Lovaas (1977), in the best developed program among this group, did not begin to train his autistic children in language skills until after the psychological behaviorism (PB) program to be described had provided the foundation. Although Skinner is widely thought to have worked with children’s behavior, that is not the case. He constructed a crib for infants that was air conditioned and easy to clean, but the crib had no learning or behavioral implications or suggestions. He also worked with programmed learning, but that was a delimited technology and did not involve behavior analyses of the intellectual repertoires taught, and the topic played out after a few years. Skinner’s experimental analysis of behavior did not indicate how to research functional human behaviors or problems of behavior or how they are learned.

Behavior Therapy

The original impetus for the development of behavior therapy (which in the present usage includes behavior modification, behavior analysis, cognitive behavior therapy, and behavior assessment) does not derive from Hull, Skinner, Tolman, or Rotter, although they and Dollard and Miller (1950) helped stimulate a general interest in the possibility of applications. One of the original sources of behavior therapy came from Great Britain, where a number of studies were conducted of simple behavior problems treated by using conditioning principles, either classical conditioning or reinforcement. The learning framework was not taken from an American behaviorist’s theory but from European developments of conditioning principles. As an example, Raymond (see Eysenck, 1960) treated a man with a fetish for baby carriages by classical conditioning. The patient’s many photographs of baby carriages were presented singly as conditioned stimuli paired with an aversive unconditioned stimulus. Under this extended conditioning the man came to avoid the pictures and baby carriages. The various British studies using conditioning were collected in a book edited by Hans Eysenck (1960). Another of the foundations of behavior therapy came from the work of Joseph Wolpe. He employed Hull’s theory nominally and loosely in several endeavors, including his systematic desensitization procedure for treating anxiety problems. It was his procedure and his assessment of it that were important.

A third foundation of behavior therapy came from my PB approach that is described here. As will be indicated, it began with a very broad agenda, that of analyzing human behavior generally employing its learning approach, including behaviors in the natural situation. Its goal included making analyses of and treating problems of specific human behavior problems of interest to the applied areas of psychology. Following several informal applications, my first published analysis of a behavior in the naturalistic situation concerned a journal report of a hospitalized schizophrenic patient who said the opposite of what was called for. In contrast to the psychodynamic interpretation of the authors, the PB analysis was that the abnormal behavior was learned through inadvertent reinforcement given by the treating doctors. This analysis suggested the treatment—that is, not to reinforce the abnormal behavior, the opposite speech, on the one hand, and to reinforce normal speech, on the other (Staats, 1957). This analysis presented what became the orientation and principles of the American behavior modification field: (a) deal with actual behavior problems, (b) analyze them in terms of reinforcement principles, (c) take account of the reinforcement that has created the problem behavior, and (d) extinguish abnormal or undesirable behavior through nonreinforcement while creating normal behavior by reinforcement.

Two years later, my long-time friend and colleague Jack Michael and his student Teodoro Ayllon (see Ayllon & Michael, 1959), used this analysis of psychotic behavior and these principles of behavior modification to treat behavioral symptoms in individual psychotic patients in a hospital. Their
The study provided strong verification of the PB behavior modification approach, and its publication in a Skinnerian journal had an impact great enough to be called the “seeds of the behavioral revolution” by radical behaviorists (Malott, Whaley, & Malott, 1997, p. 175). Ayllon and Michael’s paper was written as though this approach derived from Skinnerian behaviorism and this error was repeated in many works that came later. For example, Fordyce (see 1990) followed Michael’s suggestion both in using the PB principles and in considering his pain theory to be Skinnerian.

The study of child behavior modification began similarly. Following my development of the behavior modification principles with simple problems, I decided that a necessary step was to extend behavior analysis to more complex behavior that required long-term treatment. At UCLA (where I took my doctoral degree in general experimental and completed clinical psychology requirements) I had worked with dyslexic children. Believing that reading is crucially important to human adjustment in our society, I selected this as a focal topic of study—both remedial training as well as the original learning of reading. My first study—done with Judson Finley, Karl Minke, Richard Schutz, and Carolyn Staats—was exploratory and was used in a research grant application I made to the U.S. Office of Education. The study was based on my view that the central problem in dyslexia is motivational. Children fail in learning because their attention and participation are not maintained in the long, effortful, and nonreinforcing (for many children) learning task that involves thousands and thousands of learning trials. In my approach the child was reinforced for attending and participating, and the training materials I constructed ensured that the child would learn everything needed for good performance. Because reading training is so extended and involves so many learning trials, it is necessary to have a reinforcing system for the long haul, unlike the experimental analysis of behavior studies with children employing simple responses and M&Ms. I thus introduced the token reinforcer system consisting of poker chips backed up by items the children selected to work for (such as toys, sporting equipment, and clothing). When this token reinforcer system was adopted for work with adults, it was called the token economy (see Ayllon & Azrin, 1968) and, again, considered part of Skinner’s radical behaviorism.

With the training materials and the token reinforcement, the adolescents who had been poor students became attentive, worked well, and learned well. Thus was the token methodology born, a methodology that was to be generally applied. In 1962 and 1964 studies we showed the same effect with preschool children first learning to read. Under reinforcement their attention and participation and their learning of reading was very good, much better than that displayed by the usual four-year-old. But without the extrinsic reinforcement, their learning behavior deteriorated, and learning stopped. In reporting this and the treatment of dyslexia (Staats, 1963; Staats & Butterfield, 1965; Staats, Finley, Minke, & Wolf, 1964; Staats & Staats, 1962), I projected a program for using these child behavior modification methods in studying a wide variety of children’s (and adults’) problems. The later development of the field of behavior modification showed that this program functioned as a blueprint for the field that later developed. (The Sylvan Learning Centers also use methods similar to those of PB’s reading treatments, with similar results.)

Let me add that I took the same approach in raising my own children, selecting important areas to analyze for the application of learning-behavior principles to improve and advance their development as well as to study the complex learning involved. For example, in 1960 I began working with language development (productive and receptive) when my daughter was only several months old, with number concepts at the age of a year and a half, with reading at 2 years of age. I have audiotapes of this training with my daughter, which began in 1962 and extended for more than 5 years, and videotapes with my son and other children made in 1966. Other aspects of child development dealt with as learned behaviors include toilet training, counting, number operations, writing, walking, swimming, and throwing and catching a ball (see Staats, 1996). With some systematic training the children did such things as walk and talk at 9 months old; read letters, words, sentences, and short stories at 2.5 years of age; and count unarranged objects at 2 years (a performance Piaget suggested was standard at the age of 6 years). The principles were also applied to the question of punishment, and I devised time-out as a mild but effective punishment, first used in the literature by one of my students, Montrose Wolf (Wolf, Risely, & Mees, 1964).

Traditional behaviorism was our background. However, the research developed in Great Britain and by Wolpe and by me and a few others constituted the foundation for the field of behavior therapy. And this field now contains a huge number of studies demonstrating that conditioning principles apply to a variety of human behavior problems, in children and adults, with simple and complex behavior. There can be no question in the face of our behavior therapy evidence that learning is a centrally important determinant of human behavior.

**The State of Personality Theory and Measurement in the Field of Behavior Therapy**

Behaviorism began as a revolution against traditional psychology. The traditional behaviorist aim in analyzing psychology’s studied phenomena was to show behaviorism’s
superiority and that psychology’s approach should be abandoned. In radical behaviorism no recognition is given still that work in traditional psychology has any value or that it can be useful in a unification with behaviorism. This characteristic is illustrated by the Association of Behavior Analysis’s movement in the 1980s to separate the field from the rest of psychology. It took a PB publication to turn this tide, but the isolationism continues to operate informally. Radical behaviorism students are not trained in psychology, or even in the general field of behaviorism itself. While many things from the “outside” have been adopted by radical behaviorism, some quite inconsistent with Skinner’s views, they are accepted only when presented as indigenous developments. Radical behaviorism students are taught that all of their fundamental knowledge arose within the radical behaviorism program, that the program is fully self-sufficient.

Psychological behaviorism, in conflict with radical behaviorism, takes the different view: that traditional psychology has systematically worked in many areas of human behavior and produced valuable findings that should not be dismissed sight unseen on the basis of simplistic behaviorist methodological positions from the past. Psychology’s knowledge may not be complete. It may contain elements that need to be eliminated. And it may need, but not include, the learning-behavior perspective and substance. But the PB view has been that behaviorism has the task of using traditional psychology knowledge, improving it, and behaviorizing it. In that process, behaviorism becomes psychologized itself, hence the name of the present approach. PB has aimed to discard the idiosyncratic, delimiting positions of the radical behaviorism tradition and to introduce a new, unified tradition with the means to effect the new developments needed to create unification.

An example can be given here of the delimiting effect of radical behaviorism with respect to psychological measurement. Skinner insisted that the study of human behavior was to rest on his experimental analysis of behavior (operant conditioning) methodology. Among other things he rejected self-report data (1969, pp. 77–78). Following this lead, a general position in favor of direct observation of specific behavior, not signs of behavior, was proposed by Mischel, as well as Kanfer, and Phillips, and this became a feature of the field of behavioral assessment. The view became that psychological tests should be abandoned in favor of Skinner’s experimental analysis of behavior methodology, an orientation that could not yield a program for unification of the work of the fields of personality and psychological measurement with behavior therapy, behavior analysis, and behavioral assessment.

It may be added that PB, by contributing foundations to behavior therapy, had the anomalous effect of creating enthusiasm for a radical behaviorism that PB in good part rejects. For example, PB introduced the first general behavioral theory of abnormal behavior and a program for treatment applications (see Staats, 1963, chaps. 10 & 11), as well as a foundation for the field of behavioral assessment:

Perhaps [this] rationale for learning [behavioral] psychotherapy will also have to include some method for the assessment of behavior. In order to discover the behavioral deficiencies, the required changes in the reinforcing system [the individual’s emotional-motivational characteristics], the circumstances in which stimulus control is absent, and so on, evaluational techniques in these respects may have to be devised. Certainly, no two individuals will be alike in these various characteristics, and it may be necessary to determine such facts for the individual prior to beginning the learning program of treatment.

Such assessment might take a form similar to some of the psychological tests already in use. . . . [H]owever. . . a general learning rationale for behavior disorders and treatment will suggest techniques of assessment. (Staats, 1963, pp. 508–509)

At that time there was no other broad abnormal psychology-behavioral treatment theory in the British behavior therapy school, in Wolpe’s approach, or in radical behaviorism. But PB’s projections, including creation of a field of behavioral assessment, were generally taken up by radical behaviorists. Thus, despite its origins within PB (as described in Silva, 1993), the field of behavioral assessment was developed as a part of radical behaviorism. However, the radical behaviorism rejection of traditional psychological measurement doomed the field to failure.

That was quite contrary to the PB plan. In the same work that introduced behavioral assessment, PB unified traditional psychological testing with behavior assessment. Behavior analyses of intelligence tests (Staats, 1963, pp. 407–411) and interest, values, and needs tests (Staats, 1963, pp. 293–306) were begun. The latter three types of tests were said to measure what stimuli are reinforcing for the individual. MacPhailamy and Lewinsohn (1971) later constructed an instrument to measure reinforcers that actually put the PB analysis into practice. Again, despite using traditional rating techniques that Skinner (1969, pp. 77–78) rejected, they replaced their behavioral assessment instrument in a delimiting radical behaviorism framework. Thus, when presented in the radical behaviorism framework, this and the other behavioral assessment works referenced earlier were separated from the broader PB framework that included the traditional tests of intelligence, interests, values, and needs and its program for general unification (Staats, 1963, pp. 304–308).

The point here is that PB’s broad-scope unification orientation has made it a different kind of behaviorism in various fundamental ways, including that of making it a behaviorism
with a personality. The PB theory of personality is the only one that has been constructed on the foundation of a set of learning-behavior principles (Staats, 1996). Advancing in successive works, with different features than other personality theories, only in its later version has the theory of personality begun to arouse interest in the general field of behavior therapy. It appears that some behavior therapists are beginning to realize that behaviorists “have traditionally regarded personality, as a concept, of little use in describing and predicting behavior” (Hamburg, 2000, p. 62) and that this is a liability. Making that realization general, along with understanding how this weakens the field, is basic in effecting progress.

As it stands, behavior therapy’s rejection of the concept of personality underlies the field’s inability to join forces with the field of psychological measurement. This is anomalous because behavior therapists use psychological tests even while rejecting them conceptually. It is anomalous also because Kenneth Spence (1944), while not providing a conceptual framework for bringing behaviorism and psychological testing together, did provide a behavioral rationale for the utility of tests. He said that tests produce R-R (response-response) laws—in which a test score (one response) is used to predict some later performance (the later response). It needs to be added that tests can yield knowledge of behavior in addition to prediction as we will see.

This, then, is the state of affairs at present. Not one of the other behavioral approaches—radical behaviorism, Hullian theory, social learning theory, cognitive-behavioral theory—has produced or projected a program for the study of personality and its measurement. That is a central reason why traditional psychology is alienated from behaviorism and behavioral approaches. And that separation has seriously disadvantaged both behaviorism and traditional psychology.

**THE STATE OF THEORY IN THE FIELD OF PERSONALITY**

Thus far a critical look has been directed at the behaviorism positions with respect to the personality and psychological testing fields. This is not to say that those two fields are fulfilling their potential or are open to unification with any behavioral approach. Just as behaviorism has rejected personality and psychological measurement, so have the latter rejected behaviorism. Part of this occurs because traditional behaviorism does not develop some mutuality of interest, view, or product. But the fields of psychological testing and personality have had a tradition that considers genetic heredity as the real explanation of individual differences. Despite lip service to the contrary, these fields have never dealt with learning. So there is an ingrained mutual rejection. Furthermore, the lack of a learning approach has greatly weakened personality theory and measurement, substantively as well as methodologically, as I will suggest.

To continue, examination of the field of personology reveals it to be, at least within the present philosophy-methodology, a curiosity of science. For this is a field without guidelines, with no agreement on what its subject matter—personality—is and no concern about that lack of stipulation. It is accepted that there will be many definitions in the operating field. The only consensus, albeit implicit, is that personality is some process or structure within the individual that is a cause of the individual’s behavior. Concepts of personality range from the id, ego, and superego of Sigmund Freud, through the personal constructs of George Kelly and Carl Rogers’s life force that leads to the maintenance and enhancement of self, to Raymond B. Cattell’s source traits of sociability, intelligence, and ego strength, to mention a very few.

Moreover, there is no attempt to calibrate one concept of personality with respect to another. In textbooks each personality theory is described separately without relating concepts and principles toward creating some meaningful relationships. There are no criteria for evaluating the worth of the products of the field, for comparing them, for advancing the field as a part of science. Each author of a theory of personality is free to pursue her or his own goals, which can range from using factor analytic methods by which to establish relationships between test items and questionnaires to running pigeons on different schedules of reinforcement. There will be little criticism or evaluation of empirical methods or strategies. All is pretty much accepted as is. There will be no critical consideration of the kind of data that are employed and evaluation of what the type of data mean about the nature of the theory. Other than psychometric criteria of reliability and validity, there will be no standards of success concerning a test’s provision of understanding of the trait involved, what causes the trait, or how it can be changed. Also, the success of a personality theory will not be assessed by the extent to which it provides a foundation for constructing tests of personality, therapies, or procedures for parents to employ. It is also not necessary that a personality theory be linked to other fields of study.

Moreover, a theory in this field does not have the same types of characteristics or functions as do theories in the physical sciences. Those who consider themselves personality theorists are so named either because they have created one of the many personality theories or because they have studied and know about one or more of the various existent theories. They are not theorists in the sense that they work on
the various personality theories in order to improve the theory level of the field. They are not theorists in the sense that they study their field and pick out its weaknesses and errors in order to advance the field. They do not analyze the concepts and principles in different theories in order to bring order into the chaos of unrelated knowledge. They do not, for example, work on the large task of weaving the theories together into one or more larger, more advanced, and more general and unified theories that can then be tested empirically and advanced.

An indication of the mixed-up character of the field of personality theory is the inclusion of Skinner’s experimental analysis of behavior as a personality theory in some textbooks on personality theory. This is anomalous because Skinner has rejected the concept of personality, has never treated the phenomena of personality, has had no program for doing so, and his program guides those who are radical behaviorists to ignore the fields of personality and its measurement. His findings concerning schedules of reinforcement are not used by personologists, nor are his students’ findings using the experimental analysis of behavior with human subjects nor his philosophy-methodology of science. His approach appears to be quite irrelevant for the field. What does it say about the field’s understanding of theory that the irrelevance of his theory does not matter? From the standpoint of the philosophy and methodology of PB, the field of personality is in a very primitive state as a science.

To some extent the following sections put the cart before the horse because I discuss some theory needs of the field of personality before I describe the approach that projects those needs. That approach involves two aspects: a particular theory and a philosophy-methodology. The latter is the basis for the projections made in this section. This topic needs to be developed into a full-length treatment rather than the present abbreviation.

**The Need for Theorists Who Work the Field**

One of the things that reveals that the field of personality theory is not really part of a fully developed science is the lack of systematic treatment of the theories in the field. Many study the theories of the field and their empirical products. But that study treats the field as composed of different and independent bodies of knowledge to be learned. There is not even the level of integration of study that one would find in humanities, such as English literature and history, where there is much comparative evaluating of the characteristics of different authors’ works.

If the field of personality theory is to become a real scientific study, we need theorists who work the field. Theories have certain characteristics. They contain concepts and principles, and the theories deal with or derive from certain empirical data. And those concepts, principles, and data vary in types and in functions. With those differences, theories differ in method and content and therefore in what they can do and thus how they fit together or not. We need theorists who study such things and provide knowledge concerning the makeup of our field. What can we know about the field without such analysis?

We need theorists who work the field in other ways also. For example, two scientific fields could be at the same level in terms of scientific methods and products. One field, however, could be broken up by having many different theorists, each of whom addresses limited phenomena and does so in idiosyncratic theory language, with no rules relating the many theories. This has resulted in competing theories, much overlap among theories and the phenomena they address, and much redundancy in concepts and principles mixed in with real differences. This yields an unorganized, divided body of knowledge. Accepting this state provides no impetus for cooperative work or for attaining generality and consensus.

The other hypothetical field has phenomena of equal complexity and difficulty, and it also began with the same unorganized growth of theory. But the field devoted part of its time and effort in working those theories, that is, in assessing what phenomena the various theories addressed, what their methods of study were, what types of principles and concepts were involved, and where there was redundancy and overlap, as well as in comparing, relating, and unifying the different theory-separated islands of knowledge. The terms for the concepts and principles were standardized, and idiosyncrasy was removed. The result was a simpler, coherent body of knowledge that was also more general. That allowed people who worked in the field to speak the same language and to do research and theory developments in that language in a way that everyone could understand. In turn, researchers could build on one another’s work. That simplifying consensus also enabled applied people to use the knowledge better.

It can be seen that although these two sciences are at the same level with respect to much of their product, they are quite different with respect to their theory advancement and operation. The differences in the advancement of knowledge in science areas along these lines have not been systematically considered in the philosophy of science. There has not been an understanding that the disunified sciences (e.g., psychology) operate differently than do the unified sciences (e.g., physics) that are employed as the models in the philosophy of science. Thus, there has been no guide for theorists to work the fields of personality theory and psychological testing to produce the more advanced type of knowledge. So this remains a crying need.
We Need Theory Constructed in Certain Ways and With Certain Qualities and Data

We need theorists to work the field of personality. And they need to address certain tasks, as exemplified earlier and later. This is only a sample; other characteristics of theory also need to be considered in this large task.

Commonalities Among Theories

In the field of personality theory there is much commonality, overlap, and redundancy among theories. This goes unrecognized, however, because theorists are free to concoct their own idiosyncratic theory language. The same or related phenomena can be given different names—such as ego, self, self-concept, and self-efficacy—and left alone as different. Just in terms of parsimony (an important goal of science), each case of multiple concepts and unrecognized full or partial redundancy means that the science is unnecessarily complex and difficult, making it more difficult to learn and use. Unrecognized commonality also artificially divides up the science, separating efforts that are really relevant. Personality theorists, who are in a disunified science in which novelty is the only recognized value, make their works as different as possible from those of others. The result is a divided field, lacking methods of unification.

We need theorists who work to remove unnecessary theory elements from our body of knowledge, to work for simplicity and standardization in theory language. We need to develop concepts and principles that everyone recognizes in order to build consistency and consensus. It is essential also for profundity; when basic terms no longer need to be argued, work can progress to deeper levels.

Data of Theories and Type of Knowledge Yielded

A fundamental characteristic of the various theories in personality is that despite overlap they address different sets of phenomena and their methods of data collection are different. For example, Freud’s theory was drawn to a large extent from personal experience and from the stated experiences of his patients. Carl Rogers’s data was also drawn from personal experience and clinical practice. Gordon Allport employed the lexical approach, which involved selecting all the words from a dictionary that descriptively labeled different types of human behavior. The list of descriptive words was whittled down by using certain criteria and then was organized into categories, taken to describe traits of personality. This methodology rests on large numbers of people, with lay knowledge, having discriminated and labeled different characteristic behaviors of humans. Raymond Cattell used three sources of data. One consisted of life records, as in school or work. Another source was self-report in an interview. And a third could come from objective tests on which the individual’s responses could be compared to the responses of others. These data could be subjected to factor analytic methods to yield groupings of items to measure personality traits.

What is not considered systematically to inform us about the field is that the different types of data used in theories give those theories different characteristics and qualities. To illustrate, a theory built only on the evanescent and imprecise data of personal and psychotherapy experience—limited by the observer’s own concepts and flavored by them—is unlikely to involve precisely stated principles and concepts and findings. Moreover, any attempt by the client to explain her behavior on the basis of her life experience is limited by her own knowledge of behavior and learning and perhaps by the therapist’s interpretations. The naturalistic data of self-description, however, can address complex events (e.g., childhood experiences) not considered in the same way in an experimental setup. Test-item data, as another type, can stipulate behaviors while not including a therapist’s interpretations. However, such items concern how individuals are, not how they got that way (as through learning).

Let us take as an example an intelligence test. It can predict children’s performance in school. The test was constructed to do this. But test data do not tell us how “intelligence” comes about or what to do to increase the child’s intelligence. For in constructing the test there has been no study of the causes of intelligence or of how to manipulate those causes to change intelligence. The theory of intelligence, then, is limited by the data used. Generally, because of the data on which they rest, tests provide predictive variables but not explanatory, causal variables. Not understanding this leads to various errors.

The data employed in some theories can be of a causal nature, but not in other theories. Although data on animal conditioning may lack other qualities, it does deal with cause-effect principles. Another important aspect of data used involves breadth. How many different types of data does a theory draw on or stimulate? From how many different fields of psychology does the theory draw its data? We should assess and compare theories on the types of data on which they are based. Through an analysis of types of data we will have deeper knowledge of our theories, how they differ, how they are complementary, the extent to which they can be developed to be explanatory as well as predictive, and also how they can or cannot be combined in organizing and unifying our knowledge.
**Precision of Theories**

There are also formal differences in theories in terms of other science criteria, for example, in the extent of precision of statement. A known example of imprecision was that of Freud’s *reaction formation*. If the person did not do as predicted, then the reaction formation still allowed the theory always to be “right.” Another type of difference lies in the precision or vagueness of definition concepts. Hull aimed to define his *habit strength* concept with great precision. Rogers’s concept of the life force does not have such a precise definition. Science is ordinarily known for its interest in considering and assessing its theory tools with respect to such characteristics. The field of personality needs to consider its theories in this respect.

**Unifying and Generality Properties of Theories**

Hans Eysenck showed an interest in applications of conditioning principles to problems of human behavior. He also worked on the measurement of personality, in traits such as intelligence and extroversion-introversion. Moreover, he also had interest in variations in psychic ability as shown in experiments in psychokinisis. (During a six-month stay at the Maudsley Hospital in 1961, the author conveyed the spread of our American behavioral applications and also argued about psychic phenomena, taking the position that selecting subjects with high “psychic” ability abrogated the assumptions for the statistics employed.) Theorists vary in the number of different research areas to which they address themselves. And that constitutes an important dimension; other things equal, more general theories are more valuable than narrow theories.

Another property of a theory is that of unifying power. The example of Eysenck can be used again. That is, although he was interested in behavior therapy, personality measurement, and experimental psychic ability, he did not construct a theory within which these phenomenal areas were unified within a tightly reasoned set of interrelated principles. Both the generality and the unifying power of theories are very important.

Freud’s psychological theory was more general than Rogers’s. For example, it pertains to child development, abnormal psychology, and clinical psychology and has been used widely in those and other fields. And Freud’s theory—much more than other theories that arise in psychotherapy—also was high in the goal of unification. John Watson began behaviorism as a general approach to psychology. The behavioral theories of personality (such as that of Rotter, and to some extent the other social learning theories) exhibit some generality and unification. The present theory, PB, has the most generality and unification aims of all. None of the personality theories, with the exception of the present one, moreover, has a systematic program for advancing further in generality and unification.

In general, there are no demands in the field of personality to be systematic with respect to generality or unification, and there are no attempts to evaluate theories for success in attaining those goals. Again, that is different from the other more advanced, unified sciences. That is unfortunate, for the more a theory of personality has meaning for the different areas of psychology, employs products of those fields, and has implications for those fields, the more valuable that theory can be.

This view of the field of personality and its personality theories is a byproduct of the construction of the theory that will be considered in the remaining sections. The perspective suggests that the field of personality will continue to stagnate until it begins to work its contents along the lines proposed.

**PERSONALITY: THE PSYCHOLOGICAL BEHAVIORISM THEORY**

More than 45 years ago, while still a graduate student at UCLA, I began a research program that for some years I did not name, then called social behaviorism, later paradigmatic behaviorism, and finally PB. I saw great importance in the behaviorism tradition as a science, in fundamental learning principles, and in experimentation. But I saw also that the preceding behaviorisms were incompletely developed, animal oriented, and too restricted to laboratory research. They also contained fundamental errors and had no plan by which to connect to traditional psychology, to contribute to it, and to use its products. Very early in the research program I began to realize that animal conditioning principles are not sufficient to account for human behavior and personality. In my opinion a new behavioral theory was needed, it had to focus on human behavior systematically and broadly, it had to link with traditional psychology’s treatments of many phenomena of human behavior, and it had to include a new philosophy and methodology.

**Basic Developments**

The early years of this program consisted of studies to extend, generally and systematically, conditioning principles to samples of human behavior. This was a new program in behaviorism. Some of the studies were informal, some were formal publications, and many involved theoretical analyses of behaviors—experimental, clinical, and naturalistic—that had been described in the psychology literature. One of the goals
was to advance progressively on the dimension of simple-complex with respect to behavior. The low end of the dimension involved establishment of basic principles, already begun with the animal conditioning principles. But those principles had to be verified with humans, first with simple behaviors and laboratory control. Then more and more complex behaviors had to be confronted, with the samples of behavior treated becoming more representative of life behaviors. The beginning of this latter work showed convincingly the relevance of learning-behavior principles for understanding human behavior and progressively indicated that new human learning principles were needed to deal with complex human behavior. Several areas of PB research are described here as historical background and, especially, to indicate how the theory of personality arose in an extended research-conceptual development.

Language-Cognitive Studies

My dissertation studied how subjects' verbal responses to problem-solving objects were related to the speed with which they solved the problem. It appeared that people learn many word labels to the objects and events of life. When a situation arises that involves those objects and events, the verbal responses to them that individuals have learned will affect their behavior. The research supported that analysis.

There are various kinds of labeling responses. A child's naming the letters of the alphabet involves a labeling repertoire. Studies have shown that children straightforwardly learn such a repertoire, as they do in reading numbers and words. The verbal-labeling repertoire is composed of various types of spoken words controlled by stimulus events. The child learns to say “car” to cars as stimulus events, to say “red” to the stimulus of red light, to say “running” to the visual stimulus of rapidly alternating legs that produce rapid movement, and to say “merrily” to people happily reveling. Moreover, the child learns these verbal labeling responses—like the nouns, adjectives, verbs, and adverbs just exemplified—in large quantities, so the verbal-labeling repertoire becomes huge. This repertoire enables the person to describe the many things experienced in life, but it has other functions as well. As discussed later, this and the other language repertoires are important components of intelligence.

As another aspect of language, the child also learns to make different motor responses to a large number of words. The young child learns to look when hearing the word “look,” to approach when hearing the word “come,” to sit when told the word “sit,” and to make a touching response when told to “touch” something. The child will learn to respond to many words with motor responses, constituting the verbal-motor repertoire. This repertoire enables the person to follow directions. It is constituted not only of a large number of verbs, but also of adverbs, nouns, adjectives, and other grammatical elements. For example, most people could respond appropriately to the request to “Go quickly, please, to the top-left drawer of my dresser and bring me the car keys” because they have learned motor responses to the relevant words involved. Important human skills involve special developments of the verbal-motor subrepertoire. As examples, ballet dancers, violinists, NFL quarterbacks, mechanics, and surgeons have special verbal-motor repertoires that are essential parts of their special skills.

Another important part of language is the verbal-association repertoire. When the word salt is presented as a stimulus in a word-association task, a common response is pepper or water. However, an occasional person might respond by saying wound or of the earth or something else that is less usual. Years ago it was believed that differences in associations had personality implications, and word-association tests were given with diagnostic intent. Analysis of word associations as one of the subrepertoires of the language-cognitive repertoire suggests more definitively and specifically that this constitutes a part of personality. Consider a study by Judson, Cofer, and Gelfand (1956). One group of subjects learned a list of words that included the sequence rope, swing, and pendulum. The other group learned the same list of words, but the three words were not learned in sequence. Both groups then had to solve a problem by constructing a pendulum from a light rope and swinging it. The first group solved the problem more quickly than did the second. Thus, in the present view the reasoning ability of the two groups depended on the word associations they had learned.

Word associates are central to our grammatical speech, the logic of our speech and thought, our arithmetic and mathematical knowledge, our special area and general knowledge, our reasoning ability, our humor, our conversational ability, and our intelligence. Moreover, there are great individual differences in the verbal-association repertoire such that it contributes to differences on psychological tests. Additional repertoires are described in the PB theory of language-cognition (see Staats, 1968, 1971, 1975, 1996).

Emotional-Motivational Studies

An early research interest of PB concerned the emotional property of words. Using my language conditioning method I showed subjects a visually presented neutral word (nonsense syllable) paired once each with different auditorily presented words, each of which elicited an emotional response, with one group positive emotion and with another group negative in a
classical conditioning procedure. The results of a series of experiments have showed that a stimulus paired with positive or negative emotional words acquires positive or negative emotional properties. Social attitudes, as one example, are emotional responses to people that can be manipulated by language conditioning (Staats & Staats, 1958). To illustrate, in a political campaign the attempt is made to pair one’s candidate with positive emotional words and one’s opponent with negative emotional words. That is why the candidate with greater financial backing can condition the audience more widely, giving great advantage.

Skinner’s theory is that emotion (and classical conditioning) and behavior (and operant conditioning) are quite separate, and it is the operant behavior that is what he considers important. In contrast, PB’s basic learning-behavior theory states that the two types of conditioning are intimately related and that both are important to behavior. For one thing, a stimulus is reinforcing because it elicits an emotional response. Thus, as a stimulus comes to elicit an emotional response through classical conditioning, it gains potential as a reinforcing stimulus. My students and I have shown that words eliciting a positive or negative emotional response will function as a positive or negative reinforcer. In addition, the PB learning-behavior theory has shown that a stimulus that elicits a positive or negative emotional response will also function as a positive or negative incentive and elicit approach or avoidance behavior. That is a reason why emotional words (language) guide people’s behavior so ubiquitously. An important concept from this work is that humans learn a very large repertoire of emotion-eliciting words, the verbal-emotional repertoire. Individual differences in this repertoire widely affect individual differences in behavior (see Staats, 1996).

One other principle should be added for positive emotional stimuli: They are subject to motivational (deprivation-satiation) variations. For example, food is a stimulus that elicits a positive emotional response on a biological basis; however, the size of the response varies according to the extent of food deprivation. That also holds for the reinforcement and incentive effect of food stimuli on operant behavior. These three effects occur with stimuli that elicit an emotional response through biology (as with food) or through learning, as with a food word.

The human being has an absolutely gargantuan capacity for learning. And the human being has a hugely complex learning experience. The result is that in addition to biologically determined emotional stimuli, the human learns a gigantic repertoire that consists of stimuli that elicit an emotional response, whether positive or negative. There are many varieties of stimuli—art, music, cinema, sports, recreations, religious, political, manners, dress, and jewelry stimuli—that are operative for humans. They elicit emotion on a learned basis. As a consequence, they can also serve as motivational stimuli and act as reinforcers and incentives. That leads to a conclusion that individual differences in the quantity and type of emotional stimuli will have great significance for personality and human behavior.

**Sensorimotor Studies**

Following its human-centered learning approach, PB studied sensorimotor repertoires in children. To illustrate, consider the sensorimotor response of speech. Traditional developmental norms state that a child generally says her first words at the age of 1 year, but why there are great individual differences is not explained, other than conjecturing that this depends on biological maturation processes. In contrast, PB states that speech responses are learned according to reinforcement principles, but that reinforcement depends on prior classical conditioning of positive emotion to speech sounds (Staats, 1968, 1996). I employed this theoretical analysis and learning procedures in accelerating the language development of my own children, in naturalistic interactions spread over a period of months, but adding up to little time expenditure. Their speech development accelerated by three months, which is 25% of the usual 12-month period (Staats, 1968). I have since validated the learning procedures with parents of children with retarded speech development. Lovaas (1977) has used this PB framework. Psychological behaviorism also systematically studied sensorimotor skills such as standing, walking, throwing and catching a ball, using the toilet, writing letters, paying attention, counting objects, and so on in systematic experimental-longitudinal research (see Staats, 1968, 1996).

In this theory of child development, PB pursued its goal of unification with traditional psychology, in this case with the field of child development. The PB position is that the norms of traditional child developmentists provide valuable knowledge. But this developmental conception errs in assuming biological determination and in ignoring learning. Prior to my work, the reigning view was that it was wasteful or harmful to attempt to train the child to develop behaviors early. For example, the 4-year-old child was said to be developmentally limited to an attention span of 5 min to 15 min and thus to be incapable of formal learning. We showed that such preschoolers can attend well in the formal learning of reading skills for 40-min periods if their work behaviors are reinforced (Staats et al., 1964). When not reinforced, however, they do not attend. My later research showed that children learn progressively to attend and work well for longer periods by having been reinforced for doing so.
Rather than being a biologically determined cognitive ability, attention span is actually a learned behavior. The same is true with the infant’s standing and walking, the development of both of which can be advanced by a little systematic training. The child of 2 years also can be straightforwardly trained to count unarranged objects (Piaget said 6 years). Writing training can be introduced early and successfully, as can other parts of the sensorimotor repertoire. I also developed a procedure for potty training my children (see Staats, 1963) that was later elaborated by Azrin and Foxx (1974). Such findings have changed society’s view of child development.

What emerges from this work is that the individual learns the sensorimotor repertoire. Without the learning provided in the previous cases, children do not develop the repertoires. Moreover, the human sensorimotor repertoire is, again, vast for individuals. And over the human community it is infinitely varied and variable. There are skills that are generally learned by all, such as walking and running. And there are skills that are learned by only few, such as playing a violin, doing surgery, or acting as an NFL quarterback. As such there are vast individual differences among people in what sensorimotor skills are learned as well as in what virtuosity.

Additional Concepts and Principles

Human Learning Principles

As indicated earlier, a basic assumption of traditional behaviorism is that the animal learning principles are the necessary and sufficient principles for explaining human behavior. Psychological behaviorism’s program has led to the position that while the animal conditioning principles, inherited through evolution, are indeed necessary for explaining human behavior, they are far from sufficient. I gained an early indication of that with my research on the language conditioning of attitudes, and later findings deepened and elaborated the principles.

What the traditional behaviorists did not realize is that human learning also involves principles that are unique to humans—human learning principles. The essential, new feature of these principles is that much of what humans learn takes place on the basis of what they have learned before. For example, much human learning can occur only if the individual has first learned language. Take two children, one of whom has learned a good verbal-motor repertoire and one of whom has not. The first child will be able to follow directions and therefore will be able to learn many things the second child cannot because many learning tasks require the following of directions. The goodness of that verbal-motor repertoire distinguishes children (as we can see on any intelligence test for children). In PB, language is considered a large repertoire with many important learning functions. Learning to count, to write, to read, to go potty, to form attitudes, to have logic and history and science knowledge and opinions and beliefs, to be religious, to eat healthily and exercise, and to have political positions are additional examples in which language is a foundation. A child of 18 months can easily learn to name numbers of objects and then to count if that child has previously learned a good language repertoire (see Staats, 1968). On the other hand, a child of 3 years who has not learned language will not be able to learn those number skills. The reason for the difference is not some genetic difference in the goodness of learning. Rather, the number learning of the child is built on the child’s previous language learning. It is not age (biology) that matters in the child’s learning prowess; it is what the child has already learned.

Cumulative-Hierarchical Learning

Human learning is different from basic conditioning because it typically involves learning that is based on repertoires that have been previously learned. This is called cumulative-hierarchical learning because of the building properties involved—the second learning is built on the first learning but, in turn, provides the foundation for a third learning. Multiple levels of learning are typical when a fine performance is involved. Let us take the learning of the language repertoire. When the child has a language repertoire, the child can then learn to read. When the child has a reading repertoire, the child can learn more advanced number operations, after which the child can learn an algebra repertoire, which then is basic in learning additional mathematics repertoires, which in turn enable the learning of physics. Becoming a physicist ordinarily will involve in excess of 20 years of cumulative-hierarchical learning.

Cumulative-hierarchical learning is involved in all the individual’s complex characteristics. A sociopath—with the complex of language-cognitive, emotional-motivational, and sensorimotor repertoires this entails—does not spring forth full-blown any more than being a physicist. Understanding the sociopathic personality, hence, requires understanding the cumulative-hierarchical learning of the multiple repertoires that have been involved.

The Basic Behavioral Repertoire: A Cause as Well as an Effect

And that brings us to another concept developed in PB, that is, the basic behavioral repertoire (BBR). The BBRs are those repertoires that provide the means by which later learning can
occur, in the cumulative-hierarchical learning process. In providing foundations for further learning, the three major BBRs—the emotional-motivational, language-cognitive, and sensorimotor—also grow and elaborate through cumulative-hierarchical learning.

The learning of the basic behavioral repertoires changes the individual. The BBRs thus act as independent variables that determine what the individual experiences, how the individual behaves, and what the individual learns. The cumulative-hierarchical learning of such repertoires is fundamental in child development; in fact, the PB theory is that the study of that learning should be the primary objective of this field, as it should be in the field of personality.

**The Concept of Personality**

It is significant in comparing the PB theory to other personality theories to note differences in such things as the type of data involved and the specificity, precision, systematicity, and empirical definition of principles and concepts. It is such characteristics that determine the functions that a theory can have. Another characteristic of the PB approach concerns the schism between traditional psychology and traditional behaviorism. Traditional psychology infers personality as a unique internal process or structure that determines the individual’s unique behavior. That makes study of personality (and related concepts) very central. Traditional behaviorism, in opposition, and according to its fundamental methodology, cannot accept an inferred concept as the cause of behavior. So, while almost every personologist considers learning to be important in personality, traditional behaviorism, which should be concerned with how learning affects personality, cannot even consider the topic. The schism leaves personality theories incomplete and divides psychology.

**The PB Definition of Personality**

The PB program has led to the development of a theory of personality that can resolve that schism in a way that is valuable to both sides. The PB definition of personality is that it is composed of the three basic behavioral repertoires that the individual has learned. That definition harmonizes with behaviorism, for the PB program is to study the behaviors in those repertoires and how they are learned, as well as how they have their effects on the individual’s characteristic behavior. At the same time, that definition is very compatible with the traditional view of personality as an internal process or structure that determines behavior. As such, the PB concept of personality can link with traditional work on personality, including personality tests, and can also contribute to advancement of that work. How the three BBRs compose personality is described next.

**The Emotional-Motivational Aspects of Personality**

There are many concepts that refer to human emotions, emotional states, and emotional personality traits. As examples, it may be said that humans may feel the responses of joy or fear, may be in a depressed or euphoric state, and may be optimistic or pessimistic as traits. The three different emotional processes are not usually well defined. PB makes explicit definitions. First, the individual can experience specific, ephemeral emotional responses depending on the appearance-cessation of a stimulus. Second, multiple emotion-eliciting events can yield a series of related emotional responses that add together and continue over time; this constitutes an emotional state. Third, the individual can learn emotional responses to sets of stimuli that are organized—like learning a positive emotional response to a wide number of religious stimuli. That constitutes an emotional-motivational trait (religious values); that is, the individual will have positive emotional responses to the stimuli in the many religious situations encountered. And that emotional-motivational trait will affect the individual’s behavior in those many situations (from the reinforcer and incentive effects of the religious stimuli). For these reasons the trait has generality and continuity. There are psychological tests for traits such as interests, values, attitudes, and paranoid personality. There are also tests for states such as anxiety and depression and moods. And there are also tests for single emotional responses, such as phobias or attitudes.

Personality theories usually consider emotion. This is done in idiosyncratic terminology and principles. So how one theory considers emotion is not related to another. Theories of emotion at the personality level are not connected to studies of emotion at more basic levels. Many psychological tests measure emotions, but they are not related to one another. Psychological behaviorism provides a systematic framework theory of emotion that can deal with the various emotional phenomena, analyze many findings within the same set of concepts and principles, and thus serve as a unifying overarching theory. Psychological behaviorism experimentation has shown that interest tests deal with emotional responses to occupation-related stimuli, that attitude tests deal with emotional responses to groups of people, and that values tests deal with emotional responses to yet other stimuli, unifying them in the same theory.

In the PB theory, beginning with the basic, the individual has emotional responses to stimuli because of biological structure, such as a positive emotional response to food stimuli,
certain tactile stimuli, warm stimulation when cold, and vice versa, and a negative emotional response to aversive, harmful stimuli of various kinds. Conditioning occurs when any neutral stimulus is paired with one of those biological stimuli and comes to elicit the same type of emotional response. Conditioning occurs also when a neutral stimulus is paired with an emotion-eliciting stimulus (e.g., an emotional word) that has gained this property through learning. The human has a long life full of highly variable, complex experiences and learns an exceedingly complex emotional-motivational repertoire that is an important part of personality. People very widely have different emotional learning. Not everyone experiences positive emotional responses paired with religious stimuli, football-related stimuli, or sex-related stimuli. And different conditioning experiences will produce different emotional-motivational repertoires. Because human experience is so variegated, with huge differences, everyone’s hugely complex emotional-motivational personality characteristics are unique and different.

That means, of course, that people find different things reinforcing. What is a reward for one will be a punishment for someone else. Therefore, people placed in the same situation, with the same reinforcer setup, will learn different things. Consider a teacher who compliments two children for working hard. For one child the compliment is a positive reinforcer, but for the other child it is aversive. With the same treatment one child will learn to work hard as a consequence, whereas the other will work less hard. That is also true with respect to incentives. If one pupil has a positive emotional response to academic awards and another pupil does not, then the initiation of an award for number of books read in one semester will elicit strong reading behavior in the one but not in the other. What is reinforcing for people and what has an incentive effect for them strongly affects how they will behave. That is why the emotional BBR is an important personality cause of behavior.

The Language-Cognitive Aspects of Personality

Each human normally learns a huge and fantastically complex language repertoire that reflects the hugely complex experience each human has. There is commonality in that experience across individuals, which is why we speak the same language and can communicate. But there are gigantic individual differences as well (although research on language does not deal with those). Those differences play a central role in the individual differences we consider in the fields of personality and personality measurement.

To illustrate, let us take intelligence as an aspect of personality. In PB theory intelligence is composed of basic behavioral repertoires, largely of a language-cognitive nature but including important sensorimotor elements also. People differ in intelligence not because of some biological quality, but because of the basic behavioral repertoires that they have learned. We can see what is specifically involved at the younger age levels, where the repertoires are relatively simple. Most items, for example, measure the child’s verbal-motor repertoire, as in following instructions. Some items specifically test that repertoire, as do the items on the Stanford-Binet (Terman & Merrill, 1937, p. 77) that instruct the child to “Give me the kitty [from a group of small objects]” and to “Put the spoon in the cup.” Such items, which advance in complexity by age, also test the child’s verbal-labeling repertoire. The child can only follow instructions and be “intelligent” if he or she has learned the names of the things involved.

The language-cognitive repertoires also constitute other aspects of personality, for they are important on tests of language ability, cognitive ability, cognitive styles, readiness, learning aptitude, conceptual ability, verbal reasoning, scholastic aptitude, and academic achievement tests. The tests, considered to measure different facets of personality, actually measure characteristics of the language-cognitive BBR. The self-concept also heavily involves the verbal-labeling repertoire, that is, the labels learned to the individual’s own physical and behavioral stimuli. People differ in the labels they learn and in the emotional responses elicited by those verbal labels. We can exemplify this using an item on the MMPI (Dahlstrom & Welsh, 1960, p. 57): “I have several times given up doing a thing because I thought too little of my ability.” Individuals who have had different experience with themselves will have learned different labels to themselves (as complex stimuli) and will answer the item differently. The self-concept (composed of learned words) is an important aspect of personality because the individual reasons, plans, and decides depending on those words. So the learned self-concept plays the role of a cause of behavior. As another example, the “suspiciousness” of paranoid personality disorder heavily involves the learned verbal-labeling repertoire. This type of person labels the behaviors of others negatively in an atypical way. The problem is that the unrealistic labeling affects the person’s reasoning and behavior in ways that are not adjustive either for the individual or for others.

These examples indicate that what are traditionally considered to be parts of personality are conceived of in PB as parts of the learned language-cognitive BBR.

The Sensorimotor Aspects of Personality

Traditionally, the individual’s behavior is not considered as a part of personality. Behavior is unimportant for the
personologist. Everyone has the ability to behave. It is personality that is important, for personality determines behavior. Even when exceptional sensorimotor differences are clearly the focus of attention, as with superb athletes or virtuoso musicians, we explain the behavior with personality terms such as “natural athlete” or “talent” or “genius” each of which explains nothing.

Psychological behaviorism, in contrast, considers sensorimotor repertoires to constitute learned personality traits in whole or part. And there are very large individual differences in such sensorimotor repertoires. Part of being a physically aggressive person, for example, involves sensorimotor behaviors for being physically aggressive. Being a natural athlete, as another example, involves a complex set of sensorimotor skills (although different body types can be better suited for different actions). Being dependent, as another example, may also involve general deficits in behavior skills. Moreover, sensorimotor repertoires impact on the other two personality repertoires. For example, a person recognized for sensorimotor excellence in an important field will display language-cognitive and emotional-motivational characteristics of “confidence” that have been gained from that recognition.

A good example of how sensorimotor repertoires are part of personality occurred in a study by Staats and Burns (1981). The Mazes and Geometric Design tests of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) (Wechsler, 1967) were analyzed into sensorimotor repertoire elements. That analysis showed that children learn that repertoire—of complex visual discrimination and other sensorimotor skills—when exposed to learning to write the letters of the alphabet. The expectation, thus, is that children trained to write letters will thereby acquire the repertoire by which to be “intelligent” on the Mazes and Geometric Design tests, as confirmed in our study. As other examples, on the Stanford-Binet (Terman & Merrill, 1937) the child has to build a block tower, complete a line drawing of a man, discriminate forms, tie a knot, trace a maze, fold and cut a paper a certain way, string beads a certain way, and so on. These all require that the child have the necessary sensorimotor basic behavioral repertoire. This repertoire is also measured on developmental tests. This commonality shows that tests considered measures of different aspects of personality actually measure the same BBR. Such an integrative analysis would be central in conceptualizing the field and the field needs many such analyses.

The sensorimotor repertoire also determines the individual’s experiences in ways that produce various aspects of personality. For example, the male who acquires the skills of a ballet dancer, painter, carpenter, center in the NBA, symphony violinist, auto mechanic, hair dresser, professional boxer, architect, or opera singer will in the learning and practice of those skills have experiences that will have a marked affect on his other personality repertoires. Much emotional-motivational and language-cognitive learning will take place, and each occupational grouping will as a result have certain common characteristics.

As final examples, being physically aggressive is generally seen as an aspect of personality, a part of some inner psychological process. However, a person cannot be physically aggressive without the sensorimotor skills for being so. It is true that more is involved than just those skills. But those sensorimotor skills are an important part. Likewise, part of a person’s being caring and nurturing resides in the sensorimotor skills for being so. A person cannot be a “natural” athlete without having learned the repertoire of sensorimotor skills that enables him or her to learn new sports easily, rapidly, gracefully, and very well. One cannot be a mechanical, athletic, artistic, or surgical genius, or a musical or dance virtuoso, without the requisite sensorimotor repertoire. Are sensorimotor differences part of personality? And are those differences learned? The PB theory answer to both questions is yes.

The PB analyses that show tests measure BBRs provide a whole new way of viewing psychological tests, with a large new agenda for research, as will be indicated.

**Definition of the Personality Trait**

The personality trait is thus a particular feature of one or more of the three basic behavioral repertoires. Traits involve complex repertoires. For example, liking a religious song involves an isolated emotion. But if the person also has a positive emotional response to many religious stimuli—to the stated beliefs, history, rituals, holidays, personages, and tenets of religion, generally and particularly—this constitutes a personality trait, an important part of the emotional-motivational BBR (as well as of language-cognitive and sensorimotor repertoires). That emotional-motivational repertoire will have general effects on the individual’s behavior, life experiences, and further learning, both for normal and abnormal traits.

In PB the personality trait, as a complex repertoire of responses, is considered a universe from which the various situations of life sample. To illustrate, the individual’s language repertoire includes many different behaviors. A question like “How much are two and two?” is a life situation that samples the language-cognitive repertoire in eliciting the one response “Four.” Many items on intelligence tests sample individuals’ language repertoires. That sample is representatives of how rich that particular universe is. The entire universe is the total BBR, that is, the personality repertoire.

Personality traits are constituted of particular repertoires that produce types of experience, learning, and behavior. For
example, a person with a trait of religiosity will display coincident knowledge (language) of religious material, will experience religious situations with positive emotion and be motivated by such situations, as well as exhibit the specialized ritualistic behaviors of the religion.

The Principles of the Personality Theory

Figure 6.1 schematizes and makes more explicit the concepts and principles of the PB theory of personality. Personality is composed of the individual’s basic behavioral repertoires. As a consequence of previous learning, depicted as S₁, the individual learns BBRs. At a later time the individual is confronted with an environmental situation, S₂, which elicits (samples) elements from the individual’s BBRs. Those elements make up the individual’s behavior (B) in that situation. Personality does not equate with the individual’s behavior. For example, many individuals learn words that are never uttered. So the individual’s language-cognitive BBR can never be ascertained from observing behavior; the individual’s potential for behavior is greater than that which is exhibited.

Traditional behaviorism never established how biology works its effects in the explanation of behavior. In contrast, in PB’s personality theory the individual’s biological character plays an important role at different times. First, the learning of the basic behavioral repertoires takes place by virtue of the brain and peripheral nervous system, muscles, tendons, emotional response organs, and such. The organic state at the time of learning is thus an important independent variable. This includes permanent biological conditions such as brain damage as well as ephemeral biological conditions such as those of deprivation-satiation, illness, and drug and alcohol effects. These biological conditions that are influential at the time of learning the BBRs are designated as O₁.

In addition, however, at the time the individual experiences a later situation certain biological conditions, O₂, are operating in ways that affect the state of the individual’s BBRs. For the BBRs to be operative they have to be retained (remembered). Any temporary conditions, such as drugs or a fever, that affect the brain mechanisms that house the BBRs will be important, as will more permanent conditions such as brain damage that has deleted BBRs in whole or part. In addition, the biological mechanism plays a third role. Even though the individual has retained the BBRs, other biological conditions, O₃, may affect the ability of S₂, the later situation, to elicit them. For example, the individual’s sensory systems may be affected by drugs or other organic conditions that limit or distort the sensory responses, as occurs with a person who because of poor hearing cannot respond emotionally to a touching dialogue in a movie.

In this theoretical conception environmental conditions play two roles in the determination of the individual’s behavior. Separating these environmental events enables a more explicit consideration of both environmental and biological effects on personality and behavior. In both of these ways the definition of personality becomes more explicit. Several additional specifications can be added.

Plasticity and Continuity in Personality

There has been an issue of whether individuals behave the same across time and situations or whether their behavior is situationally determined. Watson’s behaviorism raised the issue, which was argued to a stalemate in his era. Mischel’s 1968 book revivified the contest by arguing for the situational determinism position and against the conception that the individual has a personality that acts across situations. A number of pro and con works were then published until, as generally happens in such issues, interest for the moment was exhausted. A deeper analysis can be made, however, that can resolve the issue.

To begin, Figure 6.1 has various implications. Behavior is certainly situational, for the situation does indeed play an important role in selecting the elements of behavior displayed in that situation. For example, people generally act boisterously at a football game or wrestling match and sedately in a place of worship, a library, or a museum.

But there is generality to personality also. A particular BBR over time can be relevant to various situations, and the individual’s behavior can thereby show characteristic features across those situations. For example, a person with a large repertoire of skilled singing behaviors will have learned a repertoire whose elements are called out in many later environmental situations. Compared to others the individual will sing more generally and more skillfully than others lacking that repertoire. Clearly that will be a characteristic, general, and stable feature of the individual’s behavior, considered to reflect a personality trait.

Personality typically produces stability over time and situations. For example, a person who has learned positive values (emotion) to positions on the conservative side of many
political-social-economic events (issues) will tend to display conservative behavior in the books and magazines that are read, the television programs that are watched, the lectures that are attended, the church that is attended, the voting choices that are made, the person who is married, the opinions that are expressed, and so on. As this example shows, a general trait—emotional-motivational, language-cognitive, or sensorimotor—promulgates additional trait development by ensuring additional experience of the same type that originally produced the BBR. In the abnormal area, for example, once the individual has learned negative emotional responses to people generally, the individual will display negative behaviors (such as suspicion) to people. They in turn will typically respond in negative ways that will further condition the individual to have negative emotional responses to people. That can become a general, deep, and continuing abnormal trait.

Stability in personality is produced in these ways. Thus, the BBRs, once formed, tend to ensure continuity of experience, learning, and behavior. But personality can also exhibit change. For the process of personality development never ends. Learning goes on for the whole life span. In unusual cases something may happen to change a fundamental direction in life. To illustrate, a conservative, conventional man may experience the horrors and immorality of war and thereby read things and participate in activities and meet people he otherwise would not. And these continuing experiences may ultimately provide him with new BBRs—new personality traits—that change his behavior drastically. The cumulative-hierarchical learning involved smacks of a chaos theory effect.

**The Multilevel Nature of the Theory and the Implications**

Simplification is a goal of science, and oversimplification is common. The traditional approach to personality involves this; that is, personality is conceptually simpler than myriad behaviors. Specification of personality, thus, could make it unnecessary to study all those behaviors. Furthermore, if one takes personality to be the cause of behavior, one need only study personality and not all the other fields of psychology, like animal learning principles and cognitive things (such as language), child development, social interaction principles, educational psychology, and so on.

But PB differs here. Explaining human behavior is not considered a two-level task, with one basic theory level, the study of personality, which explains the second level, behavior. Psychological behaviorism says that psychology is divided into fields that have a general hierarchical relationship with one another. The field of animal learning is basic to a field like developmental psychology because much of development depends on learning. The field of developmental psychology, on the other hand, is basic to the field of personality because important aspects of personality develop in childhood. In turn, knowledge of personality is relevant to psychological measurement, abnormal psychology, and clinical and educational psychology.

This multilevel relationship has many exceptions, and there is a bidirectional exchange between areas (levels). But the present position is that a personality theory that does not take into account the various major fields (levels) of psychology can only be a part theory. Learning, for example, is important to personality, as most personologists would agree. That being the case, the field should demand that a personality theory indicate how it links to and draws from the study of learning. The same is true of the fields of child development, experimental psychology (in studying language-cognition, emotion-motivation, and sensorimotor behavior), biology, and social interaction. Personality theory on the other side should be basic to personality measurement and to abnormal, clinical, educational, and industrial psychology. Personality theories should be evaluated comparatively for the extent to which they have a program for drawing from and contributing to the various fields of psychological knowledge (see Staats, 1996, for PB’s most advanced statement of its multilevel approach.)

The traditional oversimplified view of the study of personality needs change that broadens and deepens its scope as well as its analytic powers.

**PERSONALITY THEORY FOR THE TWENTY-FIRST CENTURY**

The PB theory of personality says the phenomena of personality—what it is, how it is learned, and the effects it has—are complex and require a theory capable of dealing with that complexity. And that complex theory suggests many more things to do than the traditional approach envisages. For one thing, there is a large task of specifying what the personality repertoires are, how they are learned, and how they operate. Psychological behaviorism says it has begun the study, but the task is huge, and the program for the twenty-first century must be suitably huge. It should be added that PB, while showing the task to be more complex than traditionally considered, provides a foundation that simplifies the task. For all the studies made within its framework will be related and meaningful to one another. They all add together and advance toward explaining personality. Doing that permits research becoming progressively more profound, unimpeded by the
necessity of arguing perennially about basics. The fact is the traditional framework allows for a seeming simplicity; personality theories can be created that are simple, but they have very little scope. Worse, however, the traditional framework allows for the creation of an infinity of such approaches to personality, all of them unrelated. The result is a large and chaotic fund of unrelated knowledge, set forth in many different and competing theory languages, impossible to work with as a student, researcher, or practitioner. This constitutes irre-solvable complexity. And the framework only guides the field to multiply its complexity with new and unrelated works. Generally, there is no advancement of knowledge in terms of parsimony, profundity, organization, non-redundancy, relatedness, and explanatory value.

Some of the implications of the PB theory of personality for study in the twenty-first century will be sketched.

Biology and Personality

Biological characteristics do indeed play an important role in human behavior and in individual differences in behavior. But in the present view, without a good conception of personality, biological research is presently not of the type needed. The traditional search is for the biological mechanisms that produce personality traits, which PB considers the wrong path. Rather, the PB position is that the individual’s biology provides the mechanism by which the learning of the BBRs can take place, be stored, and be selectively activated by the stimuli of the later environmental situations the individual encounters. Biological studies of various kinds are needed to specify the biological events involved in these processes.

Learning and Personality

While biological conditions are the most basic level of study proposed, it is the field of learning that is the most important basic level. Anomalously, however, especially since most every personologist would agree that personality is in good measure learned, personologists generally have not studied how learning-behavior principles are involved in the acquisition or function of personality. There seems to be an implicit view that learning is not that much different for people except in extreme cases.

The PB position, on the contrary, is that the personality repertoires are learned, that there are wide individual differences in the learning conditions involved, and that those differences produce infinitely varied personality characteristics. Psychological behaviorism says that the first major task of a personality theory is formulating a basic theory of learning-behavior and a theory of human learning. No other existing personality theory does this.

Human Learning and Personality

The basic animal-conditioning principles are not sufficient for dealing with the learning of personality. There have been studies, long since abandoned, employing human subjects that dealt with more complex learning situations and produced principles such as mediated generalization, sensory precondi-tioning, and verbal associations. But there has not been a conceptual framework to guide the field to study what is necessary, that is, to study how humans learn complex, functional repertoires in an advancing cumulative-hierarchical way. There has been no systematic goal of studying the basic behavioral repertoires that are important to humans. Although there are research fields that study language, emotion, and sensorimotor behavior, these fields do not systematically address how these behaviors are important for human adjustment. Studies should be conducted that indicate how such repertoires function to (a) change the individual’s experience, (b) change the individual’s behavior, and (c) change the individual’s ability to learn. Such knowledge is needed to provide foundations for advancing the study of personality. For constructing theory, personology needs fundamental knowledge of cumulative-hierarchical learning, the BBRs, their content, and how the BBRs work to affect experience, learning, and behavior.

Developmental Psychology

Some of the theories of personality include reference to how personality develops in childhood. Freud’s psychoanalytic theory initiated this and has had great influence on some other personality theories in this respect. But Freud’s theory of learning was lacking: He had no understanding of human learning principles or what is learned via those principles, no concepts of the BBRs, how they are learned, how important they are for further learning of personality, and so on. So his treatment (and others in this tradition) of child development in personality formation had to be limited and lacking.

The PB position is that the learning experiences of child-hood set the individual’s basic personality (BBRs) to a great extent so that what follows typically continues in the same line of development. This conceptual position and its empirical findings indicate that the field of child development should be an essential study. The focus of the field in the PB view should become the study of the central BBRs that are learned in childhood—a large agenda. This position recognizes the value of traditional research, such as longitudinal study of
behavioral development, but also sets new avenues of research. To illustrate, it is important to know that children stand unaided at the age of 6 months and walk at the age of 1 year. But that type of knowledge needs to be joined with a behavioral analysis of the behavior involved, how the behavior is learned, and what the function of the behavior is in later development.

Moreover, research needs to be conducted with respect to how repertoires are learned in a cumulative-hierarchical manner to constitute progressively more complex entities that constitute personality. Language development, for example, needs progressive study from the time when the repertoires are simple to the time when they are more complex, both in their features that are general to most children as well as in features personal to individuals. The manner in which different repertoires in language provide the springboards for later learning needs study. To illustrate, the verbal-motor repertoire (by which the child follows directions) is elaborated throughout childhood. How is that BBR basic in the learning of elements in other language, sensorimotor, and emotional repertoires? Such very essential subject matter is not being studied today.

This is to say that the theory of personality as BBRs projects a new framework for research in developmental psychology that will make developmental psychology fundamental for the fields of personality and personality measurement (see Staats, 1966).

Social Psychology

The basic principles of learning behavior and the human learning principles pertain to single individuals. But much learning of humans and much human behavior occur in social interaction. While learning and behavior follow the basic principles, principles of social interaction can be abstracted that are useful in understanding personality formation and function.

Take the child’s learning of the personality repertoires. Very central elements are formed in the parent-child interactions. And that process will be influenced greatly by the BBRs the child learns to the parent (as a stimulus object), as well as the reverse. To illustrate, the parent ordinarily provides for the child’s needs, which means the presentation of positive emotional stimuli (food, warmth, caresses) paired with the parent. The parent comes thereby to elicit a very positive emotional response (love) in the child. And that is important to the child’s further learning, for the more positive emotion the parent elicits, the more effective the parent will be in promoting the child’s learning. That follows from PB’s social psychological principle that the stronger a person elicits a positive emotional response in another individual the more effective the person will be as a reinforcing and directive stimulus for the individual. That means that the parent who is more loved will be more effective in rewarding the child for a desired behavior or in admonishing the child for an undesirable behavior. The more loved parent will also be a stronger “incentive” for the child to follow in learning via imitation. Moreover, generalization will occur to other people so the child has learned a general personality trait.

The point is that the PB framework calls for research that concerns how social interaction principles (see Staats, 1996) are involved in personality formation and function.

Personality Tests and Measurement

There is not room in this chapter to deal with the nature of the field of psychological measurement as a science. However, it shares the same weakness as the field of personality already described and repairing those weaknesses calls for many studies of different types, including linking psychological measurement to other fields of psychology, such as that of learning. Traditional behaviorism never made sense of how the concepts and methods of psychological testing are related to behaviorism concepts and methods (see Skinner, 1969, pp. 77–78). The conceptual gap between the two sets of knowledge is just too wide. To understand tests and test construction methods in behavioral terms, it is necessary to have the concepts and principles of a behavioral theory of personality, so the developments made by PB are necessary for bridging the gap. PB introduces the position that tests can provide information about behavior and personality.

Let me begin by making a behavioral analysis of test construction methods, in a manner that answers the question of why psychological tests can predict later behavior. Traditionally, tests are thought to predict behavior because they measure an unobservable process-entity of personality. Rather, tests can predict behavior because that is what they are constructed to do. That is, the test constructor first gathers a group of items. But in test construction only those items that do predict the behavior of interest are retained. Sometimes the test constructor first selects items without any justifying rationale. Sometimes, however, the test constructor first selects items that are believed to be measures of the personality trait. But this selection difference does not matter, for in both cases the test constructor discards and retains items on the basis of which ones relate to (predict) the behavior of interest.

The next question is why items are related to behaviors. Some, influenced by radical behaviorism, have assumed that the test item and the predicted behavior are, and should be, the same. However, in most cases that is not true. One real reason
for item-behavior relationships is that the test measures an element of a BBR or the verbal labeling of that repertoire. For example, we would find that a group of people who affirmatively answered the item “I am an excellent athlete” would also display more athletic ability than would a group who answered negatively. The two behaviors are in the same repertoire. People generally learn to describe their own behavior with some accuracy (but there are variations in that respect).

It can also be the case that a test, because of how it was constructed, measures a BBR that is necessary for the learning of the predicted behavior. Intelligence tests are a prime example. Behavioral analysis of IQ test items reveals that many test whether the child has the language-cognitive elements. Most of the items, for example, test for the child’s verbal-motor repertoire that is necessary for following instructions. Others test the number concept repertoire, the counting and other arithmetic repertoires, and the verbal-labeling repertoire. The manner in which items on the WPPSI (Wechsler, 1967) measure aspects of the sensorimotor repertoire has been described earlier. Why do such items predict later school performance? The answer is that the items measure basic behavioral repertoires the child needs to be successful in learning materials that are later presented in school. So the items correlate with school performance.

Other tests measure the emotional-motivational BBR. Consider an interest test. Constructing the test involves gathering a number of items together that are thought to represent a range of interests that are occupationally relevant. But the important part involves the standardization procedures. The items are given to different occupational groups, and those that distinguish the groups are retained and organized (keyed). When the test is used, it can be ascertained whether the individual answers the items in a manner that is like some particular occupational group. What does this mean in the PB analysis? The answer is that the items measure emotional responses (indicated, e.g., by like-dislike) to different life stimuli. So the individual’s test responses reveal life stimuli to which the individual has positive and negative emotional responses. Remember that those life stimuli the individual likes or dislikes will also serve as positive or negative reinforcers and incentives. Thus, if the individual has emotional responses to life stimuli that are like people who are successful in some occupation, then the individual should be happy in that same work situation. Moreover, the individual should be reinforced by that work and be attracted to it incentively. That means that other things equal, the individual should work harder in the job, study relevant material more, and so on. That is why interest tests predict job success.

It is important to bridge the psychological testing-behaviorism gap, for unifying the two traditions produces new knowledge. For example, in terms of the present theory of personality, the various existent psychological tests are an invaluable source of knowledge for defining the basic behavioral repertoires. PB’s basic experimental studies, developmental studies, and behavior therapy studies have been important avenues of definition. But the manner in which psychological tests have been constructed means that their items measure elements of BBRs that constitute aspects of personality. The extensive work of behaviorally analyzing the items of psychological tests can be expected to tell us much about the content of personality (see Staats, 1996). And, as indicated, those analyses will then yield directives for conducting research on how the BBRs involved are learned and how they function in producing the individual’s behavior. We have already trained children to be more intelligent (Staats & Burns, 1981) by training them in basic behavioral repertoires. In addition, interest and values (see Staats, 1996) tests have been shown to measure aspects of the learned emotional-motivational BBR. Those findings merely open the way.

Other positive avenues of development emerge from the conceptual unification of tests and PB theory. For one thing, the unified theory enables us to understand what tests are. That should be valuable in constructing tests. The approach provides an avenue for defining in objective, stipulable terms just what personality is. That should be valuable in using tests, namely, that test items—not just total scores—when analyzed behaviorally, describe the content of personality traits of the individual. This conception of tests, moreover, says that tests can yield more than prediction; they can describe the contents of personality traits and thus the nature of the individual’s BBR being measured. With study of how people come to learn those personality traits we will have knowledge on how to avoid doing things that will give children undesirable traits, while doing things to give them desirable traits. And behaviorally-analyzed tests will also give specific information regarding what remedial treatment needs to do.

Many studies are needed that analyze existing tests in terms of the behavioral repertoires they assess, as already demonstrated in PB experiments. With that knowledge tests could be compared to one another in a way that would make sense of the field. At present tests are independent entities; they are not related to each other. Many tests of different aspects of personality are actually redundant and share types of items (e.g., interest, values, and needs tests, on the one hand, and fears, anxiety, and stress tests, on the other).

The field of testing does not relate itself to the content areas of psychology or to personality theories. The analysis of tests in terms of BBRs provides a means for doing so. Studying
how the repertoires on tests are learned and function will lead to studies that are relevant in different areas of psychology. There is a vast amount of research to be conducted within this framework. The results of that research will help organize the presently chaotic knowledge of the field. That research will help relate the field of testing to the other fields of psychology. That research will render theoretically meaningful many works that exist in this field. And the knowledge produced should also enable the field to construct better tests.

Abnormal Psychology

The PB position is that a theory of personality should contain principles and concepts for formulating a theory of abnormal psychology. Freud’s psychoanalytic theory was composed to have that potentiality, and this was elaborated in others works. Radical behaviorism has not produced such a theory, nor has the traditional behavioral field.

Psychological behaviorism, however, began a new development in behaviorism when it analyzed the opposite speech of the schizophrenic patient (Staats, 1957). Not only was the abnormal symptom considered as a behavior, but the analysis also indicated how the symptom was learned and how it could be extinguished and replaced with normal behavior. In the early presentation of PB (see Staats, 1963) one chapter was devoted to further formulation of its theory of abnormal behavior. This theory was employed in the social learning theory of abnormal behavior (Bandura, 1968) and in later behavioral works of various kinds. However, the PB theory was developed a good deal further after its theory of personality was systematically formulated (Staats, 1975) and then further extended (Staats, 1989, 1996).

The PB theory of abnormal behavior follows the theory of personality schematized in Figure 6.1. However, each term in the causal circumstances can be normal or abnormal and result in abnormal behavior. With respect to biological conditions, O₁, O₂, and O₃ may be abnormal in some way. For example, because of organic conditions a child with Down syndrome does not learn normally and will display deficits in the BBRs and thus not behave normally in various life situations, such as school. The same is true of the O² and O³. When they are abnormal, they will produce abnormal behavior.

In addition, the behavioral variables in the schematized theory of personality can be either normal or abnormal. In this case abnormal can mean either deficits in what should be or inappropriate conditions that should not occur. The original learning, for example, S₁, may be deficit or inappropriate and produce deficit or inappropriate BBRs that will result in deficit or inappropriate behavior in later situations such that the individual will be diagnosed as abnormal. The deficit or inappropriate conditions can also occur in S₂ and produce behavior that will be judged as abnormal.

The task is to analyze, for the various diagnostic categories, these various behavioral or organic conditions that produce abnormal behavior. Each such analysis constitutes a theory of the disorder involved that can be employed by therapists or parents. For example, if the deficit or inappropriate conditions occur at S₁, the analysis can be used to instruct parents how to see to it that the child does not develop abnormal BBRs. The analysis will also provide the practitioner with knowledge about how to correct the abnormal conditions and treat the behavior disorder after it has occurred. For example, PB works have presented analyses of developmental disorders, developmental reading disorder, autism, and mental retardation (Staats, 1996). In addition, PB theories of depression, the anxiety disorders (phobic disorder, generalized anxiety disorder, panic disorder, obsessive-compulsive disorder, and posttraumatic stress disorder) have also been presented. Various other behavior therapists have produced other analyses of behavior disorders that use elements from PB theory. However, typically they do not employ the full approach, and there remains a general need to stipulate the elements of abnormal BBRs further, how they are learned, and how they have their effects in a general theory of abnormal behavior.

The PB theory of abnormal behavior takes the position that traditional descriptions of categories of abnormal behavior (see DSM-IV; American Psychiatric Association, 1994) are valuable. PB analyses of behavior disorders yield extensive implications for research that PB suggests for the twenty-first century. So, in addition to those already made, many analyses of the various behavior disorders are needed. Such analyses need to be empirically verified, their implications for prevention and treatment assessed, and their implications for test construction exploited. Centrally, research is needed that gathers observations of the development of abnormal behavior through learning that, strangely enough, have never been made.

Application of the Personality Theory

From the beginning the PB position has been that basic and applied work should be closely related in psychology but presently are not. For example, the field of animal learning has ceased providing useful information to the various areas of human study because the field needs input from those areas concerning important things to study. As an example in the other direction, a personality theory in the PB view should have implications for the improvement of psychology’s fields of practice. To illustrate, my own personal experience has exposed me to cases of disadvantageous parenting that re-
sulted from following psychoanalytic theory or traditional developmental (biologically oriented) theory. Such cases of applied failure represent disconfirmation of the theory. A good theory should yield good applications. An important part of PB’s development, thus, has been directed toward practice, as will be briefly mentioned.

**Clinical Applications**

The analysis of the opposite speech of the schizophrenic patient contained clinical directives. The analysis said that the opposite speech was learned and maintained via the inadvertent reinforcement provided by the professional staff (Staats, 1957). That analysis led directly to applications (Ayllon & Michael, 1959). As another example, PB’s token reinforcer system was employed as the token economy in dealing with hospitalized psychotic patients (Ayllon & Azrin, 1968). Psychological behaviorism analyses and reinforcement methods have been used to train mentally retarded children (Bijou, 1965; Birnbrauer, Bijou, Wolf, & Kidder, 1965) and autistic children (Lovaas, 1977), to toilet train children (Azrin & Foxx, 1974), and to treat juvenile delinquents in different settings (see Staats & Butterfield, 1965; Wolf, et al., 1976). Wolf, Risely, and Mees (1964) used the PB approach of working in the naturalistic situation, including PB’s time-out procedure, in their seminal study to treat an autistic child’s behavior problems. Many of the other extensions of PB’s methods, as suggested for a wide variety of children’s problems (see Staats, 1963; Staats & Butterfield, 1965; Staats & Staats, 1962) were accomplished by others, creating the body of works contributed to the establishment of the field of behavior analysis. As another example, the PB theory of language provided a basis for understanding why traditional verbal psychotherapy could be used to change behavior therapeutically laying a foundation for the field of cognitive behavior therapy (Staats, 1972). Radical behaviorism, however, rejected for some 16 years. Finally, verbal therapy was later accepted as though it were a derivative of radical behaviorism (Hamilton, 1988; Hayes & Wilson, 1994). Additional projections of clinical research and treatment have been outlined based on the additional developments of PB (Staats, 1996, chap. 8).

**Educational Psychology Applications**

The PB research on reading and treatment of nonreading has already been mentioned. Reading was conceptualized as a later elaboration of the language-cognitive BBR. Learned on the foundation of the repertoires of language, it is a complex repertoire that requires long-term training and a huge number of training trials. The subrepertoires of reading, when they have been acquired, serve various learning functions for the individual in later school learning (Staats, 1975). The PB theory of reading focuses on this extensive learning and denies the existence of biological defects responsible for learning disabilities such as dyslexia because the children have normal intelligence, which means normal language BBRs. PB research and analysis thus states the definitive principle that if the child has developed normal language, then the child has all the cognitive ability needed to learn reading perfectly well because no additional abilities are required for reading (see Staats, 1975).

Dyslexia arises because there is inadequate reinforcement to maintain the child’s attention and participation in the long task. I designed the token reinforcer system to solve the motivation problem by providing reinforcement for the child’s attention and participation. The system works widely, as shown by its use in the multitude of studies and programs designed to treat reading and other developmental academic disorders (see Burns & Kondrick, 1998; Sulzer-Azoroff & Mayer, 1986). The Sylvan Learning Centers enterprise by its use of the token reinforcer system validates the system as well as the PB theories of developmental academic disorders (see Staats, 1963, 1968, 1975, 1996). The PB theories of the various academic repertoires (reading, writing, counting, number operations, math) provide the foundation for deriving a large body of additional research to understand school learning and to solve the problems of school learning. The educational field’s absorption with cognitive psychology stands in the way of the vast research and application that would advance education so much.

**CONCLUSION**

The PB theory of personality is set in a general theory that goes from the study of basic learning, including the biology of that learning, through the multiple levels of study that provide its principles and concepts. The theory of personality, thus, is sunk into general psychology, making connections to various fields in psychology. It is specific, objective, and empirical. It draws widely on various areas of study, and it has implications for conducting large amounts of additional research and application in various areas and fields of study. The theory provides a philosophy of science and methodology of theory construction. This is the only theory of personality that claims it can be employed to establish or to change personality, a claim that if fulfilled would have enormous importance. It is the only theory that is unified and has comprehensive scope—sorely needed developments for the field and psychology generally. It is a theory that ties together personality and personality measurement on a broad front. And it projects new areas and
topics of research. An important need for the twenty-first century is to compare this theory with others as part of the general comparison and evaluation of personality theories called for by PB. Another is to exploit the theory in the various areas of theoretical analysis and empirical research it suggests.

REFERENCES


Cognitive-experiential self-theory (CEST) is a broadly integrative theory of personality that is compatible with a variety of other theories, including psychodynamic theories, learning theories, phenomenological self-theories, and modern cognitive scientific views on information processing. CEST achieves its integrative power primarily through three assumptions. The first is that people process information by two independent, interactive conceptual systems, a preconscious experiential system and a conscious rational system. By introducing a new view of the unconscious in the form of an experiential system, CEST is able to explain almost everything that psychoanalysis can and much that it cannot, and it is able to do so in a scientifically much more defensible manner. The second assumption is that the experiential system is emotionally driven. This assumption permits CEST to integrate the passionate phallic-and-tooth unconscious of psychoanalysis with the “kinder, gentler” affect-free unconscious of cognitive science (Epstein, 1994). The third assumption is that four basic needs, each of which is assumed in other theories to be the one most fundamental need, are equally important according to CEST.

In this chapter, I review the basic assumptions of CEST, summarize the research conducted to test the theory, and note the implications of the theory for research and psychotherapy.

TWO INFORMATION-PROCESSING SYSTEMS

According to CEST, humans operate by two fundamental information-processing systems: a rational system and an experiential system. The two systems operate in parallel and are interactive. CEST has nothing new to say about the rational system, other than to emphasize the degree to which it is influenced by the experiential system. CEST does have a great deal to say about the experiential system. In effect, CEST introduces a new system of unconscious processing in the experiential system that is a substitute for the unconscious system in psychoanalysis. Although like psychoanalysis, CEST emphasizes the unconscious, it differs from psychoanalysis in its conception of how the unconscious operates. Before proceeding further, it should be noted that the word rational as used in the rational system refers to a set of

This chapter includes material from several other chapters and articles as well as new information. The research reported here was supported by National Institute of Mental Health (NIMH) Research Grant MH 01293 and NIMH Research Scientist Award 5 KO5 MH 00363.
analytical principles and has no implications with respect to the reasonableness of the behavior, which is an alternative meaning of the word.

It is assumed in CEST that everyone, like it or not, automatically constructs an implicit theory of reality that includes a self-theory, a world-theory, and connecting propositions. An implicit theory of reality consists of a hierarchical organization of schemas. Toward the apex of the conceptual structure are highly general, abstract schemas, such as that the self is worthy, people are trustworthy, and the world is orderly and good. Because of their abstractness, generality, and their widespread connections with schematic networks throughout the system, these broad schemas are normally highly stable and not easily invalidated. However, should they be invalidated, the entire system would be destabilized. Evidence that this actually occurs is provided by the profound disorganization following unassimilable experiences in acute schizophrenic reactions (Epstein, 1979a). At the opposite end of the hierarchy are narrow, situation-specific schemas. Unlike the broad schemas, the narrower ones are readily susceptible to change, and their changes have little effect on the stability of the personality structure. Thus, the hierarchical structure of the implicit theory allows it to be stable at the center and flexible at the periphery. It is important to recognize that unlike other theories that propose specific implicit or heuristic rules of information processing, it is assumed in CEST that the experiential system is an organized, adaptive system, rather than simply a number of unrelated constructs or so-called cognitive shortcuts (e.g., Tversky & Kahneman, 1974). As it is assumed in CEST that the experiential system in humans is the same system by which nonhuman animals adapt to their environments, it follows that nonhuman animals also have an organized model of the world that is capable of disorganization. Support for this assumption is provided by the widespread dysfunctional behavior that is exhibited in animals when they are exposed to emotionally significant unassimilable events (e.g., Pavlov, 1941).

Unlike nonhuman animals, humans have a conscious, explicit theory of reality in their rational system in addition to the model of reality in their experiential system. The two theories of reality coincide to different degrees, varying among individuals and situations.

Comparison of the Operating Principles of the Two Systems

The experiential system in humans is the same system with which other higher order animals have adapted to their environments over millions of years of evolution. It adapts by learning from experience rather than by logical inference, which is the exclusive domain of the rational system. The experiential system operates in a manner that is preconscious, automatic, rapid, effortless, holistic, concrete, associative, primarily nonverbal, and minimally demanding of cognitive resources (see Table 7.1 for a more complete comparison of the two systems). It encodes information in two ways: as memories of individual events, particularly events that were experienced as highly emotionally arousing, and also in a more abstract, general way.

Although the experiential system is a cognitive system, its operation is intimately related to the experience of affect. It is, in fact, inconceivable that a conceptual system that learns from experience would not be used to facilitate positive affect and avoid negative affect. According to CEST, the experiential system both influences and is influenced by affect. Not only does the experiential system direct behavior in a manner anticipated to achieve pleasurable outcomes and to avoid unpleasurable ones, but the cognitions themselves are influenced by affect. As noted previously, the experiential conceptual

**TABLE 7.1 Comparison of the Experiential and Rational Systems**

<table>
<thead>
<tr>
<th>Experiential System</th>
<th>Rational System</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Emotional; pleasure-pain oriented (what feels good).</td>
<td>2. Logical; reason oriented (what is sensible).</td>
</tr>
<tr>
<td>5. Behavior mediated by vibes from past experience.</td>
<td>5. Behavior mediated by conscious appraisal of events.</td>
</tr>
<tr>
<td>7. More rapid processing; oriented toward immediate action.</td>
<td>7. Slower processing; oriented toward delayed action.</td>
</tr>
<tr>
<td>8. Slower to change; changes with repetitive or intense experience.</td>
<td>8. Changes more rapidly; changes with speed of thought.</td>
</tr>
<tr>
<td>9. More crudely differentiated; broad generalization gradient; categorical thinking.</td>
<td>9. More highly differentiated; dimensional thinking.</td>
</tr>
<tr>
<td>11. Experienced passively and preconsciously; seized by our emotions.</td>
<td>11. Experienced actively and consciously; in control of our thoughts.</td>
</tr>
</tbody>
</table>

system, according to CEST, is emotionally driven. After this is recognized, it follows that the affect-free unconscious proposed by cognitive scientists is untenable. The automatic, preconscious experiential conceptual system that regulates everyday behavior is of necessity an emotionally driven, dynamic unconscious system. Because affect determines what is attended to and what is reinforced, without affect there would be neither schemas nor motivation in the experiential system, and, therefore, no experiential system. It follows that CEST is as much an emotional as a cognitive theory.

In contrast to the experiential system, the rational system is an inferential system that operates according to a person’s understanding of the rules of reasoning and of evidence, which are mainly culturally transmitted. The rational system, unlike the experiential system, has a very brief evolutionary history. It operates in a manner that is conscious, analytical, effortful, relatively slow, affect-free, and highly demanding of cognitive resources (see Table 7.1).

Which system is superior? At first thought, it might seem that it must be the rational system. After all, the rational system, with its use of language, is a much more recent evolutionary development than is the experiential system, and it is unique to the human species. Moreover, it is capable of much higher levels of abstraction and complexity than is the experiential system, and it makes possible planning, long-term delay of gratification, complex generalization and discrimination, and comprehension of cause-and-effect relations. These attributes of the rational system have been the source of humankind’s remarkable scientific and technological achievements. Moreover, the rational system can understand the operation of the experiential system, whereas the reverse is not true.

On the other side of the coin, carefully consider the following question: If you could have only one system, which would you choose? Without question, the only reasonable choice is the experiential system. You could exist with an experiential system without a rational system, as the existence of nonhuman animals testifies, but you could not exist with only a rational system. Even mundane activities such as crossing a street would be excessively burdensome if you had to rely exclusively on conscious reasoning. Imagine having to estimate your walking speed relative to that of approaching vehicles so that you could determine when to cross a street. Moreover, without a system guided by affect, you might not even be able to decide whether you should cross the street. Given enough alternative activities to consider, you might remain lost in contemplation at the curb forever.

The experiential system also has other virtues, including the ability to solve some kinds of problems that the rational system cannot. For example, by reacting holistically, the experiential system can respond adaptively to real-life problems that are too complex to be analyzed into their components. Also, there are important lessons in living that can be learned directly from experience and that elude articulation and logical analysis. Moreover, as our research has demonstrated, the experiential system is more strongly associated with the ability to establish rewarding interpersonal relationships, with creativity, and with empathy than is the rational system (Norris & Epstein, 2000b). Most important is that the experiential system has demonstrated its adaptive value over millions of years of evolution, whereas the rational system has yet to prove itself and may yet be the source of the destruction of the human species as well as all other life on earth.

Fortunately, there is no need to choose between the systems. Each has its advantages and disadvantages, and the advantages of one can offset the limitations of the other. Besides, we have no choice in the matter. We are they, and they are us. Where we do have a choice is in improving our ability to use each and to use them in a complementary manner. As much as we might wish to suppress the experiential system in order to be rational, it is no more possible to accomplish this than to stop breathing because the air is polluted. Rather than achieving control by denying the experiential system, we lose control when we attempt to do so: By being unaware of its operation, we are unable to take its influence into account. When we are in touch with the processing of the experiential system, we can consciously decide whether to heed or discount its influence. Moreover, if, in addition, we understand its operation, we can begin to take steps to improve it by providing it with corrective experiences.

How the Experiential System Operates

As noted, the operation of the experiential system is intimately associated with the experience of affect. For want of a better word, I shall use the word vibes to refer to vague feelings that may exist only dimly (if at all) in a person’s consciousness. Stating that vibes often operate outside of awareness is not meant to imply that people cannot become aware of them. Vibes are a subset of feelings, which include other feelings that are more easily articulated than vibes, such as those that accompany standard emotions. Examples of negative vibes are vague feelings of agitation, irritation, tension, disquietude, queasiness, edginess, and apprehension. Examples of positive vibes are vague feelings of well-being, gratification, positive anticipation, calmness, and light-heartedness.

When a person responds to an emotionally significant event, the sequence of reactions is as follows: The experiential system automatically and instantaneously searches its memory banks for related events. The recalled memories and
feelings influence the course of further processing and of behavioral tendencies. If the recalled feelings are positive, the person automatically thinks and has tendencies to act in ways anticipated to reproduce the feelings. If the recalled feelings are negative, the person automatically thinks and has tendencies to act in ways anticipated to avoid experiencing the feelings. As this sequence of events occurs instantaneously and automatically, people are normally unaware of its operation. Seeking to understand their behavior, they usually succeed in finding an acceptable explanation. Insofar as they can manage it without too seriously violating reality considerations, they will also find the most emotionally satisfying explanation possible. This process of finding an explanation in the rational system for what was determined primarily by the experiential system and doing so in a manner that is emotionally acceptable corresponds to what is normally referred to as rationalization. According to CEST, such rationalization is a routine process that occurs far more often than is generally recognized. Accordingly, the influences of the experiential system on the rational system and its subsequent rationalization are regarded, in CEST, as major sources of human irrationality.

The Four Basic Needs

Almost all of the major theories of personality propose a single, most basic need. CEST considers the four most often proposed needs as equally basic. It is further assumed in CEST that their interaction plays an important role in behavior and can account for paradoxical reactions that have eluded explanation by other theoretical formulations.

Identification of the Four Basic Needs

In classical Freudian theory, before the introduction of the death instinct, the one most basic need was the pleasure principle, which refers to the desire to maximize pleasure and minimize pain (Freud, 1900/1953). Most learning theorists make a similar implicit assumption in their view of what constitutes reinforcement (e.g., Dollard & Miller, 1950). For other theorists, such as object-relations theorists, most notably Bowlby (1988), the most fundamental need is the need for relatedness. For Rogers (1951) and other phenomenological psychologists, it is the need to maintain the stability and coherence of a person’s conceptual system. For Allport (1961) and Kohut (1971), it is the need to enhance self-esteem. (For a more thorough discussion of these views, see Epstein, 1993.) Which of these views is correct? From the perspective of CEST, they are all correct, because each of the needs is basic—but they are also all incorrect because of their failure to recognize that the other needs are equally fundamental. They are equally fundamental in the sense that each can dominate the others. Moreover, there are equally serious consequences, including disorganization of the entire personality structure, when any one of the needs is insufficiently fulfilled.

Interactions Among the Basic Needs

Given four equally important needs that can operate simultaneously, it follows that behavior is determined by the combined influence of those needs that are activated in a particular situation. An important adaptive consequence of such influence is that the needs serve as checks and balances against each other. When any need is fulfilled at the expense of the others, the intensity of the others increases, thereby increasing the motivation to satisfy the other needs. However, under certain circumstances the frustration of a need may be so great that frustration of the other needs is disregarded, which can have serious maladaptive consequences. As is shown next, these assumptions about the interaction of basic needs can resolve some important, otherwise paradoxical findings.

The finding that normal people characteristically have unrealistic self-enhancing and optimistic biases (Taylor & Brown, 1988) has evoked considerable interest because it appears to contradict the widely held assumption that reality awareness is an important criterion of mental health. From the perspective of CEST, this finding does not indicate that reality awareness is a false criterion of mental health, but only that it is not the only criterion. According to CEST, a compromise occurs between the need to realistically assimilate the data of reality into a stable, coherent conceptual system and the need to enhance self-esteem. The result is a modest self-enhancing bias that is not unduly unrealistic. It suggests that normal individuals tend to give themselves the benefit of the doubt in situations in which the cost of slight inaccuracy is outweighed by the gain in positive feelings about the self. Note that this assumes that the basic need for a favorable pleasure-pain balance is also involved in the compromise.

There are more and less effective ways of balancing basic needs. A balance that is achieved among equally unfulfilled competing needs is a prescription for chronic distress—not good adjustment. Whereas poorly adjusted people tend to fulfill their basic needs in a conflictual manner, well-adjusted people fulfill their basic needs in a synergistic manner, in which the fulfillment of one need contributes to rather than conflicts with the fulfillment of the other needs. They thereby maintain a stable conceptual system, a favorable pleasure-pain balance, rewarding interpersonal relationships, and a high level of self-esteem.

Let us first consider an example of a person who balances her basic needs in a synergistic manner and then consider an opposite example. Mary is an emotionally stable, happy person with high self-esteem who establishes warm, rewarding
relationships with others. She derives pleasure from helping others. This contributes to her self-esteem, as she is proud of her helpful behavior and others admire and appreciate her for it. As a result, Mary’s behavior also contributes to favorable relationships with others. Thus, Mary satisfies all her basic needs in a harmonious manner.

Now, consider a person who fulfills his basic needs in a conflictual manner. Ralph is an unhappy, unstable person with low self-esteem who establishes poor relationships with others. Because of his low self-esteem, Ralph derives pleasure from defeating others and behaving in other ways that make him feel momentarily superior. Not surprisingly, this alienates people, so he has no close friends. Because of his low self-esteem and poor relationships with others, he anticipates rejection, from which he protects himself by maintaining a distance from people. His low self-esteem and poor relationships with others contribute to feelings of being unworthy of love as well as to an unfavorable pleasure-pain balance. Because his conceptual system is failing to fulfill its function of directing his behavior in a manner that fulfills his basic needs, it is under the stress of potential disorganization, which he experiences in the form of anxiety. The more his need for enhancing his self-esteem is thwarted, the more he acts in a self-aggrandizing manner, which exacerbates his problems with respect to fulfilling his other basic needs.

By viewing their problems in living as resulting from persecution by others, paranoid people with delusions of persecution can focus all their attention and resources on defending themselves. Such focus and mobilization provide a highly unifying state that serve as an effective defense against disorganization. Delusions of persecution also contribute to self-esteem because the perception of the persecutors as powerful or prestigious, which is invariably the case, implies that the target of the persecution must also be important. The basic needs that are sacrificed are the pleasure principle, as being persecuted is a terrifying experience, and the need for relatedness, as others are either viewed as enemies or repelled by the unrealistic behavior.

Schizophrenic disorganization can be understood as the best bargain available for preventing extreme misery under desperate circumstances in which fulfillment of the basic needs is seriously threatened. Ultimate disorganization is a state devoid of conceptualization and (relatedly) therefore of feelings. Although its anticipation is dreaded, its occurrence corresponds to a state of nonbeing, a void in which there are neither pleasant nor unpleasant feelings (Jefferson, 1974). Thus, what is gained is a net improvement in the pleasure-pain balance (from a negative to a zero value). What is sacrificed are the needs to maintain the stability of the conceptual system, to maintain relatedness, and to enhance self-esteem.

**Imbalances in the Basic Needs as Related to Specific Psychopathologies**

Specific imbalances among the basic needs are associated with specific mental disorders. For present purposes, it will suffice to present some of the more obvious examples.

Paranoia with delusions of grandeur can be understood as a compensatory reaction to threats to self-esteem. In a desperate attempt to buoy up self-esteem, paranoid individuals disregard their other needs. They sacrifice their need to maintain a favorable pleasure-pain balance because their desperate need to maintain their elevated self-esteem is continuously threatened. They sacrifice their need to maintain relationships because their grandiose behavior alienates others who do not appreciate being treated as inferiors and who are repelled by their unrealistic views. The situation is somewhat more complicated with respect to their need to realistically assimilate the data of reality into a coherent, stable, conceptual system. They sacrifice the reality aspect of this need but not the coherence aspect. In both of these respects they are similar to paranoid individuals with delusions of persecution, considered in the next example.

Paranoia with delusions of persecution can be understood as a desperate attempt to defend the stability of a person’s conceptual system and, to a lesser extent, to enhance self-esteem.

**The Four Basic Beliefs**

The four basic needs give rise to four corresponding basic beliefs, which are among the most central constructs in a personal theory of reality. They therefore play a very important role in determining how people think, feel, and behave in the world. Moreover, as previously noted, because of their dominant and central position and their influence on an entire network of lower-order beliefs, should any of them be invalidated, the entire conceptual system would be destabilized. Anticipation of such disorganization would be accompanied by overwhelming anxiety. The disorganization, should it occur (as previously noted) would correspond to an acute schizophrenic reaction.

The question may be raised as to how the four basic needs give rise to the development of four basic beliefs. Needs, or motives, in the experiential system, unlike those in the rational systems, always include an affective component. They therefore determine what is important to a person at the experiential level and what a person is spontaneously motivated to pursue or avoid. Positive affect is experienced whenever a need is fulfilled, and negative affect is experienced whenever the fulfillment of a need is frustrated. Because people wish to experience positive affect and to avoid negative affect, they automatically attend to whatever is associated with the
fulfillment or frustration of a basic need. As a result, they develop implicit beliefs associated with each of the basic needs. Let us examine this idea in greater detail.

Depending on a person’s history in fulfilling the need to maximize pleasure and minimize pain, a person tends to develop a basic belief about the world along a dimension varying from benign to malevolent. Thus, if a person experienced an environment that was predominantly a source of pleasure and security, the person will most likely develop the basic belief that the world is a good place in which to live. If a person has the opposite experiences, the person will tend to develop the opposite basic belief. The basic belief about the benignity versus malevolence of the world is the core of a network of related beliefs, including optimistic versus pessimistic views about future events.

Corresponding to the basic need to represent the data of reality in a stable and coherent conceptual system is a basic belief about the world that varies along a dimension of meaningful versus meaningless. Included in the network of related beliefs are beliefs about the predictability, controllability, and justness of the world versus its unpredictability, uncontrollability, and lack of justice. Corresponding to the basic need for relatedness is a basic belief about people that varies along a dimension from helpful and trustworthy to dangerous and untrustworthy. Included in the network of related beliefs are beliefs about the degree to which people are loving versus rejecting and trustworthy versus untrustworthy. Corresponding to the basic need for self-enhancement is a basic belief about the self that varies along a dimension from worthy to unworthy. Included in the network of related beliefs are beliefs about how competent, moral, worthy of love, and strong the self is compared to how incompetent, immoral, unworthy of love, and weak it is.

**Interaction of the Experiential and Rational Systems**

As previously noted, according to CEST, the experiential and rational systems operate in parallel and are interactive.

**The Influence of the Experiential System on the Rational System**

As the experiential system is the more rapidly reacting system, it is able to bias subsequent processing in the rational system. Because it operates automatically and preconsciously, its influence normally occurs outside of awareness. As noted previously, this prompts people to search for an explanation in their conscious rational system, which often results in rationalization. Thus, even when people believe their thinking is completely rational, it is often biased by their experiential processing.

The biases that influence conscious, rational thinking in everyday life are, for the most part, adaptive, as the experiential system operates according to schemas learned from past experience. In some situations, however, the experientially determined biases and their subsequent rationalizations are highly maladaptive. An extreme case is the life-long pursuit of “false goals.” Such goals are false in the sense that their achievement is followed by disappointment and sadness, rather than by the anticipated happiness, enhanced self-esteem, or security that was the reason for their pursuit. It is noteworthy that the achievement of a false goal is experientially disappointing although at the rational level, it is viewed as a significant achievement about which the individual is proud. The following passage from Tolstoi (1887), in which he describes his thoughts during a period of depression, provides a poignant example of such a reaction:

When I thought of the fame which my works had gained me, I used to say to myself, ‘Well, what if I should be more famous than Gogol, Pushkin, Shakespeare, Moliere—than all the writers of the world—well, and what then? I could find no reply. Such questions demand an answer, and an immediate one; without one it is impossible to live, but answer there was none.

My life had come to a sudden stop. I was able to breathe, to eat, to drink, to sleep. I could not, indeed, help doing so; but there was no real life in me. I had not a single wish to strive for the fulfillment of what I could feel to be reasonable. If I wished for something, I knew beforehand, that were I to satisfy the wish, nothing would come of it, I should still be dissatisfied.

Such was the condition I had come to, at the time when all the circumstances of my life were preeminently happy ones, and when I had not yet reached my fiftieth year. I had a good, a loving, and a well-beloved wife, good children, a fine estate, which, without much trouble on my part, continually increased my income; I was more than ever respected by my friends and acquaintances; I was praised by strangers, and could lay claim to having made my name famous . . .

The mental state in which I then was seemed to me summed up into the following: my life was a foolish and wicked joke played on me by I knew not whom . . .

Had I simply come to know that life has no meaning, I could have quietly accepted it as my allotted position. I could not, however, remain thus unmoved. Had I been like a man in a wood, about which he knows that there is no issue, I could have lived on; but I was like a man lost in a wood, and, who, terrified by the thought, rushes about trying to find a way out, and though he knows each step can only lead him farther astray, can not help running backwards and forwards.

Two features of Tolstoi’s situation are of particular interest. One is that he experiences deep despair after achieving his life goals. This suggests that his achievements, although viewed
as successes in his rational system, failed to fulfill a basic need or needs in his experiential system. His success, therefore, can be said to be success at the rational level but failure at the experiential level. This raises the question of what the deeply frustrated need in his experiential system might be. In the absence of additional information, it is, of course, impossible to know, and one can only speculate. One possibility within the framework of CEST is that the frustrated need was for unconditional love in early childhood. Such a need, of course, cannot be satisfied by material rewards or accomplishments.

The other interesting observation is that Tolstoi is distressed not only because of his feelings of emptiness and meaninglessness, but that, try as he might, he cannot solve the problem of why he should be unhappy when all the conditions of his life suggest that he should be happy. It follows from CEST that the reason he cannot solve his problem, despite his considerable intelligence and motivation, is that he believes it exists in his rational system when in fact it exists in his experiential system. Moreover, assuming the speculation about frustration of unconditional love in childhood is true, its early, preverbal occurrence and its remotesness from the kinds of motives normally present in the rational systems of adults can help account for Tolstoi’s inability to articulate the source of his distress.

The influence of the experiential system on the rational system can be positive as well as negative. As an associative system, the experiential system can be a source of creativity by suggesting ideas that would not otherwise be available to the linear-processing rational system. Because the experiential system is a learning system, it can be a source of useful information, which can be incorporated into the rational system. Most important is that the experiential system can provide a source of passion for the rational system that it would otherwise lack. The result is that intellectual pursuits can be pursued with heart, rather than as dispassionate intellectual exercises.

The Influence of the Rational System on the Experiential System

As the slower system, the rational system is in a position to correct the experiential system. It is common for people to reflect on their spontaneous, impulsive thoughts, recognize they are inappropriate, and then substitute more constructive ones. For example, in a flash of anger an employee may have the thought that he would like to tell off his boss, but on further reflection may decide this course of action would be most unwise. To investigate this process, we conducted an experiment in which people were asked to list the first three thoughts that came to mind in response to reading a variety of provocative situations. The first thought was often counterproductive and in the mode of the experiential system, whereas the third thought was usually corrective and in the mode of the rational system.

The rational system can also influence the experiential system by providing the understanding that allows a person to train the experiential system so that its initial reactions are more appropriate. That is, by understanding the operating principles of the experiential system as well as its schemas, it is possible to determine how that system can be improved; this can be accomplished in a variety of ways, the most obvious of which is by disputing the maladaptive thoughts in the experiential system, a procedure widely utilized by cognitive therapists. As the experiential system learns directly from experience, another procedure is to provide real-life corrective experiences. A third procedure is to utilize imagery, fantasy, and narratives for providing corrective experiences vicariously.

The rational system can influence the experiential system in automatic, unintentional ways as well as by its intentional employment. As the experiential system operates in an associative manner, thoughts in the rational system can trigger associations and thereby emotions in the experiential system. For example, a student attempting to solve a mathematics word problem may react to the content with conscious thoughts that produce associations in the experiential system; the associations then elicit emotional reactions that interfere with performance. In this illustration, we have an interesting cycle of the rational system’s influencing the experiential system, which in turn influences the rational system.

Another unintentional way in which the rational system can influence the experiential system is through repetition of thoughts or behavior in the rational system. Through such repetition, thoughts and behavior that were originally under rational control can become habitualized or proceduralized, with the control shifting from the rational to the experiential system (Smith & DeCoster, 2000). An obvious advantage to this shift in control is that the thought and behavior require fewer cognitive resources and can occur without conscious awareness. Potential disadvantages are that the habitual thoughts and behavior are under reduced volitional control and are more difficult to change. Although this can be desirable for certain constructive thoughts and behaviors, it is problematic when the thoughts and behavior are counterproductive.

The Lower and Higher Reaches of the Experiential System

The experiential system operates at different levels of complexity. Classical conditioning is an example of the operation of the experiential system at its simplest level. In classical
conditioning, a conditioned, neutral stimulus (the CS), such as a tone, precedes an unconditioned stimulus (the UCS), such as food. Over several trials, a connection is formed between the conditioned and unconditioned stimulus, so that the conditioned stimulus evokes a conditioned response (the CR), such as salivation, that originally occurred only to the UCS. This process illustrates the operation of several of the attributes of the experiential system, including associative processing, automatic processing, increased strength of learning over trials, affective influence (e.g., emotional significance of the UCS), and arbitrary outcome-orientation (e.g., reacting to the CS independent of its causal relation to the UCS). The CS is also responded to holistically, as the animal reacts not only to the tone, but to the entire laboratory context.

A more complex operating level of the experiential system is exhibited in heuristic processing. In an article that has had a widespread influence on understanding decisional processes, Tversky and Kahneman (1974) introduced the concept of heuristics, which they defined as cognitive shortcuts that people use naturally in making decisions in conditions of uncertainty. They and other cognitive psychologists have found such processing to be a prevalent source of irrational reactions in a wide variety of situations. For example, people typically report that the protagonists in specially constructed vignettes would become more upset following arbitrary outcomes preceded by acts of commission than by acts of omission, by near than by far misses, by free than by constrained behavior, and by unusual than by usual acts. As they respond as if the protagonist’s behavior were responsible for the arbitrary outcomes, their thinking is heuristic in the sense that it is based on simple associative reasoning rather than on cause-and-effect analysis.

A vast amount of research on heuristic processing (see review in Fiske & Taylor, 1991) has produced results that are highly consistent with the principles of experiential processing. Although the data-driven views on heuristic processing derived from social-cognitive research and the theory-driven views of CEST have much in common, the two approaches differ in three important respects. One is that CEST attributes heuristics to the normal mode of operation of an organized conceptual system, the experiential system, that is contrasted with an alternative organized conceptual system, the rational system. The second is that heuristic processing and the experiential system in CEST are embedded in a global theory of personality. The third is that heuristic processing, according to CEST, has withstood the test of time over millions of years of evolution, and is considered to be primarily adaptive. In contrast to these views, social cognitive psychologists, such as Kahneman and Tversky (1973) and Nisbett and Ross (1980), regard heuristics as individual “cognitive tools” that are employed within a single conceptual system that includes both associative (experiential) and analytical (rational) reasoning. These theorists further regard heuristics as quirks in thinking that although sometimes advantageous are common sources of error in everyday life, and therefore are usually desirable to eliminate. It is of interest in this respect to note how resistant some of these blatantly nonrational ways of processing have been to elimination by training. From the perspective of CEST, given the intrinsically compelling nature of experiential processing and its highly adaptive value in most situations in everyday life, such resilience is to be expected.

Although the experiential system encodes events concretely and holistically, it is nevertheless able to generalize, integrate, and direct behavior in complex ways, some of which very likely involve a contribution by the rational system. It does this through prototypical, metaphorical, symbolic, and narrative representations in conjunction with the use of analogy and metaphor. Representations in the experiential system are also related and generalized through their associations with emotions. It is perhaps through processes such as these that the experiential system is able to make its contributions to empathy, creativity, the establishment of rewarding interpersonal relationships, and the appreciation of art and humor (Norris & Epstein, 2000b).

**PSYCHODYNAMICS**

**Psychodynamics,** as the term is used here, refers to the interactions of implicit motives and of implicit beliefs and their influence on conscious thought and behavior. The influence on conscious thought and behavior is assumed to be mediated primarily by vibes. Two major sources of vibes that are important sources of maladaptive behavior are early-acquired beliefs and needs.

**The Influence of Early-Acquired Beliefs on Maladaptive Behavior**

As you will recall, according to CEST, the implicit beliefs in a person’s experiential system consist primarily of generalizations from emotionally significant past experiences. These affect-laden implicit beliefs correspond to schemas about what the self and other people are like and how one should relate to them. Particularly important sources of such schemas are experiences with mother and father figures and with siblings. The schemas exist in varying degrees of generality. At the broadest level is the basic belief about what people in general are like, as previously discussed. At a more specific level are views about particular categories of people, such as
authority figures, maternal figures, mentors, and peers. Such implicit beliefs, both broader and narrower ones, exert a strong influence on how people relate to others, particularly to those who provide cues that are reminders of the original generalization figures. The influence of the schemas is mediated by the vibes automatically activated in cue-relevant situations.

It is understandable why implicit beliefs that contribute to a person’s happiness and security are maintained. But why should implicit beliefs that appear to contribute only to misery also be maintained? Why do they not extinguish as a result of the negative affect following their retrieval? According to the pleasure principle, they should, of course. They do not because of the influence of the need to maintain the stability of one’s conceptual system (Epstein & Morling, 1995; Hixon & Swann, 1993; Morling & Epstein, 1997; Swann, 1990). Depending on circumstances, the need for stability can override the pleasure principle. But how exactly does this operate? What do people actually do that prevents their maladaptive beliefs acquired in an earlier period from being extinguished when they are exposed to corrective experiences in adulthood?

There are three things people do or fail to do that serve to maintain their maladaptive implicit beliefs. First, they tend to perceive and interpret events in a manner that is consistent with their biasing beliefs. Biased perceptions and interpretations allow individuals to experience events as verifying a belief even when on an objective basis they should be disconfirming it. For example, an offer to help or an expression of concern can be perceived as an attempt to control one, and an expression of love can be viewed as manipulative. Second, people often engage in self-verifying behavior, such as by provoking counterbehavior in others that provides objective confirmation of the initial beliefs. For example, a person who fears rejection in intimate relationships may behave with aggression or withdrawal whenever threatened by relationships advancing toward intimacy. This predictably provokes the other person to react with counteraggression or withdrawal, thereby providing objective evidence confirming the belief that people are rejecting. Third, people fail to recognize the influence of their implicit beliefs and associated vibes on their behavior and conscious thoughts, which prevents them from identifying and correcting their biased interpretations and self-verifying behavior. As a result, they attribute the consequences of their maladaptive behavior to unfavorable circumstances or, more likely, to the behavior of others. In the event that after repeated failed relationships, they should consider the possibility that their own behavior may play a role, they are at a loss to understand in what way this could be true, as they can cite objective evidence to support their biased views. You will recall that an important maxim in CEST is that a failure to recognize the operation of one’s experiential system means that one will be controlled by it.

There is an obvious similarity between the psychoanalytic concept of transference and the view in CEST that people’s relationships are strongly influenced by generalizations from early childhood experiences with significant others. Psychoanalysts have long emphasized the importance of transference relations in psychotherapy. They have observed that their patients, after a period in therapy, react to the analyst as if the analyst were a mother or a father figure. They encourage the development of such transference reactions with the aim of providing a corrective emotional experience. Through the use of this procedure as well as by interpreting the transference, the analyst hopes to eliminate the tendency of the patient to establish similar relationships with others. Although this procedure is understandable from the perspective of CEST, it is fraught with danger, as the patient may become overly dependent on the therapist and the therapist, despite the best of intentions, may provide a destructive rather than a corrective experience. Moreover, working through a transference relationship—even when successful—may not be the most efficient way of treating inappropriate generalizations. Nevertheless, for present purposes, it illustrates how generalizations from early childhood tend to be reproduced in later relationships, including those with therapists, and how appropriate emotional experiences can correct maladaptive generalizations.

Although there are obvious similarities between the concepts of transference in psychoanalysis and of generalization in CEST, there are also important differences. Generalization is a far broader concept, which, unlike transference, is not restricted to the influence of relationships with parents. Rather, it refers to the influence of all significant childhood relationships, including in particular those with siblings as well as with parents. Schemas derived from childhood experiences are emphasized in CEST because later experiences are assimilated by earlier schemas. Also, generalizations acquired from childhood experiences are likely to be poorly articulated (if articulated at all) in the rational system. Their influence, therefore, is likely to continue to be unrecognized into adulthood.

The Influence of Early-Acquired Motives on Maladaptive Behavior

Much of what has been said about implicit beliefs in the experiential system can also be applied to implicit needs. Like implicit beliefs, implicit needs or motives are acquired from emotionally significant experiences. They are also maintained for similar reasons. As previously noted, when people
experience a positive or negative event, they automatically acquire a behavioral tendency or motive to reproduce the experience if it was favorable and to avoid experiencing it if it was unfavorable. The stronger the emotional response and the more often it occurs in the same or similar situations, the greater the strength of the motive. Although this learning procedure is adaptive most of the time, it is maladaptive when past conditions are unrepresentative of present ones. One such condition is when a child has experiences involving the deep thwarting of one or more basic needs. For example, if the need to maintain self-esteem is deeply frustrated in childhood, the child will acquire a sensitivity to threats to self-esteem and a corresponding compulsion to protect himself or herself from such threats in the future. *Sensitivities*, in CEST, refer to areas of particular vulnerability, and *compulsions* refer to rigid, driven behavioral tendencies with the aim of protecting oneself from sensitivities. Such sensitivities and compulsions are considered in CEST to be major sources of maladaptive behavior.

The following case history illustrates the operation of a sensitivity and compulsion. In this and other case histories, names, places, and details are altered to protect the anonymity of the protagonists. Ralph was the oldest child in a family that included three other children. He was extremely bright and far outshone his siblings in academic performance. However, rather than being appreciated for it, he was resented by both his parents and siblings. When he eagerly showed his mother the excellent grades on his report card, she would politely tell him that she was busy at the moment and would like to look at it later, when she had more time. Not infrequently, she would forget to do so. It gradually became evident to Ralph that she was more upset than pleased with his accomplishments, so he stopped informing her about them.

The mother’s behavior can be understood in terms of her own background. She had been deeply resentful, as a child, when her mother expressed admiration for the accomplishments of her brighter sibling and ignored her own accomplishments. Thus, her automatic reaction to cues that reminded her of such experiences was to have unpleasant vibes accompanied by resentful thoughts. Consequently, although she meant to be a good mother to Ralph, her experiential reactions undermined her conscious intent. Being unaware of her underlying experiential reactions, she could not help but react as she did. Moreover, over time she found objective reasons for considering him as her least favored child. Little did she realize that his resentful and reticent attitude toward her and others were reactions to her own behavior toward him. She simply regarded him as a stubborn, difficult child by nature.

As a result of his experience in the family, Ralph developed feelings of being unlovable and unworthy and felt depressed much of the time. As an adult, he devoted his energy to bolstering his self-esteem by working extremely hard at becoming a successful businessman. He succeeded at this to a remarkable extent, becoming wealthy at an early age. Yet despite his success and accumulation of material things that other people admired, happiness eluded him. He continued to feel unlovable and depressed no matter what his possessions were and no matter that he had a wife and children who tried hard to please him. When his wife praised the children for their accomplishments, he became resentful toward her and the children. He spent less and less time with his family and increasingly immersed himself in his business. He also began to accuse his wife and children of not loving him and said that was the reason he was spending so little time with them. In his eyes, he was the victim of rejection, not its perpetrator. The result was that he increasingly alienated his family, which verified for him that they did not love him. He became convinced that his wife would ask him for a divorce, and rather than be openly rejected by her, he asked her for a divorce first. He was sure she would be pleased to oblige, and he was extremely relieved when she protested that she did not want a divorce. She said that she wanted more than anything else for them to work together to improve their relationship. This gave a great boost to Ralph, and he tried to the best of his ability to be a more attentive husband and father. This was no easy task for him, particularly as he had no insight into the role his own behavior played in his distressing relationships with his family. It remains to be seen if he will succeed. From the perspective of CEST, it is doubtful that he will unless he gains insight into the influence of his experiential system.

This case illustrates the development, operation, and consequences of a sensitivity and compulsion. Of further interest is that it illustrates the transference of sensitivities and compulsions across generations. The mother’s sensitivity was to being outshone intellectually, and her compulsion was to get back in some way or other at whomever activated the sensitivity. In this case it was her own son, who provided cues reminiscent of her childhood experiences with her brighter sibling. Lest you blame the mother, consider that her reactions occurred automatically, outside of her awareness, and that she was no less a victim than was Ralph.

Ralph had three related sensitivities: threat to his self-esteem, lack of appreciation for his accomplishments, and rejection by a loved one. His compulsive reaction in response to the first sensitivity was to attempt to increase his self-esteem by becoming an outstanding success in business and thereby gaining the admiration of others. His compulsive reaction to the second sensitivity was again to gain the admiration of others for his success and material possessions. His compulsive reaction to the third sensitivity was to withdraw
from and reject the members of his family before they rejected him. Not surprisingly, his compulsive reactions interfered with rather than facilitated gaining the love he so desperately desired.

**RESEARCH SUPPORT FOR THE CONSTRUCT VALIDITY OF CEST**

Research generated by a variety of dual-process theories other than CEST has produced many findings consistent with the assumptions in CEST (see review in Epstein, 1994, and articles in Chaiken & Tropé, 1999). As a review of this extensive literature is beyond the scope of this chapter, here I confine the discussion to studies my associates and I specifically designed to test assumptions in CEST. Three kinds of research are reviewed: research on the operating principles of the experiential system, research on the interactions within and between the two systems, and research on individual differences in the extent and efficacy in the use of the two systems.

**Research on the Operating Principles of the Experiential System**

For some time, my associates and I have been engaged in a research program for testing the operating principles of the experiential system. One of our approaches consisted of adapting procedures used by Tversky and Kahneman and other cognitive and social-cognitive psychologists to study heuristic, nonanalytical thinking through the use of specially constructed vignettes (for examples of this research by others, see Fiske & Taylor, 1991; Tversky & Kahneman, 1974, 1983; Kahneman, Slovic, & Tversky, 1982).

**Irrational Reactions to Unfavorable Arbitrary Outcomes**

People in everyday life often react to arbitrary, unintended outcomes as if they were intentionally and causally determined. Thus, they view more favorably the proverbial bearer of good than of evil tidings despite knowing full well that the messenger is not responsible for the message. Such behavior is an example of outcome-oriented processing. It is the typical way the experiential system reacts to events—by associating outcomes with the stimuli that precede the outcomes, as in classical conditioning.

As an example of the kinds of vignettes we used, one of them described a situation in which two people, as the result of unanticipated heavy traffic, arrive at an airport 30 minutes after the scheduled departure of their flights. One learns that her flight left on time, and the other learns that her flight just left. Tversky and Kahneman (1983) found that people typically reported that the one who barely missed her flight would be more upset than the other protagonist would be, although from a rational perspective it should not matter at all as both were equally inconvenienced and neither was responsible for the outcome. We modified Tversky and Kahneman’s experiment by having the participants respond from three perspectives: how they believed most people would react; how they themselves would react based on how they have reacted to similar situations in the past, and how a completely logical person would react (Epstein, Lipson, Holstein, & Huh, 1992). The first two perspectives were considered to be mainly under the jurisdiction of the experiential system and the third to be mainly under the jurisdiction of the rational system. In order to control for and examine the influence of each of the perspectives on the effect of subsequent perspectives, we counterbalanced the order of presentation of the perspectives.

The findings supported the following hypotheses: There are two different modes of information processing, experiential and rational; the experiential system is an associative system that automatically relates outcomes to preceding situations and behavior, treating them as if they are causally related, even when the relation is completely arbitrary; the rational system is an analytical system that judges cause-and-effect relations according to logical rules; and the systems are interactive, with each influencing the other. Support for the last hypothesis is of particular interest, as it supports the important assumption in CEST that the prevalence of irrational thinking in humans can be attributed largely to the influence of their automatic, preconscious experiential processing on their conscious analytical thinking.

In research on arbitrary outcomes in which we varied the affective consequences of the outcomes, the results supported the assumption in CEST that the degree of experiential relative to rational influence varies directly with the intensity of the affect that is implicated (Epstein et al., 1992). What we found is that the greater the emotional intensity of the outcomes, the more the responses reflected experiential (vs. rational) processing.

**The Ratio-Bias Phenomenon**

Imagine that you are told that on every trial in which you blindly draw a red jellybean from a bowl containing red and white jellybeans, you will receive two dollars. To make matters more interesting, you are given a choice between drawing from either of two bowls that offer the same 10% odds of drawing a winning bean. One contains one red jellybean and nine white ones; the other contains 10 red jellybeans and 90 white ones. Which bowl would you choose to draw from, and
how much would you pay for the privilege of drawing from the bowl of your choice, rather than having the choice decided by the toss of a coin? When people are simply asked how they would behave, almost all say they would have no preference and would not pay a cent for a choice between two equal probabilities. Yet when they are placed in a real situation, most willingly pay small sums of money for the privilege of drawing from the bowl with more red jellybeans (Kirkpatrick & Epstein, 1992). This difference in response to the verbally presented and the real situation can be explained by the greater influence of the experiential than the rational system in real situations with emotionally significant consequences compared to simulated situations without consequences. According to CEST, the experiential system is particularly reactive to real experience, whereas the rational system is uniquely responsive to abstract, verbal representations.

This jellybean experimental situation, otherwise referred to as the ratio-bias experimental paradigm, is particularly interesting with respect to CEST because it pits experiential against rational processing. The conflict between the two modes of processing arises because the experiential system is a concrete system that is less responsive to abstractions such as ratios than to the numerosity of objects. Comprehension of numerosness, unlike comprehension of ratios, is an extremely fundamental ability that is within the capacity of 3-year-old children and nonhuman animals (Gallistel & Gelman, 1992).

Even more impressive than the irrational behavior exhibited by people paying for the privilege of choosing between bowls that offer equal probabilities are the results obtained when unequal probabilities are offered by the bowls. If our reasoning is correct, a conflict between the two systems can be established by having one bowl probability-advantaged and the other numerosity-advantaged. In one study, the probability-advantaged bowl always contained 1 in 10 red jellybeans, whereas the numerosity-advantaged bowl offered between 5 and 9 red jellybeans out of 100 jellybeans, depending on the trial (Denes-Raj & Epstein, 1994). Under these circumstances, many adults made nonoptimal responses by selecting the numerosity-advantaged bowl against the better judgment of their rational thinking. For example, they often chose to draw from the bowl that contained 8 of 100 (8%) in preference to the one that contained 1 of 10 (10%) red jellybeans. Some sheepishly commented that they knew it was foolish to go against the probabilities, but somehow they felt they had a better chance of drawing a red jellybean when there were more of them. Of additional interest, participants made nonoptimal responses only to a limited degree, thereby suggesting a compromise between the two systems. Thus, although many selected a numerosity-advantaged 8% option (8 of 100 red jellybeans) over a 10% probability-advantaged one (1 of 10 red jellybeans), almost no one selected a 5% numerosity-advantaged option (5 of 100 red jellybeans) over a 10% probability-advantaged option (1 of 10 red jellybeans). Apparently, most people preferred to behave according to their experiential processing only up to a point of violating their rational understanding. To be sure, there were participants who always responded rationally. What was impressive about the study, however, was the greater number who responded irrationally despite knowing better (in their rational systems).

To determine whether children who have not had formal training in ratios have an intuitive understanding of ratios, we conducted a series of studies in which we examined children’s responses to the ratio-bias experimental paradigm (Yanko & Epstein, 2000). We were also interested in these studies in determining whether children who have only an intuitive understanding of ratios exhibit compromises between the two systems. We found that children without formal knowledge of ratios had only a rudimentary comprehension of ratios. They responded appropriately to differences between ratios only when the magnitude of the differences was large. Like adults, children exhibited compromises, but their compromises were more in the experiential direction. For example, many children but no adults selected a 5% numerosity-advantaged bowl over a 10% probability-advantaged one. However, very few of the same children selected a 2% numerosity-advantaged bowl over a 10% probability-advantaged one.

We also used the ratio-bias experimental paradigm to test the assumption in CEST that the experiential system responds to visual imagery in a way similar to the way it does to real experience (Epstein & Pacini, 2001). We presented participants in an experimental group with a verbal description of the ratio-bias experimental paradigm after training them to vividly visualize the situation. Participants in the control group were given only the verbal description. In support of the assumption, the visual-imaging group but not the control group exhibited the ratio-bias phenomenon in a manner similar to what we have repeatedly found in real situations but not in simulated situations.

The overall results from the many studies we conducted with the ratio-bias paradigm (Denes-Raj & Epstein, 1994; Denes-Raj, Epstein, & Cole, 1995; Kirkpatrick & Epstein, 1992; Pacini & Epstein, 1999a, 1999b; Yanko & Epstein, 2000) provided support for the following assumptions and hypotheses derived from CEST. There are two independent information-processing systems. Sometimes they conflict with each other, but more often they form compromises. With
increasing maturation from childhood to adulthood, the balance of influence between the two processing systems shifts in the direction of increased rational dominance. The experiential system is more responsive than is the rational system to imagery and to other concrete representations than the rational system, whereas the rational system is more responsive than is the experiential system to abstract representations. Engaging the rational system in children who do not have formal knowledge of ratios by asking them to give the reasons for their responses interferes with the application of their intuitive understanding of ratios, resulting in a deterioration of performance.

We have also used the ratio-bias phenomenon to elucidate the thinking of people with emotional disorders. In a study of depressed college students (Pacini, Muir, & Epstein, 1998), the ratio-bias phenomenon helped to clarify the paradoxical depressive-realism phenomenon (Alloy & Abramson, 1988). The phenomenon refers to the finding that depressed participants are more rather than less accurate than are nondepressed participants in judging contingencies between events. We found that the depressed participants made more optimal responses than did their nondepressed counterparts only when the stakes for nonoptimal responding were inconsequential. When we raised the stakes, the depressed participants responded more experientially and the control participants responded more rationally, so that the groups converged and no longer differed. We concluded that the depressive-realism phenomenon can be attributed to an overcompensatory reaction by subclinically depressed participants in trivial situations to a more basic tendency to behave unrealistically in emotionally significant situations. We further concluded that normal individuals tend to rely on their less demanding experiential processing when incentives are low, but increasingly engage their more demanding rational processing as incentives are increased.

The Global-Evaluation Heuristic

The global-person-evaluation heuristic refers to the tendency of people to evaluate others holistically as either good or bad people rather than to restrict their judgments to specific behaviors or attributes. Because the global-person-evaluation heuristic is consistent with the assumption that holistic evaluation is a fundamental operating principle of the experiential system (see Table 7.1), it follows that global-person-evaluations tend to be highly compelling and not easily changed. The heuristic is particularly important because of its prevalence and because of the problems that arise from it—such as when jurors are influenced by the attractiveness of a defendant’s appearance or personality in judging his or her guilt. An interesting example of this phenomenon was provided in the hearing of Clarence Thomas for appointment to the United States Supreme Court. The testimony by Anita Hill about the obscene sexual advances she alleged he made to her was discredited in the eyes of several senators because of the favorable testimony by employees and acquaintances about his character and behavior. It seemed inconceivable to the senators that an otherwise good person could be sexually abusive.

We studied the global-person-evaluation heuristic (reported in Epstein, 1994) by having participants respond to a vignette adapted from a study by Miller and Gunasegaram (1990). In the vignette, a rich benefactor tells three friends that if each throws a coin that comes up heads, he will give each $100. The first two throw a heads, but Smith, the third, throws a tails. When asked to rate how each of the protagonists feels, most participants indicated that Smith would feel guilty and the others would feel angry with him. In an alternative version with reduced stakes, the ratings of guilt and anger were correspondingly reduced. When asked if the other two would be willing, as they previously had intended, to invite Smith to join them on a gambling vacation in Las Vegas, where they would share wins and losses, most participants said they would not “because he is a loser.” These responses were made both from the perspective of how the participants reported they themselves would react in a real situation and how they believed most people would react. When responding from the perspective of how a completely logical person would react, most participants said a logical person would recognize that the outcome of the coin tosses was arbitrary, and they therefore would not hold it against Smith. They further indicated that a logical person would invite him on the gambling venture.

This study indicates that people tend to judge others holistically by outcomes, even arbitrary ones. It further indicates that people intuitively recognize that there are two systems of information processing that operate in a manner consistent with the principles of the experiential and rational systems. It also supports the hypotheses that experiential processing becomes increasingly dominant with an increase in emotional involvement and that people overgeneralize broadly in judging others on the basis of outcomes over which the person has no control, even though they know better in their rational system.

Conjunction Problems

The Linda conjunction problem is probably the most researched vignette in the history of psychology. It has evoked a great deal of interest among psychologists because of its
paradoxical results. More specifically, although the solution to the Linda problem requires the application of one of the simplest and most fundamental principles of probability theory, almost everyone—including people sophisticated in statistics—gets it wrong. How is this to be explained? As you might suspect by now, the explanation lies in the operating principles of the experiential system.

Linda is described as a 31-year-old woman who is single, outspoken, and very bright. In college she was a philosophy major who participated in antinuclear demonstrations and was concerned with issues of social justice. How would you rank the following three possibilities: Linda is a feminist, Linda is a bank teller, and Linda is a feminist and a bank teller? If you responded like most people, you ranked Linda as being a feminist and a bank teller ahead of Linda’s being a bank teller alone. In doing so, you made what Tversky & Kahneman (1982) refer to as a conjunction error (CE). It is an error or fallacy because according to the conjunction rule, the occurrence of two events cannot be more likely than the occurrence of only one of them.

The usual explanation of the high rate of CEs that people make is that they either do not know the conjunction rule or they do not think of it in the context of the Linda vignette. They respond instead, according to Tversky and Kahneman, by the representativeness heuristic, according to which being both a bank teller and a feminist is more representative of Linda’s personality than being just a bank teller.

In a series of studies on conjunction problems, including the Linda problem (Donovan & Epstein, 1997; Epstein, Denes-Raj, & Pacini, 1995; Epstein & Donovan, 1995; Epstein, Donovan, & Denes-Raj, 1999; Epstein & Pacini, 1995), we concluded that the major reason for the difficulty of the Linda problem is not an absence of knowledge of the conjunction rule or a failure to think of it. We demonstrated that almost all people have intuitive knowledge of the conjunction rule, as they apply it correctly in natural contexts, such as in problems about lotteries. Nearly all of our participants, whether or not they had formal knowledge of the conjunction rule, reported that winning two lotteries, one with a very low probability of winning and the other with a higher probability, is less likely than is winning either one of them (Epstein et al., 1995). This finding is particularly interesting from the perspective of CEST because it indicates that the experiential system (which knows the conjunction rule intuitively) is sometimes smarter than the rational system (which may not be able to articulate the rule). We also found that when we presented the conjunction rule among other alternatives, thereby circumventing the problem of whether people think of it in the context of the Linda problem, most people selected the wrong rule. They made the rule fit their responses to the Linda problem rather than the reverse, thereby demonstrating the compelling nature of experiential processing and its ability to dominate analytical thinking in certain situations.

The conclusions from our series of studies with the Linda problem can be summarized as follows:

- The difficulty of the Linda problem cannot be fully accounted for by the misleading manner in which it is presented, for even with full disclosure about the nature of the problem and the request to treat it purely as a probability problem, a substantial number of participants makes CEs. Apparently, people tend to view the Linda problem as a personality problem rather than as a probability problem, no matter what they are told.
- The difficulty of the Linda problem can be explained by the rules of operation of the experiential system, which is the mode employed by most people when responding to it. Thus, people tend to reason associatively, concretely, holistically, and in a narrative manner rather than abstractly and analytically when responding to the problem. For example, a number of participants explained their responses that violated the conjunction rule by stating that Linda is more likely to be a bank teller and a feminist than just a feminist because she has to make a living.
- The essence of the difficulty of the Linda problem is that it involves an unnatural, concrete presentation, where an unnatural presentation is defined as one that differs from the context in which a problem is normally presented. We found that concrete presentations facilitate performance in natural situations (in which the two processing systems operate in synchrony) and interfere with performance in unnatural situations (in which the two systems operate in opposition to each other).
- Processing in the experiential mode is intrinsically highly compelling and can override processing in the rational mode even when the latter requires no more effort. Thus, many people, despite knowing and thinking of the conjunction rule, nevertheless prefer a representativeness solution.
- Priming intuitive knowledge in the experiential system can facilitate the solution to problems that people are unable initially to solve intellectually.

Interaction Between the Two Processing Systems

An important assumption in CEST is that the two systems are interactive. Interaction occurs simultaneously as well as sequentially. Simultaneous interaction was demonstrated in the compromises between the two systems observed in the
studies of the ratio-bias phenomenon. Sequential interaction was demonstrated in the study in which people listed their first three thoughts and in the studies of conjunction problems, in which presenting concrete, natural problems before abstract problems facilitated the solution of the abstract problems.

There is also considerable evidence that priming the experiential system subliminally can influence subsequent responses in the rational system (see review in Bargh, 1989). Other evidence indicates that the form independent of the content of processing in the rational system can be influenced by priming the experiential system. When processing in the experiential mode is followed by attempts to respond rationally, the rational mode itself may be compromised by intrusions of experiential reasoning principles (Chaiken & Maheswaren, 1994; Denes-Raj, Epstein, & Cole, 1995; Edwards, 1990; Epstein et al., 1992).

Sequential influence does not occur only in the direction of the experiential system influencing the rational system. As previously noted, in everyday life sequential processing often proceeds in the opposite direction, as when people react to their irrational, automatic thoughts with corrective, rational thoughts. In a study designed to examine this process, we instructed participants to list the first three thoughts that came to mind after imagining themselves in various situations described in vignettes (reported in Epstein, 1994). The first response was usually a maladaptive thought consistent with the associative principle of the experiential system, whereas the third response was usually a more carefully reasoned thought in the mode of the rational system. As an example, consider the responses to the following vignette, which describes a protagonist who fails to win a lottery because she took the advice of a friend rather than follow her own inclination to buy a ticket that had her lucky number on it. Among the most common first thoughts were that the friend was to blame and that the participant would never take her advice again. By the third thought, however, the participants were likely to state that the outcome was due to chance and no one was to blame.

**Interaction Between the Basic Needs**

You will recall that a basic assumption in CEST is that behavior often represents a compromise among multiple basic needs. This process is considered to be particularly important, as it provides a means by which the basic needs serve as checks and balances against each other, with each need constrained by the influence of the other needs. To test the assumption about compromises, we examined the combined influence of the needs for self-enhancement and self-verification. Swann and his associates had previously demonstrated that the needs for enhancement and verification operate sequentially, with the former tending to precede the latter (e.g., Swann, 1990; Hixon & Swann, 1993). We wished to demonstrate that they also operate simultaneously, as manifested by compromises between them. Our procedure consisted of varying the favorableness of evaluative feedback and observing whether participants had a preference for feedback that matched or was more favorable to various degrees than their self-assessments (Epstein & Morling, 1995; Morling & Epstein, 1997). In support of our hypotheses, participants preferred feedback that was only slightly more favorable than their own self-assessments, consistent with a compromise between the need for verification and the need for self-enhancement.

**Research on Individual Differences**

**Individual Differences in the Intelligence of the Experiential System**

If there are two different systems for adapting to the environment, then it is reasonable to suspect that there are individual differences in the efficacy with which people employ each. It is therefore assumed in CEST that each system has its own form of intelligence. The question remains as to how to measure each. The intelligence of the rational system can be measured by intelligence tests, which are fairly good predictors of academic performance. To a somewhat lesser extent, they also predict performance in a wide variety of activities in the real world, including performance in the workplace, particularly in situations that require complex operations (see reviews in Gordon, 1997; Gottfredson, 1997; Hunter, 1983, 1986; Hunter & Hunter, 1984). However, intelligence tests do not measure other kinds of abilities that are equally important for success in living, including motivation, practical intelligence, ego strength, appropriate emotions, social facility, and creativity.

Until recently, there was no measure of the intelligence of the experiential system; one reason for this is that the concept of an experiential system was unknown. Having established its theoretical viability, the next step was to construct a way of measuring it, which resulted in the Constructive Thinking Inventory (CTI; Epstein, 2001). The measurement of experiential intelligence is based on the assumption that experiential intelligence is revealed by the adaptiveness of the thoughts that tend to spontaneously occur in different situations or conditions.

People respond to the CTI by reporting on a 5-point scale the degree to which they have certain common adaptive and maladaptive automatic or spontaneous thoughts. An example of an item is *I spend a lot of time thinking about my mistakes, even if there is nothing I can do about them* (reverse scored).
The CTI provides a Global Constructive Thinking scale and six main scales, most of which have several facets, or subscales. The six main scales are Emotional Coping, Behavioral Coping, Categorical Thinking, Esoteric Thinking, Naive Optimism, and Personal Superstitious Thinking. The main scales all have high internal-consistency reliability coefficients and evidence for their validity in numerous studies. They are predictive of a wide variety of criteria related to success in living. A review of the extensive literature supporting the construct validity of the CTI is beyond the scope of this chapter, but is available elsewhere (Epstein, 2001). For present purposes, it will suffice to note that favorable CTI scores have been found to be significantly associated with performance in the workplace and in the classroom, social competence, leadership ability, ability to cope with stress, emotional adjustment, physical well-being, and an absence of drug and alcohol abuse.

The relation of constructive thinking to intellectual intelligence is of considerable interest for theoretical as well as practical reasons. According to CEST, the experiential and rational systems operate independently, each by its own set of principles (see Table 7.1). One would therefore expect the intelligence or efficacy of the two processing systems to be independent. This is exactly what we have repeatedly found in several studies that have compared scores on the Global CTI scale with measures of intellective intelligence (Epstein, 2001). Of additional interest, constructive thinking and intellective intelligence were found to exhibit opposite courses of development across the life span. Constructive thinking is at its nadir in adolescence, when intellectual intelligence is at its peak, and it gradually increases throughout most of the adult years when intellectual intelligence is gradually declining. Unlike intellectual intelligence, constructive thinking is only negligibly related to academic achievement tests. Yet it adds significant variance in addition to the contribution of intellective intelligence to the prediction of performance in the classroom, as indicated by grades received and class rank (Epstein, 2001). Apparently, good constructive thinkers are able to capitalize on their knowledge and obtain appropriate recognition for their achievements, whereas poor constructive thinkers are more likely to engage in counterproductive behavior such as antagonizing their teachers, resulting in their being downgraded.

Individual Differences in Rational and Experiential Thinking Styles

If people process information by two different systems, the extent to which they employ each should be an important personality variable. To investigate this aspect of personality, we constructed a self-report test, the Rational-Experiential Inventory (REI). The REI has main scales of rational and experiential processing. Each of the main scales has subscales of self-assessed effectiveness and of frequency in use of the thinking style.

The REI scales have internal-consistency reliabilities of .87–.90 for the main scales and .79–.84 for the subscales. There is considerable evidence in support of their construct validity. The major findings from several studies (Epstein et al., 1996; Norris & Epstein, 2000a, 2000b; Pacini & Epstein, 1999b; Pacini, Muir, & Epstein, 1998; Rosenthal & Epstein, 2000) can be summarized as follows:

- In support of the assumption in CEST of independent rational and experiential processing systems, the two main scales are independent.
- In support of the inclusion of the subscales, they exhibit factorial, discriminant, and convergent validity.
- The rational and experiential scales are coherently associated with objective measures of heuristic processing. As expected, the relation of the rational scale with heuristic processing is inverse, and the relation of the experiential scale with heuristic processing is direct.
- Although the rational and experiential main scales are uniquely associated with some variables, they make independent, supplementary contributions to the prediction of other variables. The rational scale is more strongly positively associated than is the experiential scale with intellective performance, as measured by SAT scores and grade point average, and with adjustment, including measures of ego strength and self-esteem, and with measures of openness, conscientiousness, favorable beliefs about the self and the world, and physical well-being. The rational scale is more strongly negatively associated than the experiential scale with measures of neuroticism, depression, anxiety, stress in college life, subtle racism, extreme conservatism, alcohol abuse, and naive optimism. The experiential scale is more strongly positively associated than the rational scale with measures of extroversion, agreeableness, favorable interpersonal relationships, empathy, creativity, emotionality, sense of humor, and art appreciation, and it is more strongly negatively associated than the rational system with distrust and intolerance.

When introducing a new measure, it is important to demonstrate that the measure provides information that is not readily available from existing instruments. In order to determine whether the REI is redundant with more standard personality measures, we conducted a study (Pacini & Epstein, 1999b) in which we compared the REI to the NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1989), the most popular measure of the Big Five personality traits.
The two inventories contributed independent, supplementary variance to the prediction of many of the same variables and unique variance to the prediction of other variables. Moreover, when the five NEO-FFI scales were entered into a regression equation as predictors of the REI scales, they accounted for only 37% of the variance of the Rationality scale and 11% of the variance of the Experientiality scale. This is of interest not only because it demonstrates that the REI is mainly independent of the NEO-FFI, but also because of the information it provides about the NEO-FFI. It suggests that the NEO-FFI mainly measures attributes associated with the rational system and is relatively deficient in measuring attitudes and behavior associated with preconscious, automatic information processing.

Consistent with gender stereotypes, women report significantly greater appreciation of and engagement in experiential processing than men, and men report greater appreciation of and engagement in rational processing. However, the mean gender differences are small, and there is a great deal of overlap between the groups.

Given fundamentally different ways of processing information, it might reasonably be expected that people with different thinking styles would be receptive to different kinds of messages. To test this hypothesis, Rosenthal and Epstein (2000) conducted a study with the REI in which they compared the reactions of women with high scores on rationality and low scores on experientiality with women with the opposite pattern. The groups were subdivided according to whether they received messages on the danger of breast cancer and the importance of self-examination in the form of information designed to appeal to the rational or the experiential mode of information processing. The rational message emphasized actuarial and other objective information, whereas the experiential message included personal appeals and vivid individual cases. The dependent variable was the intention to regularly conduct breast self-examinations. Both groups reported a greater intention to conduct breast examinations when the message they received matched their own thinking style.

**Individual Differences in Basic Beliefs About the Self and the World**

The Basic Beliefs Inventory (BBI; Catlin & Epstein, 1992) is a self-report questionnaire that measures the four basic beliefs proposed in CEST. It includes a global scale of overall favorability of basic beliefs and separate scales for measuring each of the basic beliefs. The internal-consistency reliabilities (coefficients alpha) of the scales are between .77 to .91. The scales are moderately intercorrelated with a median correlation of .42, thereby justifying combining them into an overall scale of favorability of beliefs as well as considering them individually.

You will recall that according to CEST, a person’s basic beliefs are primarily derived from emotionally significant personal experiences. To test this hypothesis, Catlin and Epstein (1992) examined the relations of scores on the BBI and self-reports of two kinds of highly significant emotional experiences. The two kinds of experiences were extreme life events, such as loss of a loved one, and the quality of relationships with parents during early childhood. In support of this hypothesis, both kinds of experiences were significantly and coherently related to basic beliefs. Often, the two kinds of experience made independent, supplementary contributions to the prediction of the same basic belief. Of additional interest, the self-reported quality of childhood relationships with parents moderated the influence of extreme life events on basic beliefs.

**Summary and Conclusions Regarding Research Support for CEST**

In summary, the program of research on CEST has provided impressive support for its construct validity. The following basic assumptions of CEST have all been verified: There are two independent information-processing systems that operate in parallel by different rules. The systems are interactive, with each influencing the other, and the interaction occurs both sequentially and simultaneously. The influence of experiential processing on rational processing is of particular importance, as it identifies a process by which people’s automatic, preconscious, experiential processing routinely biases their conscious rational thinking. The experiential system is an associative, rapid, concretist, primarily nonverbal system that is intrinsically highly compelling to the extent that it can override the rational system, leading people to “behave against their better judgment.”

When people are aware of the maladaptive thoughts generated by their automatic experiential processing, they often correct the thoughts through more deliberative reasoning in their rational systems. There are reliable individual differences in the efficacy or intelligence of the experiential system. The intelligence of the experiential system is dependent on the intelligence of the rational system and is measured by a variety of indexes of success in living than is the intelligence of the rational system. Included are work success, social facility, absence of drug and alcohol abuse, and mental and physical well-being. There are reliable individual differences in experiential and rational thinking styles. The two thinking styles exhibit coherent patterns of relations with a variety of criterion variables. There are also reliable
individual differences in the four basic beliefs proposed by CEST. As the basic beliefs influence behavior simultaneously in the form of compromises, they serve as checks and balances against each other.

**IMPLICATIONS OF COGNITIVE-EXPERIENTIAL SELF-THEORY FOR PSYCHOTHERAPY AND RESEARCH**

**Implications for Psychotherapy**

For psychotherapy to be effective, it is necessary according to CEST for changes to occur in the experiential system. This is not meant to imply that changes in the rational system are of no importance, but rather to suggest that changes in the rational system are therapeutic only to the extent that they facilitate changes in the experiential system.

There are three basic ways of producing changes in the experiential system. These include the use of the rational system to correct and train the experiential system, the provision of emotionally significant corrective experiences, and communicating with the experiential system in its own medium—namely, fantasy, imagery, metaphor, concrete representations, and narratives. These three approaches provide a unifying framework for a wide variety of approaches in psychotherapy, including insight approaches, cognitive-behavioral approaches, and experiential approaches, including gestalt therapy and psychosynthesis (Epstein, 1994, 1998).

**Using the Rational System to Correct the Experiential System**

The rational system has an important advantage over the experiential system in that it can understand the experiential system, whereas the reverse is not true. Thus, one way the rational system can be used to improve the functioning of the experiential system is by teaching people to understand the operation of their experiential systems. Almost everyone is aware of having conflicts between the heart and the head as well as having unbidden distressing thoughts that they cannot consciously control. These are not deep, dark, inaccessible thoughts, but rather ones of which people are acutely aware. Beginning with a discussion of such reactions, it should not be difficult to convince people that they operate by two independent systems. The next step is to teach them about the operating principles of the experiential system and the manner in which it influences their behavior and biases their conscious thought. They then can be helped to understand that their problems are almost always in their automatic experiential processing, not in their conscious thinking. Not only is such knowledge useful for correcting and training the experiential system, but it also provides a useful foundation for the other two approaches.

One of the important advantages of clients’ recognizing that their problems lie primarily in their experiential and not their rational systems is that it reduces resistance and defensiveness because they no longer have to defend the reasonableness of their behavior. For example, if a client engages in excessive rational discourse and feels compelled to defend his or her behavior as reasonable, the therapist can remind the client that the experiential system does not operate by logic. Rather, what is important is to uncover the maladaptive beliefs and needs in the experiential system and ultimately change them in a constructive way.

Uncovering implicit beliefs in the experiential system can be accomplished in several ways. One way is by noting repetitive behavior patterns, and in particular becoming aware of sensitivities, compulsions, and ego-alien behavior, and becoming aware of the situations in which they arise. A second way is by using fantasy to vicariously explore reactions to different situations. A third way is by attending to emotional reactions, vibes, and the kinds of automatic thoughts that instigate them.

Emotional reactions are particularly revealing according to CEST because they provide a royal road to the important schemas in people’s implicit theories of reality. They do this in two ways. First, whenever an event elicits a strong emotional response, it indicates that a significant schema in a person’s implicit theory of reality has been implicated. Accordingly, by noting the events that elicit emotional responses, some of the more important schemas in a person’s theory of reality can be determined. Second, emotions can be used to infer schemas through knowledge of the relation between specific thoughts and specific emotions (e.g., Averill, 1980; Beck, 1976; Ellis, 1973; Epstein, 1983, 1984; Lazarus, 1991). This relation has been well documented by the clinical observations of cognitive-behavioral therapists (e.g., Beck, 1976; Ellis, 1973) and by research that has examined the relation of thoughts and emotions in everyday life (e.g., Averill, 1980; Epstein, 1983). It follows from the relation of automatic thoughts to emotions that people who characteristically have certain emotions characteristically spontaneously think in certain ways. For example, angry people can be assumed to have the implicit belief that people often behave badly and deserve to be punished, frightened people can be assumed to have the implicit belief that the world is dangerous and they should be prepared for flight, and sad people can be assumed to have the implicit belief that they have sustained an irreplaceable loss, or that they are inadequate, bad, or unlovable-they consider themselves, and there is nothing they can do about it.
As its name implies, the essence of the experiential system is that it is a system that learns from experience. It follows that the most direct route for changing maladaptive schemas in the experiential system is to provide corrective experiences. One way to accomplish this is through the relationship between client and therapist. This procedure is particularly emphasized in psychoanalytic transference relationships. Another way to learn directly from experience is by having corrective emotionally significant experiences in everyday life. As previously noted, it can be very useful in this respect for clients to gain insight into their biasing interpretations and self-verifying behavior. In the absence of such insight, potentially corrective experiences can be misinterpreted in a way that makes them contribute to the reinforcement rather than extinction of their destructive thoughts and behavior patterns. Having emphasized the contribution of insight, a caveat is in order concerning valuing it too highly and considering it a necessary condition for improvement. Although insight can be very useful, it is not a necessary condition for improvement. It is quite possible for change to occur in the experiential system in the absence of intellectual understanding of the process, which, of course, is the way nonhuman animals as well as people who are not in therapy normally learn from experience. Many a novel has been written about cures through love. In fact, for clients who are nonintellectual, corrective experiences in the absence of insight may be the only way to proceed in therapy. In the absence of recognizing the limited value of intellectual insight, there is the danger that therapists will insufficiently attend to the experiential aspects of therapy.

**Communicating with the Experiential System in Its Own Medium**

Communicating with the experiential system in its own medium refers to the use of association, metaphor, imagery, fantasy, and narrative. Within the scope of this chapter, it is impossible to discuss all of these procedures or even to discuss any in detail. It is important to recognize in this regard that there is no single kind of therapy that is specific to CEST. Rather, CEST is an integrative personality theory that provides a framework for placing into broad perspective a variety of therapies. For present purposes, it will suffice to present both a simple and a more complicated example of how communication with the experiential system in its own medium can be used therapeutically.

The simple example concerns a person who under the guidance of a therapist visualizes a situation to learn how he might react to the situation in real life. The procedure is based on the assumption that the experiential system reacts to visualized events in a similar way as to real events, an assumption supported by research expressly designed to test it (Epstein & Pacini, 2001).

Robert exhibited a life pattern of ambivalence about getting married. Recently, the woman he had been dating for several years gave him an ultimatum. She demanded that either he pronounce his intention to marry her or she would leave him. Robert loved her dearly, but he did not feel ready for marriage. He had always assumed he would settle down and raise a family, but somehow whenever he came to the
point of committing himself, something went wrong with the relationship, and he and his partner parted ways. At first, Robert attributed the partings to failings in his partners, but after repeated reenactments, it occurred to him that he might be ambivalent about marriage. Because this made no sense to him, he decided to seek the help of a therapist. The therapist instructed and trained Robert to vividly imagine being married and coming home to his wife and children after work. When he had the scene clearly in mind, he was asked to carefully attend to his feelings. To his surprise, he felt irritated and burdened when his wife greeted him at the door and the children eagerly began relating the events of the day. The therapist then instructed Robert to imagine another scene in which he had the very same feelings. His mind turned to his childhood, and he had an image of taking care of his younger siblings when his parents went out for entertainment. He deeply resented having to take care of them frequently and not being able to play with his peers. The result was that he learned to dislike interacting with children at the experiential level, but had never articulated this feeling at the rational level.

As an adult, although Robert believed in his conscious, rational mind that he wanted to get married and raise a family, in his experiential mind, the thought of being in the company of children produced unpleasant vibes. He and his therapist discussed whether he should follow his heart or his mind. In order to help him to decide, the therapist pointed out that following his heart would be the path of least resistance. He added that if Robert decided to follow his mind, it would be important for him to work on overcoming his negative feelings toward children. When Robert decided that is what he wanted to do, he was given an exercise to practice in fantasy that consisted of scenes in which Robert engaged in enjoyable activities with children. He was also encouraged to visualize whatever occasions he could remember from his childhood in which he enjoyed being with his siblings. He was given other scenes to imagine, including being pleased with himself for behaving as a better parent to his imaginary children than his parents had behaved to him.

The more complex example is taken from a book by Alice Epstein (1989) in which she described her use of fantasy and other procedures designed to communicate with her experiential system. She attributed a surprisingly rapid reorganization of her personality to this procedure. She also believed that the change in her feelings that accompanied the change in her personality contributed to a dramatic recovery from a life-threatening illness against all odds.

Alice began psychotherapy after receiving a diagnosis of terminal cancer and being informed that she would not likely live more than three months. The statistics at that time of her diagnosis on the outcome of a metastasized hypernephroma, the form of kidney cancer that she had, indicated that no more than 4 in 1000 cases experienced remission from the disease, let alone cure. Now, many years after that diagnosis, Alice has no detectable signs of cancer and has been considered cured for more than 15 years. Whether her belief that the psychotherapy actually saved her life is correct is not at issue here. What is of primary interest is the rapid resolution of deep-seated problems through the use of fantasy that usually require a prolonged period of intensive psychotherapy. However, given increasing evidence of the relation of emotions to the immune system, it would be unwise to summarily reject her belief that her psychological recovery contributed to her physical recovery. It is possible that the experiential system has a relation to physical well-being much stronger than orthodox medicine recognizes.

The following is one of the early fantasies described by Alice in her book: In the session preceding the fantasy, she had expressed hostility toward her mother for her mother’s behavior to her during a period of extended turmoil in the household. During that period, the mother surprisingly gave birth to Alice’s younger sister after denying being pregnant and attributing the change in her appearance to a gain in weight from eating too much. During the same period, the mother’s mother, who shared the household with the family, and to whom the mother was deeply attached, was dying of cancer. After the session in which Alice Epstein (1989) expressed her hostility to her mother, she experienced a prolonged feeling of isolation and loneliness that lasted until she reported and discussed the following fantasy with her therapist.

My therapist and I decided to try the same technique to try to understand my intense discomfort at being alone. Visualizing isolation was much more difficult than visualizing pain. After many attempts that we both rejected as trivial, I finally caught the spirit of what I was experiencing. I saw some figures with shrouds—very unclear. Then as they took on a more distinct form, I saw that they were witches standing around a fire. My therapist told me to ask them to come over to talk to us. They were frightening to me in the light of the fire, but they were more horrible as they came closer. They laughed at me and started to poke at me with their sticks. The visualization was so real and their presence was so chilling to me that I burst into tears over the interaction with them.

My therapist told me to ask them what I could do to get rid of the awful fear of isolation. Finally they revealed their price. It was that I make a sacrifice so that they could become beautiful and mingle with other people. When I heard their price I began to tremble. In an almost inaudible voice I
of loneliness and isolation. The source of the feeling and
reached awareness before the fantasy was an enduring feel-
comment. First, it is noteworthy that the only aspect that
There are several aspects of this fantasy that warrant further
with a weapon that would destroy them. I refused to kill them myself, but said that the wings of the
would fan the flames of their fire, which would turn back
on them and destroy them.
There was only one problem with this scenario—the horse
and I were one now and I couldn’t get airborne. The wings were
so heavy that I couldn’t flamp them hard enough to catch the
breeze. The harder I tried, the more I failed and the more the
witches laughed at me. My therapist . . . told me that another
horse would fan the flames of their fire, which would turn back
them. I refused to kill them myself, but said that the wings of the
horse would fan the flames of their fire, which would turn back
on them and destroy them.
There was only one problem with this scenario—the horse
and I were one now and I couldn’t get airborne. The wings were
so heavy that I couldn’t flamp them hard enough to catch the
breeze. The harder I tried, the more I failed and the more the
witches laughed at me. My therapist . . . told me that another
horse who loved the first horse very much would join her and to-
gether they would destroy the witches. The other horse flew
above me and made a vacuum into which I could take off. Once
in the air, I flew effortlessly and fanned the fire into a huge blaze.
The witches ran here and there trying to avoid the flames but in
the end they were consumed by the fire.
I practiced the scene over and over again until it became
easy, but I never enjoyed it. I liked to fly, but I felt sorry for the
witches, no matter how mean they were to me. My therapist felt
that it was a mistake to feel sympathy for them because they
would take advantage of any mercy that I displayed. He felt they
would use any deception and illusion they could to control me.
I was not so sure but I did agree with him that I must assume the
right to soar into the world and be free of their influence. After
the session, my therapist and I discussed the meaning of the im-
gees. Although I had begun with the concept of isolation in
mind, I knew that the witches related to my mother, particularly
the way she would poke at me and shame me. They probably
represented my fear of isolation if I did not acquiesce to her de-
mands. My therapist added that in destroying the witches I was
only destroying the hostile part of our relationship, the witch
part of it, and leaving the loving part intact. This was necessary
for me to be free, autonomous, and no longer ensnared by fear
of abandonment.
The concept that I had a great deal of conflict between the
need for association and the need for autonomy was not new. I
believed I had to buy affection and that no one would love me if
I were myself, i.e., if I attended to my own wants. I knew also
that I felt that I had to carry the burden of being responsible for
my mother’s well-being, that she would die at some level if I
broke the bond with her. (pp. 45–47)

There are several aspects of this fantasy that warrant further
comment. First, it is noteworthy that the only aspect that
reached awareness before the fantasy was an enduring feel-
ing of loneliness and isolation. The source of the feeling and
its associations remained unconscious until they were dealt
with at the experiential level and perhaps assimilated at the
rational level.
Second, the insight represented in the fantasy—namely,
that Alice had a conflict between autonomy and related-
ness—was not new to her. As she noted, she had been con-
sciously aware of this conflict before. What, then, did the
fantasy accomplish? What it accomplished was to produce a
vicarious corrective emotional experience that had a pro-
found effect at the experiential level. The previous intellec-
tual insight in the absence of involvement of the experiential
system had accomplished little. To make a therapeutic contri-
bution, the same information had to be felt and processed
experimentally.
Third, the fantasy provided useful diagnostic clues for the
psychotherapist. Alice, apparently, could not free herself from
the hold of the bad mother figure until a loving figure sup-
ported her independence, after which she could soar freely.
This suggested that what she needed to resolve her conflict
was to be convinced at a compelling experiential level that it
is possible to be autonomous and loved at the same time. This
was duly noted by her therapist, who made a point of encour-
aging its implementation in her family, as well as supporting it
himself in the therapeutic relationship.
Fourth, the fantasy illustrates the usefulness of vicarious
symbolic experience as a therapeutic tool. Alice sponta-
aneously began to practice in fantasy enjoying the feeling of
soaring freely into space, and as a result she was able to gain
a newfound freedom without guilt or fear of abandonment.
What she learned through the fantasy at a deep experiential
level suggests a therapeutic technique that may be more gen-
erally useful—namely, the practice in symbolic form of cop-
ing with a deep-seated problem that cannot be resolved by
intellectual insight. Of additional general value of this exam-
ple of a spontaneous fantasy is that it indicates how such fan-
tasies can provide diagnostic information that can be useful
in therapy.
There is, of course, no way of knowing the extent to which
the use of fantasy relative to other factors, such as having a
highly supportive environment, played in Alice’s rapid
progress. It is very likely that both contributed. However, it
should be considered in this respect that the equally favor-
able environment before the therapy was insufficient to re-
solve Alice’s conflict between autonomy and relatedness. As
she reported, the love and affection that were abundantly
available to her from her husband, her children, her extended
family, and her deeply caring friends could not penetrate, so
long as she felt that the price of love was the sacrifice of au-
tonomy. Having developed a lifelong pattern of self-sacrifice
in order to maintain relationships, she had no way of learning before therapy that it was unnecessary.

Implications for Research

If there are two different information-processing systems, it can only be a source of confusion to conduct research as if there were only one, which is the customary practice. As an example, given the existence of two different systems, it is meaningless to investigate “the” self-concept because a person’s self-concept in one system may not conform to the self-concept in the other system. Moreover, the difference between the two self-concepts can be of considerable importance in its own right. The problem of treating the two self-concepts as if there were only one has been particularly evident in research on self-esteem, in which individuals are typically classified as high or low in self-esteem based on self-report questionnaires. Yet if there are two self-concepts, then it is quite possible for people to be high in self-esteem in one system and low in the other. For example, a person might be high in self-esteem in the rational system, as measured by a self-report test, yet low in self-esteem in the experiential system, as inferred from behavior (Savin-Williams & Jaquish, 1981).

There has been much disagreement concerning whether elevating students’ self-esteem by treating them as successful no matter what their performance is desirable or undesirable. In order to resolve this issue, from the perspective of CEST it is necessary to recognize that high self-esteem at the conscious, rational level may coexist with low self-esteem at the experiential level. It is one thing to teach students to consciously believe they have high self-regard and another to have them acquire the quiet confidence that comes from feelings of mastery and competence that are a consequence of real accomplishment. The former in the absence of the latter can be considered to be no more than self-deception and a potential source of disillusionment in the future. It follows that not only is it important to examine self-esteem separately in each of the two systems, but it is equally important to conduct research on their convergence. What is obviously true of self-esteem in this respect is equally true of other personality variables, including basic needs and beliefs.

Although the importance of four basic needs and corresponding beliefs is emphasized in CEST, this is not meant to imply that lower-level beliefs and needs are not also very important. Recently, social and personality psychologists have emphasized midlevel motivational constructs (e.g., Emmons, 1986; Markus & Nurius, 1986; Mischel & Shoda, 1995). It is assumed in CEST that personality is hierarchically organized, with broad, basic needs subsuming midlevel motives, which in turn subsume narrower, situation-specific motives. It would therefore be desirable to examine the organization of such needs and beliefs, and to determine in particular the kinds of relations the different levels establish with each other, as well as with other variables. It might reasonably be expected that the lowest-order needs and beliefs are most strongly associated with situationally specific behaviors, and the higher-order beliefs and needs are more strongly associated with broad dispositions, or traits. The higher-order beliefs and needs can also be expected to be more weakly but more extensively associated with narrow behavioral tendencies. Midlevel motives can be expected to have relations that fall in between those of the higher- and lower-order needs. A particularly important hypothesis with regard to CEST is that higher-order needs and beliefs are more resistant to change than are lower-order needs and beliefs, but should they be changed they have greater effects on the overall personality structure. Moreover, any major changes, including positive changes, are disorganizing and anxiety-producing because of the basic need to maintain the stability of the conceptual system.

Although considerable research has recently been conducted on midlevel needs that has demonstrated their theoretical importance and predictive value (e.g., Emmons, 1986; Markus & Nurius, 1986; Mischel & Shoda, 1995), the question remains as to how the midlevel needs can best be designated and measured. The most thorough and compelling list of midlevel needs to date still appears to be the list proposed by Henry A. Murray (1938) many years ago. It is interesting from the perspective of CEST that Murray measured midlevel needs both explicitly via direct self-report and implicitly through the use of the Thematic Apperception Test (TAT; Murray, 1943). A more psychometrically advanced procedure for measuring the Murray midlevel needs at the explicit level has since become available in the form of the Edwards Personal Preference Schedule (Edwards, 1959).

There is a need for research to further explore the TAT as a measure of implicit needs and to also examine additional measures of implicit needs. Included could be older procedures such as word association and sentence completion, as well as promising new procedures such as priming techniques and subthreshold measures (see Bargh & Chartrand, in press, for a review of such techniques). It would be interesting to relate the various implicit measures to each other to determine whether they have enough in common to combine them into an overall measure. The implicit measure (or measures) of needs could then be related to explicit measures of needs, and both could be related to external criteria. Through such procedures it should be possible to determine in what ways implicit and explicit measures are similar and different. It could also be determined whether they contribute in a supplementary way to the prediction of the same variables and
SUMMARY AND CONCLUSIONS

Cognitive-experiential self-theory (CEST) is a psychodynamic global theory of personality that substitutes a different kind of unconscious processing for the Freudian unconscious. Unlike the maladaptive Freudian unconscious, the unconscious of CEST is an adaptive, associative learning system. It is the same system with which higher-order animals have increasingly effectively adapted to their environments over millions of years of evolution. Because it is a system that learns from experience, it is referred to as the experiential system. In addition to an experiential system, humans uniquely have a rational system. The rational system is a logical, inferential system that operates with the aid of language. The experiential system can account for the widespread irrationality in the thinking of humans despite their unique capacity for reasoning rationally by recognizing that it biases conscious thinking automatically and outside of awareness.

The operating principles of the experiential system were described and contrasted with those of the rational system. Although the systems are independent in the sense that they operate by different rules, they nevertheless are highly

whether the degree to which they coincide in individuals is an important personality variable, as assumed in CEST. It would be informative to determine what kinds of combinations of implicit needs usually result in compromises, what kinds usually result in conflict, and how this differs among individuals. Such research would not only be of theoretical importance, but would also have important implications for the diagnosis of sources of distress and for therapy.

Although considerable research has been done with the CTI that has supported its construct validity (see review in Epstein, 2001), there are many areas that could profit from further research with it. One such area is the predictive value of the CTI for success in a variety of work situations that have not yet been investigated. It would be interesting, for example, to conduct a study comparing the contribution of intellectual intelligence, as measured by a standard intelligence test, and experiential intelligence, as measured by the CTI, for predicting performance in graduate school and beyond. A hypothesis derived from CEST and consistent with previous research (Epstein, 2001) is that intellectual intelligence is a stronger predictor of grades and scores on paper-and-pencil tests, whereas constructive thinking is a stronger predictor of practical performance. The latter could be indicated by demonstrations of research productivity and creativity, by length of time to complete the PhD degree, and by successful professional performance after obtaining the PhD degree.

As noted previously, with the aid of a newly constructed instrument, the Rational-Experiential Inventory (REI; Epstein et al., 1996; Norris & Epstein, 2000a, 2000b; Pacini & Epstein, 1999b), it is possible to study the effects of individual differences in processing in each of the two modes. Of particular interest is the independent contribution of each of the modes for predicting well-being and performance in different kinds of activities. Although a promising beginning has been made in this area, there is a need for more extensive research, particularly with the objective rather than self-reported dependent variables.

An important area of research with both practical and theoretical implications is the relation of the two thinking styles to receptivity to different kinds of messages. The one research project that has been completed on this issue (Rosenthal & Epstein, 2000) has produced interesting results consistent with CEST and suggests that it is a promising area for further research. It remains to be determined how each of the processing styles—separately and in combination—is related to receptivity to messages regarding politics, advertising, and health-related behaviors such as smoking and sexual risk-taking.

An area of particular theoretical and practical importance is the influence of the experiential system on the rational system. As previously noted, this relation can account for the paradoxical irrationality exhibited by humans despite their unique capacity for rational reasoning. The influence of experiential on rational processing is assigned an extremely important role in CEST, equivalent to the influence of the Freudian unconscious in psychoanalysis. It is therefore important from the perspective of CEST to conduct further research to demonstrate the influence of experiential on rational processing under various conditions. Relatively, it is important to test the hypotheses that such influence is often mediated by feelings, the identification of which, accordingly, can be helpful as a first step in controlling the influence of the experiential on the rational system.

Research is needed on the positive contributions of experiential processing to creativity, wisdom, and physical and mental well-being. It is important in this respect to determine how people can most effectively influence and learn from their experiential systems by communicating with these systems in their own medium, as illustrated in the case history that was presented. You will recall that Alice, by practicing soaring freely and unaided in fantasy, was able to accept the belief, at a deep experiential level, that it is possible to be an autonomous being without fear of rejection in a way that intellectual insight was unable to accomplish. It will be interesting to determine how effective such symbolic rehearsal is more generally as a way of resolving deep-seated conflicts at the experiential level.
interactive. The two systems usually operate in synchrony and produce compromises between them, but sometimes they conflict with each other, resulting in what are commonly referred to as conflicts between the heart and the head. A research program was described that provided support for many of the assumptions in CEST. The implications of CEST were discussed for psychotherapy and psychological research.

It was noted that neither system is superior to the other. They are simply different ways of understanding the world and behaving in it. The experiential system is intimately associated with emotions and adapts by learning from outcomes. The rational system is a affect-free and adapts by logical inference. Each has its advantages and disadvantages. Although the rational system is responsible for remarkable achievements in science and technology, it is less well suited for everyday living than is the experiential system. Moreover, the experiential system can intuitively and holistically solve some problems that are beyond the capacity of the analytical, rule-based reasoning of the rational system (Hammond, 1996). The experiential system is also a source of some of humankind’s most desirable attributes, including the capacity for passion, compassion, love, creativity, and appreciation of aesthetics. However, it is also a source of serious difficulties, including superstitious thinking, prejudice, violence, and—perhaps most important—undermining people’s ability to think rationally. Thus, the experiential system is a mixed blessing; it is difficult to live with it, but it would be impossible to live without it.

REFERENCES


Epstein, A. (1989). Mind, fantasy, and healing. New York: Delacorte. (This book is out of print, but copies can be obtained from Amazon.com or from Balderwood Books, 37 Bay Road, Amherst, MA 01002 by enclosing a check for $18.00, which includes postage.)


Personality is a difficult concept to pin down. By necessity it is a very broad concept because personality impinges on virtually all aspects of human behavior. This breadth is viewed differently by different theorists, however. As a result, many different approaches have been taken to thinking about and conceptualizing personality. The diversity in focus among the chapters in the first part of this volume attests very clearly to that fact.

We were both trained as personality psychologists. Throughout our careers, however, our research interest has focused on a set of issues regarding the structure of behavior. These issues link the concept of personality and its functioning to a set of themes that might be regarded as representing the psychology of motivation. Our interest in how behavior occurs has taken us into a number of specific research domains—most recently health-related behavior and responses to stress (Carver & Scheier, 2001; Scheier, Carver, & Bridges, 2001). However, these specific explorations have almost always occurred in service to a more general interest in the structure of behavior.

What we mean by “the structure of behavior” is reflected in the issues underlying questions such as these: What is the most useful way to think about how people create actions from their intentions, plans, and desires? Once people have decided to do something, how do they stay on course? What is the relation between people’s values and their actions?
What processes account for the existence of feelings as people make their way through the world?

As we have tried to address such questions, we have consistently returned to the idea that people are self-regulatory entities. That is, human behavior is an attempt to make something occur in action that is already held in mind. Similarly, affects serve as self-regulatory controls on what actions take place and with how much urgency.

The self-regulatory principles we emphasize in our writings were not conceived as being a model of personality. However, the principles do turn out to provide an interesting perspective on personality. They suggest some implications about how personality is organized and expressed in people’s actions. These principles also point to some of the issues that are involved in successfully negotiating the world. The principles we emphasize deal most explicitly with the “process” aspect of personality—the functions that make everyone a little bit alike—but they can also be seen to have implications for the individual differences that are part of personality psychology.

This chapter is organized as a series of conceptual themes that reflect this self-regulatory perspective on personality. We start with basic ideas about the nature of behavior and some of the processes by which we believe behavior is regulated. We then turn to emotion—how we think it arises and a way in which two classes of affects differ from each other. This leads to a discussion of the fact that people sometimes are unable to do what they set out to do and of what follows from that problem. The next sections are more speculative and reflect emerging themes in thinking about behavior. They deal with dynamic systems, connectionism, and catastrophe theory as models for behavior and how such models may influence how people such as ourselves view self-regulation.

**BEHAVIOR AS GOAL DIRECTED AND FEEDBACK CONTROLLED**

The view we take on behavior begins with the concept of goal and the process of feedback control, ideas we see as intimately linked. Our focus on goals is in line with a growing re-emergence of goal constructs in personality psychology (e.g., Austin & Vancouver, 1996; Elliott & Dweck, 1988; Miller & Read, 1987; Pervin, 1989), constructs known by a variety of labels such as current concern (Klinger, 1975, 1977), personal strivings (Emmons, 1986), life task (Cantor & Kihlstrom, 1987), and personal project (Little, 1983). The goal construct is at its core very simple. Yet these theories all emphasize that it has room for great diversity and individualization. For example, any life task can be achieved in diverse ways. People presumably choose paths for achieving a given life task that are compatible with other aspects of their life situation (e.g., many concerns must usually be managed simultaneously) and with other aspects of their personality.

Two goal constructs that differ somewhat from those named thus far are the possible self (Markus & Nurius, 1986) and the self-guide (Higgins, 1987, 1996). These constructs were intended to bring a dynamic quality to conceptualization of the self-concept. In contrast to traditional views, but consistent with other goal frameworks, possible selves are future oriented. They concern how people think of their as-yet-unrealized potential, the kind of people they might become. Self-guides similarly reflect dynamic aspects of the self-concept.

Despite differences among these various constructs (see Austin & Vancouver, 1996; Carver & Scheier, 1998), they are the same in many ways. All include the idea that goals energize and direct activities; all implicitly convey the sense that goals give meaning to people’s lives (cf. Baumeister, 1989). Each theory emphasizes the idea that understanding the person means in part understanding the person’s goals. Indeed, the view represented by these theories often implies that the self consists partly of the person’s goals and the organization among them.

**Feedback Processes**

How are goals used in behaving? We believe that goals serve as reference values for feedback loops (Wiener, 1948). A feedback loop, the unit of cybernetic control, is a system of four elements in a particular organization (cf. MacKay, 1956; Miller, Galanter, & Pribram, 1960). The elements are an input function, a reference value, a comparator, and an output function (see Figure 8.1).

![Figure 8.1 Schematic depiction of a feedback loop, the basic unit of cybernetic control. In such a loop a sensed value is compared to a reference value or standard, and adjustments are made in an output function (if necessary) to shift the sensed value in the appropriate direction.](image-url)
An input function is a sensor. Think of it as perception. The reference value is a bit of information specified from within the system. Think of it as a goal. A comparator is something that makes continuous or repeated comparisons between the input and the reference value. The comparison yields one of two outcomes: values being compared either are or are not discriminably different from one another. Following the comparison is an output function. Think of this as behavior (although the behavior sometimes is internal). If the comparison yielded “no difference,” the output function remains whatever it was. If the comparison yielded “discrepancy,” the output changes.

There are two kinds of feedback loops, corresponding to two kinds of goals. In a discrepancy-reducing loop (a negative feedback loop), the output function is aimed at diminishing or eliminating any detected discrepancy between input and reference value. It yields conformity of input to reference. This conformity is seen in the attempt to approach or attain a valued goal.

The other kind of feedback loop is a discrepancy-enlarging loop (a positive feedback loop). The reference value here is not one to approach, but one to avoid. Think of it as an “anti-goal.” An example is a feared possible self. Other examples would be traffic tickets, public ridicule, and the experience of being fired from your job. This loop senses present conditions, compares them to the anti-goal, and tries to enlarge the discrepancy. For example, a rebellious adolescent who wants to be different from his parents senses his own behavior, compares it to his parents’ behavior, and tries to make his own behavior as different from theirs as possible.

The action of discrepancy-enlarging processes in living systems is typically constrained in some way by discrepancy-reducing loops (Figure 8.2). To put it differently, avoidance behaviors often lead into approach behaviors that are compatible with the avoidance. An avoidance loop creates pressure to increase distance from the anti-goal. The movement away occurs until it is captured by the influence of an approach loop. This loop then serves to pull the sensed input into its orbit. The rebellious adolescent, trying to be different from his parents, soon finds other adolescents to conform to, all of whom are actively deviating from their parents.

Our use of the word orbit in the last paragraph suggests a metaphor that may be useful for those to whom these concepts do not feel very intuitive. You might think of feedback processes as metaphorically equivalent to gravity and anti-gravity. The discrepancy-reducing loop exerts a kind of gravitational pull on the input it is controlling, pulling that input closer to its ground zero. The discrepancy-enlarging loop has a kind of antigravitational push, moving sensed values ever farther away. Remember, though, that this is a metaphor. More is involved here than a force field.

Note that situations are often more complex than the one in Figure 8.2 in that there often are several potential values to move toward. Thus, if several people try to deviate from a mutually disliked reference point, they may diverge from one another. For example, one adolescent trying to escape from his parents’ values may gravitate toward membership in a rock band, whereas another may gravitate toward the army. Presumably, the direction in which the person moves will depend in part on the fit between the available reference values and the person’s preexisting values, and in part on the direction the person takes initially to escape from the anti-goal.

Feedback processes have been studied for a long time in a variety of physical systems (cf. Wiener, 1948). With respect to living systems, they are commonly invoked regarding physiological systems, particularly those that maintain the equilibriums that sustain life. We all know of the existence of homeostatic systems that regulate, for example, temperature and blood pressure. It is a bit of a stretch to go from homeostatic maintenance processes to intentional behavior, but the stretch is not as great as some might think (see Miller et al., 1960; MacKay, 1956; Powers, 1973).

One key to this extrapolation is the realization that reference values for feedback loops need not be static. They can change gradually over time, and one can be substituted quickly for another. Thus, a feedback system need not be purely homeostatic. It can be highly dynamic—chasing (and avoiding) moving targets and changing targets. This is not too far from a description (albeit a very abstract one) of the events that make up human life.

Some years ago we argued that the comparator of a psychological feedback process is engaged by self-focused...
attention (Carver, 1979; Carver & Scheier, 1981). Indeed, the similarity between self-focus effects and feedback effects was one thing that attracted us to the feedback model in the first place. Self-focused attention leads to more comparisons with salient standards (Scheier & Carver, 1983) and to greater conformity to those standards. On the avoidance side, self-focus has led to rejection of attitudinal positions held by a negative reference group (Carver & Humphries, 1981) and to stronger reactance effects (Carver & Scheier, 1981).

The literature of self-awareness is not the only one in personality–social psychology that fits well the structure of the feedback loop. Another good example (Carver & Scheier, 1998) is the literature of social comparison. People use upward comparisons to help them pull themselves toward desired goals. People use downward comparisons to help them force themselves farther away from (upward from) those who are worse off than they are.

**Re-emergent Interest in Approach and Avoidance**

Our interest in the embodiment of these two different kinds of feedback processes in behavior is echoed in the recent emergence of interest in two modes of regulation in several other literatures. One of the most prominent of these literatures stems from a group of theories that are biological in focus. Their research base ranges from animal conditioning and behavioral pharmacology (Gray, 1982, 1987b) to studies of human brain activity (Davidson, 1992a, 1992b; Tomarken, Davidson, Wheeler, & Doss, 1992). These theories assume that two core biological systems (sometimes more) are involved in regulating behavior.

One system, managing approach behavior, is called the behavioral activation system (Cloninger, 1987; Fowles, 1980), behavioral approach system (Gray, 1987a, 1990), behavioral engagement system (Depue, Krauss, & Spoont, 1987), or behavioral facilitation system (Depue & Iacono, 1989). The other, dealing with withdrawal or avoidance, is usually called the behavioral inhibition system (Cloninger, 1987; Gray, 1987a, 1990), and sometimes a withdrawal system (Davidson, 1992a, 1992b). The two systems are generally regarded as relatively independent, with different portions of the brain being most involved in their functioning.

Another literature with a dual-motive theme derives from self-discrepancy theory (Higgins, 1987, 1996). This theory holds that people relate their perceptions of their actual selves to several self-guides, particularly ideals and oughts. Ideals are qualities the person desires to embody: hopes, aspirations, positive wishes for the self. Living up to an ideal means attaining something desired. An ideal is clearly an approach goal.

**Oughts**

Oughts, in contrast, are defined by a sense of duty, responsibility, or obligation. An ought is a self that one feels compelled to be, rather than intrinsically desires to be. The ought self is a positive value, in the sense that people try to conform to it. However, living up to an ought also implies acting to avoid a punishment—self-disapproval or the disapproval of others. In our view, oughts are more complex structurally than ideals. Oughts intrinsically imply both an avoidance process and an approach process. Their structure thus resembles what was illustrated earlier in Figure 8.2. Recent work has demonstrated the avoidance aspect of the dynamics behind the ought self (Carver, Lawrence, & Scheier, 1999; Higgins & Tykocinski, 1992).

A similar theme can be seen in the literature of self-determination theory (e.g., Deci & Ryan, 1985, 2000; Ryan & Deci, 2000). That theory focuses on the importance of having a sense of self-determination in one’s actions. Actions that are self-determined are engaged in because they are of intrinsic interest or because they reflect values that are incorporated within the self. Such behavior clearly represents a voluntary approach of positive goal values. In contrast to this is what is termed controlled behavior, meaning that the behavior occurs in response to some sort of coercive force. The coercion can be from outside, or it can be self-coercion. An illustration of the latter is doing something because you feel you have to do it in order not to feel guilty. Such introjected values are very similar to the oughts of self-discrepancy theory, and we have suggested that they similarly involve an avoidance process along with the approach (Carver & Scheier, 1999a, 2000).

**Hierarchicality Among Goals**

Another key issue in the translation of goals into behavior reflects the obvious fact that some goals are broader in scope than others. How to think about the difference in breadth is not always easy to put your finger on. Sometimes it is a difference in temporal commitment. Sometimes, though, it’s more than that: It’s a difference in the goal’s level of abstraction.

**Differentiating Goals by Levels of Abstraction**

The notion that goals differ in their level of abstraction is easy to illustrate. You may have the goal of being an honorable person or a self-sufficient person—goals at a fairly high level of abstraction. You may also have the goal of avoiding a person at work who gossips or of making dinner for yourself, which are at a lower level of abstraction. The first set concerns being a particular kind of person, whereas the second...
set concerns completing a particular kind of action. You can also think of goals that are even more concrete, such as the goal of walking quietly to your office and closing the door without being noticed, or the goal of slicing vegetables into a pan. These goals (which some would call plans or strategies) are closer to specifications of individual acts than was the second set, which consisted more of summary statements about the desired outcomes of intended action patterns.

As you may have noticed, the examples used to illustrate concrete goals relate directly to the examples of abstract goals. We did this to show how abstract goals join with concrete goals in a hierarchy of levels of abstraction. In 1973 William Powers argued that a hierarchical organization of feedback loops underlies the self-regulation of behavior, thus proposing a model of hierarchicity among goals.

His line of thought ran as follows: In a hierarchical organization of feedback systems, the output of a high-level system consists of the resetting of reference values at the next-lower level of abstraction. To put it differently, higher order systems “behave” by providing goals to the systems just below them. The reference values are more concrete and restricted as one moves from higher to lower levels. Each level regulates a quality that contributes to (though not entirely defining) the quality controlled at the next-higher level. Each level monitors input at the level of abstraction of its own functioning, and each level adjusts output to minimize its discrepancies. Structures at various levels presumably handle their concerns simultaneously.

Powers (1973) focused particularly on low levels of abstraction. He said much less about the levels we’re most interested in, though he did suggest labels for several of them. What he called sequences are strings of action that run off directly once cued. Programs, the next-higher level, are activities involving conscious decisions at various points. Programs are webs of sequences with an overall purpose that synthesizes the goals of the constituent sequences. The next level is principles, qualities that are abstracted from (or implemented by) programs. These are the kinds of qualities that are represented by trait labels. Powers gave the name system concepts to the highest level he considered. Goals there include the idealized sense of self, relationship, or group identity.

A simple way of portraying this hierarchy is shown in Figure 8.3. This diagram omits the loops of feedback processes, using lines to indicate only the links among goal values. The lines imply that moving toward a particular lower goal contributes to the attainment of some higher goal (or even several at once). Multiple lines to a given goal indicate that several lower level action qualities can contribute to its attainment. As indicated previously, there are goals to be a particular way and goals to do certain things (and at lower levels, goals to create physical movement).

Although the Powers hierarchy per se has not been studied empirically, research has been done from the perspective of another theory that strongly resembles it—Vallacher and Wegner’s (1985) action identification theory. This model is framed in terms of how people think about their actions, but it also conveys the sense that how people think about their actions is informative about the goals by which they are guiding the actions. People can identify a given action in many different ways, and the identifications can vary in level of abstraction. High-level identifications are abstract, whereas lower level identifications are more concrete. Low-level identifications tend to convey a sense of how an activity is done, whereas high-level identifications tend to convey a sense of why.

The Vallacher and Wegner (1985) model does not specify what qualities define various levels but simply assumes that where there is a potential emergent property, there is the potential for differing levels of identification. However, the examples used to illustrate the theory tend to map onto levels of the Powers hierarchy: sequences of acts, programs of actions (with variations of smaller scale and larger scale programs), and principles of being. Thus, work on action identification tends to suggest the reasonableness of these particular levels of abstraction in thinking about behavior.

Step back from this hierarchy for a moment to consider its broader implications. Our present interest is in linking these ideas to the construct of personality. It should be clear that this model provides a way to talk about how the values that are embedded in a person’s personality are manifested in
that person’s actions. Values are the source of intentions to
take certain patterns of actions, and those programmatic
action plans are realized in an extended series of sequences of
movement. This view also provides for a mechanism by
which the actions themselves take place, which is not typi-
cally the case in models of personality.

Multiple Paths to High-Level Goals, Multiple Meanings
from Concrete Acts

This hierarchy also has implications for several further issues
in thinking about behavior (for more detail see Carver &
Scheier, 1998, 1999a). In this view, goals at a given level can
often be attained by a variety of means at lower levels. This
addresses the fact that people sometimes shift radically the
manner in which they try to reach a goal when the goal itself
has not changed. This happens commonly when the emergent
quality that is the higher order goal is implied in several
lower order activities. For example, a person can be helpful
by writing a donation check, picking up discards for a recy-
cling center, volunteering at a charity, or holding a door open
for someone else.

Just as a given goal can be obtained via multiple path-
ways, so can a specific act be performed in the service of di-
verse goals. For example, you could buy someone a gift to
make her feel good, to repay a kindness, to put her in your
debt, or to satisfy a perceived holiday-season role. Thus, a
given act can have strikingly different meanings depending
on the purpose it’s intended to serve. This is an important
subtheme of this view on behavior: Behavior can be under-
stood only by identifying the goals to which behavior is ad-
dressed. This is not always easy to do, either from an
observer’s point of view (cf. Read, Druian, & Miller, 1989)
or from the actor’s point of view.

Goals and the Self

Another point made by the notion of hierarchical organiza-
tion concerns the fact that goals are not equivalent in their
importance. The higher you go into the organization, the
more fundamental to the overriding sense of self are the qual-
ities encountered. Thus, goal qualities at higher levels would
appear to be intrinsically more important than those at lower
levels.

Goals at a given level are not necessarily equivalent to
one another in importance, however. In a hierarchical system
there are at least two ways in which importance accrues to a
goal. The more directly an action contributes to attainment
of some highly valued goal at a more abstract level, the more
important is that action. Second, an act that contributes to
the attainment of several goals at once is thereby more im-
portant than an act that contributes to the attainment of only
one goal.

Relative importance of goals returns us again to the concept
of self. In contemporary theory the self-concept has several
aspects. One is the structure of knowledge about your personal
history; another is knowledge about who you are now. Another
is the self-guides or images of potential selves that are used to
guide movement from the present into the future. As stated
earlier, a broad implication of this view is that the self—
indeed, personality—consists partly of a person’s goals.

FEEDBACK LOOPS AND CREATION OF AFFECT

We turn now to another aspect of human self-regulation:
emotion. Here we add a layer of complexity that differs
greatly from the complexity represented by hierarchicality.
Again, the organizing principle is feedback control. But now
the control is over a different quality.

What are feelings, and what makes them exist? Many have
analyzed the information that feelings provide and situations
in which affect arises (see, e.g., Frijda, 1986; Lazarus, 1991;
Ortony, Clore, & Collins, 1988; Roseman, 1984; Scherer &
Ekman, 1984). The question we address here is slightly dif-
ferent: What is the internal mechanism by which feelings
arise?

Velocity Control

We have suggested that feelings arise within the functioning
of another feedback process (Carver & Scheier, 1990). This
process operates simultaneously with the behavior-guiding
process and in parallel to it. One way to describe this second
function is to say that it is checking on how well the behavior
loop is doing at reducing its discrepancies. Thus, the input for
this second loop is a representation of the rate of discrepancy
reduction in the action system over time. We focus first on
discrepancy-reducing loops and turn later to enlarging loops.

We find an analogy useful here. Because action implies
change between states, think of behavior as analogous to dis-
tance. If the action loop deals with distance, and if the affect-
relevant loop assesses the progress of the action loop, then
the affect loop is dealing with the psychological equivalent of
velocity, the first derivative of distance over time. To the ex-
tent that the analogy is meaningful, the perceptual input to
this loop should be the first derivative over time of the input
used by the action loop.
This input does not in itself create affect because a given rate of progress has different affective consequences under different circumstances. As in any feedback system, this input is compared against a reference value (cf. Frijda, 1986, 1988). In this case, the reference is an acceptable or desired rate of behavioral discrepancy reduction. As in other feedback loops, the comparison checks for a deviation from the standard. If there is one, the output function changes.

We suggest that the result of the comparison process in this loop (the error signal generated by its comparator) appears phenomenologically in two forms. One is a nonverbal sense of confidence or doubt (to which we turn later). The other is affect, feeling, a sense of positivity or negativity.

Research Evidence

Because this idea is relatively novel, we should devote some attention to whether any evidence supports it. Initial support came from Hsee and Abelson (1991), who arrived independently at the velocity hypothesis. They conducted two studies of velocity and satisfaction. In one, participants read descriptions of paired hypothetical scenarios and indicated which they would find more satisfying. For example, they chose whether they would be more satisfied if their class standing had gone from the 30th percentile to the 70th over the past 6 weeks, or if it had done so over the past 3 weeks. Given positive outcomes, they preferred improving to a high outcome over a constant high outcome; they preferred a fast velocity over a slow one; and they preferred fast small changes to slower larger changes. When the change was negative (e.g., salaries decreased), they preferred a constant low salary to a salary that started high and fell to the same low level; they preferred slow falls to fast falls; and they preferred large slow falls to small fast falls.

We have since conducted a study that conceptually replicates aspects of these findings but with an event that was personally experienced rather than hypothetical (Lawrence, Carver, & Scheier, in press). We manipulated success feedback on an ambiguous task over an extended period. The patterns of feedback converged such that block 6 was identical for all subjects at 50% correct. Subjects in a neutral condition had 50% on the first and last block, and 50% average across all blocks. Others had positive change in performance, starting poorly and gradually improving. Others had negative change, starting well and gradually worsening. All rated their mood before starting and again after block 6 (which they did not know ended the session). Those whose performances were improving reported mood improvement, whereas those whose performances were deteriorating reported mood deterioration, compared to those with a constant performance.

Another study that appears to bear on this view of affect was reported by Brunstein (1993). It examined subjective well-being among college students over the course of an academic term, as a function of several perceptions, including perception of progress toward goals. Of particular interest at present, perceived progress at each measurement point was strongly correlated with concurrent well-being.

Cruise Control Model

Although the theory may sound complex, the system we have proposed functions much the same as another device that is well known to many people: the cruise control on a car. If you are moving too slowly toward a goal, negative affect arises. You respond to this condition by putting more effort into your action, trying to speed up. If you are going faster than you need to, positive affect arises, and you pull back effort and coast. A car’s cruise control is very similar. You come to a hill, which slows you down. The cruise control responds by feeding the engine more gas to bring the speed back up. If you pass the crest of a hill and roll downhill too fast, the system pulls back on the gas, which eventually drags the speed back down.

This analogy is intriguing because it concerns regulation of the very quality that we believe the affect system is regulating: velocity. It is also intriguing that the analogy incorporates a similar asymmetry in the consequences of deviating from the set point. That is, both in a car’s cruise control and in human behavior, going too slow calls for investing greater effort and resources. Going too fast does not. It calls only for pulling back on resources. That is, the cruise control does not apply the brakes; it just cuts back on the gasoline. In this way it permits the car to coast gradually back to its velocity set point. In the same fashion, people do not respond to positive affect by trying to make it go away, but just by easing off.

Does positive affect actually lead people to withdraw effort? We are not aware of data that bear unambiguously on the question. To do so, a study must assess coasting with respect to the same goal as lies behind the affect. Many studies that might otherwise be seen as relevant to the question created positive affect in one context and assessed its impact on another task (see, e.g., Isen, 2000). The question thus seems to remain open, and to represent an important area for future work (for broader discussion of relevant issues see Carver, in press).
Affect from Discrepancy-Enlarging Loops

Thus far we have restricted ourselves to issues that arise in the context of approach. Now we turn to attempts to avoid a point of comparison, attempts to not-be or not-do: discrepancy-enlarging loops.

Our earlier discussion should have made it clear that behavior regarding avoidance goals is just as intelligible as behavior regarding approach goals. We think the same is true of the affective accompaniments to behavior. Our model rests on the idea that positive affect comes when a behavioral system is doing well at what it is organized to do. Thus far we have considered only systems organized to close discrepancies. There seems no obvious reason, however, why the principle should not apply just as well to systems organized to enlarge discrepancies. If the system is doing well at what it is organized to do, positive affect should arise. If it is doing poorly at what it is organized to do, negative affect should arise.

That much would seem to be fully comparable across the two types of systems. But doing well at moving toward an incentive is not exactly the same experience as doing well at moving away from a threat. Both have the potential to induce positive feelings, by doing well. Both also have the potential to induce negative feelings, by doing poorly. Yet the two positives may not be quite the same as each other, nor the negatives quite the same as each other.

Our view of this difference derives partly from the insights of Higgins and his colleagues (Higgins, 1987, 1996). Following their lead, we suggest that the affect dimension relating to discrepancy reduction is (in its purest form) the dimension that runs from depression to elation (Figure 8.4). The affect dimension relating to discrepancy enlargement is (in its purest form) the dimension from anxiety to relief or contentment. As Higgins and his colleagues have noted, dejection-related and agitation-related affect may take several forms, but these two dimensions capture the core qualities behind those two classes of affect. Similarly, Roseman (1984) has argued that joy and sadness are related to appetitive (moving-toward) motives, whereas relief and distress are related to aversive (moving-away-from) motives.

Merging Affect and Action

Theories about emotion typically emphasize the idea that emotion is related to action. How do affect and action relate in this model? We see the regulation provided by these systems as forming a two-layered array, with both simultaneously at work (Carver & Scheier, 1998, 1999a, 1999b). The two layers are analogous to position and velocity controls in a two-layered engineering control system (e.g., Clark, 1996). Such a two-layered system in engineering has the quality of responding both quickly and accurately (without undue oscillation). There is reason to believe that the simultaneous functioning of the two layers has the same broad consequence for human behavior.

Another way of addressing the relation between affect and action is to ask about the nature of the output of the affect loop. Earlier we described affect as reflecting the error signal of a loop that has as input a perception of rate of progress. The resulting output thus must be an adjustment in rate of progress. This output therefore has a direct link to behavior because it means changing its pace.

What does it mean to adjust the rate of progress? In some cases it means literally changing velocity. If you are behind, go faster. Some adjustments are less straightforward. The rates of many behaviors in which personality–social psychologists are interested are not defined in terms of literal pace of motion. Rather, they are defined in terms of choices among actions, even potential programs of action. For example, increasing your rate of progress on a reading assignment may mean choosing to spend a weekend working rather than playing. Increasing your rate of manifestation of kindness means choosing to perform an action that reflects that value. Thus, adjustment in rate must often be translated into other terms, such as concentration or reallocation of time and effort.

Despite this complexity in implementing changes in rate, it should be apparent from this description that the action system and the velocity system are presumed to work in concert with one another. Both are involved in the flow of action. They influence different aspects of the action, but both are
always involved. Thus, this view incorporates clear links between behavior and affect.

**Comparison with Biological Models of Bases of Affect**

It is useful to compare this model with the group of biologically focused theories mentioned earlier in the chapter. As indicated earlier, those theories assume that two separate systems regulate approach and avoidance behavior. Many assume further that the two systems also underlie affect. Given cues of impending reward, the activity of the approach system creates positive feelings. Given cues of impending punishment, the avoidance system creates feelings of anxiety.

Data from a variety of sources fit this picture. Of particular relevance is work by Davidson and collaborators involving electroencephalography (EEG) recordings assessing changes in cortical activation in response to affective inducing stimuli. Among the findings are these: Subjects exposed to films inducing fear and disgust (Davidson, Ekman, Saron, Senulis, & Friesen, 1990) and confronted with possible punishment (Sobotka, Davidson, & Senulis, 1992) show elevations in right frontal activation. In contrast, subjects with a chance to obtain reward (Sobota et al., 1992), subjects presented with positive emotional adjectives (Cacioppo & Petty, 1980), and smiling 10-month olds viewing their approaching mothers (Fox & Davidson, 1988) show elevations in left frontal activation. From findings such as these, Davidson (1992a, 1992b) concluded that neural substrates for approach and withdrawal systems (and thus positive and negative affect) are located in the left and right frontal areas of the cortex, respectively.

Thus far the logic of the biological models resembles the logic of our model. At this point, however, there is a divergence. The key question is what regulatory processes are involved in—and what affects result from—failure to attain reward and failure to receive punishment. Gray (1987b, 1990) holds that the avoidance system is engaged by cues of punishment and cues of frustrative nonreward. It thus is responsible for negative feelings in response to either of these types of cues. Similarly, Gray holds that the approach system is engaged by cues of reward or cues of escape from (or avoidance of) punishment. It thus is responsible for positive feelings in response to either of these types of cues. In his view, then, each system creates affect of one hedonic tone (positive in one case, negative in the other), regardless of its source. This view is consistent with a picture of two unipolar affective dimensions, each linked to a distinct behavioral system. Others have taken a similar position (see Cacioppo, Gardner, & Berntson, 1999; Lang, 1995; Lang, Bradley, & Cuthbert, 1990; Watson, Wiese, Vaidya, & Tellegen, 1999).

Our position is different. We argue that both approach and avoidance systems can create affects of both hedonic tones because affect is a product of doing well or doing poorly. We think that the frustration and eventual depression that result from failure to attain desired goals involve the approach system (for similar predictions see Clark, Watson, & Mineka, 1994, p. 107; Cloninger, 1988, p. 103; Henriques & Davidson, 1991). A parallel line of reasoning suggests that relief, contentment, tranquility, and serenity relate to the avoidance system rather than to the approach system (see Carver, 2001).

Less information exists about the bases of these affects than about anxiety and happiness. Consider first relief-tranquility. We know of two sources of evidence, both somewhat indirect. The first is a study in which people worked at a laboratory task and experienced either goal attainment or lack of attainment (Higgins, Shah, & Friedman, 1997, Study 4). Participants first were given either an approach orientation to the task (to try to attain success) or an avoidance orientation (to try to avoid failing). After the task outcome (which was manipulated), several feeling qualities were assessed. Among persons given an avoidance orientation, success caused an elevation in calmness, and failure caused an elevation in anxiety. These effects on calmness and anxiety did not occur, however, among those who had an approach orientation. This pattern suggests that calmness is linked to doing well at avoidance, rather than doing well at approach.

Another source is data reported many years ago by Watson and Tellegen (1985). In their analysis of multiple samples of mood data, they reported “calm” to be one of the 10 best markers (inversely) of negative affect (which was defined mostly by anxiety) in the majority of the data sets they examined. In contrast, “calm” never emerged as one of the top markers of positive affect in those data sets. This suggests that these feelings are linked to the functioning of a system of avoidance.

The same sources also provide information on the momentary experience of sadness. In the study by Higgins et al. (1997), failure elevated sadness and success elevated cheerfulness among persons with an approach orientation. These effects did not occur, however, among participants who had an avoidance orientation. The pattern suggests that sadness is linked to doing poorly at approach, rather than doing poorly at avoidance. Similarly, Watson and Tellegen (1985) reported “sad” to be one of the 10 best markers (inversely) of the factor that they called positive affect in the majority of the data sets they examined. In contrast, “sad” never emerged as one of the top markers of negative affect in those data sets. This pattern suggests that sad feelings are linked to the functioning of a system of approach.
This issue clearly represents an important difference among theoretical viewpoints (Carver, 2001). Just as clearly, it is not yet resolved. It seems likely that it will receive more attention in the near future.

RESPONDING TO ADVERSITY: PERSISTENCE AND GIVING UP

In describing the genesis of affect, we suggested that one process yields two subjective experiences as readouts: affect and a sense of confidence versus doubt. We turn now to confidence and doubt—expectancies for the immediate future. We focus here on the behavioral and cognitive manifestations of the sense of confidence or doubt.

One likely consequence of momentary doubt is a search for more information. We have often suggested that when people experience adversity in trying to move toward goals, they periodically interrupt efforts in order to assess in a more deliberative way the likelihood of a successful outcome (e.g., Carver & Scheier, 1981, 1990, 1998). In effect, people suspend the behavioral stream, step outside it, and evaluate in a more deliberated way. This may happen once or often. It may be brief, or it may take a long time. In this assessment people presumably depend heavily on memories of prior outcomes in similar situations. They may also consider such things as additional resources they might bring to bear, alternative approaches that might be taken, and social comparison information (Wills, 1981; Wood, 1989).

These thoughts sometimes influence the expectancies that people hold. When people retrieve “chronic” expectancies from memory, the information already is expectancies—summaries of the products of previous behavior. In some cases, however, the process is more complex. People bring to mind possibilities for changing the situation and evaluate their consequences. This is often done by briefly playing the possibility through mentally as a behavioral scenario (cf. Taylor & Pham, 1996). Doing so can lead to conclusions that influence expectancies (“If I try doing it this way instead of that way, it should work better” or “This is the only thing I can see to do, and it will just make the situation worse”).

It seems reasonable that this mental simulation engages the same mechanism as handles the affect-creation process during actual overt behavior. When your progress is temporarily stalled, playing through a confident and optimistic scenario yields a higher rate of progress than is currently being experienced. The affect loop thus yields a more optimistic outcome assessment than is being derived from current action. If the scenario is negative and hopeless, it indicates a further reduction in progress, and the loop yields further doubt.

Behavioral Manifestations

Whether stemming from the immediate flow of experience or from a more thorough introspection, people’s expectancies are reflected in their behavior. If people expect a successful outcome, they continue exerting effort toward the goal. If doubts are strong enough, the result is an impetus to disengage from effort, and potentially from the goal itself (Carver & Scheier, 1981, 1990, 1998, 1999a; see also Klinger, 1975; Kukla, 1972; Wortman & Brehm, 1975). This theme—divergence in behavioral response as a function of expectancies—is an important one, applying to a surprisingly broad range of literatures (see Carver & Scheier, 1998, chap. 11).

Sometimes the disengagement that follows from doubt is overt, but sometimes disengagement takes the form of mental disengagement—off-task thinking, daydreaming, and so on. Although this can sometimes be useful (self-distraction from a feared stimulus may allow anxiety to abate), it can also create problems. Under time pressure, mental disengagement can impair performance, as time is spent on task-irrelevant thoughts. Consistent with this, interactions between self-focus and expectancies have been shown for measures of performance (Carver, Peterson, Follansbee, & Scheier, 1983; Carver & Scheier, 1982).

Often, mental disengagement cannot be sustained, as situational cues force the person to reconfront the problematic goal. In such cases, the result is a phenomenology of repetitive negative rumination, which often focuses on self-doubt and perceptions of inadequacy. This cycle is both unpleasant and performance-impairing.

Is Disengagement Good or Bad?

Is the disengagement tendency good or bad? Both and neither. On the one hand, disengagement (at some level, at least) is an absolute necessity. Disengagement is a natural and indispensable part of self-regulation (cf. Klinger, 1975). If people are ever to turn away from unattainable goals, to back out of blind alleys, they must be able to disengage, to give up and start over somewhere else.

The importance of disengagement is particularly obvious with regard to concrete, low-level goals: People must be able to remove themselves from literal blind alleys and wrong streets, give up plans that have become disrupted by unexpected events, even spend the night in the wrong city if they miss the last plane home. Disengagement is also important, however, with regard to more abstract and higher level goals. A vast literature attests to the importance of disengaging and moving on with life after the loss of close relationships (e.g., Orbuch, 1992; Stroebe, Stroebe, & Hansson, 1993; Weiss, 1988).
People sometimes must be willing to give up even values that are deeply embedded in the self if those values create too much conflict and distress in their lives.

However, the choice between continued effort and giving up presents opportunities for things to go awry. It is possible to stop trying too soon, thereby creating potentially serious problems for oneself (Carver & Scheier, 1998). It is also possible to hold on to goals too long, thereby preventing oneself from taking adaptive steps toward new goals. But both continued effort and giving up are necessary parts of the experience of adaptive self-regulation. Each plays an important role in the flow of behavior.

**Hierarchicity and Importance Can Impede Disengagement**

Disengagement is sometimes precluded by situational constraints. However, a broader aspect of this problem stems from the idea that behavior is hierarchically organized, with goals increasingly important higher in the hierarchy, and thus harder to disengage from.

Presumably, disengaging from concrete values is often easy. Lower order goals vary, however, in how closely they link to values at a higher level, and thus in how important they are. To disengage from low-level goals that are tightly linked to higher level goals causes discrepancy enlargement at the higher level. These higher order qualities are important, even central to one’s life. One cannot disengage from them, disregard them, or tolerate large discrepancies between them and current reality without reorganizing one’s value system (Greenwald, 1980; Kelly, 1955; McIntosh & Martin, 1992; Millar, Tesser, & Millar, 1988). In such a case, disengagement from even very concrete behavioral goals can be quite difficult.

Now recall again the affective consequences of being in this situation. The desire to disengage was prompted by unfavorable expectancies. These expectancies are paralleled by negative affect. In this situation, then, the person experiences negative feelings (because of an inability to make progress toward the goal) and is unable to do anything about the feelings (because of an inability to give up). The person simply stews in the feelings that arise from irreconcilable discrepancies. This kind of situation—commitment to unattainable goals—seems a sure prescription for distress.

**Watersheds, Disjunctions, and Bifurcations Among Responses**

An issue that bears some further mention is the divergence in the model of the behavioral and cognitive responses to favorable versus unfavorable expectancies. We have long argued for a psychological watershed among responses to adversity (Carver & Scheier, 1981). One set of responses consists of continued comparisons between present state and goal, and continued efforts. The other set consists of disengagement from comparisons and quitting. Just as rainwater falling on a mountain ridge ultimately flows to one side of the ridge or the other, so do behaviors ultimately flow to one of these sets or the other.

Our initial reason for taking this position stemmed largely from several demonstrations that self-focused attention creates diverging effects on information seeking and behavior as a function of expectancies of success. We are not the only ones to have emphasized a disjunction among responses, however. A number of others have done so, for reasons of their own.

Kukla (1972) proposed an early model that emphasized the idea of a disjunction in behavior. Another such model is the reactance–helplessness integration of Wortman and Brehm (1975): the argument that threats to control produce attempts to regain control and that perceptions of loss of control produce helplessness. Brehm and his collaborators (Brehm & Self, 1989; Wright & Brehm, 1989) developed an approach to task engagement that resembles that of Kukla (1972), but their way of approaching the description of the problem is somewhat different. Not all theories about persistence and giving up yield this dichotomy among responses. The fact that some do, however, is interesting. It becomes more so a bit later on.

**Scaling Back Aspirations and Recalibration of the Affect System**

The preceding sections dealt with the creation of affect and confidence and the concomitant effects on behavior. By implication, the time frames under discussion were quite narrow. In this section we broaden our view somewhat and indicate an important way in which reference values change across longer periods of time. These particular changes are changes in the stringency of the goals being sought after. We consider this issue both with respect to the reference values underlying the creation of affect and with respect to the goals of behavior.

**Shifts in Velocity Standards**

Reference values used by the affect system presumably can shift through time and experience. That is, as people accumulate experience in a given domain, adjustments can occur in the pacing that they expect and demand of themselves. There is a recentering of the system around the past experience, which occurs via shifts in the reference value (Carver & Scheier, 2000).
Consider first upward adjustments. As an example, a person who gains work-related skills often undertakes greater challenges, requiring quicker handling of action units. Upward adjustment of the rate standard means that the person now will be satisfied only with faster performance. Such a shift has the side effect of decreasing the potential for positive affect and increasing the potential for negative affect because there now is more room to fail to reach the rate standard and less room to exceed it. Recall, however, that the shift was induced by a gain in skills. The change in skill tends to counter the shift in regions of potential success and failure. Thus, the likelihood of negative affect (vs. positive affect or no affect) remains fairly constant.

Now consider a downward adjustment. For example, a person whose health is failing may find that it takes longer to get things done than it used to. This person will gradually come to use less stringent rate standards. A lower pace will then begin to be more satisfying. One consequence of this downward shift of standard is to increase the potential for experiencing positive affect and to decrease the potential for negative affect because there now is less room for failing to reach the rate standard and more room for exceeding it. The failing health, however, tends to counter the shift in regions of potential success and failure. Again, then, the net result is that the likelihood of negative affect (vs. positive and neutral) remains fairly constant.

**Mechanism of Shift**

Such changes in comparison value do not happen quickly or abruptly. Shifting the reference value downward is not people’s first response when they have trouble maintaining a demanding pace. First, they try harder to keep up. Only more gradually, if they continue to lag behind, does the rate-related standard shift to accommodate. Similarly, the immediate response when people’s pace exceeds the standard is not an upward shift in reference value. The more typical response is to coast for a while. Only when the overshoot is frequent does the standard shift upward.

We believe that adjustments in these standards occur automatically and involuntarily, but slowly. Such adjustments themselves appear to reflect a self-corrective feedback process (Figure 8.5). This feedback process is slower than the ones focused on thus far, involving a very gradually accumulating shift. It resembles what Solomon (1980; Solomon & Corbit, 1974) described as the long-term consequences of an opponent process system (see also Helson, 1964, regarding the concept of adaptation level).

As an illustration, assume for the moment that a signal to adjust the standard occurred every time there was a signal to change output, but that the former was much weaker than the latter—say, 5% of the latter. If so, it would take a fairly long time for the standard to change. Indeed, as long as the person deviated from the standard in both directions (under and over) with comparable frequency, the standard would never change noticeably, even over an extended period. Only with repeated deviation in the same direction could there be an appreciable effect on the standard.

This view has an interesting implication for affective experience across an extended period. Such shifts in reference value (and the resultant effects on affect) would imply a mechanism within the organism that prevents both the too-frequent occurrence of positive feeling and the too-frequent occurrence of negative feeling. That is, the (bidirectional) shifting of the rate criterion over time would tend to control pacing such that affect continues to vary in both directions around neutral, roughly as before. The person thus would experience more or less the same range of variation in affective experience over long times and changing circumstances (see Myers & Diener, 1995, for evidence of this). The organization would function as a gyroscope serving to keep people floating along within the framework of the affective reality with which they are familiar. It would provide for a continuous recalibration of the feeling system across changes in situation. It would repeatedly shift the balance point of a psychic teeter-totter so that rocking both up and down remains possible.

![Figure 8.5](https://example.com/figure8.5) A feedback loop (in this case, the postulated velocity loop) acts to create change in the input function, to shift it toward the reference value. Sometimes an additional process is in place as well (gray lines), which adjusts the reference value in the direction of the input. This additional process is presumed to be weaker or slower; thus, the reference value is stable relative to the input value. *Source: From C. S. Carver and M. F. Scheier, On the Self-Regulation of Behavior, copyright 1998, Cambridge University Press. Reprinted with permission.*
Scaling Back on Behavioral Goals

The principle of gradual adjustment of a standard also operates at the level of behavioral goals (Carver & Scheier, 1981, 1998). Sometimes progress is going poorly, expectancies of success are dim, and the person wants to quit. Rather than quit altogether, the person trades this goal for a less demanding one. This is a kind of limited disengagement in the sense that the person is giving up the first goal while adopting the lesser one. However, this limited disengagement keeps the person engaged in activity in the domain he or she had wanted to quit. By scaling back the goal—giving up in a small way—the person keeps trying to move ahead—thus not giving up, in a larger way.

Small-scale disengagement occurs often in the context of moving forward in broader ways. A particularly poignant example comes from research on couples in which one partner is becoming ill and dying from AIDS (Moskowitz, Folkman, Collette, & Vittinghoff, 1996). Some healthy participants initially had the goal of overcoming their partner’s illness and continuing active lives together. As the illness progressed and it became apparent that that goal would not be met, it was not uncommon for the healthy partners to scale back their aspirations. Now the goal was, for example, to do more limited activities during the course of a day. Choosing a more limited and manageable goal ensures that it will be possible to move toward it successfully. The result was that even in those difficult circumstances the person experienced more success than would otherwise have been the case and remained engaged behaviorally with efforts to move forward.

How does the scaling back of goals within a domain occur? We believe that the answer is the same as in the case of affect: If the loop’s output function is inadequate at moving the input toward the standard, a second (slower-acting) process moves the standard toward the input. The scaling back of behavioral goals thus would involve the same structural elements as are involved in the recalibration of the affect system.

CONFLICT AND RESTRAINT

In thinking about the self-regulation of behavior, another set of issues to be considered concerns the existence of conflict. Conflict arises whenever two incompatible goals are held simultaneously and both are salient (see also Carver & Scheier, 1998, 1999b). It sometimes is possible to move toward two goals simultaneously, but sometimes moving toward one interferes with one’s ability to move toward the other. For example, the woman who wants to develop her career and also spend time with her family faces a conflict imposed by the limited number of hours in the day and days in the week (Figure 8.6). The effort to attain one (e.g., further the career by working extra hours) can interfere with efforts to attain the other (by removing the time available for family activities).

Given this structure, the experience of conflict naturally produces negative feelings, as movement toward one of the goals is impeded. If movement toward the active goal is rapid (relative to the reference velocity) as movement toward the other goal is stifled, the person may have mixed feelings, feelings relating to each of the two goal values. It is no surprise that people typically try to balance their conflicting desires so that both goals are partly attained. It is also no surprise that this strategy often feels unsatisfying, as the person “almost” keeps up with goals in both domains but keeps up fully with neither of them.

Often there is no structural basis for viewing one goal as intrinsically more valuable than the other (as in Figure 8.6). Sometimes, however, one goal has a kind of primacy because it is reflected in an explicitly formulated intention to override efforts to move toward the other goal. Sometimes the tendencies involved are mental; sometimes they are behavioral. Often, the attempt to override works for a while (sometimes a long while), but sometimes it fails.

Ironic Processes in Mental Control

One literature bearing on this theme was developed by Wegner (e.g., 1994) and his colleagues. The study that began this work was simple. Some people were told not to think of a white bear for 5 minutes. Then they were told to think about the bear. When the thought was permitted, it came more frequently than it did for people who had not had to suppress the thought first. Something about trying not to think of the bear seemed to create pressure to think of it.

This study was followed by others. Most of this research looked not at rebounds, but at what goes on during people’s
attempts to control their thoughts. The data consistently indicate that an instruction to exert mental control yields better control if the person has no other demands. If something else is going on, however (e.g., if the person is trying to remember a 9-digit number), the instruction backfires, and people tend to do the opposite of what they are trying to do.

Wegner (1994) interprets this as follows: Trying to suppress a thought engages two processes. An intentional process tries to suppress. An ironic monitoring process looks for the occurrence of whatever is being suppressed. If it finds it, it increases the effort of the first mechanism. The ironic monitor is sensitive, but it is automatic and does not require much in the way of mental resources. The intentional process requires more resources. Thus, any reduction in mental resources (e.g., being distracted by a second thought or task) disrupts the intentional process more than it disrupts the ironic monitor. The monitor, searching for lapses, in effect invites those lapses to occur.

This theory also applies to the opposite pattern—attempts to concentrate. In this case, the intentional process concentrates, and the ironic process looks for the occurrence of distractions. As in the first case, if the person's mental resources are stretched thin, the ironic process seems to invite the undesired thought into consciousness. In this case, the thought is a distraction.

This research indicates that trying hard to do something (or suppress something) gets much harder when your mental resources are stretched thin. Not only does it get harder, but you may even begin to do the opposite of what you are trying to do.

**Lapses in Self-Control**

Another important literature bearing on this set of issues concerns what Baumeister and Heatherton (1996) termed self-regulatory failure, which we will term lapse in self-control. The potential for this kind of event arises when someone has both the desire to do something (e.g., overindulge in food or drink) and also the desire to restrain that impulse. Self-control of this sort is often especially hard, and sometimes the restrained impulse breaks free.

Consider binge eating as an example. The binge eater wants to eat but also wants to restrain that desire. If self-control lapses, the person stops trying to restrain the desire to eat, lets himself or herself go, and binges.

In characterizing the decision to quit trying to restrain, Baumeister and Heatherton noted that restraint is hard work and that mental fatigue plays a role; however, giving up the restraint attempt rarely requires that the person reach a state of total exhaustion. Rather, there is a point where the person has had enough and stops trying to control the impulse. We have suggested that confidence about resisting the impulse plays a role in whether the person stops trying (Carver & Scheier, 1998). The confident person continues the struggle to restrain. The person whose confidence has sagged is more likely to give up.

Muraven, Tice, and Baumeister (1998) have extended this line of thought to argue that self-control is a resource that not only is limited but also can become depleted by extended self-control efforts. When the resource is depleted, the person becomes vulnerable to a failure of self-control. This view also suggests that there is a shared pool of self-control resources, so that exhausting the resource with one kind of self-control (e.g., concentrating very hard for many hours on a writing assignment) can leave the person vulnerable to a lapse in a different domain (e.g., eating restraint).

It seems worthwhile to compare the cases considered in this section (lapses in self-control) with those described just earlier (mental control). Both sections dealt with efforts at self-control. In many ways the situations are structurally quite similar. Each is an attempt to override one process by another, which falters when mental resources are depleted. There even is a resemblance between the “overdoing” quality in the previously restrained behavior in Baumeister and Heatherton’s cases and the rebound quality in Wegner’s research.

One difference is that the cases emphasized by Baumeister and Heatherton explicitly involve desires that direct the person in opposing directions. In most cases studied by Wegner, there is no obvious reason why the suppressed thought (or the distractor) would be desirable. This difference between cases seems far from trivial. Yet the similarities in the findings in the two literatures are striking enough to warrant further thought about how the literatures are related.

**DYNAMIC SYSTEMS AND SELF-REGULATION**

Recent years have seen the emergence in the psychological literature of new (or at least newly prominent) ideas about how to conceptualize natural systems. Several labels attach to these ideas: chaos, dynamic systems theory, complexity, catastrophe theory. A number of introductions to this body of thought have been written, some of which include applications to psychology (e.g., Brown, 1995; Gleick, 1987; Telen & Smith, 1994; Vallacher & Nowak, 1994, 1997; Waldrop, 1992). These themes are of growing interest in several areas of psychology, including personality-social psychology. In this section we sketch some of the themes that are central to this way of thinking.
Nonlinearity

Dynamic systems theory holds that the behavior of a system reflects all the forces operating on (and within) it. It also emphasizes that the behavior of a complex system over any period but a brief one is very hard to predict. One reason for this is that the system’s behavior may be influenced by these forces in nonlinear ways. Thus, the behavior of the system—even though highly determined—can appear random.

Many people are used to thinking of relationships between variables as linear. But some relationships clearly are not. Familiar examples of nonlinear relationships are step functions (ice turning to water and water turning to steam as temperature increases), threshold functions, and floor and ceiling effects. Other examples of nonlinearity are interactions. In an interaction the effect of one predictor on the outcome differs as a function of the level of a second predictor. Thus the effect of the first predictor on the outcome is not linear.

Many personality psychologists think in terms of interactions much of the time. Threshold effects and interactions are nonlinearities that most of us take for granted, though perhaps not labeling them as such. Looking intentionally for nonlinearities, however, reveals others. For example, many psychologists now think that many developmental changes are dynamic rather than linear (Goldin-Meadow & Alibali, 1995; Ruble, 1994; Siegler & Jenkins, 1989; Thelen, 1992, 1995; van der Maas & Molenaar, 1992).

Sensitive Dependence on Initial Conditions

Nonlinearity is one reason for the difficulty in predicting complex systems. Two more reasons why prediction over any but the short term is difficult is that you never know all the influences on a system, and the ones you do know you never know with total precision. What you think is going on may not be quite what’s going on. That difference, even if it is small, can be very important.

This theme is identified with the phrase sensitive dependence on initial conditions. This means that a very small difference between two states of affairs can lead to divergence and ultimately to an absence of relation between the paths that are taken later on. The idea is (partly) that a small initial difference between systems causes a difference in what they encounter next, which produces slightly different outcomes (Lorenz, 1963). Through repeated iterations, the systems diverge, eventually moving on very different pathways. After a surprisingly brief period they no longer have any noticeable relation to one another.

How does the notion of sensitive dependence on initial conditions relate to human behavior? Most generally, it suggests that a person’s behavior will be hard to predict over a long period except in general terms. For example, although you might be confident that Mel usually eats lunch, you will not be able to predict as well what time, where, or what he will eat on the second Friday of next month. This does not mean Mel’s behavior is truly random or unlawful (cf. Epstein, 1979). It just means that small differences between the influences you think are affecting him and the influences that actually exist will ruin the predictability of moment-to-moment behavior.

This principle also holds for prediction of your own behavior. People apparently do not plan very far into the future most of the time (Anderson, 1990, pp. 203–205), even experts (Gobet & Simon, 1996). People seem to have goals in which the general form is sketched out but only a few steps toward it have been planned. Even attempts at relatively thorough planning appear to be recursive and “opportunistic,” changing—sometimes drastically—when new information becomes known (Hayes-Roth & Hayes-Roth, 1979).

The notion of sensitive dependence on initial conditions fits these tendencies. It is pointless (and maybe even counterproductive) to plan too far ahead too fully (cf. Kirschenbaum, 1985), because chaotic forces in play (forces that are hard to predict because of nonlinearities and sensitive dependence) can render much of the planning irrelevant. Thus, it makes sense to plan in general terms, chart a few steps, get there, reassess, and plan the next bits. This seems a perfect illustration of how people implicitly take chaos into account in their own lives.

Phase Space, Attractors, and Repellers

Another set of concepts important to dynamic-systems thinking are variations on the terms phase space and attractor (Brown, 1995; Vallacher & Nowak, 1997). A phase diagram is a depiction of the behavior of a system over time. Its states are plotted along two (sometimes three) axes, with time displayed as the progression of the line of the plot, rather than on an axis of its own. A phase space is the array of states that the system occupies across a period of time. As the system changes states from one moment to the next, it traces a trajectory within its phase space—a path of the successive states it occupies across that period.

Phase spaces often contain regions called attractors. Attractors are areas that the system approaches, occupies, or tends toward more frequently than other areas. Attractors exert a metaphorical gravitational pull on the system, bringing
the system into proximity to them. Each attractor has a basin, which is the attractor’s region of attraction. Trajectories that enter the basin tend to move toward that attractor (Brown, 1995).

There are several kinds of attractors, some very simple, others more complex. In a point attractor, all trajectories converge onto some point in phase space, no matter where they begin (e.g., body temperature). Of greater interest are chaotic attractors. The pattern to which this term refers is an irregular and unpredictable movement around two or more attraction points. An example is the Lorenz attractor (Figure 8.7), named for the man who first plotted it (Lorenz, 1963). It has two attraction zones. Plotting the behavior of this system over time yields a tendency to loop around both attractors, but to do so unpredictably. Shifts from one basin to the other seem random.

The behavior of this system displays sensitivity to initial conditions. A small change in starting point changes the specific path of motion entirely. The general tendencies remain the same—that is, the revolving around both attractors. But details such as the number of revolutions around one before deflection to the other form an entirely different pattern. The trajectory over many iterations shows this same sensitivity to small differences. As the system continues, it often nearly repeats itself but never quite does, and what seem nearly identical paths sometimes diverge abruptly, with one path leading to one attractor and the adjacent path leading to the other.

A phase space also contains regions called repellers, regions that are hardly ever occupied. Indeed, these regions seem to be actively avoided. That is, wandering into the basin of a repeller leads to a rapid escape from that region of phase space.

**Another Way of Picturing Attractors**

The phase-space diagram gives a vivid visual sense of what an attractor looks and acts like. Another common depiction of attractors is shown in Figure 8.8. In this view, attractor basins are basins or valleys in a surface (more technically called local minima). Repellers are ridges. This view assumes a metaphoric “gravitational” drift downward in the diagram, but other forces are presumed to be operative in all directions. For simplicity, this portrayal usually is done in two dimensions (sometimes 3), but keep in mind that the diagram often assumes the merging of a large number of dimensions into the horizontal axis.

The behavior of the system at a given moment is represented as a ball on the surface. If the ball is in a valley (points...
1 and 2 in panel A of Figure 8.8), it is in an attractor basin and will tend to stay there unless disturbed. If it is on a hill (between 1 and 2), any slight movement in either direction will cause it to escape its current location and move to an adjacent attractor.

One strength of this portrayal is that it does a good job of creating a sense of how attractors vary in robustness. The breadth of a basin indicates the diversity of trajectories in phase space that are drawn into it. The broader the basin (B-1 in Figure 8.8), the more trajectories are drawn in. The narrower the basin (B-2), the closer the ball has to come to its focal point to be drawn to it. The steepness of the valley indicates how abruptly a trajectory is drawn into it. The steeper the slope of the wall (B-2), the more sudden is the entry of a system that encounters that basin.

The depth of the valley indicates how firmly entrenched the system is, once drawn into the attractor. Figure 8.8, panel C, represents a system of attractors with fairly low stability (the valleys are shallow). One attractor represents a stable situation (valley 1), whereas the others are less so. It will take a lot more “energy” to free the ball from valley 1 than from the others.

There is a sense in which both breadth and depth suggest that a goal is important. Breadth does so because the system is drawn to the attractor from widely divergent trajectories. Depth does so because the system that has been drawn into the basin tends to stay there.

A weakness of this picture, compared to a phase-space portrait, is that it is not as good at giving a sense of the erratic motion from one attractor to another in a multiple-attractor system. You can regain some of that sense of erratic shifting, however, if you think of the surface in Figure 8.8 as a tambourine being continuously shaken (Figure 8.8, panel D). Even a little shaking causes the ball to bounce around in its well and may jostle it from one well to another, particularly if the attractors are not highly stable. An alternative would be to think of the ball as a jumping bean. These two characterizations would be analogous to jostling from situational influences and jostling from internal dynamics, respectively.

**Goals as Attractors**

The themes of dynamic systems thinking outlined here have had several applications in personality—social and even clinical psychology (Hayes & Strauss, 1998; Mahoney, 1991; Nowak & Vallacher, 1998; Vallacher & Nowak, 1997). Perhaps the easiest application of the attractor concept to self-regulatory models is to link it with the goal concept. Indeed, alert readers will have noticed that we used the same metaphor—gravity and antigravity—in describing both the goal construct at the beginning of the chapter and in describing the attractor concept just earlier.

As we said at the beginning of the chapter, goals are points around which behavior is regulated. People spend much of their time doing things that keep their behavior in close proximity to their goals. It seems reasonable to suggest, then, that a goal represents a kind of attractor. Further, if a goal is an attractor, it seems reasonable that an antigoal would represent a repeller.

This functional similarity between the goal construct and the attractor basin is very interesting. However, the similarity exists only with respect to the end product—that is, maintaining proximity to a value (or remaining distant from a value). The two views make radically different assumptions about the presence or absence of structure underlying the functions. The feedback model assumes a structure underlying and supporting the process, whereas the dynamic systems model does not necessarily incorporate such an assumption.

**CONNECTIONISM**

A related set of questions about the role of central control processes is raised by the literature of connectionism. Connectionist models simulate thought processes in networks of artificial units in which “processing” consists of passing activation among the units. As in neurons, the signal can be excitatory or inhibitory. Energy passes in only one direction (though some networks have feedback links). Processing proceeds entirely by the spread of activation—there is no higher order executive to direct traffic. In a distributed connectionist network, knowledge is not represented centrally, as nodes of information. Rather, knowledge is represented in terms of the pattern of activation of the network as a whole (Smith, 1996).

In networks with feedback relations, once the system receives input, the pattern of weights and activations is updated repeatedly across many cycles. Thus, modifications or updates are made iteratively throughout the network, both with respect to activation in each node and the weighting functions. Gradually, the various values asymptote, and the system “settles” into a configuration. The settling reflects the least amount of overall error the system has been able to create, given its starting inputs and weights.

**Multiple Constraint Satisfaction**

A useful way to think about this process is that the system simultaneously satisfies multiple constraints that the elements create on each other (Thagard, 1989; see also Kelso, 1995).
For example, two mutually inhibitory nodes cannot both be highly active at the same time. Thus they constrain one another. Constraints among multiple nodes are settled out during the repeated updating of activation levels.

This idea of multiple constraint satisfaction is now having a substantial impact on how people in social psychology think about a variety of topics (Kunda & Thagard, 1996; Read, Vanman, & Miller, 1997; Schultz & Lepper, 1996). It is an idea that has a great deal of intuitive appeal. It captures well the introspective sense that people come to conclusions and decisions not by weighing the evidence, exactly, but rather by letting the evidence sort itself until it reaches a degree of internal consistency. The conclusion then pops into mind.

Another term that goes along with this picture is self-organization (e.g., Prigogine & Stengers, 1984). The idea behind this label is that multiple causal forces which have no intrinsic relation to each other can cause the spontaneous emergence of some property of the system as a whole that does not otherwise exist. The term is used to describe emergent qualities in a variety of scientific disciplines. A number of people have begun to invoke it as a basis for emergent properties in dynamic systems (Nowak & Vallacher, 1998; Prigogine & Stengers, 1984).

Self-Organization and Self-Regulation

Some would argue that models of self-organization in dynamic systems represent a serious challenge to the viability of the type of self-regulatory model with which we began. That is, it might be asserted that behavior only seems to be self-regulated—that behavior instead self-organizes from among surrounding forces, like foam appearing on roiling surf.

Do feedback processes actually reflect self-organization—a haphazard falling together of disparate forces? Or are there structures in the nervous system (and elsewhere) in living systems that carry out true feedback functions? In considering the relation between the two sets of ideas, it is of interest that MacKay (1956) anticipated the principle of self-organization many years ago when he described a system of feedback processes that could evolve its own goals (see also Beer, 1995; Maes & Brooks, 1990). Thus, MacKay found the principle of self-organization to be useful, but he found it useful explicitly within the framework of a self-regulatory model.

Our view is, similarly, that the concepts of attractors and trajectories within phase space complement the idea that behavior is guided by feedback processes but do not replace it (Carver & Scheier, in press). There do appear to be times and circumstances in which forces converge—unplanned—and induce acts to occur that were not intended beforehand. However, there also seem to be clear instances of intentionality in behavior and its management.

It is of interest in this regard that contemporary cognitive psychologists often assume the existence of both bottom-up organizational tendencies and top-down directive tendencies (see, e.g., Holyoak & Spellman, 1993; Shastri & Ajjanagadde, 1993; Sloman, 1996; Smolensky, 1988). That view would seem to fit a picture in which self-organization of action can occur, but where actions can also be planned and executed systematically, from the top down. Similar two-mode models of regulation have also appeared in several literatures in personality-social psychology (Chaiken & Trope, 1999). In short, there seems to be some degree of consensus that human experience is part self-organization and part self-regulation.

Even when the focus is on planful behavior, the two kinds of models seem to complement each other in a different way. The feedback model provides a mechanism through which goal-directed action is managed, which the phase-space model lacks. The phase-space model suggests ways of thinking about how multiple goals exist and how people shift among those multiple goals over time, an issue that is not dealt with as easily in terms of feedback processes.

That is, think of the landscape of chaotic attractors, but with many different basins rather than just two or three. This seems to capture rather well the sense of human behavior. No basin in this system ever becomes a point attractor. Behavior tends toward one goal and then another, never being completely captured by any goal. The person does one thing for a while, then something else. The goals are all predictable—in the sense that they all influence the person—and the influence is highly predictable when aggregated across time. But the shifts from one to another occur unpredictably (thus being chaotic).

CATASTROPHE THEORY

Another set of ideas that has been around for some time but may be reemerging in influence is catastrophe theory, a mathematical model that bears on the creation of discontinuities, bifurcations, or splittings (Brown, 1995; Saunders, 1980; Stewart & Perego, 1983; van der Maas & Molenaar, 1992; Woodcock & Davis, 1978; Zeeman, 1977). A catastrophe occurs when a small change in one variable produces an abrupt (and usually large) change in another variable.

An abrupt change implies nonlinearity. This focus on nonlinearity is one of several themes that catastrophe theory shares with dynamic systems theory, though the two bodies of thought have different origins (and are seen by some as quite different from each other—see Kelso, 1995, chap. 2). The similarity is nicely expressed in the statement that the discontinuity in catastrophe theory reflects “the sudden disappearance of one attractor and its basin, combined with the dominant emergence of another attractor” (Brown, 1995, p. 51).
track these points across the surface as higher levels of while the two paths track each other closely, until suddenly levels of ple characterization of what this term means is that at some important feature of a catastrophe known as hysteresis. A sim-

The preceding description also hinted at an interesting and farther along. region of the surface, the other to the lower region. Thus, a relationship becomes even more clearly discontinuous—the out-
multiple intersection exists in which more than one value of y exists. Another way to characterize hysteresis is that two regions of this surface are attractors and one is a repeller (Brown, 1995). This unstable area is illustrated in Figure 8.10. The dashed-line portion of Figure 8.10 that lies between values a and b on the x-axis—the region where the fold is going backward—repels trajectories (Brown, 1995), whereas the areas near values c and d attract trajectories. To put it more simply, you cannot be on the dashed part of this surface.

Yet another way of characterizing hysteresis is captured by the statement that the system’s behavior depends on the system’s recent history (Brown, 1995; Nowak & Lewenstein, 1994). That is, as you move into the zone of variable x that lies between points a and b in Figure 8.10, it matters which side of the figure you are coming from. If the system is moving from point c into the zone of hysteresis, it stays on the bottom surface until it reaches point b, where it jumps to the top surface. If the system is moving from d into the zone of hysteresis, it stays on the top surface until it reaches point a, where it jumps to the bottom surface.

An Application of Catastrophe Theory

How does catastrophe theory apply to the human behaviors of most interest to personality and social psychologists? Several applications of these ideas have been made in the past decade or so, and others seem obvious candidates for future study (for broader discussion see Carver & Scheier, 1998, chap. 16).

One interesting example concerns what we believe is a bifurcation between engagement in effort and giving up. Earlier we pointed to a set of theories that assume such a

Figure 8.9 Three-dimensional depiction of a cusp catastrophe. Variables x and z are predictors, and y is the system’s “behavior,” the dependent variable. The catastrophe shows sensitive dependence on initial conditions. Where z is low, points 1 and 2 are nearly the same on x. If these points are projected forward on the surface (with increases in z), they move in parallel until the cusp begins to emerge. The lines are then separated by the formation of the cusp and project to completely different regions of the surface. Source: From C. S. Carver and M. F. Scheier, On the Self-Regulation of Behavior, copyright 1998, Cambridge University Press. Reprinted with permission.

Figure 8.10 A cusp catastrophe exhibits a region of hysteresis (between values a and b on the x axis), in which z has two stable values of y (the solid lines) and one unstable value (the dotted line that cuts backward in the middle of the figure). The region represented by the dotted line repels trajectories, whereas the stable regions (those surrounding values c and d on the x-axis) attract trajectories. Traversing the zone of hysteresis from the left of this figure results in an abrupt shift (at value b on the x-axis) from the lower to the upper portion of the surface (right arrow). Traversing the zone of hysteresis from the right of this figure results in an abrupt shift (at value a on the x-axis) from the upper to the lower portion of the surface (left arrow). Thus, the disjunction between portions of the surface occurs at two different values of x, depending on the starting point. Source: From C. S. Carver and M. F. Scheier, On the Self-Regulation of Behavior, copyright 1998, Cambridge University Press. Reprinted with permission.

Hysteresis

The preceding description also hinted at an interesting and important feature of a catastrophe known as hysteresis. A simple characterization of what this term means is that at some levels of z, there is a kind of fold-over in the middle of the x-y relationship. A region of x exists in which more than one value of y exists.
conception (Brehm & Self, 1989; Kukla, 1972; Wortman & Brehm, 1975). In all those models (as in ours), there is a point at which effort seems fruitless and the person stops trying. Earlier, we simply emphasized that the models all assumed a disjunction. Now we look at the disjunction more closely and suggest that the phenomena addressed by these theories may embody a catastrophe.

Figure 8.11 shows a slightly relabeled cross section of a cusp catastrophe similar to that in Figure 8.10. This figure displays a region of hysteresis in the engagement versus disengagement function. In that region, where task demands are close to people’s perceived limits to perform, there should be greater variability in effort or engagement, as some people are on the top surface of the catastrophe and others are on the bottom surface. Some people would be continuing to exert efforts at the same point where others would be exhibiting a giving-up response.

Recall that the catastrophe figure also conveys the sense that the history of the behavior matters. A person who enters the region of hysteresis from the direction of high confidence (who starts out confident but confronts many contradictory cues) will continue to display engagement and effort, even as the situational cues imply less and less basis for confidence. A person who enters that region from the direction of low confidence (who starts doubtful but confronts contradictory cues) will continue to display little effort, even as the cues imply a greater basis for confidence.

This model helps indicate why it can be so difficult to get someone with strong and chronic doubts about success in some domain of behavior to exert real effort and engagement in that domain. It also suggests why a confident person is so rarely put off by encountering difficulties in the domain where the confidence lies. To put it in terms of broader views about life in general, it helps show why optimists tend to stay optimistic and pessimists tend to stay pessimistic, even when the current circumstances of the two sorts of people are identical (i.e., in the region of hysteresis).

It is important to keep in mind that the catastrophe cross section (Figure 8.11) is the picture that emerges under catastrophe theory only once a clear region of hysteresis has begun to develop. Farther back, the model is more of a step function. An implication is that to see the fold-over it is important to engage the variable that is responsible for bringing out the bifurcation in the surface (i.e., axis $z$ in Figure 8.9).

What is the variable that induces the bifurcation? We think that in the motivational models under discussion—and perhaps more broadly—the control parameter is importance. Importance arises from several sources, but there is a common thread among events seen as important. They demand mental resources. We suspect that almost any strong pressure that demands resources (time pressure, self-imposed pressure) will induce bifurcating effects.

### CONCLUDING COMMENT

In this chapter we sketched a set of ideas that we think are important in conceptualizing human self-regulation. We believe that behavior is goal directed and feedback controlled and that the goals underlying behavior form a hierarchy of abstractness. We believe that experiences of affect (and of confidence vs. doubt) also arise from a process of feedback control, but a feedback process that takes into account temporal constraints. We believe that confidence and doubt yield patterns of persistence versus giving up and that these two responses to adversity form a dichotomy in behavior. These ideas have been embedded in our self-regulatory viewpoint for some time.

We have also recently begun to consider some newer ideas, addressed in the latter parts of the chapter. In those sections we described ideas from dynamic systems theory, connectionism, and catastrophe theory. We suggest that they represent useful tools for the analysis and construal of behavior. Our view is that they supplement rather than replace the ideas now in use (though not everyone will agree on this point). We see many ways in which those ideas mesh with the ideas presented earlier, though space constraints limited us to discussing that integration only briefly.

In thinking about the structure of behavior, we have tried to draw on ideas from disparate sources while continuing to follow the thread of the logical model from which we started. The result is an aggregation of principles that we think have a good deal to say about how behavioral self-regulation takes place. In so doing, they also say something about personality and how it is manifested in people’s actions.

The conceptual model presented here is surely not complete, and many avenues exist for further discussion and indeed further conceptual development. For example, this chapter included little attention to the issue of how new goals
are added to people’s hierarchies or of how to think about growth and change over time (but see Carver & Scheier, 1998, 1999a, 1999b). Similarly, the concepts addressed here bear in several ways on problems in behavior and behavior change, though space constraints prevent us from describing them in detail. For example, we suspect that many problems in people’s lives are, at their core, problems of disengagement versus engagement and the failure to disengage adaptively (Carver & Scheier, 1998). As another example, it may be useful to conceptualize problems as less-than-optimal adaptations in a multidimensional phase space, which require some jostling to bounce the person to a new attractor (Hayes & Strauss, 1998). These are all areas in which more work remains to be done.

These are just some of the ways in which we think the family of ideas described here will likely be explored in the near future. Further analyses of the self-regulation of behavior are likely to produce insights that transform the models from which the insights grew. As the models change, so will our understanding of motivational processes and of how human beings function as coherent, autonomous units. This we take to be one of the core pursuits of personality psychology.

REFERENCES


INTERPERSONAL FOUNDATIONS FOR AN INTEGRATIVE THEORY OF PERSONALITY

The origins of the interpersonal theory of personality we discuss in the present chapter are found in Sullivan’s (1953a, 1953b, 1954, 1956, 1962, 1964) interpersonal theory of psychiatry. Extensions, elaborations, and modifications have consistently appeared over the last 50 years, with landmark works appearing in each successive decade (see Table 9.1). Given this clear line of theoretical development, it might seem puzzling that in a discussion of the scope of interpersonal theory held at a recent meeting of the Society for Interpersonal Theory and Research (SITAR), it was pointed out that psychology’s expanding focus on interpersonal functioning has rendered study of interpersonal processes so fundamental that interpersonal theory risks an identity crisis (Gurtman, personal communication, June 20, 2000). In our opinion, both promising and perplexing aspects of this identity crisis are respectively reflected in two growing bodies of literature. The former body recognizes the integrative and synthetic potential of interpersonal theory to complement and enhance many other theoretical approaches to the study of personality (e.g., Benjamin, 1996c; Kiesler, 1992), whereas the latter body focuses on interpersonal functioning without any recognition of interpersonal theory.

Explicit efforts have been made toward integration of interpersonal theory and cognitive theory (e.g., Benjamin, 1986; Benjamin & Friedrich, 1991; Carson, 1969, 1982; Safran, 1990a, 1990b; Tunis, Fridhandler, & Horowitz, 1990), attachment theory (e.g., Bartholomew & L. Horowitz, 1991; Benjamin, 1993; Birchnell, 1997; Florsheim, Henry, & Benjamin, 1996; Pincus, Dickinson, Schut, Castonguay, & Bedics, 1999; Stuart & Noyes, 1999), contemporary psychodynamic theory (e.g., Benjamin, 1995; Benjamin & Friedrich, 1991; Heck & Pincus, 2001; Lionells, Fiscalini, Mann, & Stern, 1995; Pincus, 1997; Roemer, 1986), and evolutionary theory (e.g., Hoyenga, Hoyenga, Walters, & Schmidt, 1998; Zuroff, Moskowitz, & Cote, 1999). Although it might be argued that such efforts could lead to identity diffusion of interpersonal theory, we believe this points to the fundamental integrative potential of an interpersonal theory of personality. In contrast, efforts at integrating interpersonal theory with social psychological theories of human interaction and social cognition appear to be lagging despite the initial works of Carson (1969) and Wiggins (1980). We note continued expansion of a significant social psychological literature on interpersonal behavior, such as self-verification and self-confirmation theories (e.g., Hardin & Higgins, 1996; Swann & Read, 1981) and interpersonal expectancies (e.g., Neuberg, 1996), that does not incorporate interpersonal theory as reviewed here. Remarkably, recent reviews of interpersonal functioning (Reis, Collins, & Berscheid, 2000; Snyder & Stukas, 1999) did not cite any of the literature reviewed for the present chapter on interpersonal theory.

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTERPERSONAL FOUNDATIONS FOR AN INTEGRATIVE THEORY OF PERSONALITY</td>
<td>209</td>
</tr>
<tr>
<td>THE INTERPERSONAL SITUATION</td>
<td>210</td>
</tr>
<tr>
<td>THE INTERPERSONAL AND THE INTRAPSYCHIC</td>
<td>211</td>
</tr>
<tr>
<td>DESCRIBING INTERPERSONAL BEHAVIOR</td>
<td>212</td>
</tr>
<tr>
<td>The Individual Differences Approach</td>
<td>212</td>
</tr>
<tr>
<td>The Dyadic Approach</td>
<td>213</td>
</tr>
<tr>
<td>INTERPERSONAL RECIPROCITY AND TRANSACTION</td>
<td>215</td>
</tr>
<tr>
<td>Describing Reciprocal Interpersonal Patterns</td>
<td>215</td>
</tr>
<tr>
<td>Contemporaneous Analysis of Human Transaction</td>
<td>217</td>
</tr>
<tr>
<td>Parataxic Integration of Interpersonal Situations</td>
<td>219</td>
</tr>
<tr>
<td>Some Comments on Interpersonal Complementarity</td>
<td>219</td>
</tr>
<tr>
<td>Summary</td>
<td>221</td>
</tr>
<tr>
<td>THE FUTURE OF INTERPERSONAL THEORY</td>
<td>221</td>
</tr>
<tr>
<td>Interpersonal Theory and Mental Representation</td>
<td>221</td>
</tr>
<tr>
<td>Development and Motivation</td>
<td>222</td>
</tr>
<tr>
<td>Concluding Propositions</td>
<td>223</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>225</td>
</tr>
</tbody>
</table>
nor do interpersonal theorists regularly recognize the social psychological literature on interpersonal interaction in their work (cf. Kiesler, 1996).

Thus, the current state of affairs compels interpersonal theorists to take the next step in defining the interpersonal foundations for an integrative theory of personality. The initial integrative efforts provide a platform to refine the scope of interpersonal theory, and the areas in which integration is lacking indicate that further development is necessary. The goal of this chapter is to begin to forge a new identity for interpersonal theory that recognizes both its unique aspects and integrative potential; in this chapter, we also suggest important areas in need of further theoretical development and empirical research.

THE INTERPERSONAL SITUATION

I had come to feel over the years that there was an acute need for a discipline that was determined to study not the individual organism or the social heritage, but the interpersonal situations through which persons manifest mental health or mental disorder. (Sullivan, 1953b, p. 18)

Personality is the relatively enduring pattern of recurrent interpersonal situations which characterize a human life. (Sullivan, 1953b, pp. 110–111)

These statements are remarkably prescient, as much of psychology in the new millenium seems devoted in one way or another to studying interpersonal aspects of human existence. To best understand how this focus has become so fundamental to the psychology of personality (and beyond), we must clarify what is meant by an interpersonal situation. Perhaps the most basic implication of the term is that the expression of personality (and hence the investigation of its nature) focuses on phenomena involving more than one person—that is to say, some form of relating is occurring (Benjamin, 1984; Kiesler, 1996; Mullahy, 1952). Sullivan (1953a, 1953b) suggested that individuals express “integrating tendencies” that bring them together in the mutual pursuit of both satisfactions (generally a large class of biologically grounded needs) and security (i.e., self-esteem and anxiety-free functioning). These integrating tendencies develop into increasingly complex patterns or dynamisms of interpersonal experience. From infancy onward through six developmental epochs these dynamisms are encoded in memory via age-appropriate learning. According to Sullivan, interpersonal learning of social behaviors and self-concept is based on an anxiety gradient associated with interpersonal situations. All interpersonal situations range from rewarding (highly secure) through various degrees of anxiety and ending in a class of situations associated with such severe anxiety that they are dissociated from experience. Individual variation in learning occurs when maturational limits affect the developing person’s understanding of cause-and-effect logic and consensual symbols such as language (i.e., Sullivan’s prototaxic, parataxic, and syntactic modes of experience), understanding of qualities of significant others (including their “reflected appraisals” of the developing person), as well as their understanding of the ultimate outcomes of interpersonal situations characterizing a human life. Thus, Sullivan’s concept of the interpersonal situation can be summarized as the experience of a pattern of relating self with other associated with varying levels of anxiety (or security) in which learning takes place that influences the development of self-concept and social behavior. This is a very fundamental human experience for psychology to investigate, and it is a significant aspect of the efforts to integrate interpersonal theory with cognitive, attachment, psychodynamic, and evolutionary theories previously noted.

Sullivan (1954) described three potential outcomes of interpersonal situations. Interpersonal situations are resolved when integrated by mutual complementary needs and reciprocal patterns of activity, leading to “felt security” and probable recurrence. A well-known example is the resolution of an infant’s distress by provision of tender care by parents. The infant’s tension of needs evokes complementary parental needs to provide care (Sullivan, 1953b). Interpersonal situations are continued when needs and patterns of activity are not initially complementary, such that tensions persist and covert processing of possible alternative steps toward resolution emerge, leading to possible negotiation of the relationship (Kiesler, 1996). Finally, interpersonal situations are
frustrating when needs and actions are not complementary and no resolution can be found, leading to an increase in anxiety and likely disintegration of the situation.

For Sullivan, the interpersonal situation underlies genesis, development, mutability, and maintenance of personality. The continuous patterning and repatterning of interpersonal experience in relation to the vicissitudes of satisfactions and security in interpersonal situations gives rise to lasting conceptions of self and other (Sullivan’s “personifications”) as well as to enduring patterns of interpersonal relating. To us, the interpersonal situation is at the core of an interpersonal theory of personality. The power of interpersonal experiences to create, refine, and change personality as Sullivan conceived is the foundation of an interpersonal theory of personality that has been elaborated in the last half century by a wide range of theoretical, empirical, and clinical efforts.

A comprehensive theory of personality includes contemporaneous analysis emphasizing present description and developmental analysis emphasizing historical origins as well as the continuing significance of past experience on current functioning (Millon, 1996). Consistent with these approaches, the fundamental aspects of an interpersonal theory of personality should include (a) a delineation of what is meant by interpersonal, (b) the systematic description of interpersonal behavior, (c) the systematic description of reciprocal interpersonal patterns, (d) articulation of processes and structures that account for enduring patterns of relating, and (e) motivational and developmental principles. In our opinion, interpersonal theorists have reached greater consensus on contemporaneous description than on developmental concepts. This consensus may be due in part to ambiguity in the meaning of the term interpersonal.

THE INTERPERSONAL AND THE INTRAPSychIC

Where are interpersonal situations to be found? Millon’s (1996) distinction between contemporaneous and developmental analysis alludes to the dichotomy of the interpersonal and the intrapsychic. Specifically, current description evokes a view of the reciprocal behavior patterns of two persons engaged in resolving, negotiating, or disintegrating their present interpersonal situation. In this sense, we might focus on what can be observed to transpire between them. In contrast, developmental analysis implies that there is something relatively stable that a person brings to each new interpersonal situation. Such enduring influences might be considered to reside within the person—that is, they are intrapsychic. The dichotomous conception of the interpersonal and the intrapsychic as two sets of phenomena—one residing between people and one residing within a person—may have at times led interpersonal theorists to focus more attention on contemporaneous analysis with perhaps greater hesitancy to elaborate on developmental influences. In our opinion, however, we must include developmental concepts if we are to be comprehensive, and this in turn requires examination of intrapsychic structures and processes. As it turns out, Sullivan would not be opposed to such efforts.

Greenberg and Mitchell (1983) point out that Sullivan’s interpersonal theory of psychiatry was largely a response to Freud’s strong emphasis on drive-based intrapsychic aspects of personality. Because of Sullivan’s opposition to drives as the source of personality structuralization, there is a risk of simplifying interpretation of interpersonal theory as focusing solely on what occurs outside the person, in the world of observable interaction. Mitchell (1988) points out that Sullivan was quite amenable to incorporating the intrapsychic into interpersonal theory because he viewed the most important contents of the mind to be the consequence of lived interpersonal experience. For example, Sullivan (1964) states, “. . . everything that can be found in the human mind has been put there by interpersonal relations, excepting only the capabilities to receive and elaborate the relevant experiences” (p. 302; see also Stern, 1985, 1988).

Mitchell (1988) specifies several concepts associated with the dichotomization of interpersonal and intrapsychic, including perception versus fantasy and actuality versus psychic reality. Sullivan clearly viewed fantasy as fundamental to interpersonal situations. He defined psychiatry as the “study of the phenomena that occur in configurations made up of two or more people, all but one of whom may be more or less completely illusory” (Sullivan, 1964, p. 33). These illusory aspects of the interpersonal situation involve mental structures—that is, personifications of self and others. Sullivan (1953b) was forceful in asserting that personifications are elaborated organizations of past interpersonal experience, stating “. . . I would like to make it forever clear that the relation of the personifications to that which is personified is always complex and sometimes multiple; and that personifications are not adequate descriptions of that which is personified” (p. 167). Sullivan also saw subjective meaning (i.e., psychic reality) as highly important. For example, Mitchell (1988) points out that Sullivan’s conception of parataxic integration involves subjective experience of the interpersonal situation influenced by intrapsychic structure and process. Sullivan (1953a) describes parataxic integrations as occurring “when, beside the interpersonal situation as defined within the awareness of the speaker, there is a concomitant interpersonal situation quite different as to its principle integrating tendencies, of which the speaker is more or less
completely unaware” (p. 92). In discussing the data of psychiatry, Sullivan (1964) asserted that “human behavior, including the verbal report of subjective appearances (phenomena), is the actual matter of observation” (p. 34).

Thus, we can assert that interpersonal theory is not strictly an interactional theory emphasizing observable behavior; rather, the term interpersonal is meant to convey a sense of primacy, a set of fundamental phenomena important for personality development, structuralization, function, and pathology. It is not a geographic indicator of locale: It is not meant to generate a dichotomy between what is inside the person and what is outside the person. From a Sullivanian standpoint, the intrapsychic is intrinsically interpersonal, derived from the registration and elaboration of interactions occurring in the interpersonal field (Mitchell, 1988). As we will see, however, descriptions of observable interpersonal behavior and patterns of relating have generated far more consensus among interpersonal theorists than have elaboration of intrapsychic processes and concepts.

DESCRIBING INTERPERSONAL BEHAVIOR

The emphasis on interpersonal functioning in Sullivan’s work stimulated efforts to develop orderly and lawful conceptual and empirical models describing interpersonal behavior. The goal of such work was to obtain a taxonomy of interpersonal behavior—“to obtain categories of increasing generality that permit description of behaviors according to their natural relationships” (Schaefer, 1961, p. 126; see also Millon, 1991, for a general discussion of taxonomy in classification of personality and psychopathology). In contemporary terms, such systems are referred to as structural models, which can be used to conceptually systematize observation and covariation of variables of interest. If sufficiently integrated with rich theory, such models can even be considered nomological nets (Benjamin, 1996a; Gurtman, 1992).

There have been two distinct but related empirical approaches to the development of structural models describing interpersonal functioning. We refer to these as the individual differences approach and the dyadic approach (Pincus, Gurtman, & Ruiz, 1998). These authors pointed out that although each approach has unique aspects, the approaches converge in that they assert that the best structural model of interpersonal behavior takes the form of a circle or circumplex (Gurtman & Pincus, 2000; Pincus et al., 1998; Wiggins & Trobst, 1997). The geometric properties of circumplex models give rise to unique computational methods for assessment and research (Gurtman, 1994, 1997, 2001; Gurtman & Balakrishnan, 1998; Gurtman & Pincus, in press) that are not reviewed here. In the present chapter, circumplex models of interpersonal behavior are used to anchor description of theoretical concepts. The development of circumplex models of interpersonal behavior has significantly influenced contemporary developments in interpersonal theory, and vice versa (Pincus, 1994).

The Individual Differences Approach

The individual differences approach focuses on qualities of the individual, (e.g., personality traits) that are assumed to give rise to behavior that is generally consistent over time and across situations (Wiggins, 1997). From a relational standpoint, this approach involves behavior which is also generally consistent across interpersonal situations, giving rise to the individual’s interpersonal style (e.g., Lorr & Youniss, 1986; Pincus & Gurtman, 1995; Pincus & Wilson, 2001), and in cases of psychopathology, an individual’s interpersonal diagnosis (Kiesler, 1986; Leary, 1957; McLemore & Benjamin, 1979; Wiggins, Phillips, & Trapnell, 1989).

The individual differences approach led to the empirical derivation of a popular structural model of interpersonal traits, problems, and behavioral acts often referred to as the Leary circle (Freedman, Leary, Ossorio, & Coffey, 1951; Leary, 1957) or the Interpersonal Circle (IPC; Kiesler, 1983; Pincus, 1994; Wiggins, 1996). Leary and his associates at the Kaiser Foundation Psychology Research Group observed interactions among group psychotherapy patients and asked, “What is the subject of the activity, e.g., the individual whose behavior is being rated, doing to the object or objects of the activity?” (Freedman et al., 1951, p. 149). This context-free cataloging of all individuals’ observed interpersonal behavior eventually led to an empirically derived circular structure based on the two underlying dimensions of dominance-submission on the vertical axis and nurturance-coldness on the horizontal axis (see Figure 9.1).

The IPC model is a geometric representation of individual differences in a variety of interpersonal domains, including interpersonal traits (Wiggins, 1979, 1995), interpersonal problems (Horowitz, Alden, Wiggins, & Pincus, 2000), verbal and nonverbal interpersonal acts (Gifford, 1991; Kiesler, 1985, 1987), and covert interpersonal impacts (Kiesler, Schmidt, & Wagner, 1997; Wagner, Keisler, & Schmidt, 1995). Thus, all qualities of individual differences within these domains can be described as blends of the circle’s two underlying dimensions. Blends of dominance and nurturance can be located along the 360° perimeter of the circle. Interpersonal qualities close to one another on the perimeter are conceptually and statistically similar, qualities at 90°
are conceptually and statistically independent, and qualities 180° apart are conceptual and statistical opposites. Although the circular model itself is a continuum without beginning or end (Carson, 1969, 1996; Gurtman & Pincus, 2000), any segmentalization of the IPC perimeter to identify lower-order taxa is potentially useful within the limits of reliable discriminability. The IPC has been segmentalized into sixteenths (Kiesler, 1983), octants (Wiggins, Trapnell, & Phillips, 1988), and quadrants (Carson, 1969).

Although the IPC represents a model of functioning in which the individual is presumed to be in many possible interpersonal situations, the model itself is monadic. The IPC structure does not include specific structural or contextual references to the interacting other. Most often, it is used to describe qualities of the individual interacting with a “generalized other” (Mead, 1932; Sullivan, 1953a, 1953b), such as the “hostile-dominant patient” interacting with a generic “psychotherapist” (e.g., Gurtman, 1996; Horowitz, Rosenberg, & Kalehzan, 1992).

The Dyadic Approach

In contrast to the individual differences approach, a second approach assumes that the basic unit of analysis for the study of interpersonal functioning was the dyad. As is the case for the IPC, there is a long history of theoretical and empirical conceptualizations of dyadic interpersonal functioning. At the same time that Leary and his colleagues were investigating individual differences in interpersonal behavior, Schaefer (1959, 1961) began investigating mother-child dyads in an effort to develop a structural model of interpersonal behavior. His methods were similar, but he emphasized the specific dyad as the basic unit of observation: “For maternal behavior, the universe [of content] is the behavior of the mother directed toward an individual child, excluding all other behaviors of the mother” (Schaefer, 1961, p. 126). His work showed a remarkable convergence with Leary (1957)—both investigators found that a two-dimensional circular model best represented interpersonal behavior. As with the IPC, the horizontal dimension was love-hostility. However, the vertical dimension differed, and was labeled autonomy, ranging from autonomy-granting to controlling. Given a dyadic focus, Schaefer (1961) also derived a complementary circular model of children’s behavior in reaction to mothers. Although this early model failed to parallel his maternal behavior model, the notion that parent-like interpersonal behaviors and childlike interpersonal behaviors may be distinguished from each other was an important advance that led to the development of a second prominent circular model of interpersonal behavior from a dyadic point of view.

Structural Analysis of Social Behavior (SASB; Benjamin, 1974, 1984, 1996a, 1996b, 2000) is a complex three-plane circumplex that operationally defines interpersonal and intrapsychic interactions (see Figure 9.2). The dimensions underlying SASB include autonomy (i.e., enmeshment-differentiation on the vertical axis), affiliation (i.e., love-hate on the horizontal axis), and interpersonal focus (i.e., parent-like transitive actions towards others represented by the top circle, childlike intransitive reactions to others represented by the middle circle, and introjected actions directed toward the self represented by the bottom circle). Benjamin (1996c) described the development of SASB as an effort “to combine the prevailing clinical wisdom about attachment with the descriptive power of the circumplex as Schaefer had envisioned it” (p. 1204). The unique multiplane structure of SASB also incorporates Sullivan’s concept of introjection—that is, the expected impact of interpersonal situations on the self-concept—by proposing a third corresponding circle that reflects how one relates to self.

By separating parent-like and childlike behaviors into two planes, SASB incorporates both the vertical dimension of Schaefer’s model (control vs. emancipate) and that of the IPC (dominate vs. submit). The transitive surface represents the former, whereas the intransitive surface opposes submission with autonomy-taking. Thus, according to circumplex geometry, controlling and autonomy-granting are opposite interpersonal actions, whereas submitting and autonomy-taking are opposite interpersonal reactions (Lorr, 1991). Dominance and submission are placed at comparable locations on different surfaces to reflect the fact that they are complementary positions rather than opposites. Thus, SASB expands interpersonal description by including taxa reflecting friendly behavior.
and hostile differentiation (e.g., affirming, ignoring) not defined within the IPC structure, as well as describing the introjected relationship with self. Although the vertical dimensions and complexity of SASB set it apart from the IPC, the same geometric assumptions are applicable. Interpersonal behaviors located along the perimeters of the SASB circles (identified as clusters in SASB terminology) represent blends of the basic dimensions with the same geometric relations among clusters on each surface.

To complete the description, we note that attachment concepts have been incorporated into the SASB structure (Benjamin, 1993, 1996a, Florsheim et al., 1996; Henry, 1994). Boxes in Figure 9.2 denote that interpersonal elements on the right side of the circles (affirm-disclose, reciprocal love, protect-trust) represent the attachment group (AG). Interpersonal elements on the left side of the circles (blame-sulk, attack-recoil, ignore-wall off) represent the disrupted attachment group (DAG).

Using this expanded taxonomy, SASB describes a dyadic interpersonal unit—that is, a real or internalized relationship—rather than the qualities of a single interactant. For example, psychotherapy research using SASB has
focused on the therapist-patient dyad as the unit of investigation (e.g., Henry, Schacht, & Strupp, 1990). Despite these differences, we view the structural models derived from the individual differences and dyadic approaches to be highly convergent in many respects, and they should be viewed as complementary approaches rather than mutually exclusive competitors (e.g., Pincus, 1998; Pincus & Wilson, 2001).

INTERPERSONAL RECIPROCITY AND TRANSACTION

The notion of reciprocity in human relating is reflected in a wide variety of psychological concepts including repetition compulsion (Freud, 1914, 1920), projective identification (Grotstein, 1981), core conflictual relational themes (Luborsky & Cits-Cristoph, 1990), self-fulfilling prophecies (Carson, 1982), vicious circles (Millon, 1996), self-verification seeking (Swann, 1983), and object-relational enactments (Kernberg, 1976), to name a few. If we assume that an interpersonal situation involves two or more people relating to each other in ways that bring about social and self-related learning, this implies that something is happening that is more than mere random activity. Reciprocal relational patterns create an interpersonal field (Wiggins & Trobst, 1999) in which various transactional influences impact both interactants as they resolve, negotiate, or disintegrate the interpersonal situation. Within this field, interpersonal behaviors tend to pull, elicit, invite, or evoke restricted classes of responses from the other, and this is a continual, dynamic transactional process. Thus, an interpersonal theory of personality emphasizes field-regulatory processes over self-regulatory or affect-regulatory processes (Mitchell, 1988).

Sullivan (1948) initially conceived of reciprocal processes in terms of basic conjunctive and disjunctive forces that lead either to resolution or to disintegration of the interpersonal situation. He further developed this in the “theorem of reciprocal emotions,” which states that “integration in an interpersonal situation is a process in which (1) complementary needs are resolved (or aggravated); (2) reciprocal patterns of activity are developed (or disintegrated); and (3) foresight of satisfaction (or rebuff) of similar needs is facilitated” (Sullivan, 1953b, p. 129). Kiesler (1983) pointed out that although this theorem was a powerful interpersonal assertion, it lacked specificity, and “the surviving general notion of complementarity was that actions of human participants are redundantly interrelated (i.e., have patterned regularity) in some manner over the sequence of transactions” (p. 198).

Leary’s (1957) “principle of reciprocal interpersonal relations” provided a more systematic declaration of the patterned regularity of interpersonal behavior, stating “interpersonal reflexes tend (with a probability greater than chance) to initiate or invite reciprocal interpersonal responses from the ‘other’ person in the interaction that lead to a repetition of the original reflex” (p. 123). Learning in interpersonal situations takes place in part because social interaction is reinforcing (Leary, 1957). Carson (1991) referred to this as an interbehavioral contingency process whereby “there is a tendency for a given individual’s interpersonal behavior to be constrained or controlled in more or less predictable ways by the behavior received from an interaction partner” (p. 191).

Describing Reciprocal Interpersonal Patterns

Structural models of interpersonal behavior such as the IPC and SASB have provided conceptual anchoring points and lexicons upon which more systematic description of the patterned regularity of reciprocal interpersonal processes can be articulated (e.g., Benjamin, 1974; Carson, 1969; Kiesler, 1983).

The Interpersonal Circle

Carson (1969) focused on the notion of interpersonal complementarity as the patterned regularity between two people that contributed to “felt security.” This notion is directly related to Sullivan’s conception of a resolved interpersonal situation as an outcome in which both persons’ needs are met via reciprocal patterns of activity leading to its likely recurrence. Anchoring his propositions within the IPC system, Carson first proposed that complementarity was based on the social exchange of status and love, as reflected in reciprocity for the vertical dimension (i.e., dominance pulls for submission; submission pulls for dominance) and correspondence for the horizontal dimension (friendliness pulls for friendliness; hostility pulls for hostility).

Kiesler’s (1983) seminal paper on complementarity significantly expanded these IPC-based conceptions in several ways. First, he recognized the continuous nature of the circular model’s descriptions of behavior, and he noted that because all interpersonal behaviors are blends of dominance and nurturance, the principles of reciprocity and correspondence could be employed to specify complementary points along the entire IPC perimeter. Thus, beyond the cardinal points of the IPC, it was asserted that (for example) hostile dominance pulls for hostile submission, friendly dominance pulls for friendly submission, and so forth, which can be further described by the lower-level taxa in these segments of the model. Second, Kiesler also incorporated Wiggins’ (1979, 1980, 1982) conception of the IPC as a formal geometric...
model into his description of complementarity, whereby the distance from the center of the circle represents a dimension of intensity. That is, complementarity involves both the class of behaviors and their strength. Reciprocity on dominance, correspondence on nurturance, and equivalent intensity thus define complementary behaviors.

In addition, Kiesler (1983, 1996) defined two other broad classes of reciprocal interpersonal patterns anchored by the IPC model. When reciprocal interpersonal patterns meet one of the two rules of complementarity, he referred to this situation as an acomplementary pattern. In such a case, interactants may exhibit correspondence with regard to nurturance or reciprocity with regard to dominance, but not both. When interactants exhibit neither reciprocity on dominance nor correspondence on nurturance, he referred to this situation as an anticomplementary pattern. In Kiesler’s (1996) discussion of these three reciprocal patterns of interpersonal behavior, it is clear that they relate rather directly to the types of outcomes of interpersonal situations suggested by Sullivan. Complementary reciprocal patterns are considered to promote relational stability—that is, such interpersonal situations are resolved, they are mutually reinforcing, and they are recurring. Acomplementary patterns are less stable and instigate negotiation (e.g., toward or away from greater complementarity). Finally, anticomplementary patterns are the most unstable and lead to avoidance, escape, and disintegration of the interpersonal situation.

**SASB**

After developing his two circular models of maternal and child behavior, Schaefer (1961) suggested that relationships between the two surfaces could be the basis for articulating a *theory of influence* of maternal behavior on child behavior, stating

Bowlby (1951) has pointed out that both European and American investigators agree that the quality of parental care has great importance to the development of the child. Less agreement exists about how specific patterns of parent behavior are related to specific patterns of child behavior. One obstacle to the understanding of such relationships has been a lack of knowledge of the interrelations of the concepts within each universe [italics added]. For the purpose of discussion, let us accept the conceptual models presented here and attempt to develop hypotheses concerning the relationship of the two models [italics added]. (pp. 143–144)

Benjamin (1974, 1984, 1996a, 1996b) has extended Schaefer’s proposition by formally articulating a class of reciprocal interpersonal patterns defined by intersurface relationships within the SASB model, referred to as SASB *predictive principles*. The main predictive principles are complementarity, similarity, opposition, antithesis, and introjection, although others may be logically deduced (Schacht, 1994). It is important to note that these principles are not mutually exclusive from those anchored in the IPC model. The first four listed can also be articulated using the IPC. Complementarity implies the very same conditions for an interpersonal situation in both models with content (i.e., differing taxa) being the point of descriptive distinction. As Kiesler (1983) noted, similarity and opposition are specific forms of an acomplementary pattern as defined on the IPC. Antithesis is a form of anticomplementarity from the IPC perspective, again distinctly described using the SASB lexicon. Only introjection cannot be at least partially specified within the IPC model.

Complementarity is based on the relations between transitive and intransitive SASB surfaces; it reflects the typical transactional so-called pulls, bids, or invitations that influence dyadic interactants. It is defined when both members of a dyad are focused on the same person and exhibit comparable amounts of affiliation and autonomy. These can be identified by the numbers indicating the SASB surface (1, 2, or 3) and the cluster (1 through 8) as indicated in Figure 9.2. For example, a therapist focuses on her patient and empathically communicates that she notices an emotional shift (1-2: affirm). In response, the patient focuses on himself and tells the therapist of the associated perceptions, cognitions, wishes, fears, or memories associated with his current affective state (2-2: disclose). All possible complementary positions are marked by taxa appearing in the same locations on surface one and surface two (i.e., attack-recoil, blame-sulk, control-submit, protect-trust, active love-reactive love, affirm-disclose, emancipate-separate, and ignore-wall off). Like the continuous nature of the IPC, the SASB model has several versions, differing in their level of segmentalization and thus precision in terms of their descriptive taxa and predictive principles.

Similarity is exhibited when an individual imitates or acts like someone else—that is, they occupy the same points on the same SASB surface. Imitation, modeling, and observational learning (Bandura, 1977) are important mechanisms in social learning theories that can be described by similarity. However, similarity has a different meaning if it is exhibited by two interactants in an interpersonal situation. If two people rigidly maintain similar positions at the same time, the situation will be rather unproductive—negotiation must occur for there to be much progress. A familiar example is a couple planning their weekend. If both attempt to control (demand their way), there is a power struggle.
little is accomplished as the pattern of What do you want to do?—I don’t know, I’ll do whatever you want cycles and stalls. In an occupational relationship, both boss and employee tend to focus on the employee. The boss controls (in a friendly, neutral, or hostile way) and the employee complies in kind (i.e., complementarity). In contrast, an employee who consistently tries to boss the boss (i.e., similarity) will not be an employee for long!

Points 180° apart describe opposition on each SASB surface. Opposing transitive actions are attack and active love, blame and affirm, control and emancipate, and protect and ignore. Opposing intransitive reactions are recoil and reactive love, sulk and disclose, submit and separate, and trust and wall off. Opposing introjected actions are self-attack and self-love, self-blame and self-affirm, self-control and self-emancipate, and self-protect and self-neglect.

The complementary point of an opposite is its antithesis. Given a particular transitive or intransitive behavior, the antithesis is identified by first locating the behavior’s opposite on the same surface, and then identifying its complement. That is, antithetical points differ in interpersonal focus and are 180° apart. Due to the impact of complementarity (i.e., a bid or invitation), the antithesis is the response that pulls for maximal change in an interpersonal relationship. For example, a psychotherapy patient treated by the first author would frequently sulk (2-6) when she experienced the therapist as not understanding or supporting her (e.g., I don’t know why I come here, this isn’t helping me). Rather than complement this with blame (1-6; e.g., If you don’t try to tolerate not getting exactly what you want from me, this won’t work), the antithetical affirming (1-2) response was enacted, (e.g., I can see that something I have done or failed to do has left you feeling pretty upset). The complement of affirm (1-2) is disclose (2-2). The patient would often visibly relax and communicate her frustration and disappointment. Thus, the antithesis of sulk (2-6) is affirm (1-2). Other antithetical pairs are emancipate and submit, active love and recoil, protect and wall off, control and separate, blame and disclose, attack and reactive love, and ignore and trust.

Introjection is based on the relations between the transitive and introject SASB surfaces and describes the circumstance where an individual treats him- or herself as he or she has been treated by important others. This reflects Sullivan’s view that important aspects of an individual’s self-concept are derived from reflected appraisals of others. That is, the person comes to conceptualize and treat himself in accordance with the ways important others have related to him or her. Common patterns often seen in psychotherapy include depressed patients who recall chronic blame and criticism from parents and now chronically self-blame, and patients with borderline personalities who were physically or sexually abused as children (perpetrator attack) and who now chronically self-attack via cutting or burning. As with complementarity, all introjected positions are marked by clusters in the same location but reflect the pairing of transitive and introject surface descriptors. These include attack and self-attack, blame and self-blame, control and self-control, protect and self-protect, active love and self-love, affirm and self-affirm, emancipate and self-emancipate, and self-protect and self-neglect.

It is important to note that reciprocal interpersonal patterns anchored in either the IPC or SASB are neither inherently good nor inherently bad; they are value-free. In addition, we have tried to present them in their simplest form—as descriptors of behavior patterns that can be observed in interpersonal situations. A taxonomy of reciprocal interpersonal patterns is fundamental to contemporaneous analysis to account for transactional influences occurring in the interpersonal field and to developmental analysis to account for the enduring patterning of interpersonal situations that characterize a human life.

Contemporaneous Analysis of Human Transaction

In examining the immediate interpersonal situation, we may now use the taxonomies of interpersonal behavior and reciprocal interpersonal patterns to provide a contemporaneous analysis of human transaction. The most central pattern discussed previously is that of complementarity, and it is this reciprocal interpersonal pattern that anchors most theoretical discussions of interpersonal interaction. If we are to regard interpersonal behavior as influential or field regulatory, there must be some basic goals toward which our behaviors are directed. Sullivan (1953b) viewed the personification of the self to be a dynamism that is built up from the positive reflected appraisals of significant others, allowing for relatively anxiety-free functioning and high levels of felt security and self-esteem. The self-dynamism tends to be self-perpetuating due to both our awareness and organization of interpersonal experience (input), and the field-regulatory influences of interpersonal behavior (output). Sullivan proposed that both our enacted behaviors and our perceptions of others’ behaviors toward us are strongly affected by our self-concept. When we interact with others, we are attempting to define and present ourselves and trying to negotiate the kinds of interactions and relationships we seek from others. Sullivan’s (1953b) theorem of reciprocal emotion and Leary’s (1957) principle of reciprocal interpersonal relations have led to the formal view that what we attempt to regulate in the interpersonal field are the responses of the other. “Interpersonal behaviors, in a relatively unaware, automatic, and unintended
fashion, tend to invite, elicit, pull, draw, or entice from inter-
actants restricted classes of reactions that are reinforcing of,
and consistent with, a person’s proffered self-definition”
(Kiesler, 1983, p. 201; see also Kiesler, 1996). To the extent
that individuals can mutually satisfy their needs for interac-
tion that are congruent with their self-definitions (i.e., com-
plementarity), the interpersonal situation remains integrated
(resolved). To the extent that this fails, negotiation or disinte-
gration of the interpersonal situation is more probable.

As noted previously, interpersonal theory includes in-
trapsychic elements. The contemporaneous description of the
interpersonal situation utilizing either the IPC or SASB to
delineate behavior and reciprocal patterns is not limited to the
observable behaviors occurring between two people. Thus,
interpersonal complementarity (or any other reciprocal pattern)
should not be conceived of as some sort of stimulus-
response process based solely on overt actions and reactions
(Pincus, 1994). A comprehensive account of the contem-
poraneous interpersonal situation must somehow bridge the gap
between the interpersonal (or overt) and the intrapsychic (or
covert). Interpersonalists have indeed proposed many con-
cepts and processes that clearly imply a rich and meaningful
intrapsychic life (Kielser, 1996; Pincus, 1994), including per-
sonifications, selective inattention, and parataxic distortions
(Sullivan, 1953a, 1953b), covert impacts (Kiesler et al.,
1997), expectancies of contingency (Carson, 1982), fantasies
and self-statements (Brokaw & McLemore, 1991), and cog-
nitive interpersonal schemas (Foa & Foa, 1974; Safran,
1990a, 1990b; Wiggins, 1982). We agree with Safran’s
(1992) conclusion that the “ongoing attempt to clarify the rela-
tionship between interpersonal and intrapsychic levels is
what is needed to fully realize the transtheoretical implica-
tions of interpersonal theory” (p. 105). Much of the field is
moving in this direction, as the relationship between the in-
terpersonal and the intrapsychic is a common entry point for
current integrative efforts (e.g., Benjamin, 1995; Florsheim
et al., 1996, Tunis et al., 1990).

Kiesler’s (1986, 1988, 1991, 1996) interpersonal transac-
tion cycle provides the most articulated discussion of the
relations among overt and covert interpersonal behavior
within interpersonal situations. He proposes that the basic
components of an interpersonal transaction are (a) Person X’s
covert experience of Person Y, (b) Person X’s overt behavior
toward Person Y, (c) Person Y’s covert experience in re-
response to Person X’s action, and (d) Person Y’s overt behav-
ioral response to Person X. These four components are part of
an ongoing transactional chain of events cycling toward res-
olution, further negotiation, or disintegration. Within this
process, overt behavioral output serves the purpose of regu-
lat ing the interpersonal field via elicitation of complementary

overt responses in the other. The structural models of inter-
personal behavior specify the range of descriptive taxa,
whereas the motivational conceptions of interpersonal theory
give rise to the nature of regulation of the interpersonal field.

For example, dominant or controlling interpersonal behavior
(e.g., Do it this way!) communicates a bid for status (e.g.,
I am an expert) that impacts the other in ways that elicit either
complementary (e.g., Can you show me how?) or noncom-
plementary (e.g., Quit bossing me around!) responses in an
ongoing cycle of reciprocal causality, mediated by covert
and subjective experience.

In our opinion, the conceptions of covert processes medi-
at ing behavioral exchange have been a weak link in the in-
terpersonal literature, reflecting much less consensus among
theorists than do the fundamental dimensions and circular
nature of structural models. The diverse conceptualizations
proposed have not been comprehensively related to develop-
mental analyses, nor have their influences on the observable
interpersonal field been fully developed. In a significant step
forward, Kiesler (1996) has synthesized many concepts (i.e.,
emotion, behavior, cognition, and fantasy) in developing the
construct referred to as the impact message (see also Kiesler
et al., 1997). Impact messages are fundamental covert aspects
of the interpersonal situation, encompassing feelings (e.g.,
elicted emotions), action tendencies (pulls to do something;
i.e., I should calm him down or I should get away), perceived
evoking messages (i.e., subjective interpretations of the
other’s intentions, desires, affect states, or perceptions of in-
terpersonal situation), and fantasies (i.e., elaborations of
the interaction beyond the current situation). Kiesler and his col-
leagues view the link between the covert and overt aspects of
the interpersonal situation to be emotional experience. Im-
 pact messages are part of a “transactional emotion process
that is peculiarly essential to interpersonal behavior itself”
(Kiesler, 1996, p. 71). Impact messages are registered
covertly by Person X in response to Person Y’s interpersonal
behavior, imposing complementary demands on the behavior
of Person X through elicited cognition, emotion, and fantasy.
Notably, the underlying structure of impact messages paral-
lels that of the IPC (Kiesler et al., 1997; Wagner et al., 1995),
allowing for description of covert processes that are on a
metric common with the description of overt interpersonal
behavior.

In summary, contemporaneous analysis of the interper-
sonal situation accounts for the patterned regularity of inter-
actions by positing that interpersonal behavior typically
mediates a class of covert responses (impact messages) that
mediate cycles of overt behavior—that is, patterned rela-
tional behavior occurs, in part, due to the field-regulatory in-
fuences of interpersonal behavior on covert experience and
the subsequent mediation of overt action by evoked covert experience. In our opinion, this is only part of the story. Covert responses are intrapsychic phenomena that give rise to subjective experience. It is clear that the nature of such covert responses—that is, feelings, action tendencies, interpretations, and fantasies—are not evoked completely in the moment due to interpersonal behavior of another, but rather arise in part from enduring organizational tendencies of the individual, as the following example illustrates.

Parataxic Integration of Interpersonal Situations

The covert impact messages evoked within a contemporaneous interpersonal transaction cycle are primarily associated with the overt behaviors of the interactants. It is assumed that interactants are generally aware of such covert experience, as the development of the self-report Impact Message Inventory (Kiesler & Schmidt, 1993) suggests. However, Sullivan (1953a) also suggested that other integrating tendencies—beyond those that are encoded within the proximal interpersonal field—often influence the interpersonal situation. Such parataxic distortions may play a more or less significant role in the covert experience of one or the other person in an interpersonal situation.

Clinical Example

A psychotherapy patient treated by the first author entered her therapy session genuinely distraught and depressed. She reported that a person she labeled “an important friend” had ignored her during a recent social gathering and failed to attend a small celebration of her birthday. This was certainly no surprise to me, as this fellow had consistently behaved in an unreliable and invalidating manner toward my patient. However, it again appeared to be a surprise to her, and her disappointment was profound. The immediate interpersonal situation with this patient was quite familiar, and I decided our alliance was now sufficiently established to allow for an empathic effort to confront her continued unrealistic expectations of this fellow and to further examine how her attachment to him seemed to leave her vulnerable to ongoing disappointments.

I responded by saying, “I can understand that what has happened over the weekend has left you hurt, but I wonder why it is that despite repeated similar experiences with this ‘friend,’ you continue to remain attached to him and hope he will give you what you want? It seems to leave you very vulnerable.” My patient responded with sullen withdrawal, curtly remarking, “Now you’re yelling at me just like my mother always does!”

I am fairly certain that had the session been videotaped, there would be no increase in the decibel level of my voice during the intervention. And, an internal scan of my reaction to her report suggested helpful intent rather than countertransferrential punitiveness. Nonetheless, my patient’s response clearly communicated that I was now berating her and putting her down, and that therapy was not supposed to go this way. This continued for several months—any effort I made to examine my patient’s contributions to her difficulties was rebuffed in a similar way. This continued to shape my therapeutic responses, leaving me hesitant to venture in this direction when my patient reported interpersonal difficulties. In other words, the repertoire of therapeutic behaviors I could provide became more and more limited by my patient’s rather rigid behaviors in our relationship. In Kiesler’s (1988) terminology, I was “hooked.”

Several things are apparent from this example. First, using interpersonal structural models to describe the contemporaneous therapeutic transaction, we would see that the relationship was often characterized by noncomplementary responses and by a movement away from an integrated therapeutic relationship. I would try to direct her attention toward herself in an empathic way, and in response my patient would withdraw and threaten to leave. Second, my patient’s response of sullen withdrawal was, however, quite complementary to her subjective experience of me as blaming and punishing. And third, the dynamic interaction between the overt and the covert aspects of our therapeutic transactions continuously exerted field-regulatory influence that allowed the therapy to continue. Too much “yelling and blaming” on my part would lead to a quick termination. My patient did not seem particularly aware of her bids to get me to back off, instead insisting she wanted my help with her depression and interpersonal difficulties.

In our opinion, this example highlights the challenges ahead for fully developing an integrative interpersonal theory of personality. In bridging the interpersonal and the intrapsychic, there are several limitations to contemporaneous analysis, three of which we discuss further in the next section.

Some Comments on Interpersonal Complementarity

The Locus of Influence

Safran (1992) is correct in pointing out that interpersonal theory’s bridge between the overt and the covert requires further development. It is possible that many interpersonal situations generate undistorted, proximal field-regulatory influences—that is, covert experience generally is consistent with overt experience and impact messages reflect reasonably accurate
**The Problem of Complementarocentricity**

Complementarocentricity can be defined as the tendency to place complementarity at the center of interpersonal theory and research. In our opinion, this overemphasis has limited the growth of theory. Three examples of complementarocentricity are as follows:

1. What does failure to find empirical support for interpersonal complementarity mean? When empirical studies do not confirm the existence of complementarity, investigators often label it “a failure to statistically support complementarity” (e.g., Orford, 1986). Even when empirical investigations do find significant results (e.g., Gurtman, 2001), they are not indicative of 100% lawfulness. In our opinion, the answer to our question is that other reciprocal interpersonal patterns are also occurring in the interpersonal situation(s) under investigation.

2. In perhaps the most influential articulation of complementarity, Kiesler (1983) defined all reciprocal patterns in relation to complementarity. That is, other forms of reciprocal interpersonal patterns are said to take either a complementary or anticomplementary forms. We wonder if this has inadvertently promoted complementarity as a more fundamental reciprocal interpersonal pattern than it actually should be.

3. In his encyclopedic review of complementarity theory and research, Kiesler (1996) presented 11 propositions to define and clarify the nature of, scope, and generizability of complementarity, and nine counterpoints to Orford’s (1986) famous critique of complementarity. In this work, Kiesler summarized important contributions by many interpersonalists emphasizing situational, personological, and intrapsychic moderators of complementarity (e.g., see Tracey, 1999), and suggested that significant attention be directed toward articulating when and under what conditions complementarity should and should not be expected to occur. Although this is exceptionally important, it continues to reflect complementarocentric thinking in that what is not recognized is that Kiesler’s (1996) 11 propositions, nine counterpoints, and continuing investigation of moderators serve to decentralize complementarity as the fundamental reciprocal interpersonal pattern by suggesting that its occurrence is more limited and contextualized. For example, consider Proposition 11 regarding “appropriate situational parameters” from Kiesler (1996): “The condition of complementarity is likely to obtain and be maintained in a dyadic relationship only if the following conditions are operative: a) the two participants are peers, b) are of the same gender, c) the setting is unstructured, and d) the situation is reactive (the possibility of reciprocal influence exists)” (p. 104). Considered alone, complementarity is thus suggested to be most applicable to understanding the unstructured interactions of same-sex peers. This is certainly important, but is perhaps not the core phenomenon of interest for a comprehensive theory of personality.

**The Problem of Motivation**

The two core theoretical assertions associated with interpersonal complementarity are Sullivan’s theorem of reciprocal emotion and Leary’s principle of reciprocal interpersonal relations. With regard to the former, we suggest that interpersonal theorists have overemphasized Sullivan’s first point (i.e., complementary needs are resolved or aggravated) and underemphasized his second point (i.e., reciprocal patterns of activity are developed or disintegrated). It is important to note that the needs involved are left undefined, and that the nature of satisfaction in the Sullivanian system involves a global sense of felt security marked by the absence of anxiety. Leary’s principle provided an important extension in its emphasis on interpersonal influence and reinforcement that shapes the nature of ongoing interpersonal situations. But to what end? What is behavior’s purpose? Traditionally, the cornerstone of complementarity has been the assertion that behavior is enacted to invite self-confirming reciprocal responses from others. We believe this has also been overemphasized in the interpersonal literature.
We agree that reciprocal interpersonal influence, reinforcement, and gratification are central to understanding human personality. This is reflected in the large number of psychological concepts that in some way reflect the notion of reciprocity. That is, individuals develop some consistently sought-after relational patterns and some strategies for achieving them. However, we do not believe that a single superordinate motive such as self-confirmation will succeed in comprehensively explaining how personality develops and is expressed.

**Summary**

Our discussion of interpersonal reciprocity and transaction has highlighted many of the unique strengths of interpersonal theory, as well as areas in which significant development and synthesis are necessary. In our view, interpersonal theory emphasizes relational functioning in understanding personality; this emphasis has led to the development of well-validated structural models that provide anchors to systematically describe interpersonal behavior and the patterned regularity of human transaction. Interpersonal theory has also emphasized field-regulatory aspects of personality in addition to the more traditional drive, self, and affect-regulatory foci of most theories of personality. The combination of descriptive structural models and clear focus on the interpersonal situation provides a rich nomological net that has had a significant impact in psychology, particularly with regard to the classification of personological and psychopathological taxa and the contemporaneous analysis of human transactions and relationships. However, we also feel that the future of interpersonal theory will require continuing efforts to address (a) the intrapsychic or covert structures and processes involved in human transaction, (b) the overemphasis on complementarity as the fundamental reciprocal interpersonal pattern in human relationship, (c) the overemphasis on self-confirmation as the fundamental motive of interpersonal behavior, and (d) the lack of a comprehensive developmental theory to complement its strength in contemporaneous analysis.

**THE FUTURE OF INTERPERSONAL THEORY**

We believe the future of interpersonal theory is bright. Addressing the four major issues previously noted will require interpersonal theorists to continue efforts at integrating interpersonal theory’s nomological net with the wisdom contained in the cognitive, psychodynamic, and attachment literature. Fortunately, this is already beginning to take place.

Benjamin (1993, 1995, 1996a, 1996b) has initiated this with her interpersonal “gift of love” theory that integrates the descriptive precision of the SASB model with intrapsychic, motivational, and developmental concepts informed by attachment, cognitive, and object-relations theories.

**Interpersonal Theory and Mental Representation**

We have previously asked the question Where are interpersonal situations to be found? Our answer is that they are found both in the proximal relating of two persons and also in the minds of individuals. There are now converging literatures that suggest mental representations of self and other are central structures of personality that significantly affect perception, emotion, cognition, and behavior (Blatt, Auerbach, & Levy, 1997). Attachment theory refers to these as internal working models (Bowlby, 1969; Main, Kaplan, & Cassidy, 1985), object-relations theory refers to these as internal object relations (Kernberg, 1976), and cognitive theory refers to these as interpersonal schemas (Safran, 1990a). Notably, theorists from each persuasion have observed the convergence in these concepts (Blatt & Maroudas, 1992; Bretherton & Munholland, 1999; Collins & Read, 1994; Diamond & Blatt, 1994; Fonagy, 1999; Safran & Segal, 1990; Westen, 1992). Benjamin (1993, 1996a, 1996b) has also proposed that mental representations of self and other are central to the intrapsychic interpersonal situation. She refers to these as important people or their internalized representations, or IPIRs. Thus, whether referred to as internal working models, internal object relations, interpersonal schemas, or IPIRs, psychological theory has converged in identifying mental representations of self and other as basic structures of personality.

In our opinion, the fundamental advantage of integrating conceptions of dyadic mental representation into interpersonal theory is the ability to import the interpersonal field (Wiggins & Trobst, 1999) into the intrapsychic world of the interactants (Heck & Pincus, 2001). What we are suggesting is that an interpersonal situation can be composed of a proximal interpersonal field in which overt behavior serves important communicative and regulatory functions, as well as an internal interpersonal field that gives rise to enduring individual differences in covert experience through the elaboration of interpersonal input.

In addition, Benjamin’s conception of IPIRs retains interpersonal theory’s advantage of descriptive precision based on the SASB model (Pincus et al., 1999). Benjamin (1993, 1996a, 1996b) proposes that the same reciprocal patterns that describe the interactions of actual dyads may be used to describe internalized relationships (mental representations...
of self and other) on the common metric articulated by the SASB model (see also Henry, 1997). In our view, this adds explanatory power for interpersonal theory to account for individuals’ enduring tendencies to organize interpersonal information in particular ways. Although the concept of the impact message is extremely useful in identifying the classes of covert cognitive, affective, and behavioral experiences of individuals, it does not necessarily account for the nature of individual differences in covert experiences. Benjamin’s IPIRs provide a way to account for the unique and enduring organizational tendencies that people bring to interpersonal situations—experiences that may underlie their covert feelings, impulses, interpretations, and fantasies in relation to others. Interpersonal theory proposes that overt behavior is mediated by covert processes. Psychodynamic, attachment, and cognitive theories converge with this assertion, and they suggest that dyadic mental representations are key influences on the subjective elaboration of interpersonal input. In our opinion, Benjamin has advanced interpersonal theory by incorporating mental representations explicitly into the conception of the interpersonal situation.

Returning briefly to our clinical example, recall that the patient consistently came into therapy reporting disappointments in her interpersonal relations. In telling her sad stories, she communicated her need to be consoled and nurtured. When she was asked to reflect on her own contributions to her disappointments, she became sullen and withdrawn. This reaction was a bid at negotiation, communicating a threat to leave in an effort to reestablish a reciprocal pattern of satisfying responses from her therapist. Why was this happening, given that the therapist attempted to provide recognition and consolation of her hurt feelings? Despite good therapeutic intentions, efforts to focus her attention on her own patterns seemed unhelpful. There was a clue in her report of her subjective covert experience. When the therapist turned the focus toward the patient’s contributions to her relational difficulties, he was experienced as similar to her mother. The proximal interpersonal field was no longer the primary source of her experience. There was now a second, parataxic integration of the situation that led to a covert experience that was driven by previous lived interpersonal experiences that now influenced the patient’s subjective experience; this became the primary mediating influence on her overt behavior. Despite her requests for help and consistent attendance in therapy, the patient was having difficulty organizing her experience of the therapist independently of her maternal IPIR. In our view, this example demonstrates that noncomplementary reciprocal interpersonal responses in the proximal interpersonal field may indicate significantly divergent experiences within the internal interpersonal field that can best be described by integrating interpersonal theory’s structural models with concepts of mental representation.

**Development and Motivation**

Adding conceptions of dyadic mental representation is not sufficient for a comprehensive interpersonal theory of personality. Sullivan (1964), Stern (1988), and others have suggested that the contents of the mind are in some way the elaborated products of lived interpersonal experience. A comprehensive interpersonal theory must account for how lived interpersonal experience is associated with the development of mental representation. In our opinion, Benjamin has provided the only comprehensive developmental approach to evolve from interpersonal theory.

Using SASB as the descriptive anchor (Figure 9.2), Benjamin (1993, 1996a, 1996b) has proposed three developmental copy processes that describe the ways in which early interpersonal experiences are internalized. The first is identification, which is defined as treating others as one has been treated; this is associated with the transitive SASB surface. To the extent that individuals strongly identify with early caretakers (typically parents), there will be a tendency to act toward others in ways that copy how important others have acted toward the developing person. The second copy process is recapitulation, which is defined as maintaining a position complementary to an IPIR; this is associated with the intransitive SASB surface and can be described as reacting as if the IPIR were still there. The third copy process is introjection, which is defined as treating the self as one has been treated. This is associated with the introject SASB surface and is related to Sullivan’s conceptions of “reflected appraisals” as a source of self-personification.

Identification, recapitulation, and introjection are not incompatible with Kiesler’s conception of covert impact messages. In fact, we suggest that the proposed copy processes can help account for individual differences in covert experience by providing developmental hypotheses regarding the origins of a person’s enduring tendencies to experience particular feelings, impulses, cognitions, and fantasies in interpersonal situations. For the patient described earlier, it seems that her experience of the therapist as yelling and blaming reflects (in part) recapitulation of her relationship with her mother. This in turn leads to a parataxic distortion of the proximal interpersonal field in therapy and noncomplementary overt behavior.

Although the copy processes help to describe possible pathways in which past interpersonal experience is internalized into mental structures (IPIRs), it is still insufficient to explain why early IPIRs remain so influential. The answer to
this question requires a discussion of motivation. Whereas Sullivan’s legacy has led many interpersonal theorists to posit self-confirmation as the core motive underlying human transaction, Benjamin (1993) proposed a fundamental shift toward the establishment of attachment as the fundamental interpersonal motivation. In doing so, she has provided one mechanism to account for the enduring influence of early experience on mental representation and interpersonal behavior. Although a complete description of attachment theory is beyond the scope of the present chapter, we agree that attachment to proximal caregivers in the early years of life is both an evolutionary imperative (e.g., Belsky, 1999; Bowlby, 1969; Simpson, 1999) and a primary organizing influence on early mental representation (Beebe & Lachmann, 1988a, 1988b; Bowlby, 1980; Stern, 1985).

Infants and toddlers must form attachments to caregivers in order to survive. Benjamin has suggested that the nature of the early interpersonal environment will dictate what must be done to establish attachments. These early attachment relationships can be described using the SASB model’s descriptive taxa, predictive principles, and copy processes. The primacy of relationships to IPIRs is thus associated with the need to maintain attachment to them even when not immediately present. Benjamin (1993) refers to this as maintaining “psychic proximity” to IPIRs. The need to maintain psychic proximity is organized around wishes for love and connectedness (secure attachment or AG on the SASB model), as well as fears of rejection and loss of love (disrupted attachment or DAG on the SASB model). The primacy of early attachment patterns and mental representations influencing current experience is consistent with psychodynamic and attachment theories. Bowlby (1980) suggested that internal working models act conservatively; thus, assimilation of new experience into established schemas is typical (see also Stern, 1988). Benjamin (1996a) suggested that “psychic proximity fulfills the organizing wish to receive love from the IPIR . . . acting like the IPIR, acting like the IPIR were present, or treating the self as would the IPIR can bring about psychic proximity” (p. 189).

Returning again to the patient described earlier, it was clear that she was ambivalently but strongly attached to her mother. She consistently experienced blame any time she attempted to convey interpersonal disappointments or bad feelings. Anything that disrupted her mother’s sense of control over the world was met with the accusation that the patient was being selfish and immature—and that it was the patient’s fault, so her feelings were not valid. In addition, she was told that if she didn’t stop causing so much trouble, her parents might divorce. It became clear that the patient had internalized a critical maternal IPIR. Whenever the patient was asked about her experience of self, she would inevitably begin her response with “My mother says that I am . . . ” or “My mother says it’s bad for me to feel this way.” When the therapist would try to explore the patient’s contributions to her interpersonal difficulties, it evoked recapitulation. Despite affirming and affiliative efforts on the part of the therapist, the patient had a difficult time accommodating the new interpersonal input; instead she covertly experienced psychic proximity to the critical maternal IPIR and responded in kind. She experienced the therapeutic interpersonal situation as if the maternal IPIR were present, and she needed to back down rather than own her disappointments. To do otherwise would risk her attachment to her mother, painful as it was.

Concluding Propositions

Benjamin’s developmental and motivational extensions of interpersonal theory provide some of the richest advances to date. We see her work, along with Kiesler’s recent integration of emotion theory into the interpersonal transaction cycle, as solid evidence that interpersonal theory as originally conceived by Sullivan has a vital and promising future as a fundamental and integrative approach to personality. In this vein we would like to close this chapter with a further extension of these contemporary works.

Interpersonal theorists are interested in understanding why certain reciprocal interpersonal patterns become prominent for an individual. Benjamin has made an important start by suggesting that a basic human motivation is attachment and that the interpersonal behaviors and reciprocal interpersonal patterns (described by interpersonal theory’s unique structural models) that help achieve attachment become fundamental to personality through internalization of relationships (characterized by the copy processes). She posits that the wish for attachment and the fear of its loss are universal, and that positive early environments lead to secure attachments and normal behavior (i.e., AG). If the developing person is faced with achieving attachment in a toxic early environment, behavior will be abnormal (DAG), but will develop in the service of attachment needs and be maintained via internalization.

We would like to extend this further in an effort to generate an interpersonal theory of personality that more broadly addresses issues of basic human motivation. It is our contention that the maturational trajectory of human life allows us to conceptualize many developmentally salient motives that may function to mediate and moderate current interpersonal experience. That is, reciprocal interpersonal patterns develop in concert with emerging motives that take developmental priority, thus expanding the goals that underlie their
formation and maintenance. We can posit core issues likely to elicit the activation of central reciprocal patterns and their associated IPIRs, potential developmental deficits associated with early experiences, and unresolved conflicts that continue to influence the subjective experience of self and others. The output of such intrapsychic structures and processes for individuals are those consistently sought-after relational patterns and their typical strategies for achieving them (i.e., proximal and internal field regulation). These become the basis for the recurrent interpersonal situations that characterize a human life.

It is our view that what catalyzes and reinforces identification, recapitulation, and introjection is the organizing power of developmental achievements and traumatic stressors. Although interpersonalists have discussed differential “evoking power” of behavior due to situational constraints and the quality of interactions (i.e., moderators of complementarity), we believe such evoking power is limited in comparison to the catalyzing effects of major personality developments and their underlying motivational influences. At different points in personality development, certain motives become a priority. Perhaps initially the formation of attachment bonds and security are primary motivations; but later, separation-individuation, self-esteem, mastery of unresolved conflicts, and identity formation may become priorities (see Table 9.2). If we are to understand the reciprocity seeking, field-regulatory strategies individuals employ, we must learn what interpersonal behaviors and patterns were required to achieve particular developmental milestones. In this way, we see that what satisfies a need or achieves an important goal for a given individual is strongly influenced by his or her developmental history. In addition to developmental achievements, traumatic learning may also catalyze the internalization of patterns associated with coping responses to early loss of an attachment figure, severe physical illness in childhood, sexual or physical abuse, and so on.

Integrating the developmental and traumatic catalysts for internalization of reciprocal interpersonal patterns allows for greater understanding of current behavior. If individuals have the goal of individuating the self in the context of a current relationship in which they feel too enmeshed, they are likely to employ strategies that have been successful in the past. Some individuals have internalized hostile forms of differentiation such as walling off, whereas others have internalized friendly forms of differentiation such as asserting their opinions in an affiliative manner. The overt behavior of the other is most influential as it activates a person’s expectations, wishes, and fears associated with current goals, needs, and motives; this will significantly influence their covert experience of impact messages. In our opinion, the most important goals, needs, and motives of individuals are those that are central to personality development.

A brief example highlights this point and provides some clues as to why individuals may repeat maladaptive interpersonal behaviors over and over. Another psychotherapy patient treated by the first author was severely sexually and emotionally abused by multiple family members while she was growing up. The predictive principle of opposition to what she experienced as a child characterized her transitive actions towards others in the present. In all dealings with others she was hyper-loving and hyper-protective, even when clearly to her detriment. She compulsively exhibited such behaviors, even when treated badly by others. In therapy, it became clear that she counteridentified with her perpetrators and chronically exhibited the opposite pattern in order to maintain a conscious sense of individuation. It was as if she were saying, “If I allow myself to become even the slightest bit angry or blaming, it will escalate and I’ll be just like those who hurt me in the past.” Unfortunately, although she could shed tears for the victims of the holocaust and the victims of the recent epidemic of school shootings, she could not do so for herself. She had also introjected her early treatment within the family and continued to self-injure and ignore her own needs and basic human rights. Thus, although she consciously behaved in ways that individuated her from her abusers, she also abused and neglected herself in ways that unconsciously maintained attachment to her abusive IPIRs (see Table 9.3).

We end this chapter with a bit of speculation. A broader taxonomy of reciprocal interpersonal patterns such as SASB predictive principles and copy processes, combined with a theory of personality development and motivation, can be the basis for understanding both personality and its pathology. Obviously this approach could take many forms. From the contemporary interpersonal perspective developed in this chapter, a basic approach would be an open system with consideration of IPIR-Goal linkages associated with fundamental developmental achievements and traumatic learning. We could also

<table>
<thead>
<tr>
<th>TABLE 9.2 Some Possible Catalysts of Internalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developmental Achievements</td>
</tr>
<tr>
<td>Attachment</td>
</tr>
<tr>
<td>Security</td>
</tr>
<tr>
<td>Separation-individuation</td>
</tr>
<tr>
<td>Positive affects</td>
</tr>
<tr>
<td>Gender identity</td>
</tr>
<tr>
<td>Resolution of Oedipal issues</td>
</tr>
<tr>
<td>Self-esteem</td>
</tr>
<tr>
<td>Self-confirmation</td>
</tr>
<tr>
<td>Mastery of unresolved conflicts</td>
</tr>
<tr>
<td>Identity formation</td>
</tr>
</tbody>
</table>
consider individual differences in the influence of certain copy processes, such that personalities are classified as highly recapitulating, highly introjective, and so on. Similarly, we could consider individual differences in the tendency to enact certain reciprocal interpersonal patterns, such that personalities are differentiated by their tendencies to exhibit oppositional, complementary, antithetical, similar, or introjected behaviors. Although these final thoughts are purely speculative, we wish to emphasize our hope that the ideas presented throughout this chapter provide the interpersonal foundations for an integrative theory of personality.

REFERENCES


Tracey, T. J. G. (1999). *Moderators of complementarity: Not all examinations of complementarity are equivalent or valid*. Paper presented at the meeting of Society for Interpersonal Theory and Research, Madison, WI.


CHAPTER 10

Structures of Personality Traits

WILLEM K. B. HOFSTEE

CONSTRUCTING PERSONALITY

THROUGH QUESTIONNAIRES 232
The Hegemony of Questionnaires 232
Definitions of Personality by Self and Others 233
THE FUTURE OF THE FIVE-DIMENSIONAL MODEL 234
THE LEXICAL BASE OF THE FIVE-DIMENSIONAL MODEL 235
The Lexical Axiom 235
Operationalizations of the Lexical Approach 235
The Credentials of the Fifth Factor 236
THE LINEAR APPROACH TO THE CONCEPT OF PERSONALITY 237
The Case for Principal Component Analysis 238
A Review of the Grounds for the Number Five 239

The Person-Centered or Typological Approach 239
HIERARCHICAL AND CIRCUMPLEX STRUCTURES 242
The Principal Component Analysis Plus Varimax Taxonomic Model 242
The Double Cone Model 244
Generalized Circumplexes 245
A FAMILY MODEL OF TRAIT STRUCTURE 248
The Primordial One 249
The Two-Dimensional Level 250
Semicircumplex Spheres and Hyperspheres 250
CONCLUSION 252
REFERENCES 252

Operations reshape concepts. Over the past decades, the very concept of personality has been subject to implicit redefinition through a set of operations labeled the Big Five taxonomy or the five-factor model of personality. In a restricted sense, the number five refers to the finding that most of the replicable variance of trait-descriptive adjectives in some Western languages is caught by five principal components whose varimax rotations are named extraversion, agreeableness, conscientiousness, emotional stability, and intellect (or openness to experience, autonomy, imagination, and so on, depending on operational variations). In a wider sense, however, the five-dimensional (5-D) approach has come to represent no less than a paradigm—in particular, a revival of the individual-differences or trait conception of personality. For an evaluation of its status and future perspectives, a systematic analysis of its operational credentials is in order.

A first module of the set of operations that constitute the 5-D paradigm consists of the questionnaire construction of personality, whereby someone’s personality is defined through his or her own answers, or more exceptionally through the answers given by third persons, to standardized questions. The questionnaire approach is not confined to the 5-D tradition, but it has to a significant extent been taken over by that paradigm (the megamerger impressing some as monopolistic). Is there a viable alternative to the questionnaire method, and if so, would it change our view of personality?

A second, more specific, operational module contains ways of choosing personality descriptors. The general guiding principle in this module is the lexical approach that consists of selecting items from a corpus of language, particularly a dictionary of that language. The distinguishing characteristic of the lexical approach is its purposely inductive nature, in contrast to approaches in which the descriptor base is deduced from particular trait constructs, for example, neuroticism. Again, the leading question is about the impact of these operations on our conception of personality.

A third operational characteristic consists of reliance on the linear model, particularly, principal component analysis (PCA) of Likert item scales. This is probably the most

The author is greatly indebted to Lewis R. Goldberg, Gerard Saucier, and Jos M. F. Ten Berge for their incisive comments on a draft of this chapter.
There is an obvious alternative to the individual himself or herself as a responder, namely, others who know the person well.

The Hegemony of Questionnaires

The association between personality and questionnaires is not merely a matter of fashion or a historical coincidence. To assess someone’s personality, we have to ask questions about it—to the person himself or herself, to third parties who know the person well, to expert observers. Between the investigator or practitioner on the one hand and the person on the other, there is an indispensable assessor. So-called behavior observations, for example, are not objective in the way they would be if behavior recordings were translated into a score without the intervention of an observer; they represent answers to questions put to a human assessor. Moving from asking questions to applying a questionnaire is a small step: A systematic approach to personality requires standard questions, and thus a questionnaire. Using an unstructured interview, for example, means obtaining answers to an imperfectly standardized set of questions.

One seeming exception is self-report, in which person and assessor coincide. Failure to distinguish between the two roles, however, would amount to denying that the assessor could be someone else, thereby abandoning personality as an intersubjective phenomenon. Another more interesting apparent exception to questionnaire use is expert clinical diagnosis, in which practitioner and assessor coincide. In the first place, however, that process may be reconstructed in part as giving answers to more or less standardized questions about the person that the diagnostician has learned to ask to himself or herself. Second and more fundamental, the diagnostician could have been another individual. By virtue of that exchangeability, a case can be made for maximizing the intersubjective character of diagnoses. Actually using a standardized set of questions (e.g., a personality questionnaire phrased in the third person singular) to guide and articulate one’s diagnostic impressions would contribute to that end. This is not to deny the heuristic element in clinical diagnosis, or in any other applied setting, but to document the central place of asking questions to third persons in the systematic study of personality.

The reason for the primacy of questionnaires may of course be sought in a tendency of students of personality to take things easy: There is nothing more convenient than giving a self-report questionnaire to a client or applicant. But more valid reasons may be brought forward. There is a tension between the concepts of “test” and “personality.” Surely, we may decide to assess a person’s typical intelligent behavior by...
means of a questionnaire (e.g., Goff & Ackerman, 1992), or test the maximal introversion of which he or she is capable (see Riemann, 1997), but neither of these crossovers has appeared to be adequate or promising. Ability and tests of maximal performance, and personality and assessments of typical behavior, are associated in a nonarbitrary manner (Hofstee, 2001).

Are Questionnaires There to Stay?

The prime product of the 5-D paradigm consists of questionnaires, including most notably the Neopersonality inventives-PI-R and NEO-FFI (Costa & McCrae, 1992), and includes many other questionnaires and trait adjective lists; the model has thus given a significant boost to the questionnaire construction of personality. I have argued in brief that the relation between personality traits and the questionnaire operationalization is intimate. Should one be happy with the prospect of such an essentially monomethod definition of personality, and if not, can alternatives be foreseen?

Asking questions to third persons in order to assess personality implies a social definition of it. Surely, the field has moved beyond the stage at which personality was deemed to be merely in the eye of the beholder; cumulative behavior-genetic research (see, e.g., chapter by Livesley, Jang, & Vernon in this volume; Loehlin, 1992) has put an end to that subjective conception of traits. But the dominant conception of personality remains social in the sense of intersubjective rather than objective. Buss (1996) made a virtue of this need by explaining the Big Five as elementary social mechanisms; for example, Factor III represents the need of the perceiver to know whether the other person can be depended upon. Most students of personality, however, would have hesitations with this subordination of personality to social psychology, especially if that bondage is a side effect of a dominant operational approach.

The scientific emancipation of a subjective or intersubjective concept appears to hinge upon the discovery of objective indicators that cover the concept well. If we wish to establish how much of a fever we run, we do not use a Likert scale but measure it with a thermometer. If we want to gauge an applicant’s intelligence, we apply a test rather than asking questions to the applicant or even to a number of third persons. If the latter example is more problematic than the first, that is because there may be doubt regarding the coverage of the concept of intelligence by an IQ score. In the same vein, one may have doubts about the thermometer scale as a measure of outdoor temperature and prefer a formula that includes sunshine, humidity, and wind force. But once a certain level of coverage is secured, a return to sheer subjectivity would count as regressive. Are adequate objective indicators of personality traits in sight?

Probably the most promising indicators of personality are genes. According to estimates based on behavior-genetic research, genetic patterns will be capable of covering some 40% of the trait variance. That degree of coverage is not enough; we would not accept a thermometer that is only 40% valid. But before discarding the prospect, one should realize that the figure of .4 is heavily attenuated. An indicator need not and should not predict the error components in subjective assessments of temperature or extraversion. Heredity coefficients in the order of .4 should thus be divided by an estimate of the proportion of valid variance in questionnaire scores.

The first source of error in the self-reports that have almost invariably been used in behavior-genetic studies of personality is lack of agreement between assessors. The highest agreement coefficients between self and other in assessing personality (Hendriks, 1997; McCrae & Costa, 1987) are in the order of .7. Unless it is assumed that self is a systematically better assessor than other or vice versa, that figure may be taken as an estimate of the rater reliability of a single respondent, and some 30% of the questionnaire variance is rater error. Second, some 20% of the variance results from lack of internal consistency of the questionnaire scale, assuming alpha reliabilities in the order of .8; and third, a comparable error component results from temporal instability. Taking all these independent sources of error into account, one is left wondering how the heredity coefficients can reach .4 at all (Hofstee, 1994a).

The ironic conclusion from this crude analysis of error components in questionnaire variance is that the perspective of molecular-genetic diagnosis of personality traits cannot at all be discarded: It may well appear that whatever valid variance remains in questionnaire data can be accounted for to a satisfactory extent by genetic configurations. However, the analysis also points to the conditions for such a development. To establish links between genes and phenotypic personality traits, the assessment of the latter will have to be much more valid than it has been up to now (see also Bouchard, 1993). The central element of that program is discussed in the next paragraphs. Another aspect—optimizing the internal consistency of questionnaire data—is treated in the section on the linear approach to personality.

Definitions of Personality by Self and Others

Self-report fosters a conception of personality whereby the individual knows best how he or she is. With self-report
questionnaires, the situation is more complicated. Standardized questions aim at comparing personalities rather than capturing unique and emergent characteristics. McAdams’s (1992) criticism of the Big Five approach as a psychology of the stranger is correct in that sense (although other phrasings might be preferred if the value of scientific objectivity is stressed); it would be even more correct if the emphasis in Big Five research were on other-report rather than self-report. Self-report questionnaires embody a discordant blend of subjective and intersubjective accents.

In preparing an earlier (Hofstee, 1994a) paper on the topic, I met with unexpectedly ardent arguments in favor of self-report from prominent American Big Five researchers, the essence of which is documented in that paper. One argument pertained to personal secrets, whose content, however, would not be central to personality in most definitions. (A person might be said to be secretive, but that trait hardly even makes sense from the person’s own point of view.) Another argument was that a person might sit in a corner over a large number of consecutive parties but still consider himself or herself to be extraverted, which would be all that counts. In practice, however, most witnesses would start worrying whether that person were still in contact with reality (which is again different from the question about introversion or extraversion). In the abstract, actors are at liberty to entertain a subjective definition of personality, but in real life it does not carry them very far. The intersubjective viewpoint is not merely a matter of scientific style; it is in touch with what people think of personality.

If the intersubjective viewpoint is accepted as a proper perspective on personality and if idiosyncrasies in self-report are seen as a source of error among other sources of error, the consequence for personality research and practice is as straightforward as it is revolutionary: Multiple assessors are needed to achieve acceptable reliability and validity; self-reports, being single by definition, are inevitably deficient. Self is of course acceptable as an assessor among others; self-ratings might even contribute more to the common variance than others’ ratings do. But in any case, the road toward an eventual objective, genetic diagnosis of personality, will have to be paved with multiple assessors; good intentions will not be enough.

THE FUTURE OF THE FIVE-DIMENSIONAL MODEL

Will genetic fingerprinting in due time describe personality in terms of extraversion, agreeableness, conscientiousness, emotional stability, and some version of Factor V? In other words, will the 5-D model survive the developments that most readers may expect to witness in their lifetimes (whether they like it or not)? At the moment of writing this, the answer can hardly be unequivocal; even the question may appear to need rephrasing.

In an extensive reanalysis of several data sets, Saucier (2002a) found a three-dimensional structure containing agreeableness, conscientiousness, and extraversion to be more replicable across samples than a 5-D structure, especially in peer ratings, which in the present reasoning are more germane than self-ratings. So we might end up with a subset. Using a comparable three-dimensional solution, Krueger (2000) showed that the additive-genetic structure underlying the Multidimensional Personality Questionnaire (Tellegen, 1982) corresponded closely to the phenotypic structure. On the other hand, Jang, McCrae, Angleitner, Riemann, & Livesly (1998) demonstrated that specific factors beyond the first five have nonzero heritability coefficients.

Even supposing reliable and valid assessments of phenotypic personality traits, a routine search for indicators of, for example, conscientiousness would require enormous samples just for tracing additive polygenetic effects; for interactions, the required sizes would rise exponentially (for a discussion of strategies of molecular-genetic research on personality, see Plomin & Caspi, 1998). At the turn of the century, attempts to trace genetic polymorphisms that explain personality showed the familiar picture of high initial expectations followed by failing replications (e.g., Herbst, Zonderman, McCrae, & Costa, 2000). According to a possibly more feasible scenario, large principal components of personality traits may be expected to reappear as an aggregate result of studies searching for single genes to explain specific patterns of deviant behavior (see, e.g., Brunner, Nelen, Breakfield, Roppers, & Van Oost, 1993). Assuming continuity between the range of normal behavior and deviant extremes, the aggregate structure of a large number of such specific patterns would resemble the 5-D structure. In the process, such taxonomies of phenotypic traits would receive a status comparable to mineralogical classifications; the chemistry of individual differences would be located at the DNA level.

Decades ago, Carlson (1971) found that personality was spelled in either of two ways: social or clinical. The questionnaire conception of personality is arguably social-psychometric by its methodological nature. If the genetic approach becomes dominant, a clinical reconstruction will regain momentum; individual differences within the normal range will be seen as mitigations and moderations of personality defects constituting the chemical elements. Meanwhile, an enormous amount of work has to be done, and 5-D questionnaires filled out by several third persons and self are instrumental in that labor.
THE LEXICAL BASE OF THE FIVE-DIMENSIONAL MODEL

A basic motive of researchers involved in the 5-D paradigm is to give a systematic and comprehensive, or at least representative, account of personality traits. An accompanying notion is that the field is characterized by a proliferation (John, 1990) of concepts and instruments, which frustrates the progress of the science of personality. The signature of the 5-D paradigm is empiricist and, in a sense, antitheoretical: If theorists, in this context, are individuals bent on disseminating their idiosyncratic concepts of personality, then their collective but uncoordinated action is responsible for a chaotic state of affairs in which thousands of unrelated concepts and their operationalizations form a market rather than a science. The 5-D conception is thus a taxonomy intended to end all idiosyncratic taxonomies.

To lift personality out of its chaotic state, an Archimedean point was needed. The most obvious candidate for a point of departure at the descriptive or phenotypic (Goldberg, 1993b) level is the lexicon. Like genetics, it provides a finite set of elements on the base of which a taxonomy may be built and proliferation may be counteracted. This section contains a discussion of the lexical point of departure, its variations, and its consequences. An analysis of the different shapes of the Factor V and their operational antecedents serves as an illustration.

The Lexical Axiom

What is usually referred to as the lexical hypothesis is more like an axiom. It states that people wish to talk about whatever is important and that the terms in which they talk may be found in the lexicon. Like genetics, it provides a finite set of elements on the base of which a taxonomy may be built and proliferation may be counteracted. This section contains a discussion of the lexical point of departure, its variations, and its consequences. An analysis of the different shapes of the Factor V and their operational antecedents serves as an illustration.

An objection that is seldom voiced although it is obvious is that the reverse of the lexical axiom does not necessarily hold true: People may well be talking about unimportant things most of the time. There is something to be said for the idea that the language of normal personality does not serve much of a purpose. However, PCA (see the next section) capitalizes on redundancies among variables. That method thus retroactively introduces a corollary of the lexical axiom, namely, that redundancy is indicative of real importance. For playful purposes, we may seek rare and sophisticated terms or combinations of terms; at the level of common components, however, we mean business. Of course, this corollary, in its turn, may or may not be judged credible.

A reverse objection is that common language is not subtle enough for scientific purposes. One may philosophize at length about this proposition, which is as metaphysical as the lexical axiom itself. The historic rebuttal, however, was delivered by Digman (1990; Digman & Inouye, 1986), who recovered the Big Five structure in questionnaires, that is, in instruments designed by experts. In a similar vein, I (Hofstee, 1999) asked 40 clinicians to score a prototypical personality disorder with which they were familiar on the items of the Five-Factor Personality Inventory (FFPI; Hendriks, Hofstee, & De Raad, 1999). These items do not contain any technical terms or pathological content. Nonetheless, very distinct and extreme profiles in 5-D terms resulted, again indicating that expert categories may be well represented by ordinary language.

In principle, the lexical approach both reflects and fosters a lay definition of personality; in practice, however, the effect seems to be slight. Thus, at low conceptual costs 5-D research has succeeded in bringing a considerable measure of order to the anarchy of phenotypic traits. Any serious investigator proposing a new trait concept would now be well advised to investigate whether it has incremental validity over an optimal linear combination of the five factors; existing concepts are better understood in that framework. An example is typical intellectual engagement (Goff & Ackerman, 1992), which appears to be a label for a mixture of Factors V and III; another is the familiar concept of sociability, blending Factors I and II. As I argue later, there is nothing against using dedicated labels for blends if they are distinguished from variables that do carry considerable specific variance. But even if taken liberally, the five factors represent a taxonomic breakthrough, part of which may be credited to the lexical approach.

Operationalizations of the Lexical Approach

There is no unique and cogent operationalization of the lexical approach. It pertains to single personality-relevant words, under the tacit supposition that words do not interact, so that the meaning of any trait combination can be represented by a linear function of them. That supposition is patent false in the case of oxymora like “amiably inimical” or “quietly exuberant,” joinings of opposite terms whose meaning cannot be accounted for in a linear fashion; however, there are reasons to be wary of such seductions of literary language. In any
case, the search for single words is a defining characteristic of the lexical approach. But the question of how to select the single words has no straightforward answer; a number of decisions must be made.

A first decision concerns grammatical categories. Most investigators, from Galton (1884) on, have concentrated on adjectives (for an overview, see De Raad, 2000; Saucier, Hampson, & Goldberg, 2000). Goldberg (1982) and De Raad (1992) have studied type nouns, alphabetically running from ace to zombie in American English, but there is a consensus that this category does not add much (cf. extraverted vs. an extravert) or consists of invectives that have uses other than describing personality. A more interesting addition to adjectives are personality-descriptive verbs, which run from abandon to yield (not counting zap, zip, and zigzag) in English, denoting acts that would be more characteristic of one person than another. De Raad’s (1992) analyses of personality verbs and nouns, however, do not result in novel content over the factors found in adjectives. The focus on adjectives does not recoil significantly on the implicit definition of personality.

A second set of operations consists of exclusion categories, for example, moods (e.g., sad), body characteristics (e.g., fat), social relations (e.g., subordinate), attitudes (e.g., progressive), and effects (e.g., famous). These exclusions are unproblematic because the categories are outside the domain of personality traits. Two other categories, however, deserve special consideration. One is called mere evaluations (e.g., good). In the language of personality, content and evaluation are intimately connected: On the one hand, neutral content is hard to find; on the other, mere evaluation is equally scarce. Tellegen (1993), in particular, has argued against excluding this category and has shown that it contains variance over and above the five factors (Almagor, Tellegen, & Waller, 1995). Thus, the 5-D model entertains a conception of personality that is somewhat sterilized with respect to evaluation.

The other problematic category is one that is invariably included, containing adjectives denoting intelligence, capabilities, talents, erudition, and the like—thus, the kind of maximum-performance traits that have traditionally been distinguished from typical-behavior traits. This inclusion is not an automatic consequence of the lexical approach; Ostendorf (1990), for example, sharply distinguished between temperament and character on the one hand, and skills and talents on the other, before joining the two sets of traits under the heading of dispositions. One could simply state that the 5-D approach has opted for the broader of the two definitions of personality, including not only temperamental or stylistic aspects (most notably Factors I, extraversion-introversion, and IV, emotional stability vs. neuroticism) and character (most notably Factors II, agreeableness, and III, conscientiousness), but also intellect, erudition, and the like (Factor V; see Hofstee, 1994b). However, I voice some reservations regarding that inclusive choice when discussing Factor V later.

A final operation consists of the exclusion of technical, highly metaphorical, and otherwise difficult terms. As I argued earlier, that procedure is probably not very consequential with respect to the scientific concept of personality, even though the literary loss is considerable. In constructing the FFPI, however, Hendriks (1997) went one step further and retained only items that were found perfectly comprehensible by students of lower professional education. Of the 1,045 brief expressions (e.g., Wants to be left alone) that made up the pool from which the items were chosen, 34% met this criterion. In a set of 195 trait-descriptive adjectives carefully selected to cover the factors of the 5-D model, only 14% did. It is a sobering thought that the founding studies of the 5-D model could not have been meaningfully carried out with these respondents. Furthermore, this sharpening of the comprehensibility criterion does appear to have consequences for the content of Factor V, as is shown next.

The Credentials of the Fifth Factor

The most spectacular vindication of the 5-D model has been brought forward by Ostendorf (1990). In the introduction to his study, Ostendorf related that he viewed the model with great skepticism at first, as the available American studies were based on very small samples of trait variables that had been composed using very subjective criteria (Ostendorf, 1990, p. 9). Not only this initial skepticism, but also the fact that the replication was completely independent, started from scratch, and was carried out in another language, added to the credibility of the 5-D model. Ostendorf, however, expressly included ability adjectives; consequently, his Factor V is a clear intellect factor defined by such terms.

In our Dutch lexical project, subjects were asked whether an adjective would fit in the framing sentence “he/she is [adjective] by nature” (cf. Brokken, 1978) in order to determine an adjective’s prototypicality as a trait descriptor. Adjectives like dull, gifted, capable, brilliant, one-sided, idiotic, sharp, and ingenious received very low prototypicality ratings (along with other categories of terms, most notably social-effect adjectives like horrible, commonplace, and captivating). In a selection of terms used by De Raad (1992) to establish the replicability of the 5-D model in the Dutch language, terms with low prototypicality were excluded; consequently, no clear fifth factor appeared. In a Dutch-German-American comparison (Hofstee, Kiers, De Raad, Goldberg, & Ostendorf, 1997), the correspondence between the American and German
earlier, the narcissistic and antisocial profiles were relatively prototypical personality disorders (Hofstee, 1999) referred to judged to be loudmouthed. In the study on 5-D profiles of telligent extravert may be found eloquent; a dull one may be intelligence and other personality traits stable and psycholog-

That shift can hardly be objected to as such. Not only are both
phasis from a narrow to a broad conception of personality. Furthermore, many of the NEO items in general, and of the openness to experience scales in particular, would not pass the comprehensibility test that was outlined earlier. Brand (1994) predicted that both intellect and openness to experience would correlate substantially with measured intelligence (g) over the whole intellectual range of the population. A special reason may be that subjects of modest IQ would reject such items because they do not understand them, and thus receive low scores.

A powerful competitor—if only by virtue of the wide-
spread use of the NEO-PI-R (Costa & McCrae, 1992)—to the intellect conception of Factor V is its interpretation as openness to experience. That construct does not come out of the lexical approach; in fact, McCrae (1990, 1994) has used it repeatedly to argue the deficiency of that approach. The consequent problem with such constructs, however, is that they do not share the taxonomic status that is awarded by the lexical paradigm. Furthermore, many of the NEO items in general, and of the openess to experience scales in particular, would not pass the comprehensibility test that was outlined earlier. Brand (1994) predicted that both intellect and openness to experience would correlate substantially with measured intelligence (g) over the whole intellectual range of the population. A special reason may be that subjects of modest IQ would reject such items because they do not understand them, and thus receive low scores.

Distinguishing Personality from Ability

The 5-D model seems to have contributed to a shifting em-
phasis from a narrow to a broad conception of personality. That shift can hardly be objected to as such. Not only are both intelligence and other personality traits stable and psychologically relevant, but they also combine with each other. An intelligent extravert may be found eloquent; a dull one may be judged to be loudmouthed. In the study on 5-D profiles of prototypical personality disorders (Hofstee, 1999) referred to earlier, the narcissistic and antisocial profiles were relatively close together, but that must be because the FFPI’s Factor V has little to do with intellect: Sizable differences between the two would be expected on measured intelligence (Millon, personal communication, September 29, 1999). For a proper assessment of personality, the inclusion of intelligence is indispensable.

There is no good reason, however, to contaminate typical behavior and maximum performance. On the contrary, there are good reasons to separate the operations. One is that objective measurement of intelligence is more scientific than its assessment, however intersubjective that assessment may be. Another is that methods are not neutral: Abilities and tests of maximum performance are as closely associated as are stylistic traits and assessments of typical behavior. To include ability items in questionnaires can only obscure the view on intelligence.

With respect to concepts of temperament and character, state-of-the-art assessment would include a 5-D question-
aire as a baseline instrument, and novel concepts would have to prove their added value against that background. According to the same principle of parsimony, however, 5-D factors have to prove their added value over measured intelligence. Precisely because personality and intelligence belong together, objective measures of intelligence should be included in investigating the structure of personality. In view of the scientific primacy of intelligence, its variance should be partialled out of the questionnaire scores. While in the process, attitudinal factors, which are out of bounds in most definitions of personality, should be removed in the same manner. They, too, are empirically correlated with certain versions of Factor V, particularly with openness to experience (Saucier, 2002a). With these corrections, it is entirely conceivable that little would remain of Factor V.

THE LINEAR APPROACH TO THE CONCEPT OF PERSONALITY

The “Magical Number Five,” in the words of Goldberg (1992b), is intricately connected with applying PCA to large numbers of trait variables. Forerunners have been pinpointed, most notably Tupes and Christal’s (1961/1992) analyses. However, Tupes and Christal’s denomination of the fifth factor in terms of culture is now obsolete. On the other hand, if the magical number had been found to be six, one could have referred to another Cattellian’s (Pawlik, 1968) set consisting of I Extraversion, II Cooperativeness, III Deliberate Control, IV Emotionality, V Independence of Opinion, and VI Gefühlsbetontheit (which is difficult to translate; the order in which the factors appear has been adjusted to the present
context). These examples of imperfect historical fit could easily be expanded upon. The five factors owe their consolidation and impact to analyses of large data matrices that did not become possible until the last decades of the twentieth century.

This section starts with setting out the strongest possible case for PCA by presenting a classical (see Horst, 1965) rationale for it. Next, it examines the grounds for the magical number five. It then considers the so-named person-centered approach as an alternative to PCA in certain contexts.

### The Case for Principal Component Analysis

Applying PCA to a scores matrix is the logical consequence of performing item analysis. In the general case, the aim of item analysis is to maximize the internal consistency of one or more scales based on the items; the exception whereby items are weighted by their predictive validity is outside the present scope. The basic idea of item analysis may be expressed as follows: The investigator is aware that each single item, carefully chosen as it may be, is an imperfect operationalization of whatever construct it represents. But the investigator has no better criterion against which to gauge the validity of the item than the total score on the set of equivalent items. Item analysis is thus a bootstrapping operation.

Carrying this basic idea to its logical consequence proceeds as follows: At the first step, items are weighted according to their association with the total score. Discarding items on that basis would amount to arbitrarily assigning a zero weight. That may be defensible in extreme cases where it is evident for substantive reasons—albeit post hoc—that the item does not belong in the set. In the general case, however, all items would be retained.

By virtue of assigning weights to the items, however, the total score has been replaced by a weighted sum. The implicit rationale is that this weighted sum is a better approximation of the underlying construct than was the unweighted sum. So the logical second step would be to assign item weights according to their association with the weighted sum. Thus an iteration procedure has been started, the endpoint of which is reached when convergence of weights and of weighted sums occurs. At that point, the weighted total score is the first principal component of the item scores (Horst, 1965). If the item set is multidimensional, more than one principal component is obtained, but the reasoning is essentially the same.

Thus a particularly strong argument in favor of PCA is that it is logically inevitable. Also, since the days of computer scoring, any practical objections against calculating weighted sums have disappeared: Sooner than applying 10 hand-scoring keys to a 5-D questionnaire (five keys for positive items and five for negative items), one would put the item scores on electronic file anyway.

### Raw-Scores PCA

The present argument does not prejudice in favor of PCA as it is usually conceived, namely, PCA of z scores or correlation matrices. Rather, it refers to raw-scores PCA, with deviation scores and their covariance matrices, or standardized scores and their correlations, as special cases. Raw-scores PCA should be performed on bipolar scores; for example, scores on a five-point scale should be coded as $-2, -1, 0, +1,$ and $+2$. We (Hofstee, 1990; Hofstee & Hendriks, 1998; Hofstee, Ten Berge, & Hendriks, 1998) have argued that a bipolar representation of personality variables is appropriate, as they tend to come in pairs of opposites. Thinking in terms of all-positive numbers is a habit imported from the abilities and achievement domain, where it does not make sense to assign a negative score.

Raw-scores PCA implies an absolute-scale interpretation of the Likert scale, rather than the conventional interval-scale interpretation. These alternative interpretations have subtle consequences for our conception of personality. The first of these concerns the reference point. With relative, interval-scale scoring, the population mean is the reference point. For desirable traits, that reference point is at the positive side of the scale midpoint (0), and vice versa. Thus a person with a score of $+.8$ on a socialness scale with a population mean of $+1.1$ (most people being found social), would be said to be somewhat asocial, albeit in a relative sense, which however is the only available interpretation when using interval scaling. The unthinking adoption of interval scales from the domain of intelligence and achievement may lead to a bleak view of humankind, whereby a sizable proportion of the population is judged more or less deviant. A poor comfort is that the proportion is a bit less than 50% because the raw-score distribution is not symmetric. Taking the scale midpoint seriously solves the problem; it prevents a positive judgment from being translated into something unfavorable and vice versa, based on an inappropriate convention.

The second way in which absolute and interval scale conceptions differ concerns spread. Using a five-point scale, most items have standard deviations close to 1, as the prevalent responses are $-1$ and $+1$; thus the difference between absolute and interval scaling is not dramatic in this respect. But extremely favorable and unfavorable items obtain smaller standard deviations. The effect of standard PCA and interval scoring procedures is to increase their impact on the total score. It would seem that this is also an unintended consequence rather than a deliberate effect.
In sum, item scoring through weights obtained by raw-scores PCA deserves more consideration than it has received so far. The standard objection to treating scores on a Likert scale as absolute is that strong assumptions would be imposed on the data. I am unable to see the validity of that argument. So-called weak models may in fact be very strong: To assume that the midpoint of a personality scale has no meaning and, consequently, that respondents’ evaluations can be reversed, is about as strong as hypothesizing, for example, that a large proportion of the population cannot be trusted. At the very least, the absolute conception of Likert-scale scores is no more indefensible than the interval conception.

A Review of the Grounds for the Number Five

A way to obtain many principal components is to analyze matrices with large numbers of variables, in this case single trait descriptors. Earlier, limitations on computing capacity virtually prevented the number of trait variables from being much larger than the 35 employed by Tupes and Christal (1961/1992). With the expanding power of computers, however, it became feasible to analyze the very large numbers of variables that were needed to justify claims of representativeness if not exhaustiveness. However, the sorcerer’s apprentice problem then becomes keeping the number of factors from getting out of hand. With hundreds of variables, it will take many factors to get down to the time-honored “eigenvalue 1” threshold; for example, the 20th factor in Ostendorf’s (1990) PCA of 430 traits still has an eigenvalue of 3.

Hofstee et al. (1998) proposed a more stringent criterion based on the alpha reliability of principal components, which is approximately $1 - 1/E$ with large numbers of variables, $E$ being the eigenvalue of that principal component. Setting the minimum alpha at .75, an “eigenvalue 4” threshold results ($E = 1$ gives $\alpha = 0$). Using this criterion, Ostendorf (1990) should still have set the dimensionality of the personality sphere at about 14 rather than 5; with even larger numbers of traits, the dimensionality would only increase. There can be no doubt that the 5-D model discards linear composites of traits that are of sufficient internal consistency, and would add to the number of dimensions. It is of interest to note that the most prominent 5-D questionnaire, Costa and McCrae’s (1992) NEO-PI-R, in fact postulates 30 dimensions rather than 5, as each of the 30 subscales is deemed to have specific variance in their hierarchical model. (The five second-order factors do not add to the dimensionality, as they are linear combinations of six subscales at a time.)

An entirely valid pragmatic reason to restrict the number of factors is parsimony. The first principal component is the linear combination of traits that explains a maximum of variance; the second maximizes the explained variance in the residual, and so on. Consecutive factors thus follow the law of diminishing returns. Next, the scree test acts on the amount of drop in eigenvalue between consecutive factors; it thus signals points of increasingly diminishing returns. Using the scree test, Brokken (1978) retained 6 principal components in a set of 1,203 trait adjectives; Ostendorf (1990) retained 5. However, the scree test does not offer a unique solution; Ostendorf, for example, could have opted for an 8-factor solution on that basis. Neither PCA nor the scree test dictates the number of five.

Does replicability of factors provide a cogent criterion for the dimensionality of the space? That depends on how the term is understood. If one and the same large trait list were administered to large samples from the same population, the number of replicable factors would in all likelihood exceed five. At the other extreme, when independent, “emic” replications of the lexical approach in different languages are undertaken, the number tends to be in the order of three (De Raad et al., 1997; Saucier et al., 2000) rather than five, Ostendorf’s replication being an exception. Saucier, Hampson, and Goldberg list 18 points on which such studies might diverge and recommend methodological standardization. A familiar objection is that standardization leads to premature closure of the issue: Not only would the outcome depend on arbitrary choices, but moreover one could not tell anymore what makes a difference and what does not. It would be preferable to use these points for studies on whether and how the number varies in function of differences in approach.

In sum, the number five takes on the character of a point estimate in a Bayesian credibility function on an abscissa that runs from 0 to some fairly large number, with the bulk of the density stacked up between 3 and 7. As with other empirical constants, the uncertainty does not so much result from random error as from the interplay of diverging arguments and specifications. In any case, the number should be taken with a grain of salt.

The Person-Centered or Typological Approach

A familiar critique of trait psychology is that it loses the individual from sight (see, e.g., Block, 1995; Magnusson, 1992). A set of alternative operations is available under labels such as type or person-centered approach; it comprises Q-sorts in preference to Likert scales, longitudinal designs to assess the dynamics of personality, and cluster analysis of persons rather than PCA of variables. Recent empirical studies (e.g., Asendorpf & Van Aken, 1999; Robins, John, Caspi, Moffitt, & Stouthamer-Loeber, 1996) concentrate on the three “Block” types: resilient, overcontrolled, and undercontrolled.
I document the relativeness of the opposition between the person-centered and variable-centered paradigms (see also Millon, 1990) but try to do justice to a real difference in their ranges of application.

**Persons in Principal Component Analysis**

Unlike factor analysis proper, in which factor scores are hardly more than an afterthought, PCA offers a fairly symmetric treatment of individuals and variables. One could rotate a matrix of scores on principal components to simple structure and characterize individuals by the person factor on which they had their highest score. In an even closer approximation to the person-centered approach, factors and loadings may be rescaled such that individuals receive loadings and variables receive factor scores. That is not precisely the same thing as performing Q-factor analysis, as the scores still become standardized per variable instead of per individual as in Q-analysis; but the two operations would be mathematically identical in the case of raw-score PCA. If the argument in favor of raw-score PCA is accepted, the difference between variable-centered and person-centered analysis reduces to a set of scaling constants and rotation criteria being applied to one matrix rather than another, which is hard to get excited about.

**Variables and Types**

An orthodox typological solution may be viewed as a binary matrix of persons by types with one 1 per row, representing the type to which that person is assigned, and 0 scores in the remaining cells. At the other extreme, each person could be given a score for each type on a continuous scale, representing the extent to which that person corresponds with that type, thereby treating the types as continuous variables. The orthodox solution could be reconstructed from that matrix by selecting the highest score per row and dichotomizing accordingly. Intermediate, liberalized typological solutions (Millon, 1990) could also be derived, most notably through matrix-wise dichotomization of the continuous scores. In the liberalized solution, some persons would appear to be assigned to more than one type, whereas some others would fail to meet the threshold for any type at all. The types would no longer be orthogonal in the way they are forced to be according to the orthodox solution; so one could correlate the types, factor analyze them, and the like.

To those who would find this methodological play with types improper, there is a perfectly serious answer. In the ideal case, a diagnosis is performed by an infinite number of independent experts. Experts do not agree perfectly in all cases. Thus the sum or average of even their orthodox typological solutions would give precisely the kind of matrix of continuous scores introduced in the earlier argument. In a scientific (in the sense of intersubjective) conception of types, the continuous matrix is the primitive case, not the binary matrix. The primitive case arises not because types (or even personality variables in general) are necessarily continuous as such, but because of the tacit third dimension of the matrix.

**Q-Sorts and Likert Scales**

Investigators working in the person-centered paradigm prefer ipsative scores, as they would represent intra-individual rather than interindividual comparisons. Varieties of ipsative scoring are row standardization, which fixes the means and standard deviations, and forced distribution, whereby all moments are fixed. Q-sorts automatically result in forced-distribution scores (unless the number of items in the “most applicable” to “least applicable” categories is not fixed, in which case, however, the method is indistinguishable from using a Likert scale).

Like orthodox typologies, ipsative scores may be constructed from continuous “interactive” scores, in this case by standardizing over variables or by forcing a distribution on them. One might object that Q-sorts are different in principle from Likert-scale scores, but that remains to be seen. In the first place, judges need not respond the way we instruct them to. If I am asked, by way of intra-individual comparison, whether I am (or John is) more reliable than friendly, I may well respond against the background of people in general; it could even be argued that the question is meaningless without that background. Conversely, when confronted with a standard personality questionnaire, intra-individual considerations might well enter into my response process. It is thus arguable that all responding is interactive. In the second place, Q-sorts are used to compare people, therefore, interindividually: If John is said to be of Type A whereas Mary is not, the intra-individual level is automatically surpassed.

The effect of ipsatization is to remove interindividual differences in elevation and spread (and skewness, kurtosis, and so on) of the responses. The operation thus implies a view of personality in which such individual differences have no place. Surprisingly, that view appears to be shared by some unadulterated trait researchers, most notably Goldberg (1992a) and Saucier (1992; see, however, Saucier, 2002a). Their rationale, however, has nothing to do with an emphasis on intra-individual differences. Rather, they use ipsatization of Likert-scale data to remove differences in scale usage, in other words, response sets. Whatever the rationale is, the implication needs to be examined in detail.
Removing differences in elevation and spread prevents one person from having more traits than another, as well as from being more extreme. Correcting for elevation is quite defensible in the special case where the variable set is completely balanced (i.e., consists of opposites like reliable and unreliable). Except in a fairly poetical manner, it hardly makes sense for a person to be both more X and un-X than another; it is more parsimonious to attribute such a response pattern to excentric scale use, traditionally denoted as the acquiescent response set. Hofstee et al. (1998; see also Ten Berge, 1999) presented ways to correct for excentric responding. However, if the variables set is not balanced, correcting for elevation removes content and social desirability variance. In the most elementary case, John is prevented from being both more friendly and reliable than Mary. That consequence is infelicitous.

The person-centered approach is thus subject to an irony of fate: An intention (a proper approach to personality) materializes into an operation (ipsative scoring) that appears to cradle aversive implications (for the very concept of personality). Ipsatization would do the job in a strictly idiographic approach, but that condition is not fulfilled: By virtue of the fact that one and the same method and vocabulary is applied to more than one person, interindividual comparison automatically creeps in. It may make sense to separate ipsative and normative components of a scores matrix by representing the latter as a vector containing the person means. Discarding that vector, however, has the effect of flattening the concept of personality. Essentially the same argument applies to individual differences in spread (and other moments of the score distribution).

**Dynamics**

Analytically, a dynamic approach to personality, as advocated by person-centered investigators, may mean either of two things: taking the time or growth dimension into account, and interpreting traits as an intra-individual pattern, therefore, in a nonlinear fashion. The dynamic approach thus stands in opposition to an orthodox trait approach, which is static and linear.

However, dynamics are easily accommodated in the individual-differences paradigm. A chronological series of assessments pertaining to an individual may be conceived as an extension of the scores vector. In a multiple prediction of some criterion, the question then becomes whether, for example, last year’s emotional stability has incremental validity over today’s. Alternatively, a (fitted) growth curve may be represented by its first derivative representing growth speed, its second derivative representing growth acceleration, and so on, in addition to the overall score of that individual. Again, the derivatives function as extra traits. Similarly, pattern interpretation may be represented by introducing extra predictors, in this case, moderator or interaction terms formed by multiplication of predictors. Thoroughbred trait psychologists would argue that growth and pattern scores cannot be expected to have incremental validity, but that is not an objection of principle. What this brief analysis shows is that the two paradigms are not ideologically incompatible but appear to consist of different generalized expectations regarding the relevance of growth and moderator terms.

A final wording of the moderator issue is whether single predictors may receive different weights according to the individual in question; thus, whether Mary’s emotional stability may be less relevant in predicting her performance as a pursuit plane pilot than is John’s. Again, there is no a priori reason why the weights should be uniform. A technical problem is that the Pearson correlation is undefined in the single case; however, raw-score association coefficients like Gower’s (1971) and Zegers and Ten Berge’s (1985) can do the job. Their application to the single case also gives a precise expression to the otherwise elusive idea of intra-individual trait structure. The Gower coefficient for the general case is the mean of the single-case coefficients; it thus writes interindividual structure as the mean of intra-individual structures, thereby joining two paradigms of personality that are usually brought in opposition to each other. This integration is still another reason for taking raw scores seriously. An empirical problem, however, is that individual weights may be extremely unstable. However, the same holds for intra-individual structure.

**Ranges of Application**

After digesting a number of red herrings, what remains is a matter of conventional preference. The trait psychologist represents the person as a vector of scores on a continuous scale, whereas the typologist would prefer a single qualification on a binary (applicable vs. not applicable) scale. Taking a sophisticated trait model incorporating growth and moderator effects, the person-centered approach is a special or degenerate case of it, and can therefore not be psychometrically superior in any respect. To justify the type approach, a different perspective should be adopted. To that end, I distinguish between a context of prediction and a context of communication.

Given the same basic materials, there can be no reasonable doubt that the trait approach is superior in a predictive context. On the one hand, typing consists of discarding information that is potentially valid. On the other, it introduces dynamic predictor terms whose empirical status is highly dubious; therefore, even an orthodox trait approach may be expected to do better upon cross validation.
Ironically, the 5-D approach meets with ambivalence from the side of its very proponents in predictive respects. McCrae and Costa (1992) and Jang and others (1998) have emphasized the incremental validity of the 30 subscales of the NEO-PI-R (Costa & McCrae, 1992) over its five factor scales, thereby implicitly questioning the 5-D model as an adequate representation of personality. The psychometric value of such arguments, however, is quite limited. Principle component analysis capitalizes on the common variance in the predictor set; successive residuals follow the law of diminishing returns. So does validity, unless in some magical and unintended way specific variance would be more valid than common variance.

The value of the type approach is to be found at a different, pragmatic level, at which personality is a subject of communication between a diagnostician and a therapist (in the wide sense of someone who is going to work with the individual, possibly the individual him- or herself). Human discourse and cognition being what they are, it makes little sense in that context to exchange vectors of continuous scores. Professional communication is better served by an attempt to capture the essence of the individual’s personality in a vivid and suggestive picture. To insist on using a trait paradigm in this context is to ignore the human element at the receiving end of a communication.

In the end, the two sets of operations appear to refer to different conceptions of personality-in-context rather than personality-in-vitro. The trait approach is geared toward automated predictive procedures in which substantive considerations do not even surface. The type approach caters to human receivers of personality information. Which of the two scripts is appropriate in a particular case is difficult to say in abstract terms. A personnel selection situation, for example, may be conceived in predictive as well as in communicative terms; the same goes for a clinical intake situation. The emphasis here is on distinguishing the scripts: Predicting on the basis of types and communicating in terms of traits are both arguably deficient.

HIERARCHICAL AND CIRCUMPLEX STRUCTURES

In a hierarchical model, trait concepts are seen as specifications of broader traits, which in turn may be grouped under the heading of supertraits. In a circumplex model, trait variables appear as combinations of each other; they form a network in which all concepts define each other in a recursive manner, without subordination or superordination. In mixed models, all variables and factors are equal, but some are more equal than others because they explain more variance or are assigned privileged status for conventional reasons.

This section contains an evaluation of trait taxonomies that have been proposed or implied, and it works its way toward a family model that may be acceptable by way of integration. However, it should be kept in mind that taxonomies are subject to contradictory demands, namely, conceptual and communicative simplicity on the one hand, and adequate coverage of empirical reality on the other.

The Principal Component Analysis Plus Varimax Taxonomic Model

In its elementary form, the Big Five structure consists of a varimax rotation of the first five principal components taken from a large heterogeneous set of trait adjectives (see, e.g., Ostendorf, 1990). Whether this result is intended as a model in any proper sense is irrelevant, as it evidently functions like one: People receive scores on the Big Five, and these scores are interpreted as their personality structure—specifically, an orthogonal structure according to which these factors vary independently over persons.

Goldberg (1993a) articulated that the model in question may be viewed as hierarchical: Items specify scales, and scales specify factors. This argument presupposes simple structure, but that condition is not fulfilled. A concomitant and very widespread notion is that the Big Five are “broad” (in the sense of fuzzy) factors of personality.

The Implicit Assumption of Simple Structure

Simple structure, in which each variable loads on only one factor and factors exhaust the common variance would be hierarchical indeed: Each variable would be a specification of only that factor; a particular factor could legitimately and meaningfully be interpreted in terms of the variables that load on it. The interpretation would not surreptitiously introduce other variance common to some subset of the variables in question.

In empirical practice, however, variable structures are so overwhelmingly complex—as opposed to simple—that the hierarchical model functions as an obstacle to proper conceptualization: The practice of interpreting factors on the basis of their highest loading items, which would be appropriate under simple structure, is quite erroneous if the condition is not fulfilled. For to the extent that some of the highest loading items share other common variance, factor interpretations become contaminated. For example, an extraversion factor easily receives a social interpretation (sociability, social extraversion, and the like; for an overview, see Digman, 1990)
because many high-loading items have positive secondary loadings on agreeableness.

**The Alleged Broadness of Factors**

Under conditions of actual simple structure, factors could be called broad in a hierarchical sense, as they capture the common variance of a number of variables. Even then, factors are not broad in a conceptual sense but rather more narrow than variables, as their internal consistency is higher and their angular position in the trait space is thus more fixed. A $g$ factor of intelligence, for example, is not a broadband but a high-fidelity measure of some latent trait. A fortiori, there is nothing broad about a Big Five factor based on a particular domain of trait variables. For lack of actual simple structure, it does not encompass a sizable number of lower level items or scales. The meaning of a factor, even if latent, is much more precise in a psychometric sense than is the meaning of the variables on which it is based. In that domain of variables, a set of five rotated principal components covers more variance than does any other set of five linear combinations, but “broadness” is an inappropriate and misleading term for that.

In another terminology, to view the Big Five as broad factors is to treat them as a circumplex structure. In a regular two-dimensional circumplex, the plane is sliced into a number of angular segments (e.g., 12 segments). Variables within a segment form a homogeneous set. A special case is simple structure, in which “mixed” segments are empty, as in Figure 10.1, panel A. The actual situation, however, is closer to panel B, amounting to a circumplex with four segments, of which two are well filled. These segments contain very heterogeneous sets of variables; two of those variables may even be orthogonal to each other. The very specific meaning of the factor is thus not adequately captured by the broad array of variables that have their primary loading on it.

**Marker Variables**

The interpretation problem would be solved if stable marker variables could be found, that is, trait terms that load exclusively on one factor. Goldberg (1992a) presented such sets of psychometric synonyms, for example, extraverted, talkative, assertive, verbal, bold, and five other terms for the positive pole of the extraversion factor. A minor problem with this interpretation strategy is that markers for some factor poles are difficult to find, for example, markers for emotional stability. A major problem is that marker sets appear to be no longer orthogonal in fresh samples or upon translation. Any two homogeneous sets of traits may be expected to correlate positively if both are desirable or if both are undesirable, negatively if they are opposite in that respect; neutral sets hardly exist. Orthogonal sets may be selected in a sample, but they will regress to obliqueness upon cross validation. On the basis of a large-scale study, Saucier (2002b) has developed marker scales that appear to be robustly orthogonal within his several data sets and might thus defy the present analysis. Still, one would have to wait and see how they do in another laboratory, for example, when transported abroad.

The obliqueness problem (see, e.g., Block, 1995) cannot be answered by the truism that varimax-rotated factors are by definition orthogonal. The missed point is that they have no interpretation—not because they are broad or fuzzy, but because any interpretation in terms of sets of variables is biased. To interpret a Big Five factor properly, one would have to perform and communicate a suppression operation, such as the following: Factor I is what remains of extraversion after suppressing any connotation of agreeableness or socialness that may be associated with it, however firmly; Factor V is a residue of intellect or openness to experience after subtracting a virtually indissoluble tinge of energy, which rather belongs to Factor I. That is a bit much to ask.

![Figure 10.1](image_url)  
*Figure 10.1* Prototypical simple-structure (A) and semicircumplex (B) configurations.
In conclusion, the PCA plus varimax set of operations leads to an inadequate representation of personality. The argument is not that traits are correlated, in any metaphysical sense: For purely predictive purposes, linear regression of criteria on orthogonal factors is a perfectly defensible approach. What was stressed is the conceptual risk of starting to talk in Big Five terms, either among experts or with others. Conceivably, we could keep our mouths shut, but in practice that is too high a price to pay.

The PCA plus varimax model has been imported into personality from the domain of intelligence research. The question arises whether it is appropriate in that domain. I (Hofstee, 1994c) have argued that it is not. The empirical structure of intelligence variables is an $n$-dimensional simplex (the all-positive quadrant of an $n$-dimensional sphere) characterized by positive manifold and lack of simple structure. Treating it as an orthogonal simple structure gives rise to biased conceptualizations of the underlying dimensions and inadequate representation of the domain. Essentially the same objection holds for the domain of personality.

**The Double Cone Model**

A seminal attempt at a specific structure model of personality in the 5-D framework is Peabody and Goldberg’s (1989) double cone, based on Peabody’s (1984; see also De Boeck, 1978) work on separating descriptive and evaluative aspects of trait terms. It focuses on the first three Factors; the smaller factors IV, emotional stability, and V, intellect, are treated as separate axes orthogonal to the sphere that is formed by the bigger three: I Extraversion, II Agreeableness, and III Conscientiousness.

The double cone model may be envisaged as follows: Take a globe with desirability as its north-south axis, so that all desirable traits are on the northern hemisphere and their undesirable opposites are on the southern hemisphere in the antipode positions. Apply an orthogonal rotation to the Factors I, II, and III such that their angular distances to the desirability axis become equal, namely, 54.7 deg with cosine $1/3$. Draw a parallel of latitude at 35.3 deg (close to Kyoto and Oklahoma City) through the positive endpoints of the Factors I, II, and III, and another one (close to Sydney and Montevideo) through the negative endpoints. Connect each possible pair of antipode points on the two circles by a vector. Together, these vectors form the double cone. The model represents empirical trait variables by their projection on the closest model vector.

The double cone was designed to embody a particular taxonomic principle, informally referred to by insiders as the Peabody plot and named chiasmic structure by Hofstee and Arends (1994). A classical example of a chiasm is

<table>
<thead>
<tr>
<th>Thrifty</th>
<th>Generous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stingy</td>
<td>Extravagant</td>
</tr>
</tbody>
</table>

In Peabody’s reasoning, this configuration arises by pitting a content contrast (i.e., not spending vs. spending) against a social desirability contrast (thrifty and generous vs. stingy and extravagant). In the double cone model, chiasmic structure recurs in the shape of Xs that are formed by vertical slicings through the center of the double cone. On the Northern circle, we would have thrifty and generous at opposite longitudes; on the southern hemisphere, stingy and extravagant. More generally, descriptive and evaluative aspects are represented by longitude and latitude, respectively.

**Evaluation of the Double Cone Model**

The model is readily generalized to five dimensions, although it loses some of its aesthetic appeal in the process: Take all 10 subsets of 3 out of the 5 factors, that is, the $I \times II \times III$, $I \times II \times IV$, through $III \times IV \times V$ subsets, and treat each of these spheres in the manner just sketched. The generalized double cone thus consists of a Gordian knot of 10 three-dimensional double cones in the 5-D space sharing their vertical (desirability) axis, or 10 pairs of latitude circles. There is no valid reason why the range of the chiasmic structural principle should be restricted to a particular subset of three dimensions. But the model easily passes the generalizability test.

It is not entirely clear whether the algorithm for analyzing the data as used by Peabody and Goldberg (1989) is consistent with the model. Via Peabody (1984), the reader is referred to an algorithm proposed by De Boeck (1978). De Boeck’s procedure, however, sets the I, II, and III dimensions orthogonal to the desirability axis, rather than at 54.7 deg. Still, it is certainly possible to design an algorithm that would be consistent with the double-cone model.

The next question, however, concerns fit. That may be tested by assessing the quality of chiasms that are generated by the model. Hofstee and Arends (1994, Table 1) present chiasms derived from Peabody and Goldberg’s (1989) materials. An example is

<table>
<thead>
<tr>
<th>Forceful</th>
<th>Peaceful</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quarrelsome</td>
<td>Submissive</td>
</tr>
</tbody>
</table>

The content contrasts in this and in other examples are not convincing. The reasons are not hard to find. First and most important, the cone structure supposes an angular distance of only 109.5 deg between terms that should form a content
contrast, like forceful-peaceful and quarrelsome-submissive. Second, 5-D factors have different angles with the desirability vertical axis: II+, agreeableness, for example, is much further north than is I+, extraversion. When these angular distances are forced to be equal, as in the model, content contrasts become contaminated by a desirability contrast. In the example, peaceful is more desirable than forceful; therefore, to the extent that they are at all judged opposite, that is partly an artifact of a desirability difference.

It is fair to conclude that the double cone does not model the underlying principle of chiasmic structure in an optimal way. One could refine the model, but there is no need to do so: Hofstee and Arends (1994) showed that the Abridged Big Five circumplex (AB5C; see Hofstee, De Raad, & Goldberg, 1992) model to be discussed later can account for chiasmic structure, and generates credible chiasms:

<table>
<thead>
<tr>
<th>Daring</th>
<th>Cautious</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reckless</td>
<td>Timid</td>
</tr>
</tbody>
</table>

In two experiments, participants judged content contrasts taken from AB5C chiasms to be superior over double cone contrasts. This is not to say that chiasmic structure exists: Hofstee and Arends reiterate a point already taken by Peabody (1967) himself, namely, that desirability and content cannot be separated. So the best one can do is create a chiasmic illusion, as in the previous example. The algorithm goes as follows: Take a particular circumplex; draw a diameter separating desirable from undesirable traits; select two traits on different sides of the diameter but close to it and to each other, for example, cautious (slightly desirable) and timid (slightly undesirable); together with their opposites, they create the chiasmic illusion. It arises because in this case the alleged content contrast is formed by two terms with an angular distance that is only slightly less than 180 deg, instead of 109.5 deg as according to the double cone model.

**Do Chiasms Have a Future?**

The double cone model was shown to be generalizable; it may be possible to design a refined version by widening the angle between content opposites, amounting to oblique rotation. The more basic questions that remain, are What is the taxonomic status of the underlying principle of chiasmic structure? and What does it do to our conception of personality?

Whatever the refined model would be, it would focus on traits that are close to the equator of a hypersphere whose vertical axis is desirability: The model would focus on fairly neutral traits. They form a small minority, so the focus would be on a counterrepresentative subset of personality variables. On the one hand, there is something venerable (to use Saucier’s, 1994, term) to such a value-free approach; personality psychologists, like everybody else, would prefer practicing a discipline that is not submerged in extrascientific values. On the other, desirability is not fruitfully considered as a mere response set or other artifact that is to be separated from content: Hofstee and Arends (1994) emphasized that even in the classical example of chiasmic structure cited earlier, stinginess is not merely undesirable thrift, but an asocial version of it, whereas generousness differs from extravagance in being prosocial; therefore, the evaluation contrast is in fact one of content, as in the AB5C model. So the most realistic conclusion is that chiasmic structure and related models cannot be central to the concept of personality, even though they may have their place in specific contexts (see Saucier, 1994; Saucier, Ostendorf, & Peabody, 2001).

Central features of the double cone model, however, appear to be valuable by themselves. One is the “circular pattern” (Peabody & Goldberg, 1989, p. 556), as opposed to simple structure, that is embodied in the model. Another is orienting the trait space toward desirability as its reference axis. These points are taken up later when developing an integrative family of structure models.

**Generalized Circumples**

In circumplex models, traits are assigned to segments of a circle and are thus represented by their projection on the bisectrix of that segment. Circumplexes picture tissues or networks of traits: Contrary to hierarchies, circumplexes have no superordinate and subordinate concepts. Eysenck and Rachman (1965), for example, represented Hippocrates’ melancholic, choleric, sanguinic, and phlegmatic types as mixtures of the positive and negative poles of neuroticism and extraversion; presumably, however, Hippocrates would have preferred a rotation by which an extravert is a mixture of the choleric and sanguinic types, neuroticism is what melancholics and choleries have in common, and so on. Circles enjoy full freedom of rotation.

Circles generalize to spheres, and spheres generalize to hyperspheres—particularly, in this context, to the 5-D hypersphere. An early example of a 3-D structure is Heyman’s (1929) temperament cube. Not until the end of the twentieth century, however, did 5-D researchers (Hofstee et al., 1992; Saucier, 1992) construct circumplexes of more than two dimensions.
**Heymans’s Cube**

Heymans (1929) constructed a network model with three dimensions—emotionality, primary versus secondary function (comparable to extraversion-introversion), and activity—forming the axes of a cube. Types are located at each of the eight vertices of the cube, among which are the four Hippocratic types; for example, the sanguinic type is at the vertex where low emotionality, primary function, and high activity meet.

Heymans tended to conceive the temperament space as unipolar: The type characterized by the absence of emotionality, activity, and secondary function is named amorphous. One amendment therefore is to move the origin of the trait space to the center of the cube. Next, it is difficult to conceive of activity and primary function as orthogonal; different dimensions (and types) would be chosen in a contemporary three-dimensional model. Finally, one would prefer rounding the cube to a sphere. On the one hand, it is thus gratifying to note that time has not stood still, and that Heymans’s cube is now obsolete by reasonable standards. On the other, it is equally gratifying to recognize Heymans’s model as a forerunner of the generalized circumplexes that did not appear until the end of the twentieth century.

**Saucier’s Rhombicuboctahedron**

Saucier (1992) presented an integration of interpersonal and mood circumplexes and the Big Five Factors I, II, and IV. He drew attention to the fact that simple structure does not materialize in these domains; many variables are interstitial in that they are closer to the bisectrix of the angle between two factors than to the factors themselves. When simple structure is nonetheless imposed, interstitial variables are likely to be assigned to different factors by different investigators, even though the positions of variables and factors are closely comparable. Saucier constructed 6 bipolar scales as benchmarks for the interstitial positions, in addition to the 3 bipolar factor markers: a I+II+ versus I−II− scale (friendly vs. unfriendly), a I+II− versus I−II+ scale (dominant vs. submissive), and so on. He depicted the resulting trait structure as a rhombicuboctahedron, a prism showing the 18 (i.e., $2 \times (3 + 6)$) unipolar benchmarks as facets.

Saucier’s model may be alternatively conceived as an abridged three-dimensional circumplex, depicted by three orthogonal circles based on two of the three factors at a time. Each circle contains two bisectrices of the angles between the factor axes; in the model, a variable is represented by its projection on the vector (out of 9 bipolar or 18 unipolar vectors) to which it is closest. This representation has the advantage that it is easily carried to the fifth dimension (discussed later). Saucier showed that the $1 \times II \times IV$ space was the most interstitially structured of all 10 spheres that are contained in the 5-D hypersphere; that difference, however, is quite relative in view of the many mixtures involving Factors III or V.

Like Wiggins’s (1980) two-dimensional interpersonal circumplex, Saucier’s model uses octants, which are 45 deg wide, corresponding to a correlation of .707. Therefore, the variables assigned to such a segment may still form a fairly heterogeneous set. Hofstee et al. (1992) distinguished traits that had their primary loading on one factor and their secondary loading on another (e.g., I+II+; sociable, social) and traits with a reverse pattern (II+I−; merry, cheerful). This strategy amounts to slicing up a circle into 12 clock segments of 30 deg, corresponding to a correlation of .866. A reason for making these finer distinctions is that 30 deg is about the angular distance at which vectors are still given the same substantive interpretation (Haven & Ten Berge, 1977). If this amendment is worked into Saucier’s model, it becomes identical to a three-dimensional version of the abridged circumplex.

**The Abridged Big Five Circumplex Model**

The AB5C model consists of the 10 circumplex planes that are based on 2 of the 5 factors at a time. Thus, variables are represented by their projections on the closest plane or, more precisely, on the closest of the 6 bipolar clock vectors (running from 12 o’clock to 6 o’clock, 1 to 7, and so on) in that plane. The hypersphere contains large empty spaces between the model planes, so it may look as if the abridgement is rather drastic. However, varimax rotation puts the variables as close to the planes as possible; Hofstee et al. (1992) showed that it does a better job at this than at maximizing simple structure, which is putting the variables as close to the single factors as possible. Thus, representing traits by their two highest loadings seems acceptable; a model including tertiary loadings is entirely conceivable, but it would be much more complex and add very little.

More aptly than by a spatial configuration, the AB5C tissue is depicted by a table using the 10 factor poles (I+, I−, II+, II−, and so on) as both warp and weft, the column denoting the primary loading, and the row, the secondary loading of the traits assigned to a cell. Of the 100 cells in that table, the 10 combinations of the positive and negative poles of the same factor are void; the remaining 90 contain the unipolar facets generated by the model. The gain over the simple-structure model is enormous. That model accommodates only
relatively pure factor markers, that is, traits assigned to the 10 diagonal \((I+I+)\) to \((V-V-)\) cells of the table. If simple structure would in fact materialize, most if not all of the variables would be found in those cells. If, on the other hand, the empirical structure is essentially circumplex, only 11.1% of the variables would find their way to the diagonal cells. In Hendriks’s (1997) analysis of 914 items, 105 (11.5%) ended up in those cells. That illustration is as dramatic as is the percentage of variables that would have to be discarded in a proper application of the simple-structure model.

In the discussion of the person-centered approach, I introduced a distinction between the contexts of prediction and communication. Against that background, it should be noted that the predictive gain of the off-diagonal AB5C facets over the five principal components is nil, as the facets are linear combinations of the components. However, they do serve conceptual, interpretive, and communicative purposes. An individual’s profile of scores on the FFPI, for example, may be typified by that person’s single most characteristic facet; thus, for example, a person whose highest score is on factor \(V^+\) and whose second highest score is on \(III^+\) may be characterized by the cluster of expressions and adjectives that form the \(V^+III^+\) facet (Knows what he/she is talking about, Uses his/her brains, Sees through problems, and the many other items listed by Hendriks, 1997, for this “Tight Intelligence” facet). One or more of these catch phrases should be more effective than presenting a 5-D profile or even the subset based on the scores in question (“This person is primarily someone who Thinks quickly \([V^+]\), and secondarily someone who Does things according to plan \([III^+]\)”). Furthermore, at the theoretical level the AB5C model accounts for a large number of concepts that do not coincide with the five Factors but are quite adequately reconstructed as their mixtures.

Another way to document the flexibility of the AB5C design is in noting that it incorporates features of both oblique rotation and cluster analysis on an orthogonal basis. Oblique rotation as such does not solve the simple-structure problem when the configuration of variables is essentially circumplex. However, the insertion of oblique model vectors enables one to capture relatively homogeneous clusters of traits. That function is also served by cluster analysis procedures, but they lose sight of the dimensional fabric of the structure and the recursive definitions of clusters.

With respect to predictive purposes, the loss incurred by adopting the AB5C model is quite limited. First, the principal components base maximizes the internal consistency of the facets (Ten Berge & Hofstee, 1999), which should be beneficial to their validity. Second, if factors beyond the Big Five are needed to increase validity, the model is easily extended to include those factors. That would be more efficient than including separate scales for each additional specific concept.

**Undoing Hierarchies**

The traditional design of questionnaires is hierarchical: Items are grouped into subscales, subscales into scales. From the manuals of such questionnaires (see, e.g., Costa & McCrae, 1992) it is easily verified that subscales actually form a network; they have substantial secondary correlations with scales other than the one they are assigned to. Upon analyzing the single items of a questionnaire, a similar tissue pattern would arise; items would appear to have all sorts of promiscuous relationships, inviting circumplex analysis of the data.

Generalized (beyond two dimensions) circumplex analysis would proceed as follows: First, the item scores are subjected to PCA. The maximum number of principal components would be the number of subscales or facets (e.g., 30 in the case of the NEO-PI-R). Note that these 30 principal components extract more variance by definition than traditional scoring does. (In practice, it would soon become apparent that only a part of these principal components should be retained because the redundancies in the item tissue are captured by fewer components than the number of subscales.)

At the scale level, the optimal strategy is to proceed from the first \(m\) principal components, \(m\) being the number of superordinate scales (e.g., 5), as they make more efficient use of the data than do traditional scale scores. If, for reasons of continuity, the original interpretations of the scales are to be simulated, target rotations of the \(m\) principal components toward these scales could be carried out. If the original scales are conceived to be orthogonal, as in 5-D questionnaires, the optimal approximation procedure would be a simultaneous orthogonal target rotation of the \(m\) principal components toward the set of \(m\) scales. That procedure conserves internal consistency (Ten Berge & Hofstee, 1999); consequently, the average coefficient alpha of the rotated principal components is maximal. Most notably, it is automatically higher than the average alpha of the original scales.

Subscales of traditional questionnaires are very short; therefore, they are unreliable or consist of asking essentially the same question over and again, which is annoying to respondents and introduces unintended specific variance. If they are to be retained, their quality can be improved to a considerable extent by estimating subscale scores on the basis of (maximally) as many principal components as are postulated.
by the questionnaire model (e.g., 30). The procedure would consist of a target rotation of these principal components toward the subscale in question. Thus, using the collateral information contained in related items would generally increase the subscale’s internal consistency, even though the contribution of “bloated-specific” variance to it would be diluted, which in itself would be a desirable side effect.

The previous script, however, amounts to proving that the performance of a hierarchically conceived questionnaire can be boosted by placing a network model under its hood. Following the script would sooner or later lead to adopting the generalized circumplex approach, which implies a view of personality structure as a tissue rather than a hierarchy. At the superordinate scale level, one would hit upon a large number of interesting AB5C facets, which are linear combinations of the first \( m \) principal components; at the subscale level, a great amount of redundancy would be found, leading to a drastic reduction of the conceptual rank of the data matrix. The future of personality structure is hyperspheric.

**Pruning the AB5C Model**

The hypersphere of personality is riddled with gaps. Upon presenting the AB5C model, Hofstee et al. (1992) already noticed that its circumplexes were not evenly filled: The 1 o’clock versus 7 o’clock and 2 versus 8 segments attracted far more trait terms than did the 4 versus 10 and 5 versus 11 segments. The former segments contain consonant mixtures of either two positives factor poles (e.g., \( I+II+ \), sociable) or two negative poles (e.g., \( I-II- \), unsociable), whereas the latter contain discords combining a positive and a negative pole (e.g., \( I-II+ \), submissive, vs. \( I+II- \), dominant). The scarcity of discords reappears in Hendriks’s (1997, p. 45) results: Of the off-diagonal items, only 23% combine a positive and a negative factor pole, whereas the number of consonant and discordant cells is equal. Furthermore, discords tend to have the smaller projections in the 5-space, as may be verified from the cited studies. These results reflect the classical (e.g., Cruse, 1965) finding that neutral trait terms are scarce. The lexical axiom would imply that people find the corresponding behaviors relatively unimportant.

The rationale for introducing blends, in circumplex models in general and in the AB5C model in particular, is communicative. In the case of discords, the communicative benefits are unlikely to materialize. Rigid, for example, has a projection of .29 on the \( II-III+ \) vector in the AB5C model (Hofstee et al., 1992, p. 157); unkind and orderly have projections of .52 on \( II- \) and .67 on \( III+ \), respectively. Thus, the projection of the weighted sum of unkind and orderly on the \( II-III+ \) vector would be about twice as high as the projection of rigid itself. “John is primarily unkind and secondarily orderly” may thus be expected to communicate better than “John is rigid.” Therefore, the discordant hyperquadrants may be deleted from the AB5C model. It would thereby become semicircumplex: Of each circumplex, only the first and third quadrants would be retained. Clinicians, who tend to be sensitive to ambivalences of personality, might deplore that loss. However, the removal might well clarify intraprofessional communication, not to speak of communication with lay clients.

Extending this analysis would lead to a proposal for a somewhat different rotational positioning of the 5-D axes in order to maximize the coverage of consonant variables. In their present definition, some factors (notably, II and III) are associated more closely with desirability than are others (notably, I and IV). Thus, the vector in the \( I+ II+ \) quadrant upon which the projections of the Desirability values of the traits would be maximal is closer to the II axis than to the I axis. This asymmetry is illustrated by the fact that an undesirable trait like unrestrained (at 2:30 on the clock) has a distance of only 30 deg from that bisectrix (which is at 1:30), whereas agreeable (at about 11:20) is more than 60 deg removed from it. A counterclockwise rotation of the two factors would recognize unrestrained as a discord and agreeable as a consonant trait, which seems appropriate.

Applying this operation to all axes jointly amounts to a rotation to desirable manifold, mirrored by undesirable manifold. The resulting abridged Big Five semicircumplex (AB5SC) model is thus contained by the 10 faces of the hyperquadrant centered around the desirability axis and their 10 opposites. Each face is divided into three segments of 30 deg, placing the model vectors at clock positions of 12:30 versus 6:30, 1:30 versus 7:30, and 2:30 versus 8:30. (A more elegant representation involving a 45-deg counterclockwise rotation is presented later.) As no vectors recur in other semicircumplexes, there are now 30 bipolar model vectors.

In the following section the AB5SC model returns as a member of a family of models. That family comes about by a somewhat different rationale. The number five is no longer fixed; the emphasis shifts from 5-D models to accounts of trait structure that incorporate a number of principles that contribute to an efficient description of personality. Also, the factors as such disappear into the background, which is where they should have been from the start.

**A FAMILY MODEL OF TRAIT STRUCTURE**

What remained of hierarchical structure is the fact that each subsequent principal component explains less variance and is subordinate to its predecessors in that respect. The head of
the trait pedigree is the first principal component, named the $p$ factor of personality by Hofstee et al. (1998), in analogy to the $g$ factor of intelligence, and in distinction to Eysenck’s (1992) psychoticism or P factor, which is intended as a lower level construct. A family of models may be constructed by adding one principal component at a time. Thus at the second level we have a circumplex or semi-circumplex; at the third a three-dimensional generalized one, and so on.

The Primordial One

In search of superlatives over the Big Five and the giant three (Eysenck, 1992), primordial appears as a good label for the $p$ factor. That factor derives a mythical quality from its close association with desirability. It presents a fundamental paradoxe to students of personality, whose ultimate challenge is to manage the potent values that nourish its roots: Not until we are capable of giving an overall evaluation of an individual’s personality in a perfectly respectful manner will we have mastered that challenge.

In principle, there is nothing broad or vague about the $p$ factor. Quite to the contrary, it is by definition the most internally consistent linear combination of all traits, explaining some 10% to 15% of the total variance in unselected item sets (see, e.g., Brokken, 1978; Hendriks, 1997; Ostendorf, 1990), just like the $g$ factor does in the intelligence domain. In other words, no scale based on any subset of the items, however optimally weighted, is as internally consistent as $p$. Its location in the personality sphere is almost completely fixed in any large data set.

Fixing the interpretation of $p$ across studies is another matter. In Hendriks’s (1997) unselected set of 914 items, $p$ is best labeled as competence (Hofstee, 2001). In Saucier’s (2002a) study of representative sets of trait adjectives, it appears as a character factor taking in altruism, self-discipline, and success. The first principal component of the FFPI, whose 100 items were selected to cover the five factors equally, is an optimism factor (Hofstee et al., 1998). In view of the psychometric accuracy and the statistical reliability that was attained in these large-scale studies, the differences in interpretation cannot be attributed to chance. Differences in composition of the item pools must be responsible.

Saucier (2002a) interprets his first principal component as SD, for socially desirable qualities. On the one hand, this interpretation cannot be far off because the first principal component in any mixed set of positive and negative trait descriptors will be close to the desirability axis. On the other, it masks the fact that the first principal component bends toward whatever content is best represented in the item pool. For an extreme example, if that pool were overloaded with fairly neutral extraversion and introversion items, the first principal component would appear as extraverted desirability. Therefore, active steps have to be taken to justify the desirability interpretation.

I propose to define the $p$ factor of personality as the individuals’ Desirability. The most obvious operational definition of that variable consists of obtaining a score by weighting the items proportional to their desirability values. Both these values and the item scores are best expressed as positive and negative deviations from the neutral midpoint of the scale. In the absence of desirability values, the first principal component of a heterogeneous and representative set of traits will closely approximate the desirability variable. The desirability score reflects the extent to which an individual is assessed to have desirable versus undesirable qualities. The result will be that most people are found to be desirable, although some are more desirable than others. A few people would be assessed to be undesirable.

The implied conception of personality is literally perpendicular to the neutral view according to which, for example, there are no right or wrong answers to the items of a questionnaire, and by which all people are equally desirable, just different. One could of course use the desirability variable just to partial it out, and retain a value-free, neutrally descriptive account of personality, as in Saucier’s (1994) model. Here, on the contrary, it functions as the pivot around which the personality hypersphere revolves. The present approach is comparable to emphasizing the $g$ factor of intelligence, rather than its multidimensional conception according to which people are just differently intelligent (even though no one, to my knowledge, has gone as far as to partial out $g$). There can hardly be any doubt that capitalizing on $p$ provides the most realistic account of personality.

In the present context, the social part of social desirability is terminologically dubious. It could be used in opposition to personal desirability, but then the proper specification would be intersubjective versus subjective. In its actual use, SD refers primarily to impression management in self-report. This socially desirable responding (SDR) may be an interesting topic of study in its own right, but it is not at issue here. People have desirable and undesirable traits; they show overall differences in the extent to which that is the case; there is substantial agreement among third persons, and even between self and other, about someone’s desirability score; its heredity coefficient is undoubtedly in the same order of magnitude as with other traits, as it is a linear combination of them. Socially desirable responding is orthogonal to these individual differences: In a Persons × Assessors × Situations design with Desirability as the dependent variable, SDR is an Assessors main effect (e.g., a self-assessment may be
relatively socially desirable), and/or a Situations effect (e.g., a personnel selection context gives rise to elevated scores), and/or some interaction effect, but not a Persons or individual differences effect. The $p$ component concerns the latter.

Carrying out the slight rotation, if needed, to align the first principal component in any particular data set with the desirability variable should prove helpful in solving the vexing problem of indeterminate rotational positions of components. Saucier (2002a) has already documented that varimax rotation does not help in this respect: Across data sets, the positions of unrotated principal components were at least as replicable as were varimaxed components. Among the principal components, the first is by far the most replicable one. Across differently composed sets of variables, however, part of this stability gets lost (as discussed earlier). Anchoring $p$ at the desirability values, which are external to the studies, should enhance replicability.

The $p$-oriented model produces another taxonomic lever, namely, a measure of the representativeness of a set of personality traits or items, in the shape of the correlation between the first principal component of the set and the desirability variable. In a set overloaded with fairly neutral extraversion-introversion items attracting the first principal component, that correlation would be clearly below unity. In a heterogeneous set of neutral items, the desirability variable would be unstable, again lowering the correlation. In the spirit of the lexical axiom, such sets would be judged insufficiently representative. The proposed measure simulates that judgment.

The Two-Dimensional Level

Upon extracting $p$, a residual remains in the shape of a matrix of part scores. The first principal component of that residual matrix comes close to the second principal component of the original scores, at least in a representative set of variables. Taking $p$ as the ordinate, a 45-deg counterclockwise rotation of the two components including $p$ will produce an X structure, or a flat version of the double cone. The upper and lower segments contain the most unambiguously positive and negative, or consonant, traits; the left and right segments contain the most relatively neutral and discordant traits. The abridged circumplex structure at this level contains two bipolar facet vectors running from 11 o’clock to 5 o’clock and from 1 to 7 in addition to the 12 to 6 $p$ vector; the relatively neutral traits are left unaccounted for by the model, as their projections on the vectors will be very low.

Substantively, the plane would resemble, but not be identical to, the interpersonal circumplex (Wiggins, 1980), the $I \times II$ or Agreeableness $\times$ Extraversion slice of the 5-D structure, Digman’s (1997) $\alpha \times \beta$ plane, and the like. In a perfectly representative set of traits as defined earlier, the model plane would be identical to the plane formed by the first two (rotated) principal components; this property makes it a good candidate for a canonical or reference structure. Its suitability for that purpose is enhanced by the absence of rotational freedom at this level: The positions of the model vectors are indirectly prescribed by the desirability values of the traits. Theoretical criteria, as in the interpersonal circumplex, or the simple-structure criterion as in 5-D models, are insufficiently capable of serving that reference function.

In the rationale of the semicircumplex model, the transition from one $p$ dimension to two dimensions means a spreading of the desirability component, in the manner of the unfolding of a fan. The primordial one becomes diluted in the process, like the $g$ factor of intelligence does when it is spread over two or more dimensions. Following elementary rules of parsimony, the transition should not be made lightly; the burden of proof is on those who take the step. Psychometrics offers an adequate procedure for this proof: More-dimensional assessments of personality should be shown to have sufficient incremental validity over the $p$ component. This requirement implies that an assessment of $p$, as a baseline variable, would have to be part of any empirical study of personality.

Incremental validity of variables other than $p$ would necessarily imply that variance orthogonal to it, thus neutral variance, is valid. This implication bridges the present family of models and those that capitalize on neutral variance, like Peabody’s (1984) and Saucier’s (1994; Saucier et al., 2001). In fact, the latter model is the complement of the Semi-Circumplex, at the present and subsequent dimensional levels; it fills in what the present model leaves empty. Although the basic assumption—potential incremental validity of neutral variance—is thus necessarily the same, a strategic difference remains at the executive level. In the semicircumplex approach, neutral variance is assessed indirectly, by suppressing the $p$ variance from consonant traits rather than directly, as in Saucier (1994). The reason was given earlier: Discordant personality concepts are difficult to handle.

Semicircumplex Spheres and Hyperspheres

The three-dimensional member of the model family arises as follows: Add the second principal component of the matrix of residual scores (after removing $p$); retain the vertical orientation so that a globe is formed with the positive traits on the northern hemisphere and the negative traits on the southern one; perform an orthogonal rotation of the three axes
including \( p \) so that they become equidistant (the angles being 54.7 deg, with cosine \( \sqrt{7/3} \); further constraints are discussed later) from the vertical axis. All this is in correspondence with the double cone model. Now form three slices (circumplexes) by taking two rotated axes at a time. The projection of \( p \) on these tilted planes has the 12 o’clock to 6 o’clock direction, and the 3 o’clock and 9 o’clock positions are on the equator. Additional model vectors are constructed running from 11 to 5 and from 1 to 6, as in the two-dimensional member of the model family.

The central positions in this structure are taken by the 12 to 6 vectors—to be labeled I/II, I/III, and II/III—that are the bisectrices of the right angle between the two rotated principal components forming the circumplex. The I, II, and III axes themselves merely guard the boundaries of the model structure; as such, they have no place or name in the model. The central model vectors appear to be close to \( p \), namely, at a distance of 35.3 deg (with cosine .816 or \( \sqrt{2/3} \); generally, \( \sqrt{2/n} \), where \( n \) is the number of dimensions). Note that this oblique structure arises as a side effect of an orthogonal rotation, not through some more liberal oblique rotation procedure as such. The central model vectors are thus much more saturated with desirability than are the factors themselves; at all dimensional levels of the model, they share exactly \( \sqrt{2} \) as much variance with \( p \) as do the orthogonal factors.

What is new about this structure is that mixtures or blends of factors have stolen the central place that used to belong to the factors. Instead of being derivatives, the bisectrices of the factor pairs have become the central concepts. This play of musical chairs comes about because of the closer association of the central vectors with \( p \), which entitles them to their position. In passing, the model resolves the uneasiness of inserting orthogonal axes into an essentially oblique structure; it rigorously defines oblique axes without giving up the convenience of an orthogonal base. The only price is that the number of musical chairs has to be increased, from four dimensions onward: There are \( n(n – 1)/2 \) central vectors, with \( n \) the number of dimensions or factors. However, that extension will be welcomed by those who have always wondered whether five is all there is. The model has shaken off the last remnants of simple-structure thinking. Parenthetically, I note that the model is equally appropriate in other domains, notably, intelligence.

With four dimensions, the rotated factors are at an angle of 60 deg with respect to the \( p \) factor; the central model vectors are at 45 deg from that pivot. With five dimensions, the factors are at 63.4 deg, and the central vectors are at 50.7 deg. Still, the model rotation maximizes the sum of the correlations of the central axes with \( p \), and in that sense minimizes their average neutralness. Conversely, any other orthogonal rotation of these dimensions (e.g., varimax) is inferior in this respect: It takes in more neutral traits, which are less representative of the domain.

With three or more dimensions, the model leaves freedom of spin. A three-dimensional structure, for example, may be rotated around its vertical \( p \)-axis without violating the model. For reasons of continuity, this freedom may be used for maximizing the correspondence of the rotated factors with the current varimax factors, particularly, the 5-D model factors. This amounts to some lowering of the positive poles of the current dimensions I and III toward the hyperequator, and some lifting of the others. One may speculate, for example, that the American lexical extraversion factor loses its aggressive connotation and moves in the direction of sociability. However, it is difficult to gauge what the substantive effects of the joint rotation will be on all the versions in all the different languages (see, e.g., Saucier et al., 2000) that have been proposed. The labels of the 5-D model are probably used in a manner vague enough to permit this twisting. (Agreeableness and conscientiousness, in particular, do not even fit their present axes; see Hofstee et al., 1992.)

From the three-dimensional level on, there is some redundancy between model vectors at different levels. At the top level, there is the one vector. At the second level, two additional bipolar vectors appear, which satisfy the requirement of being 30 deg removed from \( p \). At the third, we find three semicircumplexes with three model vectors each; at the fourth, there are \( 6 \times 3 \) at the fifth, the AB5SC model with 30 vectors appears; in general, at the \( n \)th level from 3 on, there are \( 1.5n(n – 1) \) vectors specific to that level. In successively adding levels, the cumulative number of model vectors thus becomes 1, 3, 12, 30, and 60. From the third level on, it appears impossible to rotate the central vectors in such a way that all the additional vectors stay at least 30 deg away from the ones at the second level. Thus some vectors would have indistinguishable interpretations.

One strategy would be to settle for a particular dimensionality of the trait space. That would prevent overlap and would simplify things in general. The foremost drawback is that from three dimensions onward the most central trait concepts would be missed. Furthermore, that strategy would only stir up the debate on the dimensionality of the trait space, to which there is no cogent solution; it would thus frustrate the attainment of a canonical structure rather than contribute to it. The other, preferable, strategy is to adopt the model family as a whole, including as many (or as few) levels as will appear to be needed, and deleting concepts at lower levels that are virtual clones of those at higher levels. The foreseeable result of this strategy is maximal convergence of structures at each level, and maximal efficiency in communicating about
personality up to a particular level. In this manner, the family model may become a model family.

CONCLUSION

This probe into the credentials and future of the 5-D approach to personality ends in cautious optimism. Because of its reliance on the questionnaire method, the 5-D paradigm stays at the phenotypic level; however, an efficient and coherent description of personality is indispensable also for research on genotypic and other determinants of individual differences. The exploitation of the lexical axiom, with its rich history dating back into the nineteenth century, in combination with PCA of large data sets that became feasible only in the last decades of the twentieth century, has provided a firm base for efficiently and coherently describing personality differences. Of the several and diverging taxonomic models that have arisen in the 5-D tradition, I used elements to design the contours of an integrative structure that may serve scientific and applied communication.

In the process, the penetrating evaluative aspect of personality description exerted its influence. Its pervasiveness constitutes a fundamental and often frustrating problem to the field. I have chosen to adopt the strategy recommended to bridge players who find themselves in a squeeze: Relax and enjoy it. Desirability cannot be circumvented or suppressed without sacrificing the first principal component of personality description. So it might as well be squarely faced and put in the most central position. Giving in to the desirability component of personality will in all likelihood be rewarded with a more coherent, stable, and internationally replicable conception of personality structure.

REFERENCES


Goldberg, L. R. (1992b). The structure of phenotypic personality traits (or the magical number five, plus or minus zero). Keynote lecture presented at the 6th European Conference on Personality, Groningen, Netherlands.


