Social Interaction Design for Social Media: The Case of Groupthink and Aggression

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Informatics)

Michail Tsikerdekis

Brno, October 2012

Advisor:
RNDr. Jiří Zlatuška, CSc.
Department of Computer Science
DECLARATION

I hereby declare that this dissertation is my own original work and has not previously been submitted to any other institutions of higher education. I further declare that all sources cited or quoted are indicated and acknowledged by means of a comprehensive list of references.
ACKNOWLEDGEMENT

I would like to thank several people that made this dissertation possible.

First and foremost, my advisor Jiří Zlatuška for his help and support with my work and studies throughout my doctoral journey. His direction, ideas, and encouragement were essential for the development of this interdisciplinary research.

Further, I would like to thank all faculty, staff and students at Masaryk University and especially the Faculty of Informatics that have assisted me and being influential throughout my PhD candidacy. I extend a sincere thank you especially to Bernadette Nadya Jaworsky whose research insights and assistance as well as teaching support have been influential throughout my journey, Eva Hladká whose help and support has given me a greater understanding of teaching and Ada Nazarejová whose assistance with the faculty’s procedures and culture has proven invaluable.

I would like to thank every single one of the respondents in my research studies. Without all of these volunteers the data would have never been collected. I would also like to thank my mother, father and brother as well as my friends. Without your love and support throughout my life and especially my dissertation process I would not be where I am today. Lastly, I would like to thank my other half Ani, you have always stood by my side through the highs and lows and always believed in me even when I did not believe in myself. Thank you for your never ending love and support.
ABSTRACT

Social media today play an increasingly important role in computer science, the information technologies industry and society at large, changing people’s everyday communication and interaction. The domain of social media encompasses a variety of services, such as social networking services, collaborative projects, microblogging services and even virtual social worlds and virtual game worlds. There are long established principles, guidelines, and heuristics that apply to social media design and are part of the foundations of human-computer interaction (HCI). For example, in interaction design two set of goals guide the design of systems, usability goals and user experience goals. However, current design and development frameworks are still ill-equipped for the ever-changing online world. Ironically, they fail to take into account the social dimension of social media software. Cracks in the social fabric of a community operating under social media software may have devastating effects, not only to the evolution of the community but also to the longevity of the social media service. As such, social media cannot be developed in isolation, without taking into consideration the social experiences of users. Psychological and sociological principles should become part of the design process of modern social media. My research contributes to this endeavor by focusing on the design and engineering of social experiences on social media services. In my dissertation, I propose that an additional layer be added to the usability and user experience goals. The new layer includes social experience goals, which are further classified as desirable, undesirable and neutral. I produced a new definition for social interaction design that incorporates social experience goals. Building upon previously developed frameworks and models for interaction design, I demonstrated how social interaction design applies to activities such as needfinding, developing alternative designs using prototyping and modeling, developing interactive versions of design and evaluating designs. I presented the benefits of using such framework by focusing on two showcase phenomena deeply rooted in social behaviors - aggression and groupthink. The aim was to demonstrate that social media design and development could be driven by goals that aim to increase collaboration and decrease conflict in a community. I analyzed the effects of different features found in social media today in respect to aggression and groupthink, and found positive evidence to suggest that social interaction, behavior, attitudes and phenomena can be affected by social media design. By examining two vastly diverse social experience goals using quantitative as well as qualitative research methods currently used in HCI, I demonstrated the usefulness of social interaction design in various classifications of social media services such as collaborative projects, social networking sites and even virtual game worlds. In short, I argue that social experience could be engineered through software using the framework I provide for social interaction design.
CONTENTS

1 Introduction ................................................. 1
  1.1 Definitions ............................................. 3
    1.1.1 Human-Computer Interaction ......................... 3
    1.1.2 Software Engineering ................................. 5
    1.1.3 Interaction Design, User-centered and Affective Design 6
    1.1.4 Social Interaction Design ............................ 8
    1.1.5 Social media ........................................ 9
  1.2 Examples of Social Media Issues ....................... 10
  1.3 Aim of This Study ........................................ 11
  1.4 Significance ............................................ 14

2 Web 2.0 and Social Media History .......................... 15
  2.1 Introduction ............................................. 15
  2.2 Web 1.0, 2.0 and 3.0 .................................... 16
  2.3 Social Media ............................................. 19
    2.3.1 Blogs ............................................... 20
    2.3.2 Collaborative Projects ............................... 21
    2.3.3 Social Networking Sites ............................. 22
    2.3.4 Content Communities ................................. 24
    2.3.5 Virtual Social Worlds ............................... 25
    2.3.6 Virtual Game Worlds ................................. 25
    2.3.7 Microblogging ....................................... 26
  2.4 Additional Considerations .............................. 27
  2.5 Conclusion .............................................. 29

3 Groupthink Theory and Application ........................ 30
  3.1 Introduction ............................................. 30
  3.2 Defining Groupthink ...................................... 30
  3.3 Examples of Social Media Endangered by Groupthink .... 31
  3.4 The Beginnings of Groupthink and Research ............. 32
    3.4.1 The Groupthink Model ............................... 32
    3.4.2 Historical Examples of Groupthink .................. 34
    3.4.3 A Model of Controversy ............................. 36
    3.4.4 Preventing Groupthink ............................... 37
  3.5 Groupthink and Online Groups ........................... 38
    3.5.1 Researching Online Groupthink ...................... 39
  3.6 Software Features Studied for Groupthink ............... 40

4 Researching the Individual Features for Groupthink..... 42
  4.1 Introduction ............................................. 42
  4.2 Pro/con Lists ............................................ 42
    4.2.1 Method and Initial Results .......................... 43
    4.2.2 Analysis for Argument Production .................. 43
    4.2.3 Bayesian Analysis for Readability ................. 47
4.2.4 Discussion .................................................. 47

4.3 Alternatives Randomization in Voting Interfaces .............. 50
  4.3.1 Literature on Response Order and Behavior ............... 51
  4.3.2 Assessing the Appropriate Response Order ............... 52
  4.3.3 Conclusions .............................................. 54

4.4 Leaders in Social Media Affecting Opinions .................... 54
  4.4.1 On Leaders and Groupthink ................................ 55
  4.4.2 Method .................................................. 56
  4.4.3 Results .................................................. 57
  4.4.4 Discussion .............................................. 60

4.5 Anonymity States and their Effect on Non-Conformity and Groupthink .............................................. 61
  4.5.1 Method .................................................. 62
  4.5.2 Results .................................................. 62

5 Online Aggression Theory and Application .......................... 68
  5.1 Introduction .................................................. 68
  5.2 On Aggression in General ..................................... 68
    5.2.1 Definition .............................................. 68
    5.2.2 Types and Categories of Aggression ....................... 69
    5.2.3 The Nature of Aggression ................................ 70
  5.3 Research on the Causes for Human Aggression ................ 70
    5.3.1 Theories on Aggression .................................. 71
    5.3.2 General Aggression Model ................................ 72
    5.3.3 Additional Related Phenomena ......................... 73
  5.4 Research on Online Aggression .................................. 74
    5.4.1 Media Violence Research ................................ 75
    5.4.2 Online Disinhibition Effect .............................. 76
    5.4.3 Benign & Toxic Online Disinhibition Effects ............. 79
  5.5 Preventing Online Aggression .................................. 79
    5.5.1 Software Features Studies for Aggression ................. 81

6 Researching the Individual Features for Aggression... .......... 82
  6.1 Introduction .................................................. 82
  6.2 Anonymity States and Aggression ................................ 82
    6.2.1 Background Theory ...................................... 83
    6.2.2 Method .................................................. 83
    6.2.3 Findings and Results .................................... 84
    6.2.4 New Hypotheses for the Additional Analysis .......... 84
    6.2.5 Discussion .............................................. 88
  6.3 Quality of Error Messages and Suggestion of Alternatives ... 89
    6.3.1 Error Messages and Frustration .......................... 89
    6.3.2 Method .................................................. 90
    6.3.3 Results .................................................. 92
    6.3.4 Discussion .............................................. 95
  6.4 Message Submission Errors and Recovery ...................... 96
    6.4.1 Messages, Communication and Aggression ............... 96
    6.4.2 Method .................................................. 98
7 The Social Interaction Design Framework

7.1 A New Look on Interaction Design

7.1.1 Defining Social Interaction Design within Interaction Design

7.2 Needfinding: Identifying Social Experience Goals

7.2.1 Example: Social Experience Goals in UrCity

7.3 Prototyping, Conceptual Models and Physical Models

7.3.1 Example: Prototyping and Modeling for UrCity

7.4 Evaluating Social Interaction Designs

7.4.1 Example: Evaluation for UrCity’s Features of Interest

7.5 Additional Aspects for Consideration

7.5.1 Social Cognition

7.5.2 Distributed Cognition Framework

8 Conclusion

8.1 Reflecting on the Dissertation

8.2 Looking into the Future
LIST OF FIGURES

2.1 Classifications for social media 20
2.2 Google’s official blog. 21
2.3 Wikipedia’s encyclopedia page on pseudonymity. 22
2.4 IBM’s fan page on social networking site Facebook. 23
2.5 Masaryk’s University promotional video on content community, YouTube. 24
2.6 Twitter page for the Faculty of Social Studies, Masaryk University. 27

4.1 Descriptive results for pro/con lists and textual collaborative interfaces in terms of total argument production and unique arguments. 44
4.2 The prior and posterior distributions of effect size δ, when comparing total arguments between Debatepedia and Wikipedia. The top graph illustrates the unrestricted t-test for H₁, the middle graph illustrates the order-restricted test stating that Wikipedia produced more arguments than Debatepedia (H₂), and the bottom graph illustrates the t-test for the opposite order-restriction (H₃). The dots mark the height of the prior and posterior distributions at δ = 0. 46
4.3 The prior and posterior distributions of effect size δ, when comparing uniqueness of arguments between Debatepedia and Wikipedia. The top graph illustrates the unrestricted t-test (H₁), the middle graph illustrates the order-restricted test stating that Wikipedia produced more unique arguments than Debatepedia (H₂), and the bottom graph illustrates the t-test for the opposite order-restriction (H₃). The dots mark the height of the prior and posterior distributions at δ = 0. 47
4.4 Descriptive results for pro/con lists and textual collaborative interfaces in terms of readability measures FRE, Flesch-Kincaid and SMOG. 48
4.5 The prior and posterior distributions of effect size δ, when comparing readability between Debatepedia and Wikipedia. Each sub-graph illustrates the unrestricted t-tests (H₁), the order-restricted test stating that Wikipedia scored higher in readability than Debatepedia (H₂), and the Bayesian t-test for the opposite order-restriction (H₃). The dots mark the height of the prior and posterior distributions at δ = 0. 49
4.6 The first version of the voting process with the leader supporting the first solution. 57
4.7 The second version of the voting process with the leader supporting the third solution. 58
4.8 Demographic information for the leadership survey.

4.9 Simple graph for the votes for each solution per treatment.

4.10 Research model for research on anonymity states and groupthink with results. Note: * \( p < .05 \), ** \( p < .01 \), *** \( p < .001 \)

4.11 Likelihood of not conforming for each anonymity state.

4.12 Likelihood of not conforming and perception of anonymity.

4.13 Line chart displaying the gradually ascend of the perception of anonymity as progressing to higher anonymity states.

4.14 Likelihood of not conforming for all three scenarios based on the self-reported level of importance. The second graph is significantly being impacted by this dependency.

6.1 3D clustered graph depicting the responses that individuals gave under each anonymity state and their reported strength of opinion on a topic. Responses under the state of pseudonymity (here depicted with a light brown color) show a substantial effect.

6.2 Responses based on anonymity states for each one of the three scenarios. Each left column represents complete anonymity, each middle column represents using real names and each right column represents pseudonymity.

6.3 Model for level of aggression in responses under pseudonymity based on situational context and strength of opinion about a topic.

6.4 The two versions of 404 errors that respondents encountered.

6.5 Questions included in the survey investigating the effectiveness of error messages.

6.6 Graphs depicting demographic information for the sample that took part in the AG1 survey.

6.7 Results for the three aspects of the two versions of error messages evaluated by respondents.

6.8 Survey questions for the message submission error and recovery survey.

6.9 Bar chart depicting the ten-point scale for importance of a message and time availability.

6.10 The error submission model based on the results.

7.1 Social Interaction Design Goals adapted from the Interaction Design goals developed by (Rogers, Sharp, & Preece, 2011)
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCI</td>
<td>Human-Computer Interaction</td>
</tr>
<tr>
<td>SE</td>
<td>Software Engineering</td>
</tr>
<tr>
<td>ACM SIGCHI</td>
<td>Association of Computing Machinery Special Interest Group on Human-Computer Interaction</td>
</tr>
<tr>
<td>GAM</td>
<td>General Aggression Model</td>
</tr>
<tr>
<td>NHST</td>
<td>Null Hypothesis Statistical Testing</td>
</tr>
<tr>
<td>SNS</td>
<td>Social Networking Sites</td>
</tr>
<tr>
<td>lxD</td>
<td>Interaction Design</td>
</tr>
<tr>
<td>sIxD</td>
<td>Social Interaction Design</td>
</tr>
<tr>
<td>UCD</td>
<td>User-Centered Design</td>
</tr>
<tr>
<td>GDSS</td>
<td>Group Decision Support Systems</td>
</tr>
<tr>
<td>IT</td>
<td>Information Technology</td>
</tr>
<tr>
<td>Ajax</td>
<td>Asynchronous JavaScript and XML</td>
</tr>
<tr>
<td>REST</td>
<td>Representational State Transfer</td>
</tr>
<tr>
<td>CMC</td>
<td>Computer-Mediated Communication</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

Social media are a driving force of today’s world. The sheer economic potential has mesmerized the business world (Libert, 2010). The opportunity for a young generation to actively contribute and share content has affected individuals, their families and society as a whole that can now share information free from restrictions of the physical world (Palfrey & Gasser, 2010). Software designers and developers are pushing the abilities of their field to the limit and attempt to break new grounds and deliver services in an extremely competitive Internet world that resembles the movie industry. However, unlike Hollywood, this is not a traditional industry.

Sharing and interaction is the moving force for social media (Kaplan & Haenlein, 2010). The iterative process of a project’s life-cycle goes beyond revising the final product; if there can ever be a final product in social media. This is a young, rapidly growing and ever-changing field (Poynter, 2010). Software engineers and users are interacting reciprocally and passionately demand new breakthroughs. This information society is not just hungry for digesting and sharing information but also for new technologies that will allow for new levels of interaction between users. Today’s forecast shows that “the rate of change and development of new features of social media will continue for the foreseeable future” (Poynter, 2010, p. 256). This is a world where social media services can grow rapidly through word of mouth alone (Boyd & Ellison, 2007).

The role of Human-Computer Interaction (HCI) in the development of new software and technologies such as social media has been the center of discussion among computer scientists (Canny, 2006; Cogburn, 2003; Wang, Carley, Zeng, & Mao, 2007). However, evidently the field is struggling to keep up with the technologies being developed. Social media form a diverse field of various environments varying from Social Networking Sites (SNS) to video games (Kaplan & Haenlein, 2010, 2011). The sheer variability of software and interfaces along with rapid growth seems to have led to a disruption of communication between HCI scientists and developers. Examples such as the downfall of Friendster (Boyd & Ellison, 2007), examples of social media privacy issues (Fogel & Nehmad, 2009; Gross, Acquisti, & III, 2005; Nippert-Eng, 2010) and other unwanted incidents (Atkinson, 2008; Hinduja & Patchin, 2010; Wallace, 2001) paint a really grim future for social media services and their relation to HCI. Perhaps however, a
more evidently worrying picture for the future of HCI is painted in a Microsoft Research Ltd. report (Harper, Rodden, Rogers, & Sellen, 2008). The report aims, among other things, to identify the future challenges for HCI in the emerging digital world for the 2010s. A statement is made for an outdated HCI field that needs to move forward (Harper et al., 2008):

Technology is changing, people are changing, and the society is changing. All this is happening at a rapid and rather alarming rate. [...] HCI needs to extend its methods and approaches so as to focus more clearly on human values. This will require a more sensitive view about the role, function and consequences of design, just as it will force HCI to be more inventive. HCI will need to form new partnerships with other disciplines, too, and for this to happen HCI practitioners will need to be sympathetic to the tools and techniques of other trades. Finally, HCI will need to re-examine and reflect on its basic terms and concepts. Outdated notions of the ‘user’, the ‘computer’ and ‘interaction’ are hardly sufficient to encompass all that HCI will need to attend to. (p. 52)

Developing design perspectives that consider the implications of design on user interactions and social experiences is particularly important for social media development but also currently problematic. At first glance, issues and incidents in social media demonstrate that some features implemented in the software are not investigated thoroughly in terms of their impact on the users. Additionally, the current HCI framework lacks the proper tools to deal with the challenges of the social dimension in social media. Sociological and psychological theories are sparsely implemented or considered in the work and research in HCI. In turn, questions regarding the social aspects of social media software development lack adequate answers. What is the effect of software on social interactions and behavior? Through what means can one develop software in order to proactively engineer these interactions and behaviors? What social aspects should be considered relevant to social media software and how can someone identify them? What approaches can be taken to design and develop software in relation to these aspects of interest?

This dissertation attempts to make a contribution towards the resolution of these issues. Theoretically ground theories are extracted and empirical research is employed with a much needed interdisciplinary perspective in HCI. Social media should be designed and developed with goals for improving social interaction and experiences. I argue based on evidence that this is not only possible but beneficial for social media. To achieve this, I present in this dissertation original research demonstrating that changes in software design features currently being used in social media, can affect social behavior and improve the social experiences of the users. I focus on two frequently studied sociological and psychological phenomena as a showcase: groupthink and aggression. I demonstrate how designers and developers can make choices and aim to predict, manage or even alter social behavior in social media. My dissertation research has found positive evidence for several features affecting groupthink and aggression.
1. Introduction

After a literature review on groupthink (chapter 3), I have conducted original research for software features that potentially had an effect on groupthink (chapter 4). I have found compelling evidence that: (a) anonymity states empower users to contribute their alternative solutions to a discussion (Tsikerdekis, in press-b), (b) pro/con lists produce more unique arguments as compared to textual collaborative interfaces (Tsikerdekis, in press-c), (c) publicly identified leaders do not affect decision-making voting processes, (d) item randomization in voting processes has great potential for reducing groupthink as opposed to standardization, and (e) dynamic voting status indicators do not affect voting outcomes (Tsikerdekis, in press-a).

Similarly, after a literature review on aggression (chapter 5), I conducted original research for various software features that potentially affected aggression (chapter 6). I found positive evidence that: (a) anonymity states and specifically pseudonymity is a factor for aggression (Tsikerdekis, 2011, 2012), (b) message submission errors could severely impact communication, be the cause of implications and potentially aggression, (c) quality of error messages in terms of helpfulness and comprehension can reduce frustration for users and in turn aggression.

Finally, I further used my findings to develop a framework through which software designers can identify desirable as well as undesirable social interactions, behaviors, attitudes and phenomena of interest and established a methodology for identifying features which can promote or inhibit them (chapter 7). The framework describes a new layer of Social Interaction Design (SIxD) based on social experience goals which is applied to the activities currently found in Interaction Design (needfinding, prototyping, interactive versions, and evaluation) (Rogers et al., 2011). I provide guidelines for each of these activities in relation to social interaction design and provide examples of their application based on the features investigated in this study for groupthink and aggression. The aim of this framework is to provide a new definition and approach to social interaction design that distinguishes it from traditional interaction design perspectives that aim to understand interactions and behaviors between a user and a system. Social interaction design aims to understand social interactions between users through a system and find ways to engineer the system in order to affect these interactions.

In this chapter, I introduce and open a discussion for some of the definitions used throughout this study. I proceed with presenting some examples of issues on social media and consequences of their design mismanagement. I further present the aims of this study and discuss the significance of this work on the current domain of HCI and social media.

1.1 Definitions

1.1.1 Human-Computer Interaction

The world of computers prior to the first personal computer remained private for scientists, information technologies professionals and dedicated hobbyists (Carroll, 2012). As software complexity and a new graphical interface came to be, the need for an Software Engineering (SE) evolution became important. Usability is-
issues that arose with the appearance of the first computer users made it apparent to scientists that there was a need for a new interdisciplinary project that could produce software and hardware which aimed to be friendly to the end-user; a user-friendly software.

According to Myers (1998), research for various aspects of usability was not unlikely for researchers that wanted to establish evaluations for the first graphical environments as well as hardware input objects such as the mouse. Direct manipulation of objects within an interface with a pointing device was first demonstrated in a 1963 MIT PhD thesis (Sutherland, 1964). Another example that today accompanies almost every computer in various forms is the mouse which was developed at Stanford Research Laboratory in 1965 (English, Engelbart, & Berman, 1967). In fact, most of the software and hardware usability tests were set as standards by these early studies.

A general definition of Human-Computer Interaction by the Curriculum Development Group of the Association of Computing Machinery Special Interest Group on Human-Computer Interaction (ACM SIGCHI) describes it as “a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them.” (Hewett et al., 1992, p. 5). However, a unified yet simple way of defining HCI is perhaps the description given by Carroll (1997) in an article.

Human-computer interaction (HCI) is the area of intersection between psychology and the social sciences, on the one hand, and computer science and technology, on the other. HCI researchers analyze and design-specific user-interface technologies (e.g. three-dimensional pointing devices, interactive video). They study and improve the processes of technology development (e.g. usability evaluation, design rationale). They develop and evaluate new applications of technology (e.g. computer conferencing, software design environments). Through the past two decades, HCI has progressively integrated its scientific concerns with the engineering goal of improving the usability of computer systems and applications, thus establishing a body of technical knowledge and methodology. HCI continues to provide a challenging test domain for applying and developing psychology and social science in the context of technology development and use. (p. 61)

This description encapsulates the basic goals of HCI and the epistemological challenges that it faces although the article precedes major technological breakthroughs in the world of the Internet such as social media. For HCI to advance forward, the ease of use and learning soon had to be replaced by more complex measurements in order to account for qualities like fun, well-being, collective efficacy, aesthetic tension, enhanced creativity, support for human development, and many others (Carroll, 2012). Today, HCI expanded to encompass visualization, information systems, collaborative systems, the system development process, and many areas of design. Additional fields include information technology, psychology, design, communication studies, cognitive science, information science, science and technology studies, geographical sciences, management information systems, and industrial, manufacturing, and systems engineering. All
of these are used to achieve the fundamental goals of HCI “to develop or improve the safety, utility, effectiveness, efficiency, and usability of systems that include computers” (Diaper, 1989, p. 3). When designing HCI, the degree of activity that involves a user with a computer should be analyzed extensively in all of the various different user activity levels such as physical (Chapanis, 1965), cognitive (Norman, 1986, 2002), and affective (Picard, 2000).

Modern HCI seems to approach the vision of the “Man-Computer Symbiosis” (p. 4) that was once described by Licklider (1960). HCI scientific work not only was necessary for the development and launch of personal computers but helped promote many Internet technologies which are in use today. As computers such as IBM’s Watson (IBM, 2012) gain the ability to speak and process natural language, HCI’s unique role of establishing and studying interactions will ensure a peaceful symbiosis with the new superstructure of networked people and computers.

1.1.2 Software Engineering

Much more challenging is trying to define Software Engineering. In an attempt to try and dissolve the confusion that arose during the 1980s, Humphrey (1988) tried to establish a basic definition. In the article he stated that “software engineering refers to the disciplined application of engineering, scientific, and mathematical principles and methods to the economical production of quality software” and then proceeded by clarifying that “the term quality refers to the degree to which a product meets its users’ needs, This may refer to functional content, error rates, performance, extensibility, usability, or any other product characteristics which are important to the users” (Humphrey, 1988, p. 82). Today, almost twenty years since this definition was published, the Bureau of Labor Statistics from the United States Department of Labor published their own similar definition for what is a computer software engineer in their Occupational Outlook Handbook, 2010-2011 Edition (Bureau of Labor Statistics, 2011). The report states that “computer software engineers design and develop software. They apply the theories and principles of computer science and mathematical analysis to create, test, and evaluate the software applications and systems that make computers work.”

These two definitions ostensibly state that software engineering is an interdisciplinary field which includes principles and goals in line with human-computer interaction. However, even though in software design and development both approaches are expected to be implemented and work alongside, this is not always the case. Bowles and Box (2011) describe the relationship as follows:

The fields of design and development are yin and yang, complementary yet opposed. They rely on each other, and thus either support or damage the quality of each other’s work. Poor execution in development can ruin the experience you’ve worked so hard on. (p. 132)

In fact, there seems to be a fundamental difference in the perception of practitioners in the two fields. A review study concludes that human-computer interface specialists are user-centered and software engineers are system-centered (J. Brown, 1997). This sets the stage a problematic alliance between two fields that
1. Introduction

It seems to be competing rather than cooperating. The same study made a further assertion that two factors seem to be the main constituents of the current state of affairs between the two fields. First, the contributions of HCI specialists are never included in software engineering textbooks and second, HCI methodologies do not clarify the role of the SE process relative to the HCI process. In addition, both disciplines seem to disagree in the use of formal and informal methods. Software engineering relies more on formal methods while HCI values both formal and informal methods. Downton (1993) in his book tried to establish a common ground for the two fields while trying to replace traditional software engineering intuitive methods for designing user interfaces with informal evaluation techniques established by HCI. Attempts such as these are a step forward in a direction to merge the two fields together in order to produce what was originally defined by Humphrey (1988) as quality software. Based on examples from social media software during the last decade it seems that HCI findings have yet to reach software engineers to a satisfactory degree. Additionally, HCI as a field seems to still be retaining a main focus on hardware interaction. Today studies on HCI show that “the new direction of research is to replace common regular methods of interaction with intelligent, adaptive, multimodal, natural methods” (Karray, Alemzadeh, Saleh, & Arab, 2008, p. 152). While these technologies will definitely improve physical interaction between humans and computers, this direction still ignores the social dimension which is a crucial component of social media.

1.1.3 Interaction Design, User-centered and Affective Design

An evolutionary step in the development of software that takes into account user needs came with Interaction Design (IxD). This includes several other HCI fields of which two are significantly important: User-Centered Design (UCD) and affective design.

A definition of user-centered design was given by Vredenburg, Mao, Smith, and Carey (2002) in their survey of identifying the most widely used methods and practices among UCD practitioners.

UCD is herein considered, in a broad sense, the practice of the following principles, the active involvement of users for a clear understanding of user and task requirements, iterative design and evaluation, and a multi-disciplinary approach. UCD methods are modular or identifiable processes involved in UCD practice. You should NOT think of UCD as merely usability testing or software engineering. (p. 472)

Active involvement of users is essential for the practice of user-centered design.

In a similar manner, the idea of affective design became popular as it emerged from HCI (Norman, 1983). The field of affective computing was formed with its main aim to deliver affective interfaces capable of eliciting certain emotional experiences for users (McCarthy & Wright, 2004; Reynolds & Picard, 2001). Based on the ideas of affective computing, software products contain certain features
that can invoke physiological as well as psychological responses to users and can enhance the emotional experience when using the product.

Both, user-centered design as well as affective design are incorporated in the interaction design process. In fact, Rogers et al. (2011) in their book they demonstrate that interaction design is an extremely broad concept that combines many fields that relate to the design of products. In an attempt to develop a working and short description for interaction design they stated the following (Rogers et al., 2011):

A central concern of interaction design is to develop interactive products that are usable. Generally meant to provide an easy to learn, effective to use and enjoyable user experience. (p. 2)

The most important question that interaction design addresses is: how does one optimize the user interactions with a system so they match the user activities that the system was designed for? To achieve this goal one needs to take “[...] into account what people are good and bad at [...] considering what it is that might help people with the way they currently do things [...] thinking through what might provide quality user experiences [...] listening to what people want and getting them involved in the design [...] using "tried and tested" user-based techniques during the design process” (Rogers et al., 2011, p. 8). Interaction design utilizes knowledge from a variety of different scientific fields forming a new multidisciplinary field that investigates software design from multiple perspectives. This is an extremely diverse field that tries to revise old conventions for software development which generally avoid the identification of important human factors.

Interaction design has certain concretely identified goals. Rogers et al. (2011) classify these as usability goals and user experience goals. Usability goals describe a set of goals derived from the field of usability engineering such as, effectiveness, efficiency, safety, utility, learnability, memorability. These goals are located at center of interaction design. On the other hand, equally important are the user experience goals which state that the system should be, fun, enjoyable, satisfying, entertaining, helpful, motivating, aesthetically pleasing, supporting, creative, rewarding, and emotionally fulfilling. In a figure describing the goals of interaction design, Rogers et al. (2011) assigned this goal in the peripheral spectrum of interaction design.

It becomes reasonably obvious why considering both of the above goals is so important for software development. By defining them, the software engineering team has a clearly defined vision for their software. This becomes a baseline during the development of software and should always be evaluated in regular intervals in order to establish when a project deviates from these early defined goals. The downside for such a strategy is that it requires scientific expertise. Engineering based on intuition cannot work with such a model. All aspects of the software design need to be based on solid ground theory and empirical studies. As a side-effect of having the need for expertise, this approach is more costly for projects however, the positive outcome can potentially outweigh the cost. The software that is produced based on interaction design is significantly more ef-
icient and serves user needs better. Additionally, it provides an improved user experience which is good for the general marketing of the product.

1.1.4 Social Interaction Design

Interaction design is definitely a step forward towards HCI’s future described by Harper et al. (2008), however its adoption is still limited and lacks an important aspect which should be added in the initial model; the social interaction. When it comes to developing social media, social interactions and behaviors become an unavoidable aspect of the system itself which designers should also consider.

Interaction design is considered by many to be a step forward in the future of HCI (Chan, 2006; Huang & Deng, 2008; Jakobsson, 2006). However, given the young nature of the field, definitions vary. Chan (2006) states the following about the nature of a social interaction designer:

The social interaction designer’s job is to anticipate the kinds of social interaction a system will require if it is to sustain itself once launched. Interaction designers think in terms of user behaviors. Social interaction design adds social and interpersonal attributes to the existing technical and device interactions. Users must have social competencies with the community’s theme (what it’s about, e.g. dating, jobs, etc.) and have a grasp of web navigation, links, use of form pages, search, and so on. (p. 7)

A different approach was taken by Jakobsson (2006) which not only tried to detach social interaction design from the traditional Computer-Mediated Communication (CMC) research approaches but also suggested and demonstrated that the practice could in fact include ethnographic research. The following text makes this clear (Jakobsson, 2006):

As the term computer mediated communication implies, much of the research in this field has been based on [...] the context of the phenomenon [...] taken to be the physical work context to which the communication mediating technology was introduced. In the early nineties the computer mediated communication research spawned a new branch as the arena perspective started [...] to acknowledge both that the participants perceive that the technology allows them to share a common non-physical space and that this form of social interaction offers a value in itself that makes participants spend time and effort in virtual places beyond their specific goal or purpose. (p. 62)

In other words, the environment that has been created by software engineering should be treated as a separate and unique entity with its own physical rules since it is perceived as such by users. Additionally, it also pointed out the fact that many individuals made use of these environments in ways that were not originally intended by designers.

Furthermore, the approach taken by Jakobsson (2006) demonstrates several implications that come from the ontological and epistemological approach of a
scientist. Most software development today and research in computer science and informatics takes a positivist approach that accepts that reality is a given and one can observe it objectively. However, when the social dimension is taken into account, informal methods such as qualitative research methods and especially ethnography, asserts that reality is perceived, can never be objectively stated and has to be determined from within the social experiences of the researcher. This is a good demonstration that as HCI tries to borrow from other scientific fields, it also borrows many aspects that are parts of long debates that exist within these fields. By adopting, psychological and sociological perspectives in the software development process one has to side with epistemological questions that need not to be answered with an absolute computer science perspective. While this paper tries to maintain a positivist approach in its methods and interpretation of results, it also accepts that such practice may be impossible at times. Therefore, it does not advocate that one practice is better than the other, especially since HCI is by definition an interdisciplinary field. Moreover, this way of thinking can be restrictive since for example, Bryman (2006) has shown that quantitative and qualitative methods can be combined in order to provide triangulation or completeness for scientific results; a tactic that is used in this dissertation as well.

1.1.5 Social media

Social media aim to advance not only social communication but to go beyond that and evolve social interactions. The most recent and perhaps most successful definition to describe social media was given by Kaplan and Haenlein (2010). They defined social media as “a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user-generated content” (Kaplan & Haenlein, 2010, p. 61). With a revolutionary way the “development of social media has substantially changed the way organizations, communities, and individuals communicate” (Leach, Komo, & Ngugi, 2012, p. 6).

Today, social media include blogs, microblogging, social networking sites, video sharing sites, massive multiplayer games, virtual worlds, and so on (Kaplan & Haenlein, 2010, 2011). In fact, this is just a limited description of all classifications of social media services that made an appearance in the past decade and still, new pioneering types appear regularly in an attempt to encapsulate another piece of social interaction.

This potential for expanding the domain of social interactions to the digital world is a driving force for companies that also see the potential for better marketing and in turn profits (Libert, 2010; Poynter, 2010). By analyzing identity, conversations, sharing, presence, relationships, reputation, and groups, firms can develop a congruent social media strategy based on the appropriate balance of building blocks for their community (Kietzmann, Hermkens, McCarthy, & Silvestre, 2011). Whether companies developing social media or using them in order to communicate better with their clients, the success is evident in the reports of some of the most popular social media services. Contrary to the conventional marketing wisdom were one dissatisfied customer tells ten people, in the age of social media, people have the tools to tell to 10 million customers virtually
overnight (Gillin, 2007).

At present time, Facebook which is one of the biggest social networking sites, reports 800 million active users with more than 900 million objects that people interact with such as groups, fan pages, events and community pages (Facebook, 2012). Another social media giant, YouTube, currently reports 48 hours of video being uploaded every minute with 3 billion videos being viewed by its users every day (Youtube, 2012b). These numbers paint a clear picture that it is not just the companies favoring social media but users also seem to have made them an integral part of their everyday lives. Chapter 2 provides a more detailed look on social media.

Demand for new social media services that could expand to all kinds of social interaction initiated numerous of companies providing a variety of services that keep diversifying. The rapid growth of the industry has made it popular and profitable. However, issues related with HCI have always overshadowed some of these advancements.

1.2 Examples of Social Media Issues

Evidence of inadequate communication that exists between human-computer interaction and software engineering can be found in recent examples from those involved in the social media industry. Additionally, I argue that most of these issues also demonstrate the lack of HCI to develop social media by taking social interactions into account.

Perhaps one of the most prevalent examples of software mismanagement was the case of the social networking website, Friendster. Friendster preceded most major social networking services that are known today and in many ways set the standards for many that followed. Boyd (2004) was one of the first to conduct research on the highly popular website at the time. In one of her papers she suggested “how the HCI community should consider the co-evolution of the social community and the underlying technology” (Boyd, 2004, p. 1279). Indeed this suggestion could not have been any more prophetic about what followed.

The website started off as a dating website but encouraged users to join even if they were not looking for a date (Boyd, 2003). It was built with the assumption that friends of friends are more likely to be good dates. The founders have arbitrarily decided that the distance within a network in which one can find a probable soul-mate is four degrees. As a result, any individuals beyond the four degrees could not be reached by users. The choice was obviously restrictive for the community if one takes into account the theory of six degrees of separation and the findings from the small world experiment (Travers & Milgram, 1969). There are still skeptics among the scientific community unwilling to sign on to the idea of six degrees or what boundaries of a network are defined as small (Schnettler, 2009). However, fifty years of scientific studies should have been a cause for concern for Friendster’s designers and developers. At the time that was not the case.

The website launched in the fall of 2002 its beta program and by mid-August 2003 the site had 1.5 million registered accounts and was still growing exponen-
The rapid growth created technical and social difficulties for the software engineering team. The servers could not sustain the increasing traffic of the website and the main page faltered regularly which in turn frustrated users. Boyd (2003) also described some of the social issues that were created. Friendster upset the cultural balance, contributed to the collapse of social contexts and created a phenomenon of users creating accounts based on their interests. These users were called fakesters and what they were trying to do is connect beyond the restrictive four degrees of separation that the software engineers set as a rule for the community. Many fakesters had two thousand friends in their accounts which resulted in tremendous processing loads for servers that were not designed to handle such traffic. Eventually, when software engineers were faced with a decision to resolve the problem, they decided to delete the fakesters’ accounts. This resulted in a rejection of the website in the U.S. by early adopters which were faced with a combination of technical difficulties, social collisions and a rapture of trust between users and the site (Ellison, Steinfield, & Lampe, 2007). Coincidentally, many of these issues violated also the hierarchy of needs for online users as they were adapted by Kim (2000) from Maslow’s hierarchy of needs (Maslow, 1943).

This distrust between the people and social media services which signals a probable dysfunctional relationship between HCI and software engineering can be found in many social media services today. As suggested on an article, “Social networking websites should inform potential users that risk taking and privacy concerns are potentially relevant and important concerns before individuals sign-up” (Fogel & Nehmad, 2009, p. 153). The cause for concern is justified considering that studies have demonstrated the potential for harm while on social media. Social media services can become tools for stalking, re-identification, building a digital dossier and have a fragile privacy protection nature (Gross et al., 2005).

Risks have also been found for cyber-bullying which should be a serious cause for concern. Unlike face-to-face bullying, cyber-bullying has not been the subject of research thoroughly and there are indications that it might be even more severe than face-to-face bullying and have more negative long-term effects (M. A. Campbell, 2005). Today, unfortunately there is evidence that these effects can not only be long-term but they can also have a negative impact on individuals with verbal and psychological bullying leading to physical abuse or self-abuse. Such was the case when a female teenager that fell victim to a mother pretending to be a 16-year-old boy committed suicide (Atkinson, 2008).

These are not random events or accidents but testimonies of an industry that is lacking the proper scientific understanding of the effects of software on their clients. The success and excitement for the immense potential of these new technologies should not overshadow the need for a grounded understanding and a scientific-guided software development of social media services.

1.3 Aim of This Study

Creating a viable and sustainable software design framework for developing and managing social media projects is challenging. If anything, the examples that
were described in section 1.2 show some of the problems that designers and engineers face. To add to the confusion, the current knowledge pool of the field lacks proper understanding of the effects of social media software on human factors and the social dimension. Researchers lack the ability to make casual claims because of the lack of experimental studies and because the long-term effects of these social tools are still unknown (Ellison et al., 2007). This becomes increasingly complex if designers have to account additionally for the social domain. Social interaction design could be a way forward but lacks a framework for implementation. In addition, there are no guidelines for identifying what social interactions, behaviors, attitudes and phenomena are relevant for the development and sustainability of social media services.

There are several phenomena that occur in online communities on an individual level as well as on a group level. Wallace (2001) in her book described aspects varying from impression formation, group dynamics, online conflicts to even occurrences of online altruism. Other topics of particular interest deal with online disinhibition (Gackenbach, 2007; Suler, 2004a, 2004c), social media and youth development (Palfrey & Gasser, 2010; Subrahmanyam & Smahel, 2011), computer mediated communication (Thurlow, Lengel, & Tomic, 2004) and privacy (Camenisch, Bner, & Rannenberg, 2011; Nippert-Eng, 2010) which has been the focus of many scientists in the 2000s. The amount of social behaviors, phenomena and aspects that exist and should be up for consideration when designing social media is immense.

In the light of the above, I revised my study aims in order to conclude the research within the time-frame of a doctorate study. I aimed to study cases that would be of particular interest for software design related to social media, whether for creating a virtual game world, a social networking site or even a brainstorming wiki for a corporation. As a general guideline, a specific theme was set in order to demonstrate that designers can increase collaboration and decrease conflict when developing social media. Based on this theme, two showcase phenomena were selected - groupthink and aggression.

Janis (1972) originally defined groupthink as “A mode of thinking that people engage in when they are deeply involved in a cohesive ingroup, when the members’ strivings for unanimity override their motivation to realistically appraise alternative courses of action” (p. 8). It has been attributed as the cause for Watergate, the Challenger disaster (Brooke & Tyler, 2011), and the 2008 financial crisis (Shiller, 2008). There is an extensive amount of studies on preventing groupthink in groups (Hart, 1998; Jessup, Connolly, & Galegher, 1990; Kroon, Hart, & Van Kreveld, 1991; Miranda, 1994; W.-w. Park, 1990). However, these results in their current form still cannot sufficiently be implemented in software design and development processes.

Based on the above, one of the questions that this study has raised is the viability of altering an interface in order to ensure that a decreased chance for groupthink will be observed for an online group. Preventing groupthink is a challenging task. Even in the real world when preventive measures are taken, there is no guarantee for preventing groupthink. As an example, while the homogeneity of a group is believed to be a contributing factor for groupthink (Janis, 1982),
simply adding individuals with a diverse background to a group would not benefit an organization already plagued by groupthink (Brooke & Tyler, 2011). Due to this acknowledgment, the decision was made to conduct a study with broad goals for contributing factors to groupthink than isolating the study by conducting an in-depth analysis for just one of the contributing factors. This, will also be of great help to software engineers that may be dealing with social media that do not include some of the contributing factors. Developers of a research wiki that has many scientists from all over the world coming from different scientific disciplines may not have issues with the homogeneity of a group but should still consider other contributing factors such as poor information search, not examining all alternatives (Kamau & Harorimana, 2008), opinions of leaders affecting the group (Janis, 1982) and others. There is an unprecedented opportunity to study how these factors may be translated in terms of their implementation on social media design. Detailed literature review and research results on groupthink are presented on chapters 3 and 4.

The second phenomenon under investigation was aggression. Aggression in general “is the delivery of an aversive stimulus from one person to another, with intent to harm and with an expectation of causing such harm when the other person is motivated to escape or avoid the stimulus” (Geen, 2001, p. 3). This definition also applies to online aggression. Since the intent to do harm is not under the control of software engineers, other avenues that could contribute to the prevention of such behavior were studied. Aggression can be malicious but is also attributed to situational and environmental factors (Berkowitz & LePage, 1967; Castle & Hensley, 2002; Dollard, 1989). As such, changes in environmental factors can potentially impact aggression. Software seems to be a significant contributor to online aggression (Wallace, 2001). This study maintained a focus on decreasing frustration and in turn aggression (Dollard, 1989), that could be attributed to technical or bad design issues in social media. Detailed literature review and research results on aggression are presented on chapters 5 and 6.

Based on the aforementioned considerations, I itemize the main aims of this study below.

• To investigate if software can affect social behaviors in social media by conducting novel research on two showcase online behavioral phenomena - groupthink and aggression.

• To provide a new framework for social interaction design built upon current interaction design practices that aims to engineer social behavior by altering software design.

• To provide recommendations and a methodology for software designers based on the empirical evidence for the phenomena studied.

• To advocate the benefits of social interaction design in the social media development process.

• To promote scientific research and interest in the effects of software on social behavior from HCI and advocate a better way for communicating the results from these studies to software engineers.
1.4 Significance

Research in the above phenomena as they occur online, aimed to identify if indeed software designers have more control over their users in social media. This knowledge could also help develop a new channel of communication between HCI scientists and software engineers in order to attempt to produce better quality software. As such social interaction design becomes a crucial framework for social media design and development.

Social media are taking over all aspects of Internet use. Even websites that used to be purely based on broadcasting information are shifting towards producing user-generated content and allowing users to share information. Additional benefits are also found when social media are implemented in the workplace (Ehrlich & Shami, 2010; Zhao & Rosson, 2009). The future of user experience online seems to be heading into a world where software and social interaction are intertwined. For example, a collaborative tool should be considered great when it provides efficiency for project teams but also should guarantee measures for successful decision-making. The boom in new technologies has caused the development of software to lack HCI’s insights. In addition, HCI seems to be underdeveloped to handle the social dimension that is a centerpiece of social media.

This study aims to make a contribution and attempts to close the gap between HCI and social media design and development. The goal was to not only establish how software affects social behavior online but establish a framework through which this information can be used by software engineers in order to produce quality software that will allow for more efficient social interactions. Contrary to research consisting of findings that have an observatory character, I wanted to show that not only proactive prediction of online social interaction is possible but also that designers can affect or even alter it altogether. This framework for social interaction design will assist designers and developers in delivering social media services for specific social experience goals. This information establishes a more scientifically grounded approach to the architectural design and development of software. In turn, a more sound understanding of the effects of the software on the users will result in the development of more effective software. Finally the contribution of this study also helps improve collaboration in online communities and groups by providing empirical evidence for inhibiting groupthink and aggression when adapting software features and as such demonstrating methods for future investigations of similar features.
CHAPTER 2

WEB 2.0 AND SOCIAL MEDIA HISTORY

2.1 Introduction

Before presenting the study’s research on how social media design can be adapted in order to affect groupthink and aggression, important definitions, history and arguments for the philosophy of social media development need to be presented. There is a variety of terms that relate to social media and given the young nature of the field, establishing definitions and understanding the context is essential. In addition, knowing the past can be important for predicting the future, or as President Harry S. Truman put it “You can’t know where you’re going until you know where you’ve been” (Bateson & Klopfer, 1989). The essence of this sentence can be traced all over history with names such as Adlai E. Stevenson, George Santayana and others. At its core this idea demonstrates a simple fact about human history. Times change but humanity does not. The basic needs that define us as a species remain with us as we travel through space and time. These qualities made possible for Hiltz and Turoff (1978) to make observations in their book The Network Nation which as stated by Howard (2010) were prophetic and show the advent of YouTube and Hulu. Howard (2010) points to the fact that “they based their predictions on a sociological understanding of the history of communication technologies and on deep-seated understanding about fundamental social needs rather than technological possibilities” (p. 207). These are the same social qualities that underlie in all aspects of human interaction today whether online or offline.

The reality of social processes driving technologies forward is not always obvious. In the early history of technology advances and breakthroughs happened each day. Scientists became fixated on the idea that people will adapt to technology. Bush (1945), which today is considered the inventor of hypertext (Wang et al., 2007) along with other technologies that he predicted (Millard & Ross, 2006), wrote in “As we may think”, one of his most famous essays (Bush, 1945):

Our present languages are not especially adapted to this sort of mechanization, it is true. It is strange that the inventors of universal languages have not seized upon the idea of producing one which better fitted the technique for transmitting and recording speech. Mechanization may yet force the issue, especially in the scientific field;
whereupon scientific jargon would become still less intelligible to the layman. (p. 105)

The statement is extremely fascinating and the idea of a universal language enticing. However, almost seventy years after this statement we have yet to see any universal languages that are commercially sustainable and would best fit computer technologies, especially for speech recognition. Instead, today’s speech recognition is trying to adapt to different languages and a wide range of variations such as speaker, gender, speaking rate, vocal effort, regional accent, speaking style, non-stationarity (Benzeghiba et al., 2007). On the other hand, languages can be adopted by a large proportion of a population. The adoption of the English language seems to have been successful in non-English speaking countries in the European Union and it is probably still on the rise (Labrie & Quell, 1997). As stated by Labrie and Quell (1997) this increases the potential for communication among Europeans which clearly identifies one sociological need behind this process. In fact, taking into account this sociological need for communication, one can be fairly certain that this trend may increase in the future.

One sociological need that Bush (1945) saw was the need for a better communication between research scientists and for a better way of organizing the information overload that existed at the time. His brainchild that will satisfy this human-centric need was called a “memex.” He wrote the following (Bush, 1945):

Consider a future device for individual use, which is a sort of mechanized private file and library. It needs a name, and to coin one at random, “memex” will do. A memex is a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory. (p. 107)

Today, we take the ability of doing such tasks for granted but this should serve as a good reminder that such options were not always available. Technological needs are a direct implication of sociological needs and as long as this chain is followed one can visualize near future. Coincidentally, in Information Technology (IT) project management, the idea of a “vision” is the one that gives birth to the mission statement for a project which also defines the general direction for the future development of a project (Webber & Webber, 2007).

This cognitive ability to visualize future information based on past information becomes particularly useful in software design given the limits for empirical studies that could account for every situation possible. In fact, one may argue that the best tool for software designers is this past knowledge that provides them with evidence based predictions. Using past information they can potentially make sure that mental models between designers and users match.

### 2.2 Web 1.0, 2.0 and 3.0

Much has been said about the precursor of today’s web and arbitrary lines have been drawn in order to define the transitional stages. Of course, identifying ver-
sions for such a broad concept as the Web is not similar as establishing versions for traditionally produced software. It is fairly easy to version software based on the major and minor changes in the product code and features (Ramesh & Bhattachiprolu, 2006). On the other hand, the Web as a whole does not describe a single entity and therefore there was no single event where people decided to shift from version 1.0 to 2.0. With the variety of services that existed, the most likely thing to happen was a gradual adoption of new ideas and concepts. Adoption of new technologies seems to be influenced on people’s beliefs prior to adoption but even after adopting the new technology (Karahanna, Straub, & Chervany, 1999). In addition, age seems to play a significant role with the younger population being influenced just by attitude towards technology as opposed the older populations which is affected by subjective norms and perceived behavioral control (Morris & Venkatesh, 2000). The young Internet during the period of 1995-2000 saw an under-representation for the older age groups (Katz, Rice, & Aspden, 2001) and as such adoption of new technologies was more likely to happen; and it did happen!

We can today identify the transitional period when at the time new ideas and technologies were adopted by software engineers. Perhaps, a detailed description of the Web 1.0 is made by Bernal (2009) which contrasted Web 2.0 to Web 1.0 and hence defined the transition. He states that Bernal (2009):

Whereas the focus of Web 1.0 was on delivering products, Web 2.0 had created a paradigm shift to delivering services that can be used and combined with other services in new ways. Another key aspect is the growth of interactivity with end users in new ways, enabling users to drive what is important or of the most value. (p. 2)

The Web 1.0 had a sharp distinction between a user and a contributor or webmaster. Webmasters and content contributors delivered the information on the servers from which users could finally retrieve. However, users could not contribute content on the servers. This may seem primitive now but at the time people were particularly excited about the potential of such technologies. McEnery (1995) wrote an article about the potential uses of the early Internet and the one of the first browsers that were able to display images along with text. Even at the early static version of the Web 1.0 one can still identify the sociological needs for interactivity and sharing.

This need was identified correctly by DiNucci (1999) in her article Fragmented Future. She wrote (DiNucci, 1999):

The Web we know now, which loads into a browser window in essentially static screenfuls, is only an embryo of the Web to come. The first glimmerings of Web 2.0 are beginning to appear, and we are just starting to see how that embryo might develop.[...] The Web will be understood not as screenfuls of text and graphics but as a transport mechanism, the ether through which interactivity happens. It will [...] appear on your computer screen, [...] on your TV set [...] your car dashboard [...] your cell phone [...] hand-held game machines [...] maybe even your microwave oven. (p. 32)
Since the time in which this article was published, many of the features came to be. Even more importantly, even though her statement described technologies the underlying needs that led to these technological needs were still sociological needs. These are the needs that evolved Web 1.0 to 2.0.

Moving forward a few years and Web 2.0 became a fascination of researchers from various scientific disciplines. Many have attempted to define it. Bernal (2009) gave a description of what Web 2.0 encompasses:

"[...]in the Web 2.0 model, users actively participate and contribute to a website. This bidirectional approach enables users to interact with the site and each other in ways that provide for and foster a collective community. Users can create, edit, rate, and tag content at will, which provides other users with new information and guides the relevance of what is important to the overall community." (p. 3)

This bidirectional interaction was only a matter of time with new technologies that allowed for information to be exchanged in the background between browsers and servers. Bernal (2009) identified two core technologies that allowed for such interactive experience for the end user, Asynchronous JavaScript and XML (Ajax) and Representational State Transfer (REST). Ajax is a group of techniques that are used to create asynchronous web applications while REST is a style for software architecture aimed for web applications. Additional concepts are everyday being improved in order to foster this experience. Programming languages such as PHP, Ruby, Perl, and Python are constantly improving the capabilities of server software while browsers are modernized to handle Adobe Flash, JavaScript and other client frameworks. Database and database clusters are also improving the way they handle data and constantly developing load balancing capabilities. These provide engineers with numerous combinations of server and client hardware and software in order to design and develop modern Web 2.0 applications.

The idea of a Web as a platform became extremely popular by O’Reilly (2005) which saw potential for business models. The idea also of combining Web 2.0 with service oriented architecture became popular. Schroth and Janner (2007) specified that although the two have similarities as well as differences, the importance of unifying the two philosophies will help drive the development for enterprises. Shuen (2008) in his book Web 2.0: A Strategy Guide Business thinking and strategies behind successful Web 2.0 implementations demonstrates the strategies necessary to help customers build the website and create a user-driven business. However, Hoegg, Martignoni, Meckel, and Stanoevska (2006) argued that it is still debatable if Web 2.0 could ensure commercial success for all web services.

No matter its commercial aspects, many enterprises are enjoying the benefits of Web 2.0 technologies with the majority of top executives favoring such strategies (Murugesan, 2007). This is not without good reason since Web 2.0 has significant differences than Web 1.0. Murugesan (2007) argues that Web 2.0 (pp. 34-35):

* facilitates flexible Web design, creative reuse, and updates;
* provides a rich, responsive user interface;
* facilitates collaborative content creation and modification;
* enables the creation of new applications by reusing and combining different applications on the Web or by combining data and information from different sources;
* establishes social networks of people with common interests; and
* supports collaboration and helps gather collective intelligence.

Having all of the above in mind, it becomes obvious that Web 2.0 has significant advantages for communities which were not possible before. Or is this the case? As discussed previously as long as there are sociological needs these need to be satisfied by whatever technology exists at the time until new technologies will help develop better versions of interaction.

Howard (2010) states that the excitement about Web 2.0 should not take away the fact that people collaborating online and forming communities is not something new. He also expands on the idea that this could happen also on software that is not Web based such as video games. However, one could still identify a relatively new trend of providing more social tools in online games. This is not surprising if one considers that gamers value personal interaction with the software as well as social interaction with other people (Choi & Kim, 2004). Web 2.0 represents many new ideas that are not confined in just Web applications but they become part of a novel software engineering process that focuses on Service Oriented Architecture principles and is trying to provide services across multiple domains whether web services, desktop services or gaming services. At the rate that many services change, there is also an advocacy for avoiding traditional Waterfall lifecycle methods and apply Agile Software Development principles (Patterson & Fox, 2012). Approaches for a better development of software that can accommodate if not welcome change is highly appreciated by Web 2.0 services.

Having that in mind, it becomes clear that the sociological need for social interaction was not reinvented but the tools did in order to make it more efficient. The sufficient freedom gave rise to new empowering tools that allowed Internet users to explore social interactions like never before. As a result of this, the idea of social media came to life.

### 2.3 Social Media

Social media and Web 2.0 are sometimes used interchangeably but they are far from identical terms. Kaplan and Haenlein (2010) defined them as “A group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user-generated content” (p. 61). Deep into this definition lies the difference between social media and Web 2.0. Web 2.0 is a concept or philosophy while social media describe applications. To put that into perspective with a metaphor, the hippie movement during the 1960s was a culture or a set of ideologies and we can see applications
of these ideologies in Woodstock Festival in 1969 or the Sunset Strip in L.A. that became a hippie gathering area at the time. The case is similar with social media that encompass an incredible amount of various new applications.

Kaplan and Haenlein (2010) also created a classification for social media based on social presence/media richness and self-representation/self-disclosure which include, Blogs, Collaborative project (e.g., Wikipedia), Social networking sites (e.g., Facebook), Content communities (e.g., YouTube), Virtual social worlds (e.g., Second Life), Virtual game worlds (e.g., World of Warcraft). They defined these classifications based on principles that came from media theory. In short, media differ in terms of social presence which can be influenced by intimacy and immediacy (Short, Williams, & Christie, 1976), and media richness which controls the amount of information that can be transmitted in a given time (Daft & Lengel, 1986). The second combination of factors are self-representation which states that people want to control their impression formation on others in social interactions (Goffman, 1959) and self-disclosure which identifies information disclosure to others whether consciously or unconsciously (Kaplan & Haenlein, 2010). Kaplan and Haenlein (2011) expanded this classification to include micro-blogging which stands “halfway between traditional blogs and social networking sites” (p. 106). Figure 2.1 depicts the classifications of social media (Kaplan & Haenlein, 2010, 2011).

![Figure 2.1: Classifications for social media](image)

This is comprehensive list for social media. There is a great amount of literature on social media for all of these classifications. A short description for each is provided here for reference however, most of this paper addresses general features that apply to most social media classifications.

### 2.3.1 Blogs

A blog (Web-log) is a Web site that contains entries arranged in reverse chronological order and it functions as an online journal where many people can contribute (Boulos, Maramba, & Wheeler, 2006). Moreover, “Entries contain commentary and links to other Web sites, and images as well as a search facility may also be included” (Boulos et al., 2006). This function of linking between Web sites is what makes them critical in sharing information.

Blogs are extremely popular and used in a variety of different ways. However, their efficiency is debatable. Lawson-Borders and Kirk (2005) state that “blogs are not the ideal tool for either political reporting or historical analysis” (p. 557). Re-
Regardless, messages on blogs seem to have a higher longevity which makes them ideal as journalistic tools, as shown in the Elections of 2004 (Lawson-Borders & Kirk, 2005). Gumbrecht (2004) described blogs as “protected spaces” (p. 1). They become a safe haven for bloggers that can feel safer from retribution. Benefits seem to rise also with the use of blogs in distance learning. But again, certain drawbacks still exist such as feelings of isolation and alienation for the students (Dickey, 2004). As Dickey (2004) suggests, there are methods and strategies that can be employed to alter the design for these blogs in order to counter these effects. Figure 2.2 depicts an example of a blog.

![Figure 2.2: Google’s official blog.](image)

### 2.3.2 Collaborative Projects

Blogs and collaborative projects have many similarities. One of the most famous forms of a collaborative project is a wiki. This is described as “[...a set of linked Web pages created through the incremental development by a group of collaborating users (Leuf & Cunningham, 2001) as well as the software used to manage the set of Web pages” (Khosrowpour, 2008, p. 204). Students generally find wikis as good tools for project collaboration (Chao, 2007). In fact, there are numerous ways that people can use wikis especially for educational purposes (Duffy & Bruns, 2006) and also in corporate environments where people report that their groups are sustainable and wikis help them enhance their reputation, make their work easier and help the organization improve its processes (Majchrzak, Wagner, & Yates, 2006). Wikis also possess an extreme adaptability towards team goals producing a great variety of wikis that are modified to serve the needs of their communities. Wikis can be used to create collaborative research papers, encyclopedias and even facilitate debates (West & West, 2009). They allow for non-linear, evolving, complex and networked text, argument and interaction (Black, Delaney, & Fitzgerald, 2007). Ebersbach, Glaser, Heigl, and Warta (2008) have a detailed
description for installing and using Wikis especially for MediaWiki which is famous for its use in one of the most popular wikis today, Wikipedia and its sister projects. Figure 2.3 depicts Wikipedia, an encyclopedia wiki.

Figure 2.3: Wikipedia’s encyclopedia page on pseudonymity.

2.3.3 Social Networking Sites

Social networking sites arguably became the center of a media frenzy since their appearance early in the previous decade. Kaplan and Haenlein (2010) provided a definition:

Social networking sites are applications that enable users to connect by creating personal information profiles, inviting friends and colleagues to have access to those profiles, and sending e-mails and instant messages between each other. These personal profiles can include any type of information, including photos, video, audio files, and blogs. (p. 63)

They become an extension of an individual’s real life containing a detailed documentation of a person’s social network along with aspects such as their experiences, thoughts, beliefs and preferences. Social networking sites seem to be extremely helpful for people with low self-esteem and low life satisfaction and provide a tremendous advantage for managing social capital (Ellison et al., 2007). Social capital as defined by Coleman (1988), described anything that can facilitate individual or collective action generated by a network of social relationships and established three forms for social capital, obligations and expectations, information channels, and social norms. Put simply, it is the value of social relations that helps provide benefits to individuals or groups. As such, the term became a measure of well being for groups and the society. In fact, one study finds that the more frequent users were participating in friend networking sites, the more numbers
of relationships are expected to form and receive positive feedback from these relationships (Valkenburg, Peter, & Schouten, 2006). As a result of this, it was further concluded that positive feedback does not only enhance social self-esteem but well being as well. Furthermore, benefits of using social networking sites are not just limited to individuals but apply on corporations as well. Research is favoring social networking sites for supporting brand promotion and marketing campaigns (J. Brown, Broderick, & Lee, 2007; Sicilia & Palazón, 2008). Social networks can also be profitable business models (Murchu, Breslin, & Decker, 2004). Figure 2.4 shows a company’s fan page on Facebook, a social networking site.

![Figure 2.4: IBM’s fan page on social networking site Facebook.](image)

Many of the benefits of social networking sites are overshadowed by their disadvantages. One of these is the differential adoption due to the digital inequality (Hargittai, 2008). This digital divide has economic, sociological and political drivers that affect not just the adoption of social networking sites but the adoption of the Internet (Guillen & Suarez, 2005). Furthermore, while the digital divide excludes a group of people from using this technology, problems arise for the ones that actually use it. Perhaps one of the most popular issues currently on social networking sites is privacy (Barnes, 2006; Lenhart & Madden, 2007; Rosenblum, 2007). Specifically there is a tendency by young people to disclose personal information online without considering who the recipients are (Barnes, 2006). In addition, Gross et al. (2005) have argued that this tendency of disclosing personal or even sensitive information can lead to issues such as stalking, re-identification by using information from various social networking services, building a digital dossier for an individual, and general change in attitudes where people will tend to reveal information more freely. Perhaps, the magnitude of the problem becomes apparent if one considers a study that showed a fifteen percent overall in user profiles between two major social networking sites overlapped (Liu & Maes, 2005). Considering that at the time the usage of social networking sites was not as popular as today, one could perceive that these numbers are probably higher today. As an effect of this, an assailant wanting to identify a specific person that uses a pseudonym can do so by cross-referencing information across various social networking sites until he or she gets a match.
Regardless of the above dangers, people are more than willing to join social networking sites especially when their perception of enjoyment from the service is high (Lin & Lu, 2011). As seen before, as long as there is a sociological need these services are expected to keep on growing.

2.3.4 Content Communities

Content communities are mainly geared towards exchanging and sharing content between users (Kaplan & Haenlein, 2010). The content can be in any type of a multimedia format with websites allowing one or many formats. Content communities such as YouTube (depicted on figure 2.5) are used by the media industries, amateur producers and communities of various interests providing new opportunities that challenge the barrier between production and consumption (Burgess & Green, 2009). In addition, although the main focus may be different than in social networking sites, many of the benefits and drawbacks still exist here. In fact, the idea of social networks forming on content communities is not so farfetched either (Lange, 2008). Specifically, Lange (2008) conducted an one-year ethnographic research studying the participants behavior in terms of how much information they disclosed about themselves as producers and how big was their audience. The social networking effects are also discussed on another study that sees small-world characteristics in the links generated by contributors’ choices which in turn could improve service quality for other content communities if novel techniques are developed and implemented (Cheng, Dale, & Liu, 2008).

Figure 2.5: Masaryk’s University promotional video on content community, YouTube.

On the other hand, challenges seem to increase when dealing with content sharing. Issues of copyright and privacy that first started appearing with P2P applications are now expanding in content communities and leading companies need to establish a new set of guidelines (Pike, 2007). Furthermore, along with
these guidelines, new tools need to be developed. In fact, YouTube developed a technology that can process the vast amounts of data being uploaded everyday on their servers and identify which content breaches copyright rules (Youtube, 2012a). In another case, Stewart (2010) presented how the software works in front of an audience in order to ensure transparency and alleviate any worries that copyright owners may have.

2.3.5 Virtual Social Worlds

Virtual worlds are online environments similar to games but they have no specific goal or narrative for their players other than socializing. Some may have certain basic rules and limitations (e.g., users need to eat everyday) while others are less restrictive (e.g., users have the ability to fly). As Kaplan and Haenlein (2010) argue, one expects to see behavior that is similar to real life settings (Haenlein & Kaplan, 2009; Kaplan & Haenlein, 2007, 2009). Schroeder (2002) goes a step further by also defining that only certain virtual environments which he calls IVEs are giving the ability to individuals to immerse themselves into the virtual world and perceive themselves as part of them.

Individuals not only have the opportunity to immerse themselves in the virtual world but they have the power of controlling how they would achieve that and in what manner. One of the interesting phenomena that can be observed online is the practice of gender-switching where an individual assumes an online identity that has a gender that is different from their biological sex (Roberts & Parks, 1999). This control helps them form or reform their identities according to their desires and have control over their impression formation (Wallace, 2001). Players not only have a set of tools at their disposal but they can assume new roles and obtain new skills. As an example, aside from learning how to fly in Second life, users also acquire skills such as handling unstructured problems, dealing with the game’s economy and politics, socialize and sustaining relationships, negotiating and trading resources and so on (Papargyris & Poulymenakou, 2005).

Additional benefits which may not come to mind at first are for businesses as well. Reeves and Read (2009) stated that concepts of the trend known as gamification will inevitably take over the business. The idea of a workplace that encompasses virtual awards, avatars and three-dimensional environments may seem a bit farfetched but, as long as there is a sociological need for these processes one could expect to see technology evolving in order to accommodate this need.

2.3.6 Virtual Game Worlds

Virtual game worlds are a more restrictive version of a virtual environment. Users have to follow a strict set of rules and behave in a certain manner. These provide people with a tremendous potential for role-playing and entertainment guided by a fictitious story narrative. In fact research shows that “story, graphics, length, and control are highly related to enjoyment” (Wu & Boulevard, 2008, p. 219). They additionally provide strong social interaction for people to form not only relationships with each other but eventually groups and communities. Their pop-
ularity especially for the young population makes them a potential tool for not
only helping people attain and improve skills such as logical reasoning and team-
work but also design their own storyline for games and eventually implement
such process in classrooms (Robertson & Good, 2005). Designing these worlds
is of particular interest to educators, business leaders, and entertainment execu-
tives looking to benefit from these environments (Squire & A., 2006). Researchers
from all fields are also excited about the potential for investigating various so-
cial, economic, medical, political and other effects in these online worlds. Lofgren
and Fefferman (2007) investigated a pandemic in a virtual game world (World of
Warcraft) and drew parallels on what could have happened in real life in case a
similar pandemic were to emerge. These worlds are also breaking new ground
for computer science research (Bainbridge, 2007). Trying to accommodate socio-
logical needs that require advanced features for players such as being able to be
creative and develop their own items as well as being able to share these items
with others, creates severe delays in the communication of these systems. In turn,
delays in online games affect the willingness of an individual to continue using
the service (K.-T. Chen, Huang, Huang, & Lei, 2006). These results create new
technological needs that will help push the current momentum of social media
software forward.

2.3.7 Microblogging

According to Kaplan and Haenlein (2011), microblogs are a set of Internet-based
applications “which allow users to exchange small elements of content such as
short sentences, individual images, or video links” (p. 106). An example of such a
medium is Twitter which is depicted on figure 2.6. Similar to other social media,
microblogs have a great potential for individuals as well as the industry. They
are seen as potential great tools for collaborative work which can increase social
and emotional welfare in the workplace as well as the overall flow of information
within an organization (Ehrlich & Shami, 2010; Zhao & Rosson, 2009). They
also provide many opportunities for learning (Ebner & Schiefner, 2008). Another
research on Weibo, a Chinese microblogging service, reveals that the medium is
likely to increase participation in online protests (H. Li, 2012).

However, as with other social media cases, microblogging is plagued by sev-
eral issues. C. Lee and Warren (2010) discussed some of the major issues that rise
with the use of microblogging services in the workplace. These relate to the infor-
mation trust, privacy, confidentiality, and protection against scams and frauds.
Privacy and security was also a major concern discussed by Zhao and Rosson
(2009). Sensitive information that can be leaked without a user’s knowledge as
well as preset privacy settings for user accounts, should be a top concern for
designers and developers. Similarly, by having microblogging in the workplace,
sensitive work information may leak with severe consequences for organizations.
2.4 Additional Considerations

These classifications demonstrate the diversity of social media services and the necessity of properly defining the boundaries between each. However, even social media services under these classifications are not always consistent. Boundaries may overlap and types can be combined to include two or more classifications along with their features. Additionally, even within one classification differences can be found.

One such example can be found on an article written by Ellison et al. (2007). They argued that the terms *Social Networking Sites* and *Social Network Sites* are used interchangeably but the two vary. The term *Networking* is describing formations of new relationships usually by strangers. In contrast, they argue that what makes *Social Network Sites* unique is that they “enable users to articulate and make visible their social networks” (Ellison et al., 2007). As such, they identify three unique elements that make social network sites based on the ability to (Ellison et al., 2007):

- construct a public or semi-public profile within a bounded system.
- articulate a list of other users with whom they share a connection.
- view and traverse their list of connections and those made by others within the system.

These elements can vary however, depending on the structure of the social network/ing service. In addition many of these features can be found in several other social media that in effect become a hybrid.

The above idea is reinforced by the recent trend to use the term *social networks* and *online communities* interchangeably. However, Howard (2010) makes a clear distinction between the two and provides the fundamental differences.
One difference is the lack of secondary relations that can be found in online communities but not on social networking sites. But perhaps the most important difference is the fact that communities can exist on a variety of social media services while social networking sites are just one service. Put simply, online communities are not confined to just one social media service necessarily. Furthermore, social networks are also considered to be limited for certain types of activities that were described in what is coined as Shirky’s ladder (Howard, 2010; Shirky, 2009). In general, the ladder contains activities that increase in complexity and as such the relationships between individuals need to be stronger in order to achieve higher activities in the ladder. Howard (2010) concludes that while social networks are ideal for sharing activities, when it comes to activities that require cooperation or collective action they are insufficient. On the other hand, online communities containing strong secondary relationships are more likely to be able to accommodate for such more complex activities. This as an idea should not sound so absurd if one considers that online communities have a broad set of rules, ethics and social norms that form a more complex social structure than just a simple social network. Of course, one may still argue that today many social network sites, or social networking sites, provide additional services that help organize people into online communities. One such example is the famous Facebook application Causes which not only became a sustainable for-profit model of a social networking application but also collective actions taken by Facebook users raised $40 million dollars for 27,000 nonprofit organizations (Causes, 2012).

Many also have been mesmerized by the idea of merging Web 2.0 technologies with the principles of a semantic web in order to form the next platform for web, the Web 3.0 (Giustini, 2007; Hendler, 2008, 2009; Lassila & Hendler, 2007; Wahlster et al., 2006). The excitement is not unjustified considering the tremendous benefits that semantic web can bring to the current social media landscape. Some of these benefits and implementations of semantic web in current social media are demonstrated by Breslin, Passant, and Decker (2009) and Mika (2007).

Looking into the future we can now see that the ideas originally written by Bush (1945) and tackling the information overload still exist today. Web 2.0 and 3.0 are nothing but evolutionary manifestations of the Web 1.0 that were originally created to serve the exact same sociological needs. These are the same sociological needs used to establish the predictions of Hiltz and Turoff (1978) as pointed out by Howard (2010). Needs that one day may even transform our workplace by using virtual social and game worlds in our businesses (Reeves & Read, 2009).

But, as we move forward towards a brand new future with technologies that are evolving with a rapid pace, one still needs to remember that technology may change but people do not. In fact, while people are still referring to social media as something new that does not appear to be the case. Gundry (2006) points out in the year 1991 in a world prior to the World Wide Web, there was a generation of social media that was successfully used in organizations. In his paper at the time, he identifies the tremendous advantages seen by the use of collaborative technologies for educational purposes in organizations (Gundry, 1992). Coincidentally, earlier on Hiltz (1985, 1990); Hiltz and Turoff (1985) identified in
their work the tremendous potential for the networked software that one day will become part of the new Web which among its various application one can find social media.

2.5 Conclusion

This is the current state of the social media services that exist out there along with some of the major issues they are facing. This chapter aimed to demonstrate the variety of the services that exist and also the challenge of producing software that is deeply intertwined with everyday social activity. When a decision is made by an individual or a corporation to implement a social media service to accommodate for a set of interactions, the design of the service can have a great impact on the efficiency of that interaction and the overall individual and social experience. As such, developing social media that will account for these experiences requires a revolution in the way designers and developers think and progress through software development. The first goal for such a development requires one to identify what desirable and undesirable social experiences are for social media communities. This will also depend on what type of a social media service is developed. In the following chapters I demonstrate how research on software features and social behaviors can assist in a development of better social media guided towards achieving goals related to these desirable and undesirable phenomena. A deep understanding of sociology and psychology is important for understanding the effects of software on social interactions.
CHAPTER 3

GROUPTHINK THEORY AND APPLICATION

3.1 Introduction

This chapter provides an in-depth review on groupthink. The relevance of groupthink in this dissertation research was already mentioned in chapter 1. In this chapter, I elaborate more on the relevance of groupthink in social media. Definitions for groupthink are presented along with a literature review for groupthink in real world and on the Internet. Furthermore, the features under examination by this study are presented here and will be analyzed in the next chapter.

3.2 Defining Groupthink

Janis (1982) defined groupthink as:

A mode of thinking that people engage in when they are deeply involved in a cohesive in-group, when the members’ strivings for unanimity override their motivation to realistically appraise alternative courses of action. (p. 8)

This definition identifies some of the causes of groupthink. People within a group try to minimize conflict and reach a consensus without proper and critical evaluation of all the information available to them, individually as well as a collectively.

A more modern definition for groupthink was given by the Merriam-Webster Dictionary (Groupthink, 2012).

a pattern of thought characterized by self-deception, forced manufacture of consent, and conformity to group values and ethics

This definition demonstrates the effects of groupthink. It describes what happens while a group is experiencing groupthink. This behavior goes past simply conforming with the group. Individuals force themselves to accept information in order to help a group reach an agreement. Arguably, this combination is what makes groupthink severely destructive in group decision-making processes.
It is not difficult to imagine why groupthink is an unwanted behavior in groups and communities in everyday life. Furthermore, it is an unwanted behavior in online groups and communities that exist in social media, especially for services meant for collaboration. For example, it is a serious issue for the development of Group Decision Support Systems (GDSS) which include wikis (Tsikerdekis, in press-c).

### 3.3 Examples of Social Media Endangered by Groupthink

To understand the involvement of software designers for preventing groupthink in social media, two examples are presented. A team was asked to develop a social media software, and more specifically a collaborative project for a classroom. A decision was made to use wiki software which will lead students in participating in a collaborative effort to create and improve on the current learning material, but also discuss issues that exist in class and try to provide solutions. A defective decision-making process in this setting could have devastating and unforeseen consequences especially when dealing with a population of non-adults. Considering that wikis are considered to have tremendous potential for classroom settings and education in general, this scenarios is not far-fetched (Ferris & Wilder, 2006; D. Turner, 2004). Another example could be the assignment of developing software that will accommodate decision-making for an IT company. A group may be developing new projects or working towards problem resolution for issues in current projects or even their clients’ projects. It becomes clear that a solid and efficient decision-making process is essential for business success. Employees may need to hold multiple meetings and information may be stored in wiki software in order to help weight their options better. The idea of wikis for business collaboration, although requires changes in the current management culture (Edwards, 2007), is still considered beneficial (Kussmaul & Jack, 2009; Wood, 2005).

These two examples demonstrate the tremendous potential for social media being used in various settings that will involve collaboration. In fact, groups are really easy to form in any community in general. D. H. Smith (1967) gives a good definition of a group based on a couple of criteria.

A group is defined here as (1) the largest set of two or more individuals who are jointly characterized by (2) a network of relevant communications, (3) a shared sense of collective identity, and (4) one or more shared goal dispositions with associated normative strength. (p. 141)

The above statement demonstrates that groups not only are easy to form but they can vary in size. Another definition was given by Levine and Moreland (1994). They stated that a group is several people that “interact on a regular basis, have affective ties with one another, share a common frame of reference and are behaviorally independent” (Levine & Moreland, 1994, p. 306).

Based on the above, it becomes clear that social media are bound to have groups within them whether a collaborative project for a company or an online guild of individuals in a virtual game world. Wherever one can find groups,
group phenomena are bound to exist and this rule applies also to social media. However, people can alter the virtual environment produced by the software of social media, unlike the real world situations where people have limited control over the environment. In fact, the software is entirely manufactured. As such, a question that comes to mind is, could we alter the software in such a way that we can affect group interactions, and as such protect groups from dangerous group phenomena such as groupthink?

3.4 The Beginnings of Groupthink and Research

To understand how groupthink prevention became a popular idea one has to understand how the model of groupthink was formulated. Groupthink traces its roots to March 1952. Whyte (1952) wrote about groupthink:

> Groupthink being a coinage - and, admittedly, a loaded one - a working definition is in order. We are not talking about mere instinctive conformity - it is, after all, a perennial failing of mankind. What we are talking about is a rationalized conformity - an open, articulate philosophy which holds that group values are not only expedient but right and good as well. (p. 114)

Later on, Janis (1971) developed the initial model on groupthink. He wrote that groupthink leads “to a deterioration in mental efficiency, reality testing and moral judgments as a result of group pressures” (p. 43). He followed up with his research on several publications (Janis, 1972, 1982).

3.4.1 The Groupthink Model

The model of groupthink that Janis (1972) created provided explanations for groups with defective decision-making. Janis and Mann (1977) further established the groupthink model as a five-step casual model (Henningsen, Henningsen, Eden, & Cruz, 2006).

Antecedents

The first step included a set of antecedents such as high levels of cohesiveness, structural defects and provocative situational contexts (Janis & Mann, 1977). Structural defects can be analyzed further as insulation of the group, lack of impartial leadership, lack of norms that require methodological procedures, and homogeneity of the social background ideologies of members. Similarly, provocative situational contexts can appear in terms of highly stressful external threats for the group, other recent failures, significant difficulties in the decision-making task and moral dilemmas that may arise in the collaborative progress.

Cohesiveness in particular is perceived as the single most important hazard for groupthink (Janis, 1982). This seems to be in agreement with a later meta-analysis study conducted by Mullen, Anthony, Salas, and Driskell (1994) which associated high group cohesiveness with defective decision making. Also, Janis
(1982) noted that groupthink cannot occur unless all the antecedents are present. However, Hart (1991) described that high cohesion is important but groupthink cannot occur unless one or both of the other antecedents are present; asserting that groupthink could occur without the need for all antecedents to appear in a group.

Concurrence seeking

In the second step, concurrence seeking is leading members to openly agree with a group’s perceived position although they privately disagree (Janis, 1972). Coincidentally, this behavior is similar to conformity described by Asch (1958, 1992) where the distortion at the action level affects individual judgment in group decision-making. It is understandable that it takes a significant force to lead someone to consciously avoid expressing their opinion to a group. This may be why this behavior is so dangerous. If an individual withholds a key piece of information, this information will never be put up for consideration by his or her group.

Symptoms of Groupthink

As a result of the previous two steps, symptoms of groupthink start occurring. These include illusion of invulnerability, illusion of morality, collective rationalization, excessive stereotyping, self-censorship, illusion of unanimity, pressure of conformity and self-appointed mindguards (Janis & Mann, 1977). These were later revised into three categories: overestimation of the group, close-mindedness, pressures toward uniformity (Janis, 1982). They establish pressure on group members to not only conform with the group but also perceive that the decisions made by the group are productive and successful (Henningsen et al., 2006).

Decision-making defects

Consequentially, a group that starts experiencing the symptoms of groupthink eventually may experience deficiencies in the decision-making process as well. Janis and Mann (1977) identified the following defects: incomplete survey of alternatives, incomplete survey of objectives, failure to examine risks associated with the preferred choice, poor information search, selective bias in processing information, failure to reappraise alternatives, and failure to provide contingency plans. Essentially, the group is “[...] processing information ineffectively” (Henningsen et al., 2006, p. 38).

Poor decision outcomes

The final outcome of all the previous steps lead to a faulty decision by the group decision-making process (Janis, 1972, 1982, 1989). The faulty decision at this point is usually not the optimum one. Additionally, while Janis (1989) pointed out that groups may arrive in a good decision even under the influence of groupthink, Henningsen et al. (2006) reported that most accounts under which groupthink had occurred point to a negative outcome and poor decisions being taken.
3. GROUPTHINK THEORY AND APPLICATION

3.4.2 Historical Examples of Groupthink

The model of groupthink might have become popular because the above model is fairly easy to understand and follow, but also due to popular historical cases that Janis analyzed (Esser, 1998). Some of the fiascoes in human history that are attributed to groupthink are: the Bay of Pigs Invasion, the attack on Pearl Harbor, and the space shuttle Challenger disaster. These cases are going to be described below but there are additional cases such as the North Korea escalation or the Vietnam escalation (Rose, 2011) that were originally studied by Janis (1972) for groupthink as well.

Pearl Harbor

On December 7, 1941 the Imperial Japanese Navy attacked Pearl Harbor in a surprise military strike. Janis (1982) argued that the naval officers that were stationed in Hawaii had shared illusions and rationalizations which contributed to lack of being on alert. The United States intercepted messages from the Japanese side that was preparing for an offensive attack and warned the officers in Pearl Harbor, but their warnings were not taken seriously. The army officers had succumbed to social pressures and additionally were rationalizing their beliefs for how unlikely an attack on Pearl Harbor was by the Japanese. Some of the statements are illuminating (Janis, 1982, pp. 83-85):

- “The Japanese would never dare attempt a full-scale surprise assault against Hawaii because they would realize that it would precipitate an all-out war, which the United States would surely win.”
- “The Pacific Fleet concentrated at Pearl Harbor was a major deterrent against air or naval attack.”
- “Even if the Japanese were foolhardy to send their carriers to attack us, we could certainly detect and destroy them in plenty of time.”
- “No warships anchored in the shallow water of Pearl Harbor could ever be sunk by torpedo bombs launched from enemy aircraft.”

All of these statements testify to a sense of invulnerability that the army felt. This was part of symptoms that should have raised red flags. At the time nobody wanted to question or oppose mainstream beliefs. Eventually this lead to decision-making defects by dismissing information and being inactive and finally the outcome of this decision that lead to the disaster of Pearl Harbor. Arguably, had they known about groupthink and applied preventive measures, this disaster might have been limited, if not averted.

Bay of Pigs Invasion

Twenty years after Pearl Harbor, the Bay of Pigs invasion incident stands as the primary example used by Janis (1982) to describe groupthink. The invasion
plan that Kennedy used targeted to overthrow Castro’s Cuban government. Originally, the plan was initiated by the Eisenhower administration and later the Kennedy administration adopted CIA’s plan. There is enough evidence that portrays that this information was not critically evaluated (Janis, 1972) but also those that tried to raise concerns or object were questioned about their loyalty. The group believed in the morality of their plan or even self-censored their thoughts. The sense of invulnerability was also high as well as stereotyping Castro’s administration. At the end, within few hours after the invasion, all the invaders were killed or captured. Later on, in the case of Cuban Missile Crisis, president Kennedy employed preventive measures for meetings in order to avoid a similar fiasco. On the other hand, concluding that the Bay of Pigs fiasco is an event caused just by groupthink may also be presumptuous. As noted by Kramer (1998) aside from group dynamics there may have been political reasons that could also explain the Bay of Pigs event. They called for a “multiple lens” perspective when investigating such organizational fiascoes.

Space Shuttle Challenger

Another popular disaster that given the evidence is attributed to groupthink, is the case of the space shuttle Challenger disaster. The official report released by the NASA History Office (Scobee et al., 1986) after the accident is illuminating:

The decision to launch the Challenger was flawed. Those who made that decision were unaware of the recent history of problems concerning the O-rings and the joint and were unaware of the initial written recommendation of the contractor advising against the launch at temperatures below 53 degrees Fahrenheit [...] They did not have a clear understanding of Rockwell’s concern that it was not safe to launch because of ice on the pad. If the decision makers had known all of the facts, it is highly unlikely that they would have decided to launch 51-L on January 28, 1986 [...] based on incomplete and sometimes misleading information, a conflict between engineering data and management judgments, and a NASA management structure that permitted internal flight safety problems to bypass key Shuttle managers. (p. 82)

Research into this specific accident in relation to groupthink has been positive. Esser and Lindoerfer (1989) conducted quantitative research by using archive records and positively identified an increase in the observable antecedents and consequences defined by Janis (1982) in the last 24 hours prior to the accident. Another study into the accident by Moorhead, Ference, and Neck (1991) also found substantial evidence of the same antecedent conditions, the symptoms, and the decision-making defects and concluded that the Challenger disaster should be classified as a “groupthink situation.” Additionally, they revised the original model in order to include two more variables, time and leadership style.
3.4.3 A Model of Controversy

Literature review articles today seem to agree that research and experimental studies on groupthink have been limited (Esser, 1998; Rose, 2011). In general most of the offline research investigated popular cases such as the 1999 baseball umpire strike (Koerber & Neck, 2003), Kent State University gymnasium controversy (Hensley & Griffin, 1986), France’s 1940 WWII defeat (Ahlstrom & Wang, 2009), Iran Hostage Crisis (S. Smith & Jan, 1985) and other case studies which provide support for the theory. According to a review on groupthink by Rose (2011) there have been also fifteen experimental studies that investigated groupthink. Erdem (2003) showed that a high degree of trust can increase groupthink, Ahlfinger and Esser (2001) showed that promotional leadership partially supported that it can affect some of the antecedents and consequences, and M. E. Turner, Pratkanis, Probasco, and Leve (1992) found support for groupthink by testing many of its antecedents and consequences.

While it may seem that researchers have reached an agreement after decades of studies since the creation of initial model, there is still a lot of controversy over the topic. As Rose (2011) pointed out, while Mitchell and Eckstein (2009) state that groupthink has been widely accepted, W.-W. Park (2000) states that “very little consensus among researchers on the validity of the groupthink model” (p.873). This is far from an unfounded statement given that studies have also failed to make a direct connection between groupthink and the processes studied (Kramer, 1998; Maier, 2002). Some of the well tested areas such as insulation and impartial leadership show support of the model, while cohesion produced mixed results, mainly because of the diversity of perspectives in cohesion (Rose, 2011). Some attribute those mixed results in the lack of standardization for measuring the antecedents and combining the results (Esser, 1998), while others in the general lack of multiple studies that explore all the antecedents and consequences (Rose, 2011).

As Rose (2011) indicated this started a debate. The result of this is some stating that “the model represents a brilliant construction founded in part on the existing group dynamics literature” (Paulus, 1998, p. 371) while others state that “in our view, groupthink is a compelling myth” (Fuller & Aldag, 1998, p. 177) that was accepted without substantial evidence. Perhaps the middle ground could be considering positive as well as negative evidence for groupthink, and provide certain revisions to the initial model. Some modifications to the initial model are provided by Chapman (2006), Flippen (1999), and Neck (1995) with a main focus on adding more variables to the initial model. A full survey of all the proposals is covered by Rose (2011).

For now, the current model is highly popular and being applied to jury decision-making (Mitchell & Eckstein, 2009), hockey teams to improve performance (Rovio, Eskola, Kozub, Duda, & Lintunen, 2009), academia and other cases (Klein & Stern, 2009; Rose, 2011). Paying close attention for preventing groupthink in groups is not unjustified. Having looked in the popular fiascoes with case studies as well as experimental studies, if groupthink is not a myth the implications can be severe. One can survive a false positive but can never survive a false negative. Whether the model is a figment of the imagination or reality, it does not exclude
the fact that it is a model that tries to explain what happened in all of these popular cases and probably is the best model that does this to this day. The debate goes on with research trying to argue for and against the model, regardless of the initial acceptance of the model. This paper agrees with such a strategy. In the light of past and current evidence which can be found in literature reviews (Esser, 1998; Rose, 2011), some antecedents and consequences of the original groupthink model have been positively observed. It is also evident that some variables have not been tested at all, supporting the argument made by Rose (2011) for a limited research by experimental studies.

This paper argues that from an HCI perspective, this debate can also become trivial. Whether the model provides a substantial predictability of the decision outcomes of groups, or it needs revisions, the implications of such realizations are not limiting HCI research that investigates groupthink. In HCI, the focus should mainly be the application of such processes into the software development based on empirical evidence. As an example, consistent studies show that insulation (Moorhead et al., 1991) and impartial leadership (Flowers, 1977; Fodor & Smith, 1982; Leana, 1985) play a significant role to defective decision-making processes and poor decision outcomes. Hence undoubtedly, even if the groupthink model needs to be revised, the implications and goals for HCI researchers and software engineers that need to apply this knowledge in software design and development practices do not change.

This approach seems to be in accordance with the initial goal of Irving L. Janis that as indicated by M. E. Turner and Pratkanis (1998), he was always interested in the practical significance of the research. That is, the goal of the groupthink model was not just to explain why groups make bad decisions but also to establish ways that this can be prevented. This is what makes the groupthink model a valuable tool for software designers wanting to apply this knowledge in various social media services.

### 3.4.4 Preventing Groupthink

Given the variables presented from the groupthink model we can make certain speculations when attempting to prevent groupthink. In effect, by inhibiting the antecedents and consequences one can expect that the probabilities for groupthink occurring within a group will decrease.

In general, one can expect that groups with more information available to them will produce more alternatives. In addition, Janis (1982) showed that a group’s high cohesion will lead to a group making compromises and impair the decision-making ability of the group. Consensus may be reached faster but critical thinking may be impaired for the members of a highly cohesive group. These processes can be found in larger groups as well (Koerber & Neck, 2003; McCauley, 1998). Hart (1991) considered extremely important to actively combat the effects of groupthink.

Seven measures were provided by Janis that could prevent groupthink (Janis, 1972, pp. 209-215):

- One member of the group should always be assigned the role of Devil’s
advocate. This person needs to be different for each meeting.

- Each member should be assigned the role of a critical evaluator by the leaders. Objections and doubts should be presented freely.
- Leaders should not present their preference initially.
- All effective alternatives should be examined.
- Independent groups should be set up and work on the same task.
- Each member should discuss ideas with outsiders.
- Outsiders should be brought in the meetings.

When applying the following guidelines, groupthink can be averted. In fact, they were applied during the Cuban Missile Crisis by President John F. Kennedy (Hart, 1991). Of course, these are not but just rough guidelines. As described before, any given method that can inhibit antecedents or consequences can be used as a preventive measure. Each individual condition that inhibits those variables that are considered to be a cause for groupthink can help avert potential disastrous decision outcomes. Identifying conditions that have such power is essential especially when developing software where software engineers have a lot of control over the environment that their users will collaborate in.

3.5 Groupthink and Online Groups

Developing an online environment in order to reduce groupthink behavior has seen even less interest from the scientific community. However, the model is arguably, tremendously useful if one considers the benefits of implementing preventive measures that can reduce groupthink.

Back in the two original examples that were given early on in this chapter, one can consider the necessity of assigning an individual the role of the devil’s advocate or the iconoclast. An iconoclast always tries to redefine whatever is considered already as a well-accepted consensus. In effect, this role exists with a sole purpose to disagree and point weaknesses to the group’s general plans. In both examples this role can automatically be assigned based on the profile characteristics of individuals. The assignment of the role would not only be less biased since the software would do most of the work, but it could also ensure that the individual selected as the devil’s advocate fits the psychological profile of person comfortable with that role. This can also be reinforced with the use of anonymity which for certain people can be a liberating factor to voice out their opinions. However, such rules cannot be easily implemented in groups collaborating on the physical world.

On the other hand, software for online collaboration is extremely useful. In fact, one of the arguments supported by this dissertation is that developers do not consider the tremendous potential and opportunities that online software has for controlling behavior. Most of the aspects described in the previous paragraph can easily be implemented in social media software.
3. GROUPTHINK THEORY AND APPLICATION

3.5.1 Researching Online Groupthink

Most of the research that investigates online effects for groupthink, is focused on GDSS. Right from the emergence of Group Decision Support Systems researchers became excited about the potential of technologies to constrain the environment and produce designs that could actually prevent certain group phenomena. DeSanctis and Gallupe (1984) wrote about the potential benefits of GDSS in a list that stated several benefits along with the following statement:

The GDSS contains built-in mechanisms which discourage development of negative group behaviors, such as destructive conflict, miscommunication, or “groupthink”. (p. 24)

It becomes evident that since the early beginnings of collaborative technologies people thought highly of their potential for preventing negative group phenomena. This idea was at the core of some early studies that followed on group collaboration and GDSS (DeSanctis & Gallupe, 1985; Driscoll & King, 1988). Such interest was not unjustified as later on group support systems were suspected of having an effect on groupthink (Miranda, 1994). Miranda (1994) reviewed antecedents and procedural conditions in order to establish how groupthink can be prevented in groups and what the implications are for these early group support systems.

A similar research, only a few years back, investigated the effects of anonymity in GDSS. Jessup et al. (1990) conducted a study that investigated the effects of anonymity as a separate condition but it was also partially connected to groupthink. This is a versatile approach of tackling groupthink by investigating individual aspects that may contribute to groupthink. As mentioned before, groupthink may be caused by a number of variables that may or may not be present. Hence, it becomes particularly difficult to evaluate the actual occurrence of groupthink. However, one can study the individual conditions that are known or suspected to be causes for groupthink according to the case studies and afterward infer about their implications.

Indeed group support systems show promising results by giving groups advantages of accessing external information (Miranda & Saunders, 1995); an important preventive measure for groupthink. On the other hand, the same study shows that groups tend to discuss alternatives less which may lead to defective decision-making. Alternatives play a key role for the chances of observing groupthink behavior in a group. It is also evident that it is not sufficient to achieve enough contributions for alternatives if the group will not spend time considering them. In fact, one may argue that even the process through which alternatives are evaluated is important; that is, to ensure an unbiased way for evaluating those alternatives.

Troyer (2003) mentioned the importance of taking into account group phenomena when designing GDSS. Specifically one of the main arguments investigated the effects of positive and negative evaluations and their effects to groupthink. Specifically, the case of negative evaluations that is particularly important for critically assessing all alternatives and the fact that shortcomings should be
thoroughly explored. It is no wonder that Cordell (1996) described that GDSS software has the ability to prevent groupthink in a large extent, mainly because of the ability to provide anonymity. The idea and potential benefits for anonymity in GDSS is described by many (Kraemer & King, 1988; Nunamaker, Vogel, & Konsynski, 1989; Nunamaker Jr., Applegate, & Konsynski, 1987) and it becomes a core concept for this study as well.

Other research in GDSS in the past decade seems to investigate several group phenomena in relation to GDSS with a general consensus that GDSS are effectively preventing Groupthink (Lawrence, 2002; Marreiros, Ramos, & Neves, 2004; Wadhwa & Bhattacharya, 2000). However, Shirani and Lee (2008) made the following suggestion for research into computer-mediated communication technologies:

Groupthink: CMC environments eliminate or reduce verbal and social context cues. Which CMC environments, under what conditions, are better at avoiding groupthink? (p. 5)

Arguably the question stated here is valid. While most researchers and in turn software designers and developers seem to agree that groupthink is a condition that can be avoided with group collaborating technologies, it is naive to state that this applies to all GDSS software. This is particularly true if one considers the lack of experimental studies for the conditions known to cause groupthink for GDSS. In fact, even if most of the statements are grounded in theory, it is essential communicating these results to software engineers so that software can be developed properly. Evidently, it seems that the current consensus is that any group decision support system and other broader collaborative projects that are part of social media will guarantee up to a point avoidance of groupthink.

As it will become obvious during the next sections, this study showed that this is not the case. Certain software features affect antecedents and consequences that contribute to groupthink. If while designing software these are not considered properly, they can have devastating consequences and lead to groups experiencing groupthink. In addition, the current consensus about GDSS preventing groupthink can make things significantly worse by having usability engineers or software engineers acknowledging a groupthink-proof software while this may not be the case always.

3.6 **Software Features Studied for Groupthink**

Based on the literature presented here, several software features were studied to help assert the main goals of this dissertation. These were selected based on the cost of studying them and being examples that showed potential based on the literature. Since this dissertation work wanted also to encourage similar research, demonstrating an “easy to investigate” approach was of particular importance. In addition, elements had to be diverse as well. This guaranteed that the study would investigate multiple perspectives of the problem of groupthink but, it would also demonstrate a variety of ways for researching such aspects in
3. Groupthink Theory and Application

HCI. This study aimed to focus on features of social media software that had preventive potential based on the factors described in this chapter that are known to contribute to groupthink. Features that contribute to information production and reduce conformity pressures were at the core of this study. Several diverse features currently in use in social media were identified and investigated in terms on their effect on individual and social behavior. Coincidentally, many of these features were found to be easily adjusted on social media. A summary of the main aspects that were studied is presented below.

- Pro/con lists and their potential for reducing groupthink as compared to textual collaborative interfaces. How effective are they in terms of argument production, uniqueness and comprehension?
- Anonymity states and their effects on non-conformity and groupthink. Can they be implemented to improve collaboration?
- Dynamic voting interface indicators and their effects online voting processes. Is the final voting in decision-making process biased and could contribute to groupthink?
- Item randomization in voting process for improving collaboration and shield-proofing against groupthink.
- Publicly identified administrators and leaders in online groups as potential candidates to influence open voting results and contribute towards groupthink.
CHAPTER 4

RESEARCHING THE INDIVIDUAL FEATURES FOR GROUPTHINK IN SOCIAL MEDIA

4.1 Introduction

In this chapter, research analysis and findings will be presented for the social media features described in the previous chapter in relation to groupthink. Each feature being investigated is presented in a separate section. Individual conclusions are presented in each section. Additionally, all results are combined later on in the chapter where the social interaction design framework is presented. Furthermore, some of the features that were investigated for the purposes of this dissertation have been published in journals and conferences. In such cases, summaries and highlights from the published papers are presented. One exception to this is the article regarding dynamic voting interfaces and their effects on affecting decision outcomes (Tsikerdekis, in press-a), which was not included in this chapter due to size restrictions on the dissertation. However, results that came from that particular study are mentioned in the chapter 7 where the social interaction design framework is presented.

4.2 Pro/con Lists

One of the features that were explored to assert a relation between design and social behavior in terms of groupthink was the case of implementing pro/con lists as a feature that can potentially reduce groupthink. Information that can be organized in a more concise and probably more comprehensive format may be more effective than traditional textual collaborative interfaces. I conducted a comparative analysis that investigated argument production and comprehension between pro/con lists and textual collaborative interfaces. This was a previously unexplored software feature in HCI, especially in terms of its relation to groupthink. The study’s findings were published on a scientific journal (Tsikerdekis, in press-c). A short version of the process is presented here along with additional statistical analyses of the results which were not included in the published version. Finally, I present the conclusions and remarks based on the work conducted in the original paper as well as in this dissertation.
4.2.1 Method and Initial Results

In my study, I proceeded with the analysis using real world data from two popular websites that used the two interfaces. The first website was Wikipedia that uses a textual collaborative interface while the second was Debatepedia, a website that facilitates debates with the use of pro/con lists. I obtained a random sample of articles from both websites that collaborated on the same topic. Content analysis was used to code the text into quantitative data that was later used in the statistical analysis. From the qualitative analysis several additional conclusions were extracted in relation to argument production. Finally, I used readability formulas to develop quantitative data to assert the potential for comprehension between the two interfaces.

The results showed that pro/con lists may not be more advanced than textual collaborative interfaces in terms of argument production however, statistically significant differences were found for the production of unique arguments that favored pro/con lists. Additional comparisons that examined the readability of the text between the two interfaces produced no statistically significant differences. Several other results coming from the qualitative portion of the study showed that the original segmentation and categorization of an article had a significant effect on the development of articles for both interfaces.

The use of a variety of qualitative and quantitative methods helped triangulate results and validated some of the conclusions. This demonstrates the variety of approaches that can be employed by researchers to assess certain hypotheses. Human-computer interaction as an interdisciplinary field can use the diverse variety of methods that exist without being constrained to a strict research framework. In addition, the current approach demonstrates the wealth of information that is currently on the web and can be used for research. Finally, comparative content analysis is arguably a lot less costly as opposed to experimental design methods, and examines real people in real situations as opposed to having “guinea pigs” in laboratory or simulated conditions. On the other hand, the downside for the above methods is that real examples that match the conditions that need to be investigated are difficult to find.

4.2.2 Analysis for Argument Production

A detailed account of how the data were obtained and processed can be found in the original article Tsikerdekis (in press-c). In this dissertation, I present a Bayesian analysis for the original hypotheses on the article that further helps assess the overall results of this study. The analyses that follow investigate two aspects of pro/con lists over textual collaborative interfaces: production of total arguments and production of unique arguments.

Descriptive statistics are shown on figure 4.1. Total arguments favored Debatepedia ($M = 32.27, SD = 10.90$) over Wikipedia ($M = 25.47, SD = 16.88$). The result was similar with the unique arguments where Debatepedia ($M = 59.84, SD = 16.76$) showed higher measurements for unique arguments over Wikipedia ($M = 41.72, SD = 24.99$). In the original article however, a statistical analysis using independent t-tests achieved significance only for the sec-
ond comparison. Therefore, a Bayesian analysis may more efficiently help assess these results and establish the statistical likelihood of the two major hypotheses being valid. Furthermore, while in null-hypothesis statistical testing, “p-values are incapable of providing support for the null hypothesis” (Wagenmakers, Lee, Lodewyckx, & Iverson, 2008, p. 200), Bayesian hypothesis testing does quantify the probability that the null hypothesis is true (Kaptein & Robertson, 2012; Wagenmakers et al., 2008).

Figure 4.1: Descriptive results for pro/con lists and textual collaborative interfaces in terms of total argument production and unique arguments.

For the Bayesian statistical analysis, I used a model developed by Lee Michael D and Eric-Jan Wagenmakers (2010) which I briefly present here. The model es-
establishes that $X$ and $Y$ represent the observed data that follow a normal distribution with a shared variance $\sigma^2$.

$$X \sim \text{Normal}(\mu_X, \sigma^2) \quad (4.1)$$

$$Y \sim \text{Normal}(\mu_Y, \sigma^2) \quad (4.2)$$

The means for $X$ and $Y$ are defined as:

$$\mu_X = \mu + \alpha/2 \quad (4.3)$$

$$\mu_Y = \mu - \alpha/2 \quad (4.4)$$

Therefore $\alpha$ becomes their difference in their means. The effect size $\delta$ is given by the equation:

$$\delta = \alpha/\sigma \quad (4.5)$$

The priors for $\delta$, $\mu$ and $\sigma$ are established as:

$$\delta \sim \text{Cauchy}(0, 1) \quad (4.6)$$

$$\mu \sim \text{Cauchy}(0, 1) \quad (4.7)$$

$$\sigma \sim \text{Cauchy}(0, 1)^+ \quad (4.8)$$

Based on the model, the hypotheses are the following:

$$H_0 : \delta = 0 \quad (4.9)$$

$$H_1 : \delta \sim \text{Cauchy}(0, 1) \quad (4.10)$$

$$H_2(\delta < 0) : \delta \sim \text{Cauchy}(0, 1)^- \quad (4.11)$$

$$H_3(\delta > 0) : \delta \sim \text{Cauchy}(0, 1)^+ \quad (4.12)$$

The null hypothesis ($H_0$) asserts that there is no effect while the alternative ($H_1$) asserts that there is a difference between $X$ and $Y$. Further, $H_2$ supports the claim that $X$ has larger values than $Y$ while $H_3$ supports the opposite.

Using MCMC sampling I drew 100,000 samples from the posterior for the effect size $\delta$ with a burnin for the first 10,000 samples and with three chains. Next, I obtained the Bayes factor using the height of prior and posterior distributions for $\delta$ at $\delta = 0$. The height of the prior was computed directly from the Cauchy(0,1) distribution (using function $dcauchy(0)$ in R) and the posterior was estimated from the MCMC samples by applying a non-parametric estimator (Stone, Hansen, Kooperberg, & Troung, 1997) which can be found in the polspline package in R (using function $dlogspline$).
The Savage-Dickey ratio at $\delta = 0$ yielded a Bayes factor of $BF_{01} = 1.85$, indicating that the null hypothesis ($H_0$) is 1.85 more likely given the data than the alternative hypothesis $H_1$. In terms of strength of evidence this Bayes factor is classified as weak, meaning that while the null hypothesis still has to be accepted, caution is advised (Jeffreys, 1998; Kass & Raftery, 1995; Wagenmakers et al., 2008). The dots indicate the height of the two distributions at the point $\delta = 0$, the point of interest. A similar result was obtained for the order-restricted hypothesis $H_3$ with a Bayes factor of $BF_{03} = 1.05$. This indicates that the null hypothesis is 1.05 times more likely to occur than the hypothesis $H_3$. The result becomes extremely hard to interpret. In other words, given the data we can reject $H_3$ but only by an extremely small fraction. A more concrete result was obtained for $H_2$ with a Bayes factor of $BF_{02} = 8.08$. This is classified as substantial positive evidence that the null hypothesis ($H_0$) is 8.08 times more likely to happen than the $H_2$. Put simply, having strong evidence against $H_2$ means that Wikipedia did not produce more arguments than Debatepedia. Figure 4.2 shows the graphs illustrating the prior and posterior distributions obtained for all three hypotheses.

Figure 4.2: The prior and posterior distributions of effect size $\delta$, when comparing total arguments between Debatepedia and Wikipedia. The top graph illustrates the unrestricted t-test for $H_1$, the middle graph illustrates the order-restricted test stating that Wikipedia produced more arguments than Debatepedia ($H_2$), and the bottom graph illustrates the t-test for the opposite order-restriction ($H_3$). The dots mark the height of the prior and posterior distributions at $\delta = 0$.

Using the previous procedure, I evaluated the likelihood of the same unrestricted and order-restricted models but for the uniqueness of the arguments between Debatepedia and Wikipedia. The Bayes factor for $H_1$ was $BF_{10} = 2.28$ indicating that given the data $H_1$ (which corresponds to hypothesis $H_2$) is 2.28 times more likely than $H_0$. In terms of strength of evidence this is far from strong and it is different from the frequentist t-test which showed a medium effect size, $r = .404$. Nevertheless, both indicate that there is a difference in the uniqueness of the arguments between the two websites. The order-restricted testing yielded Bayes factors $BF_{02} = 11.73$ and $BF_{30} = 4.48$. These seem to provide substantial evidence that Debatepedia is more likely to produce more unique arguments than Wikipedia. Figure 4.3 shows the plots for prior and posterior distributions for all three hypotheses.
4. RESEARCHING THE INDIVIDUAL FEATURES FOR GROUPTHINK...

Figure 4.3: The prior and posterior distributions of effect size $\delta$, when comparing uniqueness of arguments between Debatepedia and Wikipedia. The top graph illustrates the unrestricted t-test ($H_1$), the middle graph illustrates the order-restricted test stating that Wikipedia produced more unique arguments than Debatepedia ($H_2$), and the bottom graph illustrates the t-test for the opposite order-restriction ($H_3$). The dots mark the height of the prior and posterior distributions at $\delta = 0$.

4.2.3 Bayesian Analysis for Readability

A similar assessment was performed for the readability data that were presented in the original paper. In the study, three readability measures were used in order to increase internal validity: FRE, Flesch-Kincaid and SMOG. Descriptive results for all three readability measures are depicted on figure 4.4. These show that Debatepedia had slightly superior results over Wikipedia. However, none of the readability comparisons between the two interfaces achieved significance which may be an indicator of readability being similar between the two interfaces. An additional Bayesian analysis can further help assert this claim.

Following a similar approach as described in the previous section, I conducted a Bayesian statistical analysis on the data. The Bayes factors for the unrestricted and order-restricted hypotheses for the FRE formula were, $BF_{01} = 1.69, BF_{02} = 8.29$ and $BF_{30} = 1.05$. In general results agree with the t-test on the original paper but substantial evidence was found only for rejecting the $H_2$ hypothesis. In contrast, more substantial results were obtained for rejecting all three hypotheses for the data based on the Flesch-Kincaid formula. Bayes factors were returned all in favor of the $H_0$ hypothesis, $BF_{01} = 3.13, BF_{02} = 2.18$, and $BF_{03} = 5.98$. Similarly, the Bayes factors for the data based on the SMOG formula also provided substantial support for the $H_0$, $BF_{01} = 3.17, BF_{02} = 2.26$, and $BF_{03} = 5.87$. Figure 4.5 shows the plots for prior and posterior distributions of all these results.

4.2.4 Discussion

The results from the Bayesian statistical analysis are in accordance with the results from the Null Hypothesis Statistical Testing (NHST) that was applied in the original paper. Results suggest that the two interfaces produce no different amount of arguments but do vary in terms of the unique arguments. The Bayesian analysis however allows for order-restricted hypothesis that can further help assert
more specific questions. Strong evidence was produced against the case of the order-restricted hypothesis that stated that Wikipedia produced more arguments than Debatepedia. Similarly, according to the order-restricted hypothesis for the amount of unique arguments between the two interfaces, Debatepedia seems to be favored by the analysis. Put simply, pro/con lists, while they do not seem to produce more arguments than textual collaborative interfaces, they are definitely producing more unique arguments.

The readability analyses also produced similar results with the original article. The Bayesian analysis is supporting the null hypothesis that states that both interfaces produced similar levels of readability for all three readability measures. However, that does not mean that articles produced similar ratings. As can be seen in the results provided in the original article, there is a great variance within the articles of Wikipedia and Debatepedia in terms of readability.

Aside from these results coming from the NHST and Bayesian analysis, there were additional features that were discovered during the qualitative analysis. One of the most prominent observations made during this analysis was the power that the original segmentation of the article had on the further development of the articles on both interfaces. In Debatepedia, pro/con lists were divided in terms of categories and claims per categories. Similarly on Wikipedia an article is segmented in terms of categories as well. For example, both interfaces occasionally skipped categories that were however included in one of them in their corre-
4. RESEARCHING THE INDIVIDUAL FEATURES FOR GROUPTHINK...

Figure 4.5: The prior and posterior distributions of effect size $\delta$, when comparing readability between Debatepedia and Wikipedia. Each sub-graph illustrates the unrestricted t-tests ($H_1$), the order-restricted test stating that Wikipedia scored higher in readability than Debatepedia ($H_2$), and the Bayesian $t$-test for the opposite order-restriction ($H_3$). The dots mark the height of the prior and posterior distributions at $\delta = 0$.

Pro/con lists had an advantage on producing unique arguments as it was found in the statistical analysis, however they are also limited. According to the original analysis, pro/con lists seem to have an upper limit for arguments per each claim (Tsikerdekis, in press-c). In the study the limit was found to be between two and three arguments per claim and rarely arguments surpassed this limit.
Recommendations

The recommendations based on the results produced by the Bayesian analysis, are still in accordance with the recommendations made in the original article. Given the evidence, pro/con lists are not superior for producing arguments that could increase information flow within a group and help prevent groupthink. However, there is strong evidence that they produce more unique arguments which is a desired feature for successful decision-making. Additionally, both interfaces are limited in the original categorization of the article that could have severe consequences in the production of arguments and is detrimental for preventing groupthink. Finally while pro/con lists seem to be favored by the results, they still have limitations when it comes to argument production. As such, pro/con lists are recommended to be used in software that supports decision-making for preventing groupthink however, a combination of the two interfaces investigated in this study may prove to be just as effective or potentially more effective.

4.3 Alternatives Randomization in Voting Interfaces

While having sufficient alternatives is essential for a decision-making process, it may not be enough to guarantee a positive outcome. After all, as indicated by Henningsen et al. (2006), groupthink as a five-step model has implications all throughout the decision-making process. One of the potential issues becomes apparent when considering that a list of alternatives is like any other list for the individuals reviewing it. In other words, the design of a list is subject to limitations which may or may not contribute to groupthink. In this study, the order in which lists represent items in collaborative social media was a feature examined to determine if it could affect the choices of individuals.

In this part, I focus specifically in the formation of voting procedures and the way alternatives are presented to individuals. This usually describes the final part of the decision-making process where all the alternatives have been introduced, analyzed and the final stage requires members to vote and reach a consensus. In web design, this would usually be displayed as a multiple choice list of items. It is tempting for designers to deliver the alternatives in terms of chronological order or to allow for leaders to design the item order in the multiple choice list. I introduce an alternative hypothesis which states that the randomization of the list of alternatives for each individual is a better choice for preventing groupthink behavior compared to a fixed-order list of items.

The choice of item presentation is far from trivial. Further, it is a realistic choice for any voting system. Whether, one designs a corporate brainstorming group or a collaborative project for educational purposes, this decision is an intricate part of the software especially if there are implications on the final outcome and could be the cause for groupthink.

For this hypothesis, I decided to follow a theoretical approach to establish whether the randomization feature for lists may reduce groupthink. There are two reasons for this choice. First, I can demonstrate the diverse approaches that
can be taken in order to evaluate software interfaces. Second, that parts of software evaluation for the new social interaction design framework can be based on existing research findings. The result demonstrates a tremendous advantage for HCI researchers that can produce cost-efficient results based on previous research and therefore avoiding having to “reinvent the wheel.”

4.3.1 Literature on Response Order and Behavior

Multiple choice is an extremely intuitive design and extremely popular for educational testing, research, elections, and other potential social media uses. In its simplest form, it requires from an individual to select one or more items from a list. It is highly popular in the United States (Phelps, 2005) but in the rest of the world as well. In fact, many social media today utilize this feature. Each option in a multiple choice design is usually called an item, and one theory governing the creation and evaluation of such items is the item response theory. This is a paradigm for the design, analysis and scoring of such representations that are measuring variables.

Van Der Linden and Hambleton (1997) pointed that in experimental designs the issue of controlling for experimental errors in responses is essential when designing tests. This is a similar issue in designing any kind of item response design such as multiple choice. They identified based on previous literature (Cox, 1958), that there are choices for coping with experimental error such as standardization or randomization. Standardization and randomization belong under the category where the experimenter (or in our case designer) intervenes and his or her decisions are expected to affect the outcome. Standardization is ideal for experimental designs (Field, 2009; Lewin, 2005) because it allows groups to be compared based on the same categories. If variance exists due to the way the items were presented, the error should be equal because the items were the same for all groups. With randomization on the other hand, this assumption would not hold due to the fact that different groups received questions in different orders.

Evidence for this can be traced in literature. Schwarz and Hippler (1991) identify the impact of response order in surveys. Specifically they state that the “presentation order may strongly influence the obtained results.” (Schwarz & Hippler, 1991, p. 51). The expected effects based on theoretical evidence are further identified as primacy and recency effects. Primacy effects create higher endorsements for items early in the list while recency effects have higher endorsements for items late in the list. Schwarz and Hippler (1991) also indicate that this effects are more likely to manifest when each response alternative represents different opinions for a topic rather when using ordered scales (Mingay & Greenwell, 1989) such as a Likert scale.

Primacy and recency effects in verbally-presented response categories are due to memory bias (Krosnick & Alwin, 1987). In verbal speech it becomes obvious that limited amount of time for reasoning and issues with retaining information from a list, may affect individual choices on items. However, today the same idea has been endorsed by literature for paper or digital surveys as well (Bishop & Smith, 1997; Jobe & Mingay, 1991; Jobe, Tourangeau, & Smith, 1993; Krosnick & Presser, 2010). In addition, there seems to be a general agreement that
such effects in surveys are more likely to have primacy than recency effects, since memory does not become an issue for the respondent in most survey designs (Schwarz & Hippler, 1991). However, Schwarz and Hippler (1991) noted that while memory may not be a factor affecting the decision-making process of an individual, contrast effects may be a cause for primacy or recency effects.

Contrast effects can manifest in any list including multiple choice as long as there is variance in the preferences of individuals for certain items. For example, when having an extremely likeable alternative for most of the group, the order that items will be positioned in relation to the extremely likeable item may affect the outcome. A moderately likeable item that will be presented after the extremely likeable solution may seem less appealing as an alternative while if presented before the popular item, it may lead people to viewing it more favorably.

4.3.2 Assessing the Appropriate Response Order

The overall likelihood of observing primacy effects and/or contrast effects becomes a serious issue in surveys as well as decision-making processes. The measurement error may lead to false decisions and conclusions for the groups involved. It is quite surprising therefore that these conditions have never been linked with the groupthink model and in turn with the design of software for social media, especially for collaborative projects.

Given my hypothesis and the information provided by literature, I compared the two competing options for designing how lists are represented in a voting interface; that is, standardization or randomization.

Standardization of Response Order

Standardization is ideal for controlling for measurement errors and seemingly obtaining a result that can be compared between individuals. In fact, one can argue that the differences that one observes between individuals are only due to their reasoning about the responses and not any other unknown factors. However, literature tends to disagree with this argument. Both empirical and theoretical evidence show that in survey designs choices may be made due to subconscious processes or as elsewhere defined, unsystematic variance (Field, 2009). Additionally, primacy and contrast effects may sufficiently contribute to such differentiations that in turn may affect the voting outcome.

One way for controlling such errors while using standardization for the response order, is to develop the order based on previous information and display items appropriately. Prior information may come from results of a pilot study or the insights of an individual. This prior information can then be used to assign the order of the items in such way that will be unfavorable for the items that were favorable before. The easiest way to achieve this is by inverting the original response order. After the final results are obtained, one can determine if there was an effect due to the response order.

However, such method is not without issues. One of the problems is that the overall selection of the response order is left to an individual which introduces
subjectivity. Even if that individual is an expert on the topic, has discussed and understands the general preferences of the group, he or she is still required to make a decision based also on intuition. Moreover, even if one detects presence of primacy or contrast effects when comparing the outcome with the prior information, there is no ideal course of action. As an example, one can revise the response order once again and repeat the voting process or merge the final results with prior information in order to control for the error. Either can be costly and does not guarantee a genuine outcome. As Schwarz and Hippler (1991) pointed out, because “the relative power of the factors involved is difficult to evaluate a priori, the emergence of response order effects is likely to remain a surprise in many specific cases” (p. 56).

Randomization of Response Order

Randomization on the other hand, is likely to produce more variability in the results but I argue that errors are more likely to cancel each other out between individuals. Especially the case of primacy effects can be diminished on a group level since even if participants have a tendency of selecting the early responses, the effect will be diffused in the overall group.

However, contrast effects will still be present in a randomized response set. In fact, one may argue that a standardized response set may at least provide some control over contrast effects. But, the variance that will be created by the randomization is likely to end up diffusing the contrast effects on a group level. Moreover, while a standardized response set ideally can reduce the contrast effects, it is extremely difficult to determine if that is the case a priori.

Randomization for multiple choice designs has been praised by McLeod, Zhang, and Yu (2003) for educational purposes. As they suggest, randomization provides advantages for preventing cheating and has no adverse effect on student performance. Coincidentally, this may provide support that contrast effects may be diluted within a sample of randomized sets. Multiple choice however, were found to also have certain negative effects. One of them is that students may actually try to guess that the more deceptive option may be the correct one (Roediger & Marsh, 2005). While this may not be an issue for decision-making processes as people have no need to guess for correct answers, it is still an issue worth considering. As advocated here, understanding the effects of the interface can only improve the design of the software.

Finally, randomization can be applied along with standardization across randomized sets. This provides a way for evaluating any effects caused by randomization and evaluating issues with voting procedures. For example, one approach may be to deliver randomized sets in a stratified manner just as it is applied in stratified random sampling in quantitative research. However, as mentioned by Hensher, Rose, and Greene (2005), research is limited in determining what would be the ideal number of randomized sets for a sample. Perhaps, such procedures can be applied in a similar way to conducting survey research. Lewin (2005) indicated minimum estimates for the subgroups within a sample to be between twenty to fifty depending on the design. Using these estimated one can determine the number of randomized sets to be delivered to a group. For example, a
4. RESEARCHING THE INDIVIDUAL FEATURES FOR GROUPTHINK...

A group of a hundred members may receive five different versions of response sets based on randomization.

4.3.3 Conclusions

Based on this literature review, I attempted to make a connection between the design of the order for alternatives and the chances of occurrence for groupthink. The choice between a standardized or a randomized design may not be clear for controlling contrast effects but randomized response sets have a clear advantage over standardized response sets for controlling primacy effects.

When designing software that may affect voting processes it is important to control for errors caused from these effects. Given the substantial advantage for randomized sets, algorithms can be developed that could create such sets for the list of alternatives. In addition, as I mention before, a designer can also have control over how many randomized sets will be created and even manufacture sets to be delivered to the end users in a stratified manner. These sets allow for post-hoc procedures to be applied so that software can indicate whether one of the randomized sets was plague by certain issues due to the response order. Algorithms can be developed to not only test that but also raise red flags to the group after the votes are in. Tests of statistical significance can be automatically applied for such tests and give an estimate of the differentiation between randomized sets.

The end result is that designers can claim that the voting process provides measures for preventing effects that could introduce bias in the voting results and guarantee that they have taken steps to prevent groupthink. Even in cases where members knowingly are willing to attempt to conform with a group’s suspected favorable position are likely to face different response sets, which in turn will produce different contrast effects. The result is that the software ensures that all alternatives are given equal consideration in terms of the response order on a group level.

4.4 Leaders in Social Media Affecting Opinions

One of the early hypotheses for groupthink described in my dissertation proposal involved the issue of leaders being publicly identified in social media and their power to affect voting procedures. Janis (1982) in his groupthink model assigned significance on the role of the leader and the decision-making process. One of the recommendations involved not allowing the leader to express their opinion at least when initially are discussing a problem. However, I argued that if left untested one can sufficiently claim that an identified leader by other members cannot affect the voting outcome even at a later stage. It was therefore important to determine how much of an effect physical design features would have for identifying leaders in a voting processes.
4.4.1 On Leaders and Groupthink

Studies on leadership show that leaders can affect behavior in several ways. George (1995) has found that a leader’s positive mood can improve group performance. Similarly, Sy, Côté, and Saavedra (2005) investigated the effects of leaders’ mood on groups. They found that a leader’s mood can affect the group members’ mood, the affective tone between group members and the coordination within a group. In turn, mood can also affect a leader’s effectiveness and attraction to a leader for group members (Bono & Ilies, 2006). Further, leaders seem to have even more power than just affecting mood and improving efficiency within corporate environments. Jones and Olken (2005) found positive evidence that leaders are detrimental to the economic development of nations and can affect policy outcomes. As such, their role in group decision-making processes is crucial and hold a great potential for affecting a group’s decision in terms of groupthink.

According to early literature, leadership is associated as a condition that could lead to groupthink. Leana (1985), in a study utilized students which were divided into groups that needed to resolve a hypothetical business problem. Some teams had leaders assigned to them and were instructed to be directive while others had leaders who encouraged member participation. The result was that groups with directive leaders discussed fewer alternative solutions. Similar studies favored these results (Z. Chen, Lawson, Gordon, & McIntosh, 1996; McCauley, 1989). Ahlfinger and Esser (2001) in their study, tested the hypothesis that groups with leaders that promoted their own preferred solutions will likely develop groupthink. The findings showed that groups with promotional leaders produced more symptoms of groupthink, discussed fewer facts and reached to a consensus much faster than groups with non-promotional leaders.

Promotional leadership is not only found in laboratory experiments but in real life scenarios as well. In one of the case studies for groupthink involving the space shuttle Challenger, Moorhead et al. (1991) indicated that top level managers actively promoted their pro-launch opinions in the face of opposition. In addition several managers ignored the warning about low temperatures and tried to push for launch.

In social media geared towards collaboration, leaders exist in a variety of ways. The terms administrator or moderator are highly popular in forums and discussion boards. Other types of social media such as virtual game worlds have been more creative with terms such as guild leaders or dungeon masters. No matter the title, these individuals are in key positions within an organization and as leaders can affect decisions. When social media software is designed one has to evaluate if these leaders will affect group decision-making in a negative manner and be a cause for groupthink.

It is challenging however, to determine aspects through which one can balance the negative effects of leaders and promote the positive effects. Rosenthal and Okie (2005) suggested that anonymous forums with no administrators overseeing the messages may be more effective and increase participation for students suffering from depression. Disabling leaders may be possible in certain cases, however, this may not be possible in other cases where the presence of a leader is necessary for increasing positive attitude and performance in a group (Bono &
Ilies, 2006; George, 1995). A better approach however, may be to avoid making leaders appear so prominent especially under crucial processes such as the final stages of a decision-making process (e.g., voting).

There is a great variety of voting mechanisms which definitely have an effect on the outcome of a voting process (Brams, 1991; Dummett, 1984; Gavish & Gerdes, 1997; Z. Li, 2003). Dynamic open voting is an attractive idea for many groups that wish to see individuals and their choices as the voting process progresses. In many cases, leaders are distinguished from users using titles in front of their names or even color variations; a choice extensively used by software designers in forums and similar discussion boards. Given the power that leaders hold over a group and the potential for groupthink, this public identification of leaders and their distinction from the rest of the group can potentially affect voting outcomes. This helped formulate the main hypothesis for this part of the study:

\[ H_1: \] Visually identified leaders can affect a voting process where individuals vote publicly.

4.4.2 Method

In order to evaluate this hypothesis, I developed a set of questions and included them in a larger survey that I had administered to evaluate other hypotheses that were part of the overall dissertation work. Hereby, the large survey is denoted as GR1 in other sections of the dissertation. The type of survey was a randomized controlled trial “in which participants are allocated truly randomly to an experimental group and a control group” (Lewin, 2005, p. 221) which strengthens the claims for internal validity.

The survey involved a scenario about global warming where people had to vote for one of four potential solutions. The simulated voting procedure involved sixteen individuals which had already voted for their solutions and the result was a draw. The respondent was called upon to vote and essentially break the tie with his or her decision. Individuals were identified publicly by name, while the administrator was distinguished from the rest of the individuals with a different color as well as a title in front of the name. Two versions were created where in the first the leader voted for the first solution while in the second the vote was placed on the third solution. Figures 4.6 and 4.7 depict these two versions.

Additionally, a further open-ended question was included in the survey in order to help understand the reasons behind the respondents’ choices and in order to help triangulate and complete the results (Bryman, 2006) (“Why did you select this solution? Please elaborate.”).

The survey was administered through Facebook. Invitations for the survey were circulated in several academic and non-academic groups and fan pages on the social networking website. Additionally, a snowball sampling method was introduced by allowing users to invite their friends in order to reach farther users within the network. The survey was originally developed in English but also translated to Greek and Spanish as well. The language was automatically adapted based on the respondents’ system settings and Facebook profile settings. The survey took place during the period of February 15 to March 19, 2012.
4. RESEARCHING THE INDIVIDUAL FEATURES FOR GROUPTHINK...

4.4.3 Results

In total, 123 people visited the application during the survey and 89 (72.4%) went through the whole survey. Most of the non-respondents quit the survey while on the first page. The sample of 89 respondents was used for the analyses. Their demographic information reveals that the sample was well balanced in terms of sex but in terms of age it was relatively young. The self-reported geographical information reveals that most participants came from Europe. Figure 4.8 summarizes the demographic data.

Looking at the results between treatments the differences in the votes did not seem to be large. Figure 4.9 shows the counts for each category per treatment. In both cases the voting outcome was the same. However, there was a small difference in the preference of the third solution. It seems, that by having the leader identified and voting for the third solution did affect people’s choice up to a degree. However, was this a significant effect?

A chi-square test for the contingency table 4.1 showed no statistical significance, \( \chi^2(3) = 4.336, p > 0.05, V = .221 \). However, two of the assumptions for the chi-square test were violated. Seventy five percent of cells had an expected count less than five and the minimum expect count was less than one. This severely impacts the accuracy of the result and therefore a Fisher’s exact test was performed in order to obtain the exact significance value. The result was also not significant, \( p > 0.05 \).

I used Bayesian statistics to determine the probability of the null hypothesis. I used a procedure developed by Albert (2009) which can be accessed from R using the LearnBayes package. The ctable function returned a Bayes Factor in favor of the null hypothesis, \( BF_{01} = 11.33 \). This indicates that the null hypothesis...
You are a member of an administrative board on environmental issues. You are participating in a forum discussion about solutions for global warming. Many arguments and solutions are proposed and four of them stood out from the rest. The administrator of the forum initiated a voting poll in order to determine which of the solutions should be put to action. Everyone can see who voted which solution. You are the last to vote and by the looks of it you will decide on the winning solution.

Solutions for global warming vote:

| Reduce fossil fuel usage, conserve and switch to renewable energy. | 4 | 25.00% |
| A man made satellite that can deflect the sun's energy. | 4 | 25.00% |
| No need to worry about global warming, the earth will recover by itself. | 4 | 25.00% |
| Wait until a more cost-efficient solution is being discovered. | 4 | 25.00% |

Voters: 16

1. What do you think is the best solution to global warming?

\( \bullet \) Reduce fossil fuel usage, conserve and switch to renewable energy.
\( \bullet \) A man made satellite that can deflect the sun's energy.
\( \bullet \) No need to worry about global warming, the earth will recover by itself.
\( \bullet \) Wait until a more cost-efficient solution is being discovered.

Figure 4.7: The second version of the voting process with the leader supporting the third solution.

<table>
<thead>
<tr>
<th>Treatment / Response</th>
<th>Reduce...</th>
<th>A man...</th>
<th>No need...</th>
<th>Wait...</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leader supporting the first position</td>
<td>42</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>48</td>
</tr>
<tr>
<td>Leader supporting the second position</td>
<td>33</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>41</td>
</tr>
</tbody>
</table>

Table 4.1: Contingency table for each treatment and votes given by the respondents

of independence was 11.33 times more likely over the alternative hypothesis of dependence. This is considered strong evidence (Jeffreys, 1998; Kass & Raftery, 1995; Wagenmakers et al., 2008). Additionally, using the bfindep function, I evaluated the case of having a hypothesis “close to independence” allowing for some variability due to random factors. This hypothesis was tested in terms of its likelihood over the hypothesis of independence for \( K \) precision parameters between 2 and 7. Having 100,000 simulations, the highest Bayes factor that was obtained was \( BF_{10} = 1.16 \). The result indicated weak evidence in favor of the hypothesis “close to independence” against the null hypothesis of independence.

Most of the respondents (86.5%) have also elaborated on their choice on the open-ended question while a small percentage (13.5%) left the question unanswered. Qualitative results seem to agree with the statistical analyses. Many respondents showed that their decision was based on their emotional reasoning. For example one respondent under the first treatment stated that he or she chose
the first solution “Because i [sic] believe is the most applicable.” Similarly, a respondent elaborated further by saying that “It’s the most active solution. Something we can do right now. It will also efficiently reduce costs.” Another individual preferred the first choice while however under the second treatment. He or she said that the choice was made “because it is what i [sic] already believe is an imperative for people.” Furthermore, an individual with a rationalized choice under the first treatment, voted for the second solution while stating that it is a better option “cause [sic] are day to day habits would not change.”

In general, almost all the answers for the qualitative question followed a similar pattern and at no place there was an indication that the position of the leader affected their decision. One answer however was particularly different from the rest. An individual under the second treatment that selected the third solution (the leader’s choice) stated that he or she was “unsure.” While a single word
barely classifies as informative it is still worth mentioning given that this was the only individual of those that had answered that they could not elaborate on his or her choice. Under the second treatment, there were additionally two respondents that preferred the third solution however did not answer the qualitative question. It may be the case that these additional individuals may have felt in a similar manner and were unsure about committing to one of the choices provided. This may indicate that in general leaders may not affect the voting process but under circumstances where individuals hold a degree of uncertainty over a topic an effect may be found.

4.4.4 Discussion

The results have been promising for social media software and especially online collaborative communities. All statistical analyses pointed to the same direction. Leaders while identified in the voting process do not have enough power to change the voting outcome. This result is promising given the popularity of having administrators and other members holding leader positions being publicly identified in many online communities. In addition, it also shows a great difference with the general consensus over promotional leadership and its relation to groupthink. According to the results, by having administrators publicly show their preference while being publicly identified would not necessarily contribute to groupthink.

However, some minor differences were observed especially for the third so-
lution. The difference between treatments indicated that some users did follow the leader’s position. This is a small difference and did not change the outcome; which could have otherwise been a cause for concern in terms of groupthink. However, it is still worth considering. This is especially important given the qualitative findings. While one can design a voting interface with leaders being identified and expect that this would not affect the voting outcome, it has the potential to distort the results and affect individual votes.

Recommendations

Given the evidence, it is safe for software designers and developers to utilize in social media open voting processes where leaders are publicly identified and distinguished by the rest of the users. Voting outcomes are likely to not be affected and therefore the potential for groupthink is small at the final voting stage of a decision-making process. However, this design should be implemented with caution. While the overall outcome is not affected by the leaders’ positions, individual votes may be affected under conditions where individuals are uncertain or confused about a topic. Additional information that can be provided about a topic and each solution may help individuals better evaluate their decisions.

Limitations

This study has similar drawbacks with any laboratory experiments. Occasionally laboratory settings cannot sufficiently reproduce real world situations. For example, in groups usually people have affections and preferences for other people within the group. If the administrator is popular, he or she may have more power over the voting process. In addition, members may not wish to make their votes public in favor of one solution that goes against the crowd and it is an issue that should also be addressed.

4.5 Anonymity States and their Effect on Non-Conformity and Groupthink

One of the features investigated by my dissertation work was the effect of the different anonymity states that can be found on social media today in relation to groupthink. These anonymity states were: pseudonymity and individuals being completely anonymous. According to literature, anonymity has long being found to be a condition that may reduce compliance pressures and is an antecedent for groupthink (Ahlfinger & Esser, 2001; McCauley, 1989; Rovio et al., 2009). However, anonymity states were never studied in regard on their effect on groupthink. My aim was to establish if there is an effect in terms of ensuring alternative solutions for situations that require problem-solving. An adequate amount of alternatives that should be explored is an important measure for preventing groupthink as stated by Janis (1972, 1982). The outcome of this investigation led to a published article on a scientific journal (Tsikerdekis, in press-b). In this section, I provide the method and major findings of this research.
4.5.1 Method

For the purposes of this study, I decided to use a survey which had the form of a randomized controlled trial (Lewin, 2005). I created three different vignettes that featured scenarios of a group making a decision about a particular problem. The majority of the group had already decided on a particular position and the respondent was required to make a choice on whether he or she would propose an alternative solution for discussion or follow the group’s opinion. Prior to that, their opinion was obtained for each problem presented on the vignettes and the software always simulated the group’s opinion to be different from the one selected by a respondent. Several additional questions were retrieved from individuals regarding how important they thought the resolution of each problem was, along with their perception of anonymity under each vignette.

For each of the three different vignettes, three versions were produced that were different only on the presentation of the anonymity state that the group was collaborating under. These were pseudonymity, complete anonymity and an additional state under which participants used their real names. Each respondent received the three vignettes in a random order. Additionally, each was randomly assigned one of the versions representing a different anonymity state.

The research was conducted in collaboration with the Wikimedia Foundation Research Committee which reviewed and approved it. A random sample of respondents was invited through the English version of Wikipedia. A detailed account of the survey design, process and overall survey information can be found in the original article (Tsikerdekis, in press-b).

4.5.2 Results

Several hypothesis were tested in the study (Tsikerdekis, in press-b)

- $H_1$: With higher levels of anonymity, the likelihood of not conforming increases.
- $H_2$: With higher levels of anonymity, the perception of anonymity increases.
- $H_3$: As the perception of anonymity increases, the likelihood of not conforming increases.
- $H_4$: There will be a relationship between a given scenario and the likelihood of not conforming.
- $H_5$: There will be a relationship between a given scenario and the perception of anonymity.
- $H_6$: As the level of importance assigned to a problem increases, the likelihood of not conforming increases.
- $H_7$: As the level of importance assigned to a problem increases, the perception of anonymity decreases.

A research model was also formed based on these hypotheses which is depicted in figure 4.10 along with the results based on the statistical analyses.
Hypotheses $H_1, H_2, H_3, H_4, H_7$ were supported while $H_6$ was conditionally supported and $H_5$ was not supported. Anonymity states had an effect on conformity and as such on groupthink. The more anonymous an individual was, the more likely he or she would not conform. Figure 4.11 depicts counts for the likelihood of not conforming for each one of the anonymity states. A statistically significant effect was found, however, it was small, $r_s = .102, p < .05$ (1-tailed).

Similarly, using the scale of perception of anonymity, a small effect was found with the likelihood of not conforming, $r_s = .101, p < .05$ (1-tailed). Figure 4.12 depicts a line chart between the two variables. In general, the medians indicate that up to the point where the perception of anonymity is reported higher, the likelihood of not conforming is one value lower.

In order to establish whether the perception of anonymity was indeed dependent on the anonymity states the two variables were analyzed using Spearman’s correlation. The result indicated that there was a statistically significant medium effect size, $r_s = .466, p < .001$ (1-tailed). Figure 4.13 clearly displays this association between the two variables.

Additionally, scenarios and level of importance had an effect either directly or indirectly through other variables. Scenarios also seemed to have had inhibiting effects for correlations. For example, the scenarios involving a noisy neighbor and a misbehaving employee showed conformity being dependent on anonymity, $r_s = .178, p < .05$ (1-tailed) and $r_s = .160, p < .05$ (1-tailed) respectively. However, the scenario involving an employee’s unpaid overtime did not show a statistically significant association between these variables. Instead, it showed that the respondent’s likelihood of not conforming depended greatly on the self-reported level of importance assigned to the problem, $r_s = .222, p < .01$ (1-tailed). Figure 4.14 depicts the likelihood of not conforming for all three scenarios in respect to the self-reported level of importance.

Qualitative evidence while agreeing with the quantitative data, in many cases showed individuals that stated anonymity was a factor for their decisions. Many respondents made clear statements that they would not propose their alternative
solutions under circumstances where real names were used and have done so. Many were afraid of the consequences especially in some of the scenarios where the penalties might have been more severe for going against the rest of the group. Additionally, among the findings discovered from the qualitative data was the issue of trust between respondents and social media services. There were statements made by the participants that indicated that they were skeptical about the anonymity provided by a service. As such, this was identified as a probable factor that may affect perceptions of anonymity and in turn the likelihood of conformity.

Recommendations
Recommendations can be found in detail in the original article (Tsikerdekis, in press-b). It is recommended based on the evidence brought forth by this study that software designers and developers should promote the use of anonymity especially in social media aimed for collaborative activities. While the effect is small, it is still significant so that all potential solutions should be examined before dismissed. Between pseudonymity and complete anonymity, the latter holds the greatest potential for reducing groupthink. An alternative course of action could also be the use of real names or pseudonyms with an additional option for anonymous information exchange. Additionally, designers should find ways to communicate with their users and alleviate any issues of trust that may exist for anonymity.
Additionally, several other factors should be given enough consideration. For example, the level of importance that participants assign for a resolution of a problem, has been found to affect the likelihood of conforming. When designing collaborative software, it may be beneficial to assess a self-reported level of importance from the collaborators in order to establish prior to a conference what would the likelihood for conformity and groupthink be. Finally, developers should try to better communicate how anonymity is being implemented and generally to achieve higher levels of trust for the anonymity provided to the final users.
Figure 4.13: Line chart displaying the gradually ascend of the perception of anonymity as progressing to higher anonymity states.
Figure 4.14: Likelihood of not conforming for all three scenarios based on the self-reported level of importance. The second graph is significantly being impacted by this dependency.
CHAPTER 5

ONLINE AGGRESSION THEORY AND APPLICATION

5.1 Introduction

In this chapter, I provide a literature review on aggression and present features that can potentially inhibit aggression. The purpose is to provide a clear view for the motivation and the importance into researching online aggression as well as what should be considered as important when designing social media. Specifically, the aim of this chapter is to demonstrate that one can engineer aggression through software features. As with the case of groupthink, having an understanding of sociological and psychological literature is important for developing procedures that establish how changes in the software may alter social behaviors.

5.2 On Aggression in General

Aggression is a phenomenon that can be seen in all aspects of life. On the other hand, online aggression as a phenomenon is relatively new compared to the span of human civilization. However, knowledge of the two is necessary in order to understand the differences between them as well as for identifying design opportunities. Many studies in aggression take a comparative approach between online and offline aggression in order to examine such differences (Dooley, Pyżalski, & Cross, 2009; Hinduja & Patchin, 2010; Schoffstall & Cohen, 2011; Yen, Yen, Wu, Huang, & Ko, 2011).

5.2.1 Definition

Defining aggression can be difficult especially when attempting to identify various classifications of aggression. As Baron and Richardson (2004) pointed out, Berkowitz (1993a) described the linguistic issues for defining aggression.

When people describe someone as being aggressive, they might be saying that he frequently attempts to hurt others, or that he is often unfriendly, or in a quite different sense, that he is typically very forceful and tries to get his own way in his dealings with others, or maybe that he assertively stands up for his beliefs, or perhaps that he usually attempts to solve problems facing him. (p. 4)
The main issue is whether aggression is defined by the intention to do harm to others or when harm is actually inflicted on others. Baron and Richardson (2004) provided a working definition that includes both intention to harm and harming someone:

Aggression is any form of behavior directed toward the goal of harming or injuring another living being who is motivated to avoid such treatment. (p. 7)

This definition identifies aggression as a behavior, not as an emotion or motive. In addition, it also requires a clear intention by the assailant while the victim should always try to avoid harm. Additionally, within this definition one can also identify violence, which is aggression that has extreme harm as its goal (Anderson & Bushman, 2002). While this definition is endorsed for the purposes of this study, it should also be acknowledged that in an online setting damage can be inflicted unintentionally due to software design. However, classifying accidental damage inflicted on someone as aggression can be problematic.

5.2.2 Types and Categories of Aggression

Aggression can be identified as physical or verbal (Baron & Richardson, 2004) but one can also observe indirect aggression (Björkqvist, 1994). Verbal aggression is often described as psychological aggression which is usually perceived to precede physical aggression (Capaldi & Crosby, 1997). This is a particularly interesting type of aggression since it is essentially an online type of aggression as well. Follingstad, Rutledge, Berg, Hauge, and Polek (1990) identified several types of psychological abuse such as threats of abuse, ridicule, jealousy, threats to change relationship, restriction and damage to property. Additionally, Kessler (1990) included insults, swearing, spiteful behavior, and abrupt withdrawal from interaction. Arguably many of these cases of psychological aggression can be found in social media today. While one may argue there are not as damaging as physical aggression, real world studies show otherwise (Follingstad et al., 1990). Hence, it is important to explore such issues and find ways that aggression can be reduced in social media.

Furthermore, aggression can be divided into two major categories; hostile aggression and instrumental aggression (Anderson & Bushman, 2002). Hostile aggression is also called affective, impulsive, reactive, or retaliatory. It can be distinguished by instrumental aggression where there is a need for premeditative means of obtaining some goal aside from harming someone (Berkowitz, 1993b; Geen, 2001). Hostile aggression is also found to be associated with people that have a lower IQ score (Vitiello, Behar, Hunt, Stoff, & Ricciuti, 1990). Ramírez and Andreu (2006) suggested based on their research that personality measures can be used to clearly differentiate aggressive subjects. Arguably, the same personality measures can be used for a population in social media to identify which category of aggression is more likely to appear given the types of users that exist in the online community.
On the other hand, while this categorical distinction between hostile and instrumental aggression is supported by some researchers (Mcellistrem, 2004; Weinshenker, 2002), others insist that such a dichotomy fails to consider aggressive acts with multiple motives and advocate against it (B. J. Bushman & Anderson, 2001). While in this dissertation, I will not elaborate more on this debate, I do consider both perspectives as likely to be better explanatory and descriptive models depending on the situational context. As mentioned previously in chapter 3, I further advocate that HCI research involved with psychological and sociological behaviors should always maintain a broad view with the concepts involved and should not try to act as the agent for resolving long-term debates from other disciplines. Instead, it should include using the knowledge coming from these fields in order to improve over the current design of software and maintain a practical perspective for the issues studied.

5.2.3 The Nature of Aggression

Aggression based on the definition presented in this chapter is not restricted in the human species. For example, as indicated by Lorenz (2002), aggression is a life-preserving mechanism for insects. It is not surprising that such mechanism can be found in other species including ours. However, the origin of aggression is debated. Researchers seem to disagree on whether the origin for human aggression has a biological or evolutionary basis (Somit, 1990).

Proponents of the evolutionary basis have indicated how some of these behaviors may have evolved (Buss & Duntley, 2006; McCall & Shields, 2008) however, there is difficulty in distinguishing which were truly selected and which have been a by-product of the process, as it is in the case with collective violence (Durrant, 2011). In addition, selection seems to have been part of the differences between sexes for aggression. The differences between male and female are consistent with sexually-selected behavioral differences (Archer, 2009) with a support that women may have evolved to be less physically dangerous and more covert (A. Campbell, 1999, 2005).

On the opposite side of the seesaw, advocates of the physiological basis identify several root causes. Studies in animals have found that electrical stimulation of the hypothalamus causes aggressive behavior (Kruk et al., 1983). Others psychological responses can be found in chimpanzees which as continuous breeders show significantly raised testosterone levels and aggression between males when females are present (Muller & Wrangham, 2004).

No matter the case, it becomes evident that aggression is an intricate part of our existence and reality.

5.3 Research on the Causes for Human Aggression

Research on aggression was conducted in a variety of contexts such as in culture (Fujihara, Kohyama, Andreu, & Ramirez, 1999; Harman, Klopf, & Ishii, 1990; Lomas, 2009), aggression in media (Aronson, Wilson, Akert, & Akert, 2009; Freedman, 2002), in children (Bongers, Koot, Van Der Ende, & Verhulst, 2008;
Network, 2004; Tremblay, 2000) and even between sexes (Björkqvist, Österman, & Lagerspetz, 1994; Card, Stucky, Sawalani, & Little, 2008; Hines & Saudino, 2003). There is an increasing interest over the recent years to understand why humans occasionally act aggressively (Anderson & Bushman, 2002). Anderson and Bushman (2002) in their review identified a number of causes for the increased aggression such as accessibility to guns (O’Donnell, 1995), global warming (Anderson, Bushman, & Groom, 1997), violence against children in schools and homes (Hyman, 1995), and the widespread exposure to media (B. Bushman & Huesmann, 2001). Explanatory theories were at the core of all of these studies that aimed to understand the nature of human aggression.

5.3.1 Theories on Aggression

Some of the most prevalent theories were presented by Anderson and Bushman (2002). These are described here as they are an important part for identifying design features that may have an effect on social interaction related to aggression.

Neoassociation Theory

The theory built upon the ideas developed by Berkowitz (1989, 1990, 1993a) where aversive events such as frustrations, provocations, loud noises and similar events produce a negative effect that automatically stimulates various thoughts and memories, expressive motor reactions and physiological responses associated with fight or flight tendencies (Anderson & Bushman, 2002). The theory was built upon the frustration-aggression theory developed by Dollard, Doob, Miller, Mowrer, and Sears (1939) which at its simplest form states that the more frustrated an individual is, the more aggressive he or she is likely to become.

Social Learning Theory

A different approach of explaining the cause of aggression is taken by the social learning theory. In this case, aggression becomes a social behavior that is obtained through observing others and in effect learning how to act in situations (Bandura, 2001a; Mischel, 1973, 1999; Mischel & Shoda, 1995). This theory is ideal for describing differences in aggressive behaviors between cultures. In addition as indicated by Anderson and Bushman (2002) the theory is ideal for explaining acquisition of aggressive behaviors as well as instrumental aggression.

Script Theory

Script theory resembles some similarities with the social learning theory. It states that as individuals observe violence in mass media they learn aggressive scripts. This was first proposed by Huesmann (1986, 1998) to explain behaviors of children. Once a script is learned it can be used at any time to guide a behavior. Further research into script theory by Abelson (1981); Anderson (1983); Anderson and Godfrey (1987); Marsh, Hicks, and Bink (1998); Schank and Abelson (1977)
indicated the power that these scripts have in dictating social behaviors. As Anderson and Bushman (2002) puts it, “a child who has witnessed several thousand instances of using a gun to settle a dispute on television is likely to have a very accessible script that has generalized across many situations. In other words, the script becomes chronically accessible” (p. 32).

Excitation Transfer Theory

Excitation transfer theory as developed by Zillmann (1983), describes how physiological arousal can dissipate slowly and be transferred between several arousal events. If the event following a previous event brings forth anger then the additional residual arousal can make the person angrier. Anderson and Bushman (2002) also has indicated that anger can stay for long periods of time with the person even after the arousal has dissipated.

Social Interaction Theory

Social interaction theory seeks to interpret coercive actions as part of social influence (Tedeschi & Felson, 1994). The choice to act in an aggressive manner falls under an individual’s judgment based on his or her cost-benefit analysis of a situation. This theory helps explain aggressive acts that are motivated by higher level goals (Anderson & Bushman, 2002).

5.3.2 General Aggression Model

As described by Anderson and Bushman (2002); DeWall, Anderson, and Bushman (2011); Parrott (2008) the General Aggression Model (GAM) becomes a unifying model for all the theories on aggression that were presented previously. As an effect of this integration of all theories into one model, the general aggression model as a social-cognitive model incorporates biological, personality development, social processes, basic cognitive processes, short-term and long-term processes, and decision processes into understanding aggression (DeWall et al., 2011). Several articles describe in detail the inner working of the model (Anderson & Bushman, 2002; DeWall & Anderson, 2011). The most important parts and structure of the model are presented here.

The model contains three critical stages in understanding a single episodic cycle of aggression: (1) person and situation inputs, (2) present internal states (i.e., cognition, arousal, affect, including brain activity), and (3) outcomes of appraisal and decision-making processes (DeWall et al., 2011). Additionally, Anderson, Buckley, and Carnagey (2008); DeWall and Anderson (2011) indicated that a feedback loop can influence future cycles of aggression and start a violence escalation cycle.

In the input stage, variables based on person and situational factors play a critical role in the overall process. Beliefs, traits, attitudes, sex, values, long-term goals, scripts are all factors related to an individual (Anderson & Bushman, 2002). Similarly, situational factors include aggressive cues, provocation, frustration, pain and discomfort, drugs, and incentives.
In turn, these input variables develop the present internal state and through this process the final outcome behavior can be affected. These internal states are: (1) cognition which includes hostile thoughts (i.e., chronically accessible prime aggressive thoughts) and scripts, (2) affect which includes mood and emotion and expressive motor responses (i.e., unexpected pain which produces angry expressions), and (3) arousal which can influence in several ways such as, (a) if coming from an irrelevant source strengthens the action tendency (especially aggressive tendencies), (b) if elicited by an irrelevant source and is mislabeled as anger in situations involving provocation, (c) if high and low levels of arousal can be aversive states and as such stimulate aggression (Anderson & Bushman, 2002). All of the above states are interconnected; effectively one causing the other, and vice versa.

The final stage describes appraisal and the decision-making process. Appraisal can be further analyzed into immediate appraisal, which is considered to involve relatively automatic information, processing and reappraisal which is considered to involve more heavily controlled information processing (Anderson & Bushman, 2002). In the cases of an immediate appraisal if the person has sufficient resources (i.e., time, cognitive capacity) and the immediate appraisal outcome is both important and unsatisfying, then an individual will seek a more cognitively demanding set of reappraisals. Otherwise, an individual exhibits impulsive actions which can be aggressive or non-aggressive depending on the content of the immediate appraisal.

Due to its unifying nature, the model has been successful in explaining several aggressive behaviors. DeWall et al. (2011) indicated how the model can be applied to understand intimate partner violence, intergroup violence, global climate change effects on violence, and suicide. Similarly, Parrott (2008) has used this model to develop a theoretical framework for understanding antigay aggression.

However, the model is not without its critics. Ferguson and Dyck (2012) reported that the model may not be satisfactory due to inconclusive evidence. They further based their claims on the fact that social science is based mainly on NHST which provides no road to falsification and that most researchers have a tendency to attribute a non-significant result in the small effect that is attempted to be detected and not considering that there may be such a small effect that could become trivial. Hence, misinterpretations of the effect sizes are considered an additional cause for supporting the general aggression model according to Ferguson (2009); Ferguson and Dyck (2012).

### 5.3.3 Additional Related Phenomena

There are several phenomena in literature that could act also as explanatory models for certain cases of aggression. Some are listed below as they were described by Anderson and Bushman (2002).
Opportunity

Opportunity or the social situation may restrict or allow for aggression (A. P. Goldstein, 1994). Schools have such restrictive rules where aggression is discouraged through strong social norms and the presence of witnesses. On the other hand, bars in which there is a presence of alcohol, aggressive cues, relative anonymity and competitiveness between males for females are more likely to develop more aggression for individuals; a model that fits well with the general aggression model as well (Anderson & Bushman, 2002).

Overriding Inhibitions

Research on overriding inhibitions identified certain patterns in which individuals that would not normally act aggressively, act in such a manner (Bandura, 2001a; Bandura, Barbaranelli, Caprara, & Pastorelli, 1996; Keltner & Robinson, 1996; Staub, 1989). The two important mechanisms for this are moral justification in which individuals rationalize a decision (i.e., when an individual sees that the end for a common good justify the means) and victim dehumanization in which moral standards are not applicable (i.e., war propaganda) (Anderson & Bushman, 2002). Within this phenomenon there has also been research about online disinhibition which becomes relevant for this study as well (Suler, 2004a).

Shared Motivations

Social groups have several common needs that may feel inclined to protect when a threat to those needs appears (Anderson & Bushman, 2002). Aggression is often the dominant response to such threats which as suggested by Anderson and Bushman (2002) positions of power are abused as well as the probable evolutionary tendency to emit aggressive behaviors when faced with pain, whether it be physical or psychological.

Role of Anger

Anger also seems to act as a factor for aggression. This view has been challenged by Berkowitz (1993a); Geen (2001) however Anderson and Bushman (2002) identified five causal roles that anger plays for aggression. These are: (1) anger reduces inhibitions against aggressing by either providing a justification for aggressive retaliation or by having anger interfering with higher-level cognitive processes, (2) anger allows for maintaining an aggressive intention over time, (3) anger is used as an information cue, (4) anger primes against thoughts, scripts etc., and (5) anger energizes behavior by increasing arousal levels.

5.4 Research on Online Aggression

It is of no surprise that people may become aggressive online. While most of the online aggression is through psychological means where people are not physi-
cally harmed, residual aggressive tendencies may be expressed in physical reality as well.

One recent example was the case where an adult male shot an adult female on a college campus and then committed suicide in 2009 (Taylor-Bonds, 2009). Both individuals had accounts on YouTube while the male had publicly expressed aggressive intentions in his videos for people opposing his religious beliefs (YouTube, 2010). In one of them, one can clearly see an overriding of inhibitions by victim dehumanization. He stated “[they] are not humans, they are monkeys, they are animals […] worms and dogs” (YouTube, 2010). While statements such as these are not a guarantee for the violent acts that followed, they should be a cause for concern and raise red flags.

Of course, not all online incidents lead to physical manifestations of violence. Some remain online however the psychological damage on the victims can be life damaging as well. Cyberbullying, a type of bullying “in which the aggression occurs through modern technological devices, and specifically mobile phones or the internet” (Slonje & Smith, 2008, p. 147), is a popular method for psychological aggression online. One of the most popular cases, is the one involving a 15-year-old Quebec boy that had one of his videos leaked on the internet by his classmates and has been edited and submitted in various forms by individuals from all over the world ever since (Brady & Conn, 2006; M. A. Campbell, 2005). According to media reports, the student’s family filed a lawsuit against the families of the students responsible for releasing the video on the Internet. In the lawsuit they stated that the student “had to endure, and still endures today, harassment and derision from his high-school mates and the public at large” (Popkin, 2007).

The above incidents while arguably rare, demonstrate that just as in real life, online life in social media has at times many unforeseeable dangers. This probably makes them more dangerous and their development and management becomes more problematic. The blowback from these incidents in online communities can be devastating with lawsuits waging against the companies as well as the individual perpetrators. These situations can become a PR’s nightmare. It is therefore justified that many researchers try to understand such online incidents and why they occur.

### 5.4.1 Media Violence Research

Research in media violence can expand in various types of media such as novels, comic books, television but it also applies in social media found today, whether online collaborative communities, social network sites or social game worlds. While some believe that there is a link between media and violence (Anderson et al., 2003; APA, 2004) others have expressed that such effects may have been overestimated (Ferguson, 2009; Freedman, 2002).

Media research has been influenced by the social learning theory (Bandura, 1978) and the social cognitive theory (Bandura, 2001b) which has many similarities with script theory and interconnectedness of socio-cognitive factors which can also be found in the GAM. Additional theories however were developed with their main focus on the media such as, the catalyst model and the moral panic theory.
The catalyst model has been developed by Ferguson et al. (2008) in order to become an explanatory model for media violence. However, the model itself advocates that media violence plays a small role as a casual influence to aggression. The model suggests that stressful environment circumstances and a genetic and early social influence creates a combination in which predisposed individuals may become aggressive. The theory is not well tested but it does represent modern ideas that media violence is not produced because of the medium alone but because of additional factors as well. In particular, this theory presents an interesting aspect which has not yet been thoroughly explored in research. The environmental stressful parameters have a role to play in aggression. These parameters are part of the social media software design.

On the other hand, moral panic theory explores the media violence from a historical and cyclical perspective. The theory was developed by Gauntlett (2005) not as an explanatory model for whether there is an increasing violence by media but whether the violence is perceived to exist in any novel media by individuals. Put simply, there is bias from researchers that have a tendency to try and confirm the preexisting belief that media causes aggression. In addition this pattern is not just temporary but repeats itself throughout history with every new invention classified as “new” media.

No matter which theory one may prefer over the other, I attempt to maintain a neutral position in respect to social media violence. My aim was not to infer that social media increase or decrease aggression but to show that there is an opportunity for software developers to intervene and play a more proactive role in preventing aggression through changes in the features provided on these services. By understanding the effects of media violence leading to aggression, one can sufficiently determine the ways that the software may play a key role in this process and alter it. The modifications can vary from providing parental control constraints to the software to reducing the blood and gore violence for young individuals. I argue that these are all valid considerations which should be part of project planning and development process of social media and that currently such constraints are only imposed mostly by the law (e.g., in the development of video games: “Game Experience May Change During Online Play”).

5.4.2 Online Disinhibition Effect

Another centerpiece for online research into aggression is the online disinhibition effect. The model was developed by Suler (2004c) as an explanatory theory for the reasons behind individuals acting in a disinhibited manner when online. Such behavior has been observed in many previous studies. One in particular, conducted by Siegel, Dubrovsky, Kiesler, and McGuire (1986) before the advent of the Internet, clearly demonstrated that computer-mediated groups produce more uninhibited behavior. In another paper, Kiesler, Siegel, and McGuire (1984) discusses the potential for sociological and psychological research that can lead to a better understanding for computer-mediated technologies. Many have attributed phenomena of online aggression such as internet violence, flaming, verbal attacks, self-disclosure and other to online disinhibition (A. Joinson & Gäckenbach, 1998; A. N. Joinson, 2001; Kiesler et al., 1984). Those negative behaviors were coined
by Suler (2004c) as toxic disinhibition while behaviors with no negative consequences or with even positive effects are classified as benign.

The concept behind the online disinhibition effect starts with a comparison between the offline and online world. In everyday life, social interactions are governed by unwritten rules for social conduct. These are not enforced by authorities but there are social consequences for those that do not follow them. On the Internet and therefore in social media these social consequences are reduced compared to face to face communication and therefore behaviors may differ. As an effect, people act with less restraint. Suler (2004b, 2004c) identified the following psychological aspects of online communication as causative: dissociative anonymity, invisibility, asynchronicity, solipsistic introjection, dissociative imagination, minimizing authority.

Dissociative Anonymity

Dissociative anonymity describes the condition where individuals behave online under the assumption that nobody knows who they are. This was coined by Suler (2012) as “You Don’t Know Me”. Dissociative anonymity creates a universe where identities are concealed and as such social consequences could be perceived by individuals as negligible. This factor is not necessarily negative. Under the same token one can see a positive result of being able to speak more freely and disclose information that would otherwise be too embarrassing for an individual. However, similarly one can act in an aggressive manner if they wish to do so. Disinhibited behaviors are more acceptable by individuals; they are not vulnerable to others since they cannot be identified.

Invisibility

Invisibility as a factor is coined by Suler (2012) as “You Can’t See Me”. It is described as a state where individuals are able to conceal their physical presence. This state provides individuals with the ability to conceal their physical characteristics should they wish to do so. An example for this could be an individual choosing to upload a picture which would not reveal his or her psychological characteristics, or even concealing their surname to avoid being ethnically identified. Additionally, this state also provides individuals with the ability to avoid having to be conscious about their standard social cues. In other words, in social media many of the standard social cues such as small changes in facial expression, tone of voice, aversion of eyes, nostril flaring and others may be absent in social interactions between users. In fact, many of the social media classifications (Kaplan & Haenlein, 2010, 2011) are missing many if not all of these social cues. Furthermore, while this factor is similar to dissociative anonymity, it is distinctive because even if people are not anonymous there is still the factor of invisibility that is ever-present in online communication. In fact, a recent study on online disinhibition by Lapidot-Lefler and Barak (2012) found that the most important factor between anonymity, invisibility and eye-contact which created flaming behaviors (a product of online disinhibition) was the lack of eye-contact.
5. **Online Aggression Theory and Application**

**Asynchronicity**

Asynchronicity accounts for the fact that online communication is rarely as synchronous as in real life. There is void between the exchange of messages which leads individuals to never expect immediate feedback and therefore being less afraid about their actions. Wallace (2001) has sarcastically described the web as the “World Wide Wait” (p. 113) in order to demonstrate the seriousness of the issue. While real life exchange of messages vary between 1.35 to 3.21 seconds (Zimmerman & West, 1975), online communications even in synchronous chat rooms could reach significantly higher delays. Today, most social media still prefer asynchronous means of communications for their users with text communication being arguably the most preferable. Bandwidth constrains still prevent technologies from providing synchronous communication to their users and many virtual game worlds (or virtual social worlds) still avoid it as an option. In fact, the expectations of users for online communications have been transformed to such degree that not only immediate feedback is not expected but it has become optional. Many individuals do not expect a response from others. A “See You Later” (Suler, 2012) attitude seems to have been embedded into individuals while online.

**Solipsistic Introjection**

Solipsistic introjection can be seen as the result of all previous factors presented. It involves a projection of one’s identity onto media. Simply put, lacking any kind of input (e.g., standard social cues), the human mind assigns its own traits and characteristics to another individual when interacting online. This factor acts as a form of stereotyping where by having someone looking at lines of text one can reach to certain conclusions about the author’s character and traits without necessarily those conclusions being true. As Suler (2012) further points, individuals may end up playing stories in their heads and writing as if they were authors writing their own novel without being fully aware of it. It becomes apparent, how meanings can be distorted without individuals being aware of such things.

**Dissociative Imagination**

“It’s All a Game” (Suler, 2012) is the feeling that individuals may have due to the effects of dissociative imagination. Cyberspace creates a sense of escapism from a mundane reality to a world where different rules may apply. In such a place, characters become alive in a make-believe dimension which is separate from the responsibilities and limitations of one’s reality. This is the key difference between dissociative imagination and dissociative anonymity. While in the latter identities are simplified or reduced, in dissociative imagination there is a disconnect from one’s personality and instead creativity and role-playing commence to form a new online identity in a “new” world.
Minimizing Authority

The Internet and social media tend to be looked upon by individuals as egalitarian communities. Individuals in social media have a “We’re equals” (Suler, 2012) state of mind. This in turn, leads to them expressing opinions and ideas to people with otherwise intimidating authority. Everyone regardless of their status, wealth, gender or race is on a level playing field in social media. As Suler (2012) points out, since the Internet was engineered with no centralized control, ideas about online communities of equal individuals became the core of the Internet experience.

5.4.3 Benign & Toxic Online Disinhibition Effects

As mentioned previously, the online disinhibition effect does not necessarily have online negative effects for online individuals. This is a pattern that can be observed for many social and psychological phenomena where their existence seems to be agnostic towards a group; in a non-deterministic setting they cannot be identified as neither good nor bad. This point is important for understanding how HCI research should behave when examining research phenomena but also for software developers that need to develop software in order to affect social and psychological behaviors.

Suler (2004c) provided a dichotomy for the online disinhibition effect, with effects being described as either benign or toxic. Benign effects of disinhibition do not have negative consequences or can even prove to have a positive effect in social interactions. On the other hand, toxic instances are those perceived to be harmful to participants involved.

Examples of benign effects can be seen in online collaborative projects (e.g., wikis, GDSS, online support systems) where invisibility along with anonymity provides a “complete” anonymity. Participants in online alcoholic anonymous groups are enjoying these effects. Another popular case is virtual game worlds or virtual social worlds in which individuals can select characters that have any qualities they desire. Such effects can help individuals that could be described as introverts to express their true selves online (Sheeks & Birchmeier, 2007).

On the other hand, toxic effects are a serious cause for concern. Online collaborative projects can also present cases where one can find toxic disinhibition such as environments for educational purposes or the workplace. An individual acting in a disinhibited manner may be facing detention or lose his or her job. In fact, toxic disinhibition effects can be found in all the spectrum of the social media world today. Content communities are suffering from the effects of cyber-bullying and flaming just as much as virtual game or social worlds do. The online nature of the environment for social interaction that is provided by social media seems to amplify disinhibition which can potentially lead to aggression.

5.5 Preventing Online Aggression

Most of the explanatory theories that were developed to explain offline and online aggression are powerful at displaying the causal influences that lead to ag-
gression but are less effective in demonstrating ways for preventing it, specifically in online settings. Their usefulness is as such isolated to fields of sociology and psychology. However, these theories hold tremendous applicative potential for other scientific fields such as HCI.

I argue that there is a need for a new perspective. Most strategies put the individual at the center of interest for investigation. While this is definitely efficient for gaining understanding, it is of little use for practical applications without taking into account that this is a human-computer symbiosis as described by Licklider (1960). Therefore, it is unwise to assert that factors of individuals are the cause for behaviors when computers are an intricate part of the system. Theories on aggression have successfully identified situational and environmental factors playing a key role in affecting aggressive behavior but did not establish preventive measures. One probable explanation for this is mainly the observational nature of sociology and psychology, as well as the general lack of interest by HCI to design social interaction. However, this lack of interest leads to a faulty thinking pattern.

For example, one may expect that familiarity with technologies and online communication will reduce the online disinhibition effect. This strategy is deeply engraved into computer science philosophy. However, it is problematic. Just as Bush (1945) expected that people may develop and attain a language in order to fit the technology’s needs, similarly this expectation develops the idea that people will become familiar with technologies in order to work in accordance with a designer’s expected online behavior in social media; that is, avoiding being aggressive. Hence, behavior is expected to be forced on individuals through familiarization or even by law. Such a way of thinking, takes the responsibility away from software designers and puts the blame on the users. Evidence today, shows that this is truly the case. A fallacy has been created where, while software designers are not concerned with user behaviors, users seem to have tremendous issues with the software that in turn affects their behavior and the way they interact with others online.

In fact, the issue is such that Wallace (2001) has named one of her sections in the chapter describing flaming and fighting as “THE #$@@!!!& SOFTWARE” (p. 127) in order to demonstrate that software is a factor in the aggression that can be seen today. If one looks at aggression through a lens where a human-computer interaction becomes the centerpiece, then part of the environmental factors are transformed into the software itself. Rather than expecting individuals to familiarize themselves with the software and to behave in a desirable manner, the software can form guidelines that enforce behavior directly or indirectly. The outcome from this revelation, is that aggression becomes a side-effect of social interaction that can be inhibited and additionally social media can benefit from having communities that will display less toxic disinhibition effects and lesser aggression in general. Such approach is also endorsed by Suler (2004c) which stated that “The self does not exist separate from the environment in which that self is expressed” (p. 325). He also argued for the potential benefits of understanding environmental effects in different online settings and for designing software by having such knowledge (Suler, 2000).
5.5.1 Software Features Studies for Aggression

I demonstrate using HCI research methods, how software can affect aggressive behavior and how one can identify certain elements of the software and alter them to rid social media of such a negative behavior. The basis for identifying social media features with potential for affecting aggression is based on current psychological and sociological literature while merging it with the current practices for social media software development. My main aim was that the features investigated be broad enough so that results can be generalized and be applied in most if not all social media classifications. Below I itemize a list of features that were explored for which findings are presented on the next chapter.

- Different anonymity states produce different effects of online disinhibition and in turn aggression.
- When online messages are lost due to submission errors, there is a loss in the quality of messages which can indirectly lead to aggression through frustration or textual misunderstandings.
- The quality of error messages and suggestion of alternatives in them can reduce the frustration caused by the incident and therefore reduce aggression.
CHAPTER 6

RESEARCHING THE INDIVIDUAL FEATURES FOR AGGRESSION IN SOCIAL MEDIA

6.1 Introduction

In this chapter, I present the research findings for the features that were explored in relation to aggression. Each feature examination was given a separate section since most literature involved for each feature as well as its method may not apply to another. Additionally, some of the work has been published in journals and conferences and as such, summaries of the background, processes and findings are presented here as well. Finally, while individual conclusions for each section are presented here, they are all combined later on the chapter describing the social interaction design framework.

6.2 Anonymity States and Aggression

For this part of the study, I wanted to investigate the effects of anonymity states and aggression. Since the online disinhibition effect theory directly described the effects of anonymity as a factor for disinhibited behavior, it became an essential part of my investigation. Keeping in mind that the research results are required to always have had direct applications in the software design and development process, I decided to investigate how anonymity states may contribute to aggressive behavior. Establishing that there is an effect will directly give control to designers over how much disinhibited behavior may be expected in their social media services. While the choice of providing an anonymity state is an important decision when designing social media, it has never been investigated in terms of the effect that it could have on aggressive behavior for the users.

This study has been part of an investigation that resulted in a conference publication (Tsikerdekis, 2011) as well as a published paper on a scientific journal (Tsikerdekis, 2012). Both papers describe in detail my research study involving anonymity states and aggression. While specific details can be found on these papers, a summary is also presented here. Additional statistical analyses for the original hypotheses are applied in order to provide a better perspective for the
findings. Examining results from multiple perspectives could give a better insight into the overall processes involved. Furthermore, since the publication of the original paper, there was a similar study involving anonymity and online disinhibition (Lapidot-Lefler & Barak, 2012) that is worth discussing along with the results obtained from the work involved in this dissertation work. A brief theoretical introduction follows.

6.2.1 Background Theory

The theoretical background of the study revolves around two anonymity states identified by computer science literature (Pfitzmann & Hansen, 2010). While more anonymity states exist in literature, I argued that these cannot be easily identified by individuals online (Tsikerdekis, 2011, 2012). The states that can be identified by users are pseudonymity and complete anonymity. Additionally, I was interested in the state where participants used their real names which acted as my control state. Studies on anonymity have confirmed a link with aggression (A. N. Joinson, 2006; Suler, 2004a, 2012; Wallace, 2001). However, none at the time produced research that investigated if there is a difference between these anonymity states that would provide designers with control over aggression.

Anonymity is associated with positive (Ainsworth et al., 2011; Lea, Spears, & de Groot, 2001; G. B. Lee, 1996; Tanis & Postmes, 2007) as well as negative effects (Davenport, 2002; Farkas, Ziegler, Meretei, & Lörincz, 2002; Keesom, 2004; Rains, 2007; Wallace, 2001) for users in social media. It is identified as an important part of the online disinhibition effect in which dissociative anonymity, invisibility, solipsistic introjection and dissociative imagination affect the social interactions of individuals through anonymity (Suler, 2012). In my study, my main focus revolved around dissociative anonymity and dissociative pseudonymity (Tsikerdekis, 2011, 2012). I argued, that while dissociative anonymity can be a factor for the case of complete anonymity as well as pseudonymity, the same cannot be said about dissociative imagination. Dissociative imagination is a factor that involves more creativity and as such it is more likely to play a significant role in the case of pseudonymity rather than complete anonymity. Furthermore, if that is true, then there may be a difference in the disinhibition that individuals would exhibit between the states of pseudonymity and complete anonymity.

6.2.2 Method

In order to assess this hypothesis, I conducted a survey on Facebook, a social networking website in which respondents faced different situations in controversial topics. These contained problems that in turn respondents were required to respond to with one of the predetermined set of written messages. The research was an experimental design survey with three treatments, real names, pseudonymity, and complete anonymity. The messages consisted of four-point ordinal scale ranging from messages that were very polite, to very rude messages. This way standardization was achieved between all the responses which increased internal validity. Additionally, under the suspicion that the opinion on a topic may affect the outcome, participants were asked prior to facing the scenar-
ios, their opinion about each topic on a three-point ordinal scale. This additional variable would help determine if the strength of one’s opinion could play a significant role. The survey was administered in a form of a Facebook application and a volunteer/snowball sample was obtained through advertisement in academic and non-academic groups and fan pages on the social networking service. A more detailed description of the survey design and process can be found on the research article (Tsikerdekis, 2012).

6.2.3 Findings and Results

According to the analysis of data in the study, anonymity states were found to have an effect on the way individuals reacted but only conditionally to the self-reported level of opinion on a topic. This can be seen on figure 6.1. Under conditions where individuals stated that their opinion about the topic was “extremely strong” their behavior was significantly different and more disinhibited. However, that effect was found only for the condition of pseudonymity while analysis for complete anonymity did not produce statistically significant differences compared to the control state as well as pseudonymity. The finding that pseudonymity produced more aggressive responses was in accordance with the literature on online disinhibition.

The effect of an anonymity state in the way a participant would respond was also found to be greatly dependent on the scenarios that were used. The scenario involving animal rights showed a great difference compared to the other two scenarios used. Figure 6.2 depicts these differences.

The conclusions in the original paper rendered complete anonymity as a safer choice for software designers and developers that want to provide a form of anonymity for their users while wanting to avoid negative behaviors that may come with it. On the other hand, pseudonymity was not recommended for social media services under which discussion of highly controversial topics is expected to take place since this can potentially increase disinhibited behavior and in turn the chances for aggression.

6.2.4 New Hypotheses for the Additional Analysis

Several statistical analyses from the original study are worthy of examining based on a Bayesian framework. This can potentially help in reassessing some of the results as well as help determine if null hypotheses were more likely for results that did not achieve statistical significance. The following hypotheses were formed for this re-evaluation:

- $H_1$: Responses are dependent on the anonymity state through which individuals respond.
- $H_2$: There is a dependency between importance of a topic and the responses.
- $H_3$: There is a dependency between scenarios and the responses.

When examining the data as a repeated-measures design across all anonymity states for all scenarios and self-reported importance of a topic, a Friedman’s test
6. RESEARCHING THE INDIVIDUAL FEATURES FOR AGGRESSION...

Figure 6.1: 3D clustered graph depicting the responses that individuals gave under each anonymity state and their reported strength of opinion on a topic. Responses under the state of pseudonymity (here depicted with a light brown color) show a substantial effect failed to show statistical significance, $\chi^2(2) = 4.671, N = 160, p > .05$. This failed to provide support for $H_1$. However, the same data can be reexamined under a Bayesian network.

A contingency table was created using the three anonymity states and the four responses of individuals. This can be seen on table 6.1. Variables lose their ordinal nature and become categorical, however, they can still provide evidence on whether a dependency exists. $H_1$ becomes the hypothesis of dependence and $H_0$ becomes the hypothesis of independence between the two variables. Using the `ctable` function from the `LearnBayes` package for R, the Bayes factor that was produced provided support for the hypothesis of independence, $BF_{01} = 138.06$. This is considered decisive evidence that anonymity states have no effect on the answers of the respondents.

However, in this result there is unwanted variance due to other variables as it was found in the original paper. A small effect was found between the answers and the importance of a topic, $r_s = .142, p < .01$. A second dependency was found between the scenarios and the importance of a topic, $\chi^2(4) = 30.371, p < .001, V = .177$. Looking for the standardized residuals within each cell, the scenarios about abortion and animal rights had some cells with residuals higher than 1.96. Most of the respondents rated the scenario about the animal
6. RESEARCHING THE INDIVIDUAL FEATURES FOR AGGRESSION...

Figure 6.2: Responses based on anonymity states for each one of the three scenarios. Each left column represents complete anonymity, each middle column represents using real names and each right column represents pseudonymity.

rights higher in the scale of importance while the opposite occurred in their responses. Additionally, I conducted an analysis to look for a dependency between responses and scenarios which showed a statistically significant result as well, $\chi^2(6) = 56.737, p < .001, V = .242$. In general, the NHST analysis provided evidence for both $H_2$ and $H_3$.

Using Bayesian analysis the results became harder to interpret. After forming a contingency table based on the variables, Bayes factors were obtained through an analysis using the `ctable`. The test between the answers and the importance of a topic produced a Bayes Factor in favor of the null hypothesis, $BF_{01} = 1.32$. Given that the result is so close to one, it becomes harder to interpret. However, when testing for a model “close to independence” (here denoted as $H_c$ for clarity) using 100,000 simulations and a Dirichlet parameter $K$ between 2 and 7, the `bfindep` function produced the highest Bayes factor in favor of the model “close to independence” $BF_{c0} = 12.76$. This result shows that if one is willing to accept that
there are uncertainties in data, then a dependency between these two variables is unlikely. This may seem different than the result obtained by the NHST analysis but this is not the case if one takes into account the small effect that was produced by Spearman’s correlation analysis. On the other hand, if we consider that the importance of a topic was found to be dependent on the scenarios then a model close to independence is unlikely. In fact the Bayesian analysis using the ctable function produced a Bayes Factor in favor of the dependency between the importance of the topic and scenario, $BF_{10} = 11.245$. Strong evidence was also found for a dependency between the responses and scenarios, $BF_{10} = 391.556,237$. The Bayesian analysis while providing support for $H_3$ had showed little evidence for $H_2$.

While a dependency between responses and strength of opinion was established, there was a need to evaluate this in relation to anonymity states. In my original paper I proceeded by conducting correlation tests for the importance of a topic and responses within each anonymity state. I was able to trace the degree through which the importance of a topic affected each anonymity state. As the importance of a topic would rise so would the response become ruder. I conducted 1-tailed tests that produced non-significant results for the state where participants used their real names, $r_s = .068, p > .05$, and for the completely anonymous state, $r_s = .083, p > .05$. However, the state of pseudonymity produced a statistically significant result, $r_s = .288, p < .001$. This analysis was similar to the one conducted on the original paper (Tsikerdekis, 2012).

I wanted however to reevaluate these results using the Bayesian framework since I wanted power to assert conclusively that the states where participants used real names and complete anonymity, were independent of the self-reported opinion about a topic. The Bayes factors obtained using the ctable function for the states of real names, pseudonymity and complete anonymity were $BF_{01} = 11.20$, $BF_{10} = 8.04, BF_{01} = 48.83$, respectively. The results seem to agree with NHST analysis. Both, the state where respondents used their real names as well as the state that they were completely anonymous support the hypotheses of independence from the importance of a topic. On the other hand, under the state of pseudonymity, responses were dependent on the self-reported importance of a topic.

<table>
<thead>
<tr>
<th>Response Anonymity State</th>
<th>Complete Anonymity</th>
<th>Real names</th>
<th>Pseudonyms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very polite</td>
<td>82</td>
<td>62</td>
<td>67</td>
</tr>
<tr>
<td>Polite</td>
<td>47</td>
<td>53</td>
<td>49</td>
</tr>
<tr>
<td>Rude</td>
<td>15</td>
<td>23</td>
<td>29</td>
</tr>
<tr>
<td>Very Rude</td>
<td>20</td>
<td>24</td>
<td>17</td>
</tr>
</tbody>
</table>

Table 6.1: Contingency table for anonymity states and responses for all scenarios
6.2.5 Discussion

The reexamination of the original results using Bayesian analysis seemed to provide evidence to support that indeed pseudonymity as a state is different than the other two states in the study. Additionally, the level of importance did not affect how participants acted under complete anonymity or when using real names.

Additional remarks were also illuminating. One of them was the powerful dependence that was found between the importance of a topic and the scenario, as well as for the effect that scenarios had on the responses of the participants. In both cases the NHST analysis showed relatively small effects while the Bayesian analysis was definitely more favorable of dependencies.

Another major difference was found when examining a dependency between importance of a topic and responses across all anonymity states. While the NHST analysis indicated that there was a dependency, the Bayesian analysis showed that a model “close to independence” is also highly likely. However, further examination of this dependency within each anonymity state revealed that this was not the case. The dependency was not due to unsystematic variance but was caused because of the differences between pseudonymity and the other two states. Pseudonymity seems to be enhancing the dependence between the importance of a topic and the way individuals respond; a probable cause of dissociative imagination having a larger effect as speculated at the beginning of the study (Tsikerdekis, 2011).

Furthermore, the results from this study may seem to contradict the study conducted recently by Lapidot-Lefler and Barak (2012). In their study strong evidence was not found between participants that used their real names along with other personally identifiable information and their pseudonyms. However, there are two important differences between the two studies. In their study, Lapidot-Lefler and Barak (2012) automatically assigned pseudonyms to its participants, while this study requested participants to assign their own pseudonyms a priori. This may indicate that self-assigned pseudonyms have more power to produce toxic disinhibition than automatically assigned pseudonyms. Furthermore, there was no measurement of a self-reported strength of opinion on a topic which was an essential link that helped discover the effect that it had on pseudonymity.

Considering all the above and given the new results using the Bayesian analysis, a revised causal model can be created for the variables measured. Figure 6.3 depicts the model.

The following causal model is still conditioned only for the state of pseudonymity where an effect was found. According to the results, the states where respondents used their real names and completely anonymity produced similar results and did not affect responses.

Recommendations

As I described in the original paper, given the evidence, the decision for providing users with the option of pseudonymity in social media is bound to have negative consequences. In general, it is expected that it will raise the levels of aggression in the community. Unfortunately, for certain classifications of social
media, this is the only option. However, for other classifications such as online collaborative projects, software designers are not limited to just pseudonymity. A collaborative wiki can have the choice of real names, complete anonymity or even both with the latter provided as an option at any time a user wishes send an anonymous message. Since complete anonymity was not found to have the same effects as pseudonymity, it can be considered as a safer alternative and therefore can provide designers with control over aggressive behavior on social media.

6.3 **Quality of Error Messages and Suggestion of Alternatives**

One of the features suspected of contributing to aggression is the design of error messages. In social media software today, software errors can happen frequently and users are confronted with all sorts of unknown situations which they are required to resolve. To add to the confusion, a limited amount of time for finding the information that a user truly seeks could make matters worse. The final outcome of such interaction results in an increased frustration for individuals that could lead aggression. The reasons for this outcome can be found in neoassociation theory (Berkowitz, 1989, 1990, 1993a) as well as the frustration-aggression theory originally developed by Dollard et al. (1939).

This part of the study investigated if the quality of error messages as well as the suggestion of alternatives can potentially improve user perceptions about such unfortunate incidents and achieve higher levels for comprehension, helpfulness and appeal. These features can potentially reduce the negative effects of an error message. In this part of the study, I administered a survey which aimed to investigate the potential for error messages with better quality of text, graphics and suggestion of alternatives to reduce frustration on users.

6.3.1 **Error Messages and Frustration**

Nielsen (1994a, 1994b) developed several heuristic guidelines to be followed by software designers and developers. One of the most important aside from the avoidance of errors was the use of error messages that can be helpful and be
understood by the users. In particular, a message should help users recognize the problem, diagnose and recover from errors. Similarly, in the extended version of Nielsen’s heuristics, Ryu (2007) suggested that each and every error message should offer at least one solution or a link to a solution in the error page.

However, evidently these heuristics are not followed in many cases. In their observation study on the causes for user frustration during their work with computers, Ceaparu, Lazar, Bessiere, Robinson, and Shneiderman (2004) found that many of the causes were attributed to a clear violation of the guidelines discussed previously. In addition, such frustrated events also reduced the time spent when performing actual work on a computer. Error messages are among the chief causes for end-user frustration.

While in the isolated universe of yesterday, frustration and even aggression was contained to local distances, today technical issues that do not prompt for a clear path for resolution may be more costly. Hara and Kling (1999) conducted a qualitative case study of a Web-based distance education course. They discovered that some students had severe issues with error messages and they lost a lot of time trying to resolve them. One of the participants clearly linked their frustration with the anger that they felt. Their conclusion was that issues that can cause frustration inhibit opportunities for education.

In today’s complex system such issues do not just haunt the users but also system administrators. Maglio and Kandogan (2004) described two cases under which system administrators were frustrated with resolving error messages. Many errors were completely disconnected from the original causes and the error messages were not helpful. Their recommendation was to try and display messages that can be interpreted within the context of which they occur.

In social media design, all of the cases described above should be of great cause for concern. They are not just technologically complex systems but also include a social ecosystem which is fragile to events that can cause frustration and aggression. The degree where a message can be understood by a user, the aesthetics of the unfortunate incident and the proposition of alternative paths are all important factors to be considered. Are such heuristic guidelines going to have a significant impact on these factors and in turn help reduce frustration levels for the users? This research question helped formulate the three main hypotheses for this part of the study:

- $H_1$: Messages that explain the problem and provide more relevant information will result in an increased comprehension.
- $H_2$: Messages that provide alternative paths and solutions will be rated as more helpful.
- $H_3$: Messages that have better graphics will be more appealing to users.

### 6.3.2 Method

In order to assert these hypotheses I developed a set of questions which were included in a larger survey that was administered for the purposes of this dissertation work. This survey will hereby be referred to as AG1 since parts will be
used for asserting other features studied in this dissertation later on. The segment consisted of two different versions of error messages that users encountered in a hypothetical scenario. Users were randomly assigned to groups. For the purposes of realism, I decided to use already existing error messages which coincidentally were developed by the same company for two different products. These are presented side by side on figure 6.4.

![Error Message 1](image1)

(a) First version

![Error Message 2](image2)

(b) Second version

Figure 6.4: The two versions of 404 errors that respondents encountered.

The first error message uses code to state the problem and is not specific in its first statement with the type of error. Additionally, the final statement may be considered irrelevant with the problem. However, the image may add to the aesthetics and appeal and therefore the message may be seen as less dramatic to the user. On the other hand, the second version does not use code to state the error and makes a clear statement about the type of error. Additionally, there are clear suggestions and alternative paths that a user can take so that the problem can be resolved.

Users were asked a set of questions that aimed to measure comprehension, helpfulness and appeal of the messages. These questions can be seen on figure 6.5. Additionally, a qualitative question was added to the study in an attempt to uncover the reasons behind user choices.

The survey was launched on Facebook as an application, and advertised in several academic and non-academic pages and groups. Additionally a snowball sampling method was used throughout the survey allowing participants to invite their friends to take part. The survey was originally designed in English but later
6. RESEARCHING THE INDIVIDUAL FEATURES FOR AGGRESSION...

![Figure 6.5: Questions included in the survey investigating the effectiveness of error messages.](image)

Figure 6.5: Questions included in the survey investigating the effectiveness of error messages.

was translated in Greek as well as Spanish. The survey took place between May 1, 2012 and August 21, 2012.

6.3.3 Results

A total of 206 individuals visited the application that hosted the survey of which 157 successfully completed it. The demographics of the sample are presented on figure 6.6. The sample consisted of a young and middle-age population with a limited however population for the extremities. There was balance between sexes and the majority of people came from Europe (72%), although other continents also made significant contributions to the sample. There were no respondents that came from Africa in the sample.

Figure 6.7 depicts the results for the three variables. Comprehension was higher for second version of the error message ($M = 7.64, SD = 2.573$) than the first version ($M = 6.35, SD = 3.386$). Helpfulness also had a higher rating for the second version of the error message ($M = 6.49, SD = 2.873$) than the first version ($M = 2.87, SD = 2.569$). Smaller differences were found for the aesthetics variable between the second ($M = 5.09, SD = 3.003$) and the first version of error messages ($M = 4.96, SD = 2.766$).

NHST analysis was used in order to assess the hypotheses for this study.
While the ten-point scale variables could have been used as interval variables, violations of normality assumptions made it impossible to be used in parametric tests. Instead, Mann-Whitney U tests were used to determine if the groups reported significant differences for the variables of comprehension, helpfulness and appeal. The analysis showed significant differences for comprehension ($U = 2424.5$, $Z = -2.344, p < 0.05, r = .187$) as well as for helpfulness ($U = 1088$, $Z = -7.099, p < 0.001, r = .566$). These results provide support for $H_1$ and $H_2$. On the other hand, results for the appeal did show any statistically significant differences ($U = 3024$, $Z = -.202, p > 0.05, r = .016$).

Qualitative analysis

Findings from the qualitative analysis are in accordance with the quantitative analysis. The differences between the error messages are apparent in the comments of participants especially for the first error message. One respondent that received the first version of the error reported that "the message should include
a solution so I can fix the error.” Another respondent specifically reaffirmed hypotheses $H_1$ and $H_2$: “Not at all useful. Didn’t explain what the problem actually is and didn’t tell about the solution as well.” Evidently the need for solutions is extremely high among these respondents. The language used in the text also went under scrutiny as this respondent demonstrates: “It’s the ‘That’s all we know’ that sounds uncaring and makes me give up hope.” While users giving up hope because of these error messages could decrease productivity, communication and interaction on social media, frustration is a more severe consequence that could be linked to aggression in a community. A respondent clearly demonstrated this link between a faulty design of error messages and frustration by stating that it is “frustrating. If an [sic] URL is not available, the link for that URL must be taken down.” Finally, one can find even evidence for the non-significant result regarding the aesthetics of the first error message: “The picture and message is not user understandable. They [sic] can change into the website you are looking is not available not a message 404.”

On the other hand, while quantitative results seem to have favored the second error message, respondents did not seem to do so in their comments. For example, one respondent requested additional solutions for resolving a problem: “The message could indicate other ways to resolve the issue. One can use the Way Back machine to view a broken link in hopes of discovering its non-existence any longer [sic].” Another respondent stated that “The suggestions are not helpful. I generally know to try a different approach in the search engine. The only helpful information is that the link is broken.” Evidently, computer and software efficacy is an important factor that determines the usefulness of a message. Users with

---

**Figure 6.7:** Results for the three aspects of the two versions of error messages evaluated by respondents.

---

6. RESEARCHING THE INDIVIDUAL FEATURES FOR AGGRESSION...
more advanced skills will probably require more solutions. The same statement was also made by another respondent: “I think there’s something wrong with the website. Normally, if the website isn’t a first priority, I always quit and try to search other websites though there’s [sic] suggestions below.”

Finally, qualitative results contained some mixed messages about aesthetics. As an example, a respondent viewing the first error message said “cute but not helpful,” indicating that aesthetics actually mattered in this case. Similarly, another respondent viewing the second error message dismissed the solutions as not helpful and made a request for better aesthetics: “give a funny face will make me feel better.” While quantitative results did not achieve significance for the aesthetics of the message, it may be the case that an effect may exist but is probably small. However, these cases demonstrated that for some individuals, aesthetics matter even for error messages.

6.3.4 Discussion

In general, results from this study seem to reaffirm heuristics that are widely used as guidelines in human-computer interaction. More importantly however, they demonstrate that such differences in the comprehension of the text as well as the helpfulness of error messages statistically matter. In turn these features have the power to decrease negative effects such as frustration which could be the cause for aggressive behavior in social media. Aggression in social groups can be contiguous (Cairns, Cairns, Neckerman, Gest, & Gariepy, 1988; N. E. Goldstein, Arnold, Rosenberg, Stowe, & Ortiz, 2001) which could contribute towards developing flame wars online (Wallace, 2001). As such these features are not helpful only for improving user experience on social media software but also social experience.

An important find for this study is that users seem to have higher expectations than what heuristics provide. Even in the case of the second error message where alternative solutions were provided, respondents still seem to be craving for more and more radical solutions. While providing the obvious solutions will lead to an increase in helpfulness, an effort to develop solutions that go beyond the current expectations of users may be appreciated by individuals.

Additionally, while aesthetics may still play a role in error messages, strong evidence was not found for the $H_3$ hypothesis. However, some individuals seem to have an artistic sensitivity for aesthetic details according to the qualitative results.

Recommendation

It is recommended that designers and developers will avoid the use of codes, and cryptic language for error messages. Additionally, text that may not be considered necessary or related to the error messages should be omitted from the message. Error messages should clearly communicate in non-technical terms the problem and provide solutions. Solutions need not be limited to the bare obvious but attempt to provide more innovative ways for users to work around a problem. Finally, while evidence for aesthetics was not found in this study, designers
and developers should make an effort to design messages that would achieve near-universal appeal. Changes such as the ones described here can significantly decrease the potential for frustration by users and in turn aggressive behavior in social media communities.

Limitations

This study has used a mixed-methods experimental design survey to assert that the design of error messages can help decrease frustration and the potential aggression. However, this is still an indirect measurement for establishing that error message could contribute to aggression. The study relied on self-reported measures of helpfulness and comprehension to determine their effect on frustration. Additional future studies may wish to explore scenarios where aggression can actually be observed shortly after individuals face an error message.

6.4 Message Submission Errors and Recovery

Another important aspect to investigate in relation to aggression was incidents of submission errors. Packet loss is nothing new to computer science and even with the sophisticated degree of hardware and software occasionally information is lost. In many occurrences this information is part of messages being exchanged between users in social media. While the event by itself is unfortunate and definitely contributes to user frustration, I wanted to investigate the recovery events followed by users after the unfortunate incident. My main concern revolved around what was the action taken by the user since the original message was lost. Was there a rewritten message being submitted or did users abandon their effort to send the message completely? If the user decided to rewrite the message then a loss in the quality of the message could result in a less effective communication. The latter is highly likely if the maxims of the cooperative principle (Grice, 1989) are not satisfied with the rewritten message.

6.4.1 Messages, Communication and Aggression

Messages and the exchange of messages online is the subject of study of computer-mediated communication. Computer-mediated communication is defined as any “communicative transaction that takes place by way of a computer, whether online or offline” (McQuail, 2010, p. 552). The field of study investigates the effects of such communication on individuals, groups of individuals and the society at large.

Gricean maxims (Grice, 1989) or as else known the cooperative principle, have been of interest for studies in computer-mediated communication (Herring, 1999; Nastri, Peña, & Hancock, 2006; C. B. Smith, McLaughlin, & Osborne, 1997). Grice (1989) described maxims that rather than guidelines are seen as conditions for successful communication. There are four maxims: maxim of quality, maxim of quantity, maxim of relation and maxim of manner. The maxim of quality requires from individuals to be truthful, while the maxim of quantity requires
from individuals to be specific with the amount of information that they share (no more and no less than what is required of them). The maxim of relation requires for individuals to maintain relevance to the topic. Finally, the maxim of manner requires for individuals to be clear such as avoid ambiguity, obscurity of expression and be brief. These are mutually assumed from the speaker and the listener but whenever they are not followed implicatures could occur. These described meanings which are not explicitly conveyed in the actual conversation. In turn, these implicatures have a great potential for misunderstandings occurring in communication which not only affects the efficiency of communication but also can be the cause for conflict between individuals in online communities.

The effects and attempts of individuals to be in accordance with these maxims in order to successfully communicate to other their position can be seen widely on the Internet. For example, Nastri et al. (2006) studied the abbreviations used in away messages by users and argued that these were made by users in an attempt to be in accordance with the maxims of quantity and relevance. Such text abbreviations can be seen in all sorts of social media from collaborative projects for the workplace, classrooms but also video games (Peña & Hancock, 2006). Additionally, non-adherence to Gricean maxims, especially the one relating to the maxim of quality could violate social norms within a group which could become the start of flame wars (C. B. Smith et al., 1997). Herring (1999) found that in many cases of online exchange of messages, the maxim of relevance was violated which led the invitation of jokes and humorous play. While in this case Herring (1999) argued that this was a positive effect that explained the popularity of computer-mediated communication in many cases the development of implicatures could influence a conversation in a negative manner.

Several hypotheses were developed to establish if maxims have been violated and under which conditions. Specifically the maxims of interest were related to quantity and quality which was likely to be affected by a potential rewrite of the original message. Messages with shorter amounts of information or of less quality, could lead to implicatures. It was expected that two important factors, the time and the importance of the message, will affect the action that will be taken by the user and will also affect the quality of the message. Additionally, the action taken by the user will probably have an effect on the quality of the message. Based on the above, I itemize below the main hypotheses:

- $H_1$: The condition of somebody being pressed for time will affect the action that he or she will take.
- $H_2$: The importance of the message will affect the action that he or she will take.
- $H_3$: The importance of the message to be sent will be dependent on the whether someone is pressed for time.
- $H_4$: The quality of message that will be sent again will be dependent on the importance of the message.
- $H_5$: The quality of message that will be sent again will be dependent on whether someone is pressed for time.
6. RESEARCHING THE INDIVIDUAL FEATURES FOR AGGRESSION...

- $H_0$: The quality of message that will be sent again will be dependent on the action taken.

6.4.2 Method

A survey was used to assess the following hypotheses and was administered as part of the AG1 survey (see section 6.3.2). The method used for assessing these hypotheses relied on the respondent’s memory to report events where a message submission error had occurred in their past experience online. It was anticipated that this method will sufficiently be able to demonstrate the relation of external factors such as time and importance of a message which would have been difficult to be replicated in laboratory conditions.

The survey consisted of a set of five questions which are depicted on figure 6.8. The primary question asked respondents on whether they have experienced such an incident in their past. Those that answered negative did not answer the rest of the question. The second and third question aimed to develop variables for time pressure and importance of the message. The fourth and fifth questions aimed to see what action the respondents took and what the quality of the second message that they sent again was (in case there was a second message).

Please answer the following questions:

1. Have you ever written a message online and after submitting your message an error appears on the page informing you that your message was lost? (e.g. you submit a message, you get an error message on the page and then you realize that the message you have been writing is lost)
   - Yes
   - No

2. How important to you was that specific message that was lost?
   Not important at all 1 2 3 4 5 6 7 8 9 10 Extremely Important

3. Were you pressed for time to send the message?
   - Yes
   - No

4. If yes what were the implications of the above?
   - Rewrote the message almost exactly as it was and sent it again.
   - Wrote a shorter version of the message.
   - Abandoned writing the message altogether.

5. In case you wrote a new message, what did you think of the quality of your new message?
   - Much better quality
   - Better quality
   - About the same quality
   - Lesser quality
   - Much lesser quality

Figure 6.8: Survey questions for the message submission error and recovery survey.
6.4.3 Results

Demographic statistics about the consistency of the sample for the AG1 survey are presented in figure 6.6. The total number of survey respondents that went through the whole survey was 157. This was the initially used sample for the analysis. 107 out of the 157 (68.15%) of the respondents reported that they had experienced an event in their past where a message was lost due to submission errors. This sub-sample was further used for the subsequent analyses.

In general, results showed that 48.6% of the respondents reported that they chose to rewrite the message in the exact length, 30.8% rewrote a shorter message and 5.6% skipped writing the message altogether. This result is indicative of a probable violation for the maxim of quantity. There was also 15% of individuals that skipped answering this question and their answers were omitted for subsequent tests.

Most respondents reported that their message quality remained about the same (45.8%). Sixteen-point-eight percent rated that their messages had a higher quality, while 28.9% rated that their message was of lesser quality. This is another indication for a probable violation of the maxim of quality. While participants may have been truthful in their second message, omitting important arguments and information may have been part of the process that led to a perceived lower quality message. An additional 8.4% of participants skipped answering this question and were omitted for the subsequent statistical analyses.

A chi-square test was used to assess the $H_1$. A statistically significant effect was not found for the variables of time and action taken by the respondents, $\chi^2(2) = 2.773, p > .05, V = .175$. However, due to violations of assumptions for the chi-square test, a Fisher’s exact test was used to obtain an accurate significance value. The result remained non-significant, $p > .05$. Using a Bayesian analysis with the ctable function I obtained a Bayes factor that favored the null-hypothesis, $BF_{01} = 2.58$. While both results have produced evidence for rejecting $H_1$, the evidence was not strong.

In order to assess $H_2$, I converted the ten-point self-reported scale for the importance of the message to a five-point scale and developed a contingency table for the action variable and the newly created variable. The chi-square test showed no statistically significant dependency $\chi^2(8) = 9.445, p > .05, V = .228$. However, assumptions for chi-square were violated and Monte Carlo algorithm was used with confidence level at 95% and 100,000 samples which produced a non-significant value, $p > .05$. To assess the likelihood of a null hypothesis against $H_2$, Bayesian analysis using ctable produced a Bayes factor favoring the null hypothesis, $BF_{02} = 28.34$. This provided strong evidence that the importance of the message does not affect the action that will be taken in the occurrence of a submission error.

A similar procedure was followed for testing the $H_3$ hypothesis for the variable of time and the five-point scale for the importance of a message. A chi-square test found a statistically significant dependency between the two variables, $\chi^2(4) = 13.403, p < .01, V = .356$. A strong Bayes factor was also produced in support of $H_3$, $BF_{30} = 26.74$. While it is difficult to determine causality between the two variables, this is considered strong evidence that whenever the
message was important, respondents had time at their disposal. Figure 6.9 depicts the dependency for the two variables.

![Bar chart depicting the ten-point scale for importance of a message and time availability.](image)

Figure 6.9: Bar chart depicting the ten-point scale for importance of a message and time availability.

A Spearman’s correlation for the variables of the quality of message and the ten-point scale for the importance of message produced a statistically significant result, $r_s = -0.285, p < 0.01$. The result indicated that as the importance of the message got higher, the message quality lessened. Therefore, $H_4$ was supported but in a highly concerning way.

Furthermore, time was found to have a significant effect on the quality of the message, $\chi^2(4) = 12.668, p < .05, V = .360$. Due to violations of assumptions I used Monte Carlo algorithms (95% confidence level using 100,000 samples) to obtain a significance value which reaffirmed the initial result, $p < .01$. This was further reassessed using Bayesian statistics using the `ctable` function which produced a Bayes factor favoring the $H_5$ hypothesis, $BF_{50} = 7.32$.

The final hypothesis stated that a relation between the action taken and the quality of the message may exist. A chi-square between the two variables returned a result that approached significance, $\chi^2(8) = 14.174, p = .077, V = .287$. However, assumptions for chi-square were violated and so an attempt to obtain a more accurate $p$ value using Monte Carlo sampling (95% confidence level using 100,000 samples) returned a similar result, $p = .076$. The Bayesian `ctable` function also returned a Bayes factor against $H_6$ hypothesis, $BF_{06} = 16.58$. Results of the analyses for $H_6$ showed evidence that a dependency between these two variables did not exist.
6.4.4 Discussion

The results for this part of the study were illuminating as well as surprising. Support was provided for hypotheses $H_3$, $H_4$ and $H_5$. These findings indicate a strong triangle of relation between the time available for participants, the importance of their message and the quality of the new message that will be written. One of the most troubling results was the case between the quality of a message and the importance of the message. There seems to be a negative dependency which suggests that as the importance of a message rises the quality of the newly rewritten message falls. In organizations where key messages are exchanged between employees this loss in quality may have drastic implications in the efficiency of the work. Maxims of quality are genuinely affected which could be the cause for implicatures and even worse lead to aggression at the workplace. This effect also applies to other social media services as well. The more participants perceive a message as important and they have a limited amount of time to submit it, the more likely it is that the message will be of lesser quality. On the positive side of things however, a positive effect was found between time and importance which suggests that usually there is enough time available when the message is important.

On the other hand, expected hypotheses that aimed to find a dependency between the variables and the action that a user will take were not found. Specifically, there is no evidence to suggest that the importance of the message will affect the action taken by a participant. Put simply, participants have equal chances of not writing or rewriting a message in respect to the importance of a message; a troubling finding. Additional dependencies between the time available and action taken by a user as well as between the quality of a message and the action taken by a user were not found. However, the evidence was not conclusive on these dependencies and so caution is advised.

Figure 6.10 depicts the model using all the variables studied based on the results. There is a dependency between the two variables of time available and the importance of a message and in turn, these two variables affect the quality of the new message. However, their effects are reverse. Further, dependencies with inconclusive evidence were also included in the model so that further studies can establish whether there is a genuine effect between these variables.

Recommendations

The results have direct implications for social media design. Submission errors are unacceptable from a design standpoint given the technologies available. This study demonstrated that they can be the cause for implicatures which can result to more severe consequences such as aggression. Designers should take steps to ensure that the information on social media is preserved for all forms, especially the ones where messages are exchanged. Failing to do so threatens the social experience of users and effectively their efficiency of communication and collaboration.
6. RESEARCHING THE INDIVIDUAL FEATURES FOR AGGRESSION...

Figure 6.10: The error submission model based on the results.

Limitations

This study attempted to develop a working model through which one can understand the effects of a submission error on the overall communication for social media. However, there are limitations when using a survey to determine these effects. Respondents were required to recall from memory what happened in their past, which may affect the results. Additionally, this method also relied on their perception of the incident and the actions taken.
CHAPTER 7

THE SOCIAL INTERACTION DESIGN FRAMEWORK

The previous chapters presented research on groupthink and aggression that had been conducted in order to demonstrate how features on social media software can affect these phenomena. Based on these results, I argue that social behavior is not only affected by interfaces but one can proactively design interfaces to engineer behavior. This chapter helps bind together, all previous information and results based on the two showcase phenomena used in the study and demonstrates how interfaces affect behavior. This is achieved by providing a new definition for social interaction design based on social experience goals and a framework through which one can design and develop software. It accounts for desirable and undesirable social behaviors, discusses methods through which one can obtain relevant results for such behaviors and implements these results, to help produce better social media software that take social experiences into account.

7.1 A New Look on Interaction Design

In the introduction chapter, I have already discussed the benefits provided by interaction design as well as of social interaction design which is a relatively new concept. However, goals and ideas of social interaction design are not clearly established on the overall existing framework of interaction design and human-computer interaction. Having already asserted that social behavior can be altered through the means of social media design, implementing this into a framework that can be adopted and used in software design is essential.

Prior to demonstrating this, two significant issues have to be considered. As I pointed out earlier on in the introduction chapter, J. Brown (1997) discussed that in part, the fault for dysfunctional systems falls in HCI specialists that do not properly convey their results to software engineers. Results should not be presented in an abstract way but concisely and specifically. The second issue is that by introducing more terminologies into a scientific domain, this can clutter and confuse the people involved within it. If social interaction design is to be taken seriously and be implemented into the work process of social media development teams, it needs to be implemented in a seamless and painless manner. These two points served as guides for the proposal that I am presenting in this chapter.
7.1.1 Defining Social Interaction Design within Interaction Design

There are three key characteristics in the interaction design process: (1) focus on users, (2) usability and user experience goals and, (3) iteration (Rogers et al., 2011). In their book, Rogers et al. (2011) described that two main goals are essential for interaction designers. These were defined as usability goals and user experience goals. Both types of goals are not just mere white noise in the background but an essential part for effective software design. Strategies such as needfinding are an essential part of identifying development opportunities and providing a dynamic platform for design (Patnaik & Becker, 2010). By identifying the needs of the clients and stakeholders, one ends up defining the goals of the system. Usability and user experience goals help produce a fine-tuned product. These goals become the centerpiece for the core activities of interaction design: needfinding, conceptual and physical models, interactive models and evaluation.

However, the above framework does not take into account the social dimension found in social media software. While user experience goals describe certain psychological aspects for an interaction between a human and a system (e.g., computer, software, video game) it does not take into account the interconnectivity that exists between different users through social media. Individuals are interacting with the software as much as they interact with each other through it. Through these interactions behaviors emerge. As demonstrated by this study, social behaviors can be altered by the software design in a proactive manner.

According to the above, I provide a new definition for social interaction design, which in light of this evidence, forms the base for the framework presented in this chapter.

Social interaction design is a set of principles, models, methodologies and other aids that are used in the proactive design and development of systems which involve social environments, in order to satisfy a system’s social experience goals. In turn, social experience goals are a set of goals derived from the social dimension of a system and describe social behaviors, interactions, attitudes and phenomena, which are identified by designers in order to meet system and user explicit as well as implicit overall goals.

I proceed by presenting the social interaction design framework that builds upon the current interaction design framework. As such, all current models that are currently in use by interaction designers can still be used in this version of interaction design that encompasses social interaction design. To achieve this I developed an example that will demonstrate how goals for phenomena such as groupthink and aggression, can determine social behavior and should concern the decisions that designers make about features in social media. Below, I present the example that is going to be used to demonstrate the practical aspects of the framework.

A government organization is asked to develop a social media platform called UrCity to encourage an active involvement of people in their communities. They want to empower communities, states and
the country as a whole to understand issues, find solutions and help notify authorities faster about these problems. Additionally, this project helps to also evaluate the efficiency of certain authorities. Authorities that are not willing to listen get a lower rating based on how long it takes them to resolve a problem. The public not only has the power to report and find solutions for problems, but also has the power to assess whether high rank officials are effectively doing their jobs.

There are many important issues that need to be resolved from a software engineering and design standpoint but for the purposes of demonstrating how social behavior can be engineered through the design of software, I focus only on the social interaction design aspects. I attempt to demonstrate three major activities of interaction design and how they relate to the design of social experiences: (a) needfinding, (b) prototyping, conceptual and physical models, and (c) evaluation.

### 7.2 Needfinding: Identifying Social Experience Goals

During needfinding designers should make sure that they understand all of the social needs and dangers of the project. These are social behaviors of users while interacting with others and can have an impact on the general mission of the product. The process of identifying requirements in a project in terms of social needs does not significantly differ from the process used to identify other functional or non-functional requirements of the project. The same strategies can be used to account for the social dimension. Rogers et al. (2011) have discussed that identifying environmental requirements for the social aspects of the environment is particularly important. For example, one has to think about whether communication will be synchronous or asynchronous, or if all data is going to be shared.

However, this can be taken a step further. A general identification of requirements does not need to be confined in terms of how people will communicate or collaborate. Social goals can also define social behaviors, interactions, attitudes and phenomena in social media. These behaviors can later be classified as desirable, undesirable or neutral. This classification is not something to be taken lightly. Since many social behaviors may come in conflict with each other one has to determine the desirable behaviors that are necessary for an efficient interaction between users and those that are undesirable in the eyes of a designer. Additionally, there may be behaviors classified as neutral. Finally, all these goals can also be assigned different severity.

I define the new set of goals as **social experience goals**. These goals aim to identify all aspects of social interaction, behavior, attitudes and phenomena between humans through the system that is to be designed. Humans are social creatures by nature and as argued earlier on in this dissertation, sociological needs are the force that drives technological development. It therefore becomes appropriate that such needs will finally be taken into account in the design process of social media software.

Designers can set these goals based on user needs, which is a fundamental basis for interaction design work. However, identifying user needs is not sufficient
for identifying social experience goals. For example, most users will vouch for an aggression-free environment however, few if any, will actually mention groupthink as an undesired behavior. Groupthink is a goal that is likely to be identified only through the eyes of stakeholders or the designers. Therefore, social interaction design takes into account both user requirements as well as system requirements and expects that these will be adapted in order to complete each other.

The approach for classifying social experience goals is also important. While everyone can agree on some nearly universal set of usability and user experience goals, social experience goals are context-dependent to the intent of the software. As an example, this paper has treated aggression as a negative aspect on social media communities. On the other hand, Wallace (2001) has also pointed out that there is a belief by some that venting out anger may provide them with catharsis. While this remains a controversial topic, it is still worth noting that it may be an option for a software designer (e.g., one that may be asked to develop a social media platform where people come to vent their anger towards others). Finally, there are certain social phenomena that exist in social media but may be of no interest for the software designers. For example, one may notice that people exhibit signs of altruism online, but the community would not benefit, nor would be affected negatively by such behavior. Or, it may be the case that such behavior is acceptable at moderate levels while if it falls or rises above a certain threshold it becomes either desirable or undesirable.

This process of identifying desirable, undesirable and neutral social experience goals will help designers in their further development of designs. The result should help produce a comprehensive list that would cover most of the important social experience goals that are of direct interest and concern for designers. Additionally, these will be classified under certain categories that better help designers address cost and benefit issues. Finally, each one of these goals can also be assigned a severity to establish priorities.

Based on the above, I present a revised diagram describing usability, user experience goals and social experience goals based on the original diagram created by Preece, Rogers, and Sharp (2002); Rogers et al. (2011). The new model still has usability goals as a centerpiece, while user experience goals remain in the outer rim. However, I added an extra layer that describes social experience goals that builds on top of the other set of goals. Additionally, while it may seem that the most important goals for a product are usability goals, that should not be the case. User experience and social experience goals are just as important for avoiding faulty designs and costly mistakes. For example one can build a highly efficient product that does not provide an enjoyable experience, or an efficient product that however leads to groupthink behavior. While the diagram is designed in such a way that will help align the priorities for the design and development teams, it is important to note that all goals are important. Through the iterative process of design, implementation and evaluation, the team should constantly aim to assess if goals are satisfied and to what degree.

Additionally, this model also demonstrates how important is that these goals be identified early on in a project’s lifecycle. The cost for changes in the system increases over time. This may be especially true for social experience goals where
they need to be identified correctly and early on. As an example, a shift for a groupthink goal from undesirable to desirable is easier to achieve early on in the design stages, than later on during the development stages. This is an important difference between social experience goals and user experience or usability goals. While in the latter, changes may not be costly, the former will produce significantly more costs for changes later on in a project’s lifecycle.

To identify social experience goals one can use a variety of methods. Some methods can be on-site observation, interviews and participant observation that can help develop different perspectives for software needs. Questionnaires and focus groups can also be an effective way (Rogers et al., 2011). Most methods can be efficient for identifying social experience goals, however there is a tremendous potential for methods that will produce real world data. As such, participant and on-site observations may be more ideal in determining what social behaviors occur, how severe they are, and whether they should be classified as desirable, undesirable or neutral.

Having identified social experience goals, these can further be used in the development of personas. Personas are highly valuable and popular in interaction design because of their ability to demonstrate behaviors (D. M. Brown, 2010; Cooper, Reimann, & Cronin, 2012). However, personas do not have to be limi-
7. The Social Interaction Design Framework

7.2.1 Example: Social Experience Goals in UrCity

In the example of the UrCity project, one can identify two separate behaviors. People are expected to collaborate with others about various topics. Some of them may be controversial topics and disagreements may quickly evolve into flaming wars. As such, the first identified social experience goal is aggression. Furthermore, one can imagine that since users will have to discuss problems as well as propose their own solutions and vote for them, one has to evaluate how effective is the decision-making process. Groupthink is a serious phenomenon that can affect decision-making and has been the cause of many historical fiascoes (Rose, 2011). As such, a second social experience goal is groupthink.

The second step involves classifying these behaviors. Both are classified as undesirable based on the descriptions of the previous paragraph. These two however, do not necessarily have the same severity. When having wicked problems one can expect flaming wars in social media software. Low levels may be considered acceptable. On the other hand, groupthink is a problematic issue that can affect the efficiency of the decision-making process and as such it should be assigned a high severity. Through needfinding (which is an iterative process in its nature) one not only identifies the social experience goals derived by desirable, undesirable or neutral social behaviors but also assigns severity to each goal that will significantly affect the decisions that will be later taken in to the design.

7.3 Prototyping, Conceptual Models and Physical Models

Storyboarding, sketching, or wizard of oz prototyping are great ways for showing how social experience goals are satisfied by various models. They also put more focus on the needs rather than on the design of the product itself which is particularly helpful because of the complex nature of social behaviors. Identifying opportunities for alternative designs can be helpful when creating low-fidelity prototyping. It also helps to put into perspective, when there is an opportunity for a new design as well as to establish how interactive versions of the design will affect users.

Early on, ideas for prototypes that are expected to have an effect towards satisfying certain goals can be taken from published research studies or a designer’s expert knowledge. In many cases, such research may require the development of high-fidelity prototypes which can be more expensive. It becomes essential that designers have exhausted all possible options and the designs selected to be evaluated are the ones suspected to be more efficient. Additionally, when conceptual designs are created it may be beneficial to develop plus or minus scenarios.
7. The Social Interaction Design Framework

(Rogers et al., 2011, p. 292). This would help develop the most positive and most negative results of a design on social experiences. Such an approach can also help see if there is a balance between social experience goals and in general they are in accordance with the product’s mission goal and user social needs.

When designing a conceptual model that incorporates social goals, a use of multiple methods such as sketching out ideas, storyboarding, describing possible scenarios, and prototyping aspects of the proposed behavior of the system is important. This should reflect scenarios where the social needs can be put to the test. For example, a storyboard can demonstrate a frustrated user using the software to send a message to all users. Prototyping or video prototyping can then show how versions of the design can have a preventive effect on these interactions. New opportunities for exploration may be discovered during this process. Wizard of oz prototyping can also be extremely helpful in observing social behavior while having individuals within the same observable space.

Transitioning between a physical design and a conceptual design does not produce additional challenges in respect to social experience goals. Some social experience goals may be able to be satisfied early on when prototyping while for other goals a designer may be able to identify opportunities later on when dealing with the physical design. For example, affective computing research has shown that the interface can have a positive effect on reducing frustration (Prendinger & Ishizuka, 2005) and in turn aggression in social media. Certain features can have a positive effect on individuals but only be considered when dealing with the actual physical design. At this point, many different approaches can be invented so that specific goals such as the one provided in this example can be satisfied.

When it is time for interactive versions of the design to be developed, one must have a relative certainty that social experience goals are likely to be satisfied. A theoretically grounded approach is essential. Features that need to be changed at this point due to the violation of these goals can be more costly for the design team. For example, imagine designing a social media software platform meant to be used by a community for anger management. One can start creating conceptual models, create storyboard and prototypes. One of the social experience goals classified with high severity as undesirable is aggression. However, during the prototyping phase, one has not considered thoroughly that anonymity states may affect inhibition. Later on, once an interactive design has been created the opportunity of using various anonymity states and the effects on aggression is identified. However, at this later point in the design, most of the features of the system that relate to the anonymity state used need to be reevaluated in accordance to the usability, user experience and social experience goals.

A final consideration involves the long established use of metaphors in design. While useful in interaction design it may have a limited power for developing new paradigms. Preece et al. (2002) have described that it limits “the designer’s imagination in conjuring up new paradigms and models” (p. 59). Designing social experiences in an online world is an opportunity in itself to achieve things that could not be achieved otherwise. For example, using masks to hide one’s anonymity is not the same with being online and use pseudonyms or even be completely anonymous. Being restricted to having new ideas may not be ben-
7. The Social Interaction Design Framework

7.3.1 Example: Prototyping and Modeling for UrCity

Since we identified the main goals, a better understanding of the factors involved is necessary. After having read the literature on the topic one can start developing ideas that will satisfy these goals. Some may be concrete such as the randomization of the solutions for each user while others may not be as easy to determine. For example, one may suspect that anonymity states can affect aggressive behavior and as such create various designs that display different anonymity states and the way individuals communicate.

In early designs, several versions were created in using different anonymity states and their combinations. It was suspected that anonymity had a tremendous effect on aggression but also on groupthink by allowing users to discuss alternatives. Another suspicious feature was the dynamic voting indicator that may affect voting and be a factor for groupthink. Further, a feature that was added was a pro/con list representation of arguments for each solution. It was thought that potentially could be better than textual collaborative interface for increasing alternatives and reducing groupthink. Notes were sent to software developers to ensure that information submitted through forms is not lost and also provide alternatives for any errors that may appear in the system. Designs were created for having dynamic voting indicators for the voting processes for solutions. Furthermore, two separate designs were created in which leaders could be identified to vote publicly, whether they are administrators or moderators. Finally, two designs were created, displaying solutions in a randomized and standardized manner.

At this early stage, most of these design features have neutral values for satisfying the social experience goals for the UrCity project. Early prototypes are helpful and will reduce the chances of having to redesign costly elements. Several evaluations of the early designs will help to move toward the direction of developing interactive versions that can formally be evaluated and establish on whether features satisfy social experience goals.

7.4 Evaluating Social Interaction Designs

Evaluation can be used throughout the previously described activities. Evaluations were defined by Rogers et al. (2011) as formative and summative. These evaluations can help assess that the needs are satisfied throughout the course of the design. The same applies for social experience goals. Evaluation is important so that designers will know that they are on the right track. Evaluation of social behavior while can be more complex, it does not significantly differ from evaluation for usability goals or user experience goals.

Evaluation can happen in a number of ways. Rogers et al. (2011) suggested a couple of evaluation paradigms such as “quick and dirty,” usability testing, field studies, predictive. When exploring social experience goals most approaches can be efficient. For example, the "quick and dirty" paradigm was used in this dis-
sertation to compare argument production for pro/con list interfaces and textual collaborative interfaces. Usability testing was also used extensively but aimed not to measure usability goals but social behavior goals. Field studies can also be an efficient way for getting in touch with the users and finding out how the design may affect behavior. However, this technique limits designs which affect individuals without the individuals realizing that this is happening; something that has been observed during the course of this dissertation research. Moreover, the predictive paradigm, while effective for usability goals can be limiting for exploring social experience goals. What needs to be observed, is never observed and it is assumed. However, there are certain aspects of the software that given enough literature one can be fairly certain of the expected behavior. Additionally, there is a need that the experts involved have a background not only in HCI, but also in psychology and sociology. Finally, a combination of all these paradigms can also be an efficient strategy. For example, “quick and dirty” techniques produce qualitative data that can be later transformed through content analysis to quantitative data and be statistically evaluated.

Exploring social behavior affected by the interface can be easier said than done. During most of my dissertation research, I decided to establish alternative designs and measure quantitatively or qualitatively specific aspects. While some remarks were drawn from literature, in most of the cases I had to use actual users. At times I would ask them open-ended questions so that I can get a better understanding of how things work. But even this was not sufficient enough information to determine social behavior. For example, during pilot testing I asked people why they behaved the way that they did and some of the responses went along the lines of “because I felt this way.” In other words, their response stated that the interface had no effect on the way that they reacted in certain situations through different interface designs. However, in many cases, quantitative results suggested that the interface had an effect on them. A good example is the research on anonymity states and aggression. As such, if my decision would be to just ask the users and especially with faulty yes or no designs (i.e., 'Do you feel that the interface affects your reactions') then my results would have been completely different and inaccurate.

Finally, the evaluation of the data is also an issue that needs to be addressed. While NHST analysis is quite popular and has been used extensively in HCI, many argue that it is insufficient or that it may lead to misinterpreting results (Kaptein & Robertson, 2012). Ethnographic research can also be extremely affective, but others may still call upon its reliance on a researcher’s subjectivity. There are always benefits and tradeoffs. Alternative methods should always be considered as a researcher needs to understand the limits of an analysis method and its implications on the final results and conclusions.

7.4.1 Example: Evaluation for UrCity’s Features of Interest

Several features were evaluated in order to establish if some of the designs could satisfy some of the social experience goals set forth by social interaction design. These were evaluated using several quantitative as well as qualitative methods. Results established decisions that need to be made by designers.
Pseudonymity was found to play a significant role in increasing the chances for aggression while complete anonymity did not seem to have an effect on aggression. On the other hand, complete anonymity seems to have produced the highest potential for reducing groupthink compared to pseudonymity. However, from a practical perspective, designing UrCity in a way that everyone will be anonymous is not feasible for practical reasons. As such, a decision was made for users to appear with their real names, but the option of complete anonymity will also be provided when posting messages.

The feature of dynamic voting did not produce significant differences for results (Tsikerdekis, in press-a). As such it is rendered as safe and users will be allowed to participate publicly in voting processes using this mechanism. It has a limited effect on groupthink.

Pro/con lists will be employed for all collaborative sessions on the software as they were found to increase unique arguments. However, an additional decision was made to use a textual collaborative interface for a narrative and discussion on each topic along side with pro/con lists. As such, it was expected that more alternatives would be proposed and more information would be produced for users to make better judgments on the topics.

Error messages were designed to be exact, specific and short. Additionally, alternative solutions were designed for finding the information faster and had even provided options for automatically looking at Internet archives to find information that may be lost. An additional option was given to users to ask for help from the community on the forums of the website. These elements were identified as critical for achieving a higher comprehension and helpfulness for error messages, therefore decreasing unnecessary frustration for users.

Steps have also been taken to ensure that messages are not lost due to submission errors. Analysis has found that lost messages can critically affect the efficiency of communication on the community and as such, be the cause for implicatures and misunderstandings. Taking measures to always try and attempt to store information in forms on the server will significantly improve both user and social experience.

Solution lists for all the voting mechanisms will use algorithms that randomize them for every user. While an alternative design was made for delivering solutions in a stratified manner, the nature of the project did not necessitate a post-hoc evaluation of voting procedures. Therefore, the chosen design was rendered an acceptable choice.

Finally, some of the designs had administrators, moderators and community organizers distinguished graphically from the rest of the users. Since people can make their votes public if they wish to vote using their real names, people holding such positions in the community will be distinguished from the rest of the voters. According to the evaluation of these designs, no effect was found by leaders influencing individual votes especially when all information for the topic is available to users. As such, it was concluded that this is a good feature in the design and would likely not affect decision-making processes, decision outcomes and be the cause for groupthink.
7.5 Additional Aspects for Consideration

7.5.1 Social Cognition

Cognition is an important part for interaction design and the creation of mental models. Rogers et al. (2011) have shown the importance of having good knowledge of all the processes involved with cognition such as attention, perception and recognition, memory, learning, reading, speaking, listening, problem solving, planning, reasoning and decision-making. This knowledge of the physical world can be applied in the design of the digital world.

However, cognition cannot sufficiently help a designer understand its users in terms of social experience goals. For this, social cognition needs to be included in the work of interaction designers, which is used in social psychology to explain attitudes, attribution and groups dynamics. It “broadly includes the cognitive processes used to decode and encode the social world” (Beer & Ochsner, 2006, p. 98). Many have studied social cognition and identified its components (Adolphs, Tranel, & Damasio, 2001; Kihlstrom & Cantor, 2000; Kihlstrom & Klein, 1994). Beer and Ochsner (2006) used all this information to develop a better definition for social cognition. They stated that, social cognition describes perception of others, perception of self and interpersonal knowledge. Social stimuli are perceived in various degrees of complexity and along with contextual knowledge people conjure representations of responses to a situation (Beer & Ochsner, 2006). As such, the way people respond in varying settings in social media is not dependent only on a single individual, but also on the others involved in the “immediate” surroundings as well as the context in which a situation unfolds.

Social cognition can help create mental models or information processing models the same why that cognition is used. In the example presented here, social cognition processes can help identify gaps in the design that relate to undesired goals. For example, highly controversial topics are known to attract people coming from the opposing sides of debates and that may be biased towards others. Beer and Ochsner (2006) have discussed that in the process of perceiving others, biases due to stereotyping may lead to people seeing things that are not present in the behavior of others. Consequently, this is also similar to the solipsistic introjection described by Suler (2012). The result in circumstances where bias is introduced, may be a well-founded anticipated aggressive collision. This realization can help create mental models that a designer can employ for finding ways around such problems. In this case, attempts to reduce the bias regarding perceptions of others may be the best course of action.

7.5.2 Distributed Cognition Framework

An existing effective framework for interaction design when dealing with collaboration and communication that deals with the interaction of multiple individuals is the distributed cognition framework. This mainly investigates how individuals pass information through media (Rogers et al., 2011). However, it can also be expanded to incorporate and take into account social behavior. This way, the emphasis can be placed on the social aspects of cognition as well. Rogers et al. (2011)
have described that distributed cognition involves examining several aspects of a system and its individuals. I further expand the list of aims by adding one additional aspect worthy of examination that could help interaction designers take into account social interactions and behaviors. A socially distributed cognition framework should aim to examine:

- The distributed collaboration or communication that takes place (including the way people work together to solve a problem).
- The role of verbal and non-verbal behavior.
- The various coordinating mechanisms that are used (e.g., rules, procedures).
- The various communicative pathways that take place as an activity progresses.
- How knowledge is shared and accessed.
- *How the behavior of individuals affected by the design as well as exterior factors (e.g., the social environment) could influence the processes in a system.*

While using such a framework, a designer can make use of storyboards and prototypes to display behavior relative to the social experience goals. For example, according to an undesirable groupthink goal, participants willing to follow a majority opinion due to compliance pressures can be seen following a leader’s opinion. This can be implemented as a behavior on the distributed cognition model. In turn, a designer can identify issues on the outcome of the process violating the goal and attempt to find ways to minimize the impact of such a behavior. A way for working around the issue may be for the program to ask participants if they have enough information or if they would like to go through the information one more time before voting. Representation of arguments can also prove to be an important way for inhibiting such a behavior.
CHAPTER 8

CONCLUSION

8.1 Reflecting on the Dissertation

The work conducted in this dissertation and further work associated with it, has shown great promise in not only answering key questions discussed earlier on in chapter 1, but has also revealed how potentially vast the effect of software design is in social media. Many social media features have been found to have an effect on aggression such as, (a) anonymity states and specifically pseudonymity being a factor for aggression (Tsikerdekis, 2011, 2012), (b) message submission errors could severely impact communication, be the cause of implicatures and potentially lead to aggression, (c) quality of error messages in terms of helpfulness and comprehension can reduce frustration for users and in turn aggression. Similarly, features have been found to affect the potential for developing groupthink such as, (a) anonymity states empower users to contribute their alternative solutions to a discussion (Tsikerdekis, in press-b), (b) pro/con lists produce more unique arguments as compared to textual collaborative interfaces (Tsikerdekis, in press-c), (c) publicly identified leaders do not affect decision-making voting processes, (d) item randomization in voting processes has great potential for reducing groupthink as opposed to standardization, and (e) dynamic voting status indicators do not affect voting outcomes (Tsikerdekis, in press-a). These results were put into developing what is now social interaction design for social media and how the framework provided can ensure the development of social media services that satisfy social experience goals. Finally, the process of actively researching and developing social interaction design for this dissertation has clearly helped to identify some of its core elements that can positively impact the process and make it successful. I identify these as, (a) contextual awareness, (b) multidimensional apprehension, (c) creativity, (d) scientific application, and (e) multimethodology.

A broad variety of multiple social behaviors, attitudes, interactions and phenomena are affected by system features choices when designing and developing social media services. Since social media services describe multiple classifications, this calls for social interaction designers to be contextually aware and have a good understanding of the system as well as its social dimension and user
population. Further, it is now evident that many system elements can potentially impact not only one but multiple social behaviors, attitudes, interactions and phenomena on social media. A great example for this is the evidence provided by this dissertation work to demonstrate that the choice of anonymity states can impact two major social phenomena, aggression and groupthink. Hence, social interaction design not only requires a process necessitating contextual awareness but also multidimensional apprehension. Additionally, as it is common in processes involving design, creativity is an important factor for identifying social software features that impact the way users communicate, collaborate and interact with one another. For example, one not only needs to be creative in the way information disseminates and reaches individuals when designing social media services, but also to understand how this information is perceived and what system factors may affect this process. Because the process where information is perceived by users is invisible to the naked eye, creativity not only requires production of diverse alternatives but also reaching for ideas “out of the box.” Moreover, as this dissertation has shown, intuitive predictions for an effect rarely coincide with actual data. The complexity of sociotechnical systems is such that only through scientific application one can reveal the true effects hidden between interactions. However, scientific application should not be confined in a narrow set of methodologies and methods. A multimethodology approach is required to provide social interaction designers with a deeper understanding of how elements affect processes and outcomes as well as answering the questions as to why that is. The benefits for this can be seen in many of the social media elements that were investigated using a variety of methods in this dissertation.

8.2 Looking into the Future

In a broader perspective, in this research I attempted to successfully demonstrate that digital environments just as environments in the real world affect human behavior. The results have shown that not only this effect exists but also that it has drastic implications on the interaction of people that are online. Through social interaction design one can engineer these interactions. Furthermore, the idea of increasing collaboration and decreasing conflict sets a standard for how future online interactions are going to evolve.

About a decade ago, Zlatuška (1997) talked about the nature of Information Society in terms of a synergy between technological and social changes. Today, these changes are here but they are not over yet. Society is still facing some anxiety because of these changes, the infinite possibilities and the fear of the unknown. What we do in the present could ripple throughout our future. The behavioral phenomena studied in this dissertation are not a looming threat for future generations. They have been part of our societies long before the technological advances that we came to see the past decade. The threat to our future is simply ignorance. Ignorance in not identifying the opportunities that are presented by Information Technologies, as well as the conscious ignorance for not trying to improve the new virtual reality that we create. As Neil deGrasse Tyson put it, “Where Ignorance lurks, so too do the frontiers of discovery and imagination.”
In a future not far from now, HCI specialists will have at their disposal, a database which describes system features and their effects on social interaction, attitudes, behaviors, and phenomena. Decisions will be made not in terms of qualitative likelihood for satisfying goals, but pure probabilities. In a future world, people will gather at a virtual spot to discuss matters and the software will ensure that an effective decision will take place. Frustrated users may be distanced from the group or become distracted so that the group can work on their collaborative task. Users which may have a weak understanding of a topic will be detected and assisted in order to help lead to a more unbiased voting process. The software will automatically adjust its features, so that all important discussion elements will be heard and evaluated by all users. The experience will be tailored to every individual’s social, psychological and physiological state.

It may sound farfetched, but Clarke’s three laws for predicting the future have so far indicated otherwise (Clarke, 1962). Venturing through the impossible can help one achieve the greatest breakthroughs to the point of advanced technology becoming indistinguishable from magic. As we move our lives more into the digital world, decisions on how this world is designed affects the new reality which we will create for ourselves. Research defines what it is to be human, how we interact in the real world and how different we interact online. As the consciousness of a future generation moves to an online world, they can always look back at how things used to be once upon a time ago. If we understand enough about our interactions today, this online consciousness may end up living in a safer and more effective virtual world. A digital world that serves humanity’s needs better. A social world built by humans, meant for humans, and designed for humans!
REFERENCES


8. CONCLUSION


8. Conclusion


Daft, R. L., & Lengel, R. H. (1986). Organizational Information Requirements,
8. Conclusion


8. Conclusion


8. CONCLUSION


10.1023/A:1018931801461


8. Conclusion


8. Conclusion


8. CONCLUSION


Network, N. E. C. C. R. (2004). Trajectories of physical aggression from tod-


Phelps, R. P. (2005, October). Are U.S. Students the Most Heavily Tested on...
8. Conclusion


8. Conclusion


8. Conclusion


8. CONCLUSION


University Press.


8. Conclusion

243–252). New York, NY, USA: ACM.


SELECTED PUBLICATIONS OF MICHAEL TSIKERDEKIS

International Peer-Reviewed Publications


Lectures and Presentations

