
14. What kind of data is appropriate for my question? Choosing a unit of analysis

Elsy Gonzalez and Paul Poast

All research projects begin with a basic question: What is being observed? Are we interested in when two countries enter a war or, instead, are we observing when and if a larger group of countries sign a multilateral trade agreement? Or are we observing the tariff schedule between two countries in a week, or the number of troops deployed by state A to state B in a month? Are we even interested in states, or is the research instead interested in exploring the behavior of non-state actors such as human rights organizations, multinational firms, or rebel groups?

The question “What is being observed?” is intimately linked to the building block of all research studies: the unit of analysis. Questions ranging from “Why does deterrence fail?” to “Why do states form alliances?” to “Why do countries impose tariffs?” all require identifying the actors, actions, and timing that comprise an event. Before source information is gathered, variables are coded, statistical models estimated, control variables chosen, the scholar must settle on a unit of analysis that captures the event of interest. Whether conducting quantitative or qualitative analysis, one must know the unit of analysis before one can even hope to draw inferences.

This chapter aims to offer conceptual and practical guidance for assisting researchers to critically consider this essential (and yet, surprisingly, taken-for-granted) step in the research process. In the next section, we offer a history of how international relations scholarship evolved in the choice of appropriate units of analysis. In doing so, we explain how unit of analysis relates to two other concepts: unit of observation and levels of analysis. Equipped with this conceptual and historical background, we then walk through the decisions and choices that scholars must confront when selecting the unit of analysis for a study. We illustrate these choices using a study about the decision to recognize (or not) a secessionist group as a foreign sovereign. We end the chapter with a series of useful guidelines (in the form of questions) that scholars should consider in order to ensure that they adopt the appropriate unit of analysis for their research question.

UNITS OF ANALYSIS, LEVELS OF ANALYSIS, AND UNITS OF OBSERVATION

A scientific research project starts with the unit of analysis, or the subject of study. The unit of analysis is closely related to but conceptually distinct from the unit of observation (Long 2003, 1157). The unit of observation is the subject on which the data is actually collected. Those data could, in turn, be aggregated up to the unit of analysis. For example, if one is studying whether democracies are more likely to enter trade agreements with other democracies compared to non-democracies (Mansfield et al. 2002), one might create a “joint democracy” variable capturing whether both states in a state-to-state pair (or dyad) are democracies. Since regime type

is measured at the country level, this makes the country the unit of observation and the dyad the unit of analysis.

Debates over the appropriate unit of analysis are central to the development of international relations as a scientific discipline. Consider the idea of a “level of analysis” that was put forward by Waltz (1956). Focusing on the onset of war between states in the international system, Waltz lays out three “images”: the human image, where properties of the people and of human nature more generally play the critical role; the state image, which considers the regime type or material resources under the territorial control of a state in making states enter conflict; and the system image, which examines how anarchy and the global distribution of power. Reviewing Waltz’s book for *World Politics*, J. David Singer writes that the book is an examination of the assumptions that “lead into, and flow from, the level of social organization which the observer selects as his point of entry into any study of the subject” (Singer 1959, 453). Most notably, it was Singer, not Waltz, who used the phrase “levels of analysis” (Temby 2015, 723).

Yurdusev (1993) offers a corrective on Singer, by pointing out that units of analysis should be made intellectually distinct from Waltzian levels of analysis. The latter pertains to the context within which the subject under scrutiny can be analyzed, while the former pertains to the major actor(s) of the subject to be studied (Yurdusev 1993, 78). Yurdusev offers the example of the Concert of Europe (Yurdusev 1993, 79). The Concert itself is the level of analysis by which one examines how particular states’ policies operate in a post-war system. This means that the unit of analysis of such a study is the individual state within that level (for example, Russia).

Failure by scholars to keep the level of analysis conceptually distinct from the unit of analysis is likely due to a lack of clarity by Waltz himself: were his “images” levels of analysis, units of analysis, or something else?¹ Waltz was interested in studying the onset of war. One could read “war” as the level of analysis, with each of his “images” as units of analysis: (1) leaders who decide to take their state to war; (2) states that actually initiate war; or (3) a system with characteristics that make war possible in year *t*. Alternatively, one could view his unit of analysis as the state, with the “images” being the independent variables capturing characteristics within the state (the leader’s traits, such as a tendency for aggressive impulse), about the state (characteristics of the state, such as having a democratic regime), or around the state (characteristics of anarchy, such as the global distribution of power) at a point in time.

This latter interpretation of Waltz is sensible. Consider the argument of Lake (2008). Lake puts forward a case for the state remaining central to devising units of analysis in international relations. He labels it an “assumption” or “bet” that “a focus on states will yield parsimonious yet empirically powerful explanations of world politics.” Lake views state-centrism as a “good bet” because, as sovereign actors, states hold the final and ultimate authority over the inhabitants of their defined territory (*ibid.*, 43). Given that states are presently the primary means of organizing global society, it seems sensible to give states pride of place in any analysis of global society: “As Wendt famously observes, anarchy is what states—not individuals, groups, or transnational advocacy networks—make of it” (*ibid.*, 44).

Though Lake makes a case for why he views a state-centric “bet” a good one, he acknowledges that “this is a subjective choice over which reasonable scholars can disagree” (Lake 2008, 42). A notable disagreement, at least to a degree, comes from Barnett and Sikkink (2008, 62). While acknowledging that states retain “considerable power and privileges” in the international system, states are increasingly having to “share the stage with a multitude of

other actors.” These non-state actors include multinational corporations, terrorist networks, non-governmental advocacy organizations, transnational movements, and the bureaucracies of international organizations (ibid., 63). But Barnett and Sikkink recognize the intellectual and scholarly value of the state focus of international relations scholarship. As they write (ibid., 77–78):

All disciplines, if they are to have any coherence whatsoever, must have an overarching narrative. The anarchy thematic has helped to generate coherence for the discipline of international relations. It provided a common narrative that focused on states as actors that were struggling to maintain their security and generate wealth in an inhospitable environment. It helped to define the boundaries of the field and distinguish the study of international relations from the study of comparative politics. It focused scholarly attention on a manageable set of issues that could be subjected to theoretical emendation and empirical analysis. It provided a coherent account of the discipline that could be passed down from one generation to the next. The anarchy thematic served various useful functions.

But Barnett and Sikkink also point to limitations, namely the “theoretical, intellectual, and empirical myopia” of a state-centric unit of observation. On the one hand, broadening the units to be studied “risks generating disciplinary fragmentation, because there no longer exists a single, overarching story.” On the other hand, if the goal is to understand how the rules of global order and global governance “are created, produced, sustained, and refined,” and if these other actors influence those processes, then their inclusion into the set of valid units of observation or analysis is necessary and welcome (ibid., 78). Examples abound of quantitative studies of international processes that treat non-state actors as the primary object of inquiry. These include multinational firms shaping trade flows, such as through the influence of intra-industry trade (see, for example, Madeira 2016; Baccini et al. 2018), rebel groups fighting one another, thereby driving the violence of civil wars (Fjelde and Nilsson 2012), international human rights organizations taking actions that influence the prevalence of human rights violations (Allendoerfer et al. 2020), and transnational ethnic diasporas enabling the survival of terrorist organizations (Piazza 2018).

Lake acknowledges the influence of non-state actors, quoting Waltz (1979, 93–94) himself: “states are not and never have been the only transnational actors ... The importance of non-state actors and the extent of transnational activities are obvious.” But Waltz also leverages the analogy of microeconomic theory, claiming that while markets are comprised of or influenced by a host of actors (consumers, regulators, firms, consumer protection organizations, industry trade associations), economists studying markets define firms as the primary actors for determining the “structure” of the market (that is, perfectly competitive, oligopolistic, or monopolistic). Hence, because Waltz is interested in studying how the overall structure of the international system shapes and influences the outcomes observed in international politics, he views it as appropriate to adopt a state-centric approach to international politics.

Even if one continues to view the state as the central actor in international politics and, hence, the primary unit of observation, the choice of a unit of analysis remains controversial. Stuart Bremer’s (1992) seminal article “Dangerous dyads: conditions affecting the likelihood of interstate war, 1816–1965” placed concerns over the appropriate unit of analysis in a state-centered framework at the forefront of quantitative studies of international relations. As he wrote, “interstate wars arise out of the interactions between states,” but scholars study war at either the system level or the national level (Bremer 1992, 310). By system level, he was referring to studies such as Singer et al. (1972), which introduced a measure of the con-

centration of major power capabilities, CON, and then used it to test the relationship between system power concentration and war frequency.² By national level, he referenced studies such as Small and Singer (1976), which looked at the “war proneness” of countries with democratic regimes compared to countries with non-democratic regimes.³ Bremer stressed how “between” is the key word, saying that the key question for scholars of international conflict is the “dyadic question”: one must consider whether state A and state B will enter war.

Following Bremer’s piece, the dyad became the dominant unit of analysis in studies of international conflict and international relations more generally (Poast 2016). But subsequent work, particularly work over the past decade, sought to push scholars to shift to alternative units of analysis. Much of this push was driven by a recognition of two problems: the problem of interdependence and the problem of multilateral events (Poast 2016).

The problem of interdependence refers to the direct or indirect influence of third parties (either one or multiple) on the actions of a dyad (Cranmer and Desmarais 2016). For instance, the relationship between A and B could (and likely does) affect the relationship between A and C and between B and C. Work abounds in this area, largely adopting and extending the tools of network analysis (Hafner-Burton et al. 2009; Dorussen et al. 2016). In this case, one must structure their data, not in a dyadic form, but as an adjacency matrix (for example, the rows run from actor A through actor Z, the columns run from actor A through actor Z, and each cell indicates a shared value between the two actors).

The problem of multilateral events refers to instances when a single event is carried out by more than two actors, such as the signing of a multilateral trade treaty. Such events are particularly challenging for thinking about the appropriate unit of analysis. The event is not as straightforward as actor C influencing the connection between actor A and B (that is, interdependence). Instead, the event is when actors C, A, and B make a decision as a collective (Croco and Teo 2005; Poast 2010; Lupu and Poast 2016; Fordham and Poast 2016; Ausderan 2018). Hence, the unit of analysis should capture that decision. Poast (2010) suggests that the appropriate unit of analysis for analyzing such events is the k -ad, where $k \geq 2$. In other words, if an event of interest involves three states, then one must use the triad, not the dyad, as the unit of analysis. Of course, this raises the question of where to draw the line: If an event involved three states, could it have involved four? Does one need to account for all combinations of three states, four states, and two states? Does one need to consider larger combinations of states? While Poast (2010) offers suggestions for addressing these issues,⁴ these are difficult questions to answer. There is no formula for answering them.

What is at stake? At the end of the day, besides providing the structure of the research project (a worthy goal in and of itself), why does the appropriate unit of analysis matter? If two studies use different units of analysis to study the same phenomenon, can they both be right? While this can be the case, there are times where the choice of unit of analysis is critical for determining whether there is an empirical “there, there.” This is aptly demonstrated by one of the key literatures in international relations: the democratic peace.

THE UNIT OF ANALYSIS AND DEMOCRATIC PEACE

The democratic peace literature explores one of the core questions of international relations, if not the social sciences as a whole: Do democracies promote peace? By the late 1980s, Levy labeled the finding that democracies largely avoid war with one another as “the closest to an

empirical law in the study of international politics.” Subsequent studies have sought to reinforce, undermine, or modify that “law.” The work posits a host of mechanisms, control variables, and sample sizes in order to stress-test the relationship between democracy and peace.

As an empirical claim, its foundations lie in measurement and the unit of analysis. Measurement pertains to the challenges of identifying when countries are democratic. This includes selecting the appropriate data, commonly the *Polity IV* measure of regime type, and then determining the appropriate threshold for a country to be counted as democratic (Oren 1995). Discussions of measuring democracy (and the appropriateness of the *Polity* variable) can (and should) consume their own chapter (Colgan 2019; Bush 2019).

But fundamental to these studies, and the focus of this chapter, is a debate about the appropriate unit of analysis. One branch of work, and the work that has received perhaps the most attention, looks at the war propensity between democracies. Indeed, perhaps the earliest empirical work on the topic, Babst (1964), found that no democratic states were on opposing sides during the World Wars. This finding was subsequently tested, notably by Bruce Russett, using a wider range of conflicts over a longer period of time (Maoz and Russett 1993; Oneal et al. 1996; Oneal and Russett 1997; Russett et al. 1998).⁵

One point of contention within this branch is consideration of the dyads relevant to the study.⁶ Some scholars contend that in order to properly evaluate conflict, one must not include states where the opportunity and possibility of conflict initiation is virtually impossible (Most and Starr 1989). The most common means of identifying the “relevant sample” (or, more precisely, “politically relevant sample”) is to limit the sample of countries to contiguous dyads (for example, Israel and Egypt in 1967) and dyads with at least one major power (for example, the United States and Serbia in 1999). On the one hand, there is a certain sensibility to recognizing that two small and distant countries (for example, Tuvalu and Nepal) are at virtually no risk of entering military conflict with one another. As Lemke (1995, 29) writes, “relevant dyads matter because they comprise the correct referent group, and thus function as a true control group, against which war dyads are compared.” On the other hand, this non-random truncation of the sample could undermine the usefulness of the sample for drawing inferences. As Braumoeller and Goertz (2004, 201) remark, “the unfortunate fact is that dyads are rarely *completely* irrelevant.” This led some scholars, such as Xiang (2010) to propose an estimation approach to identifying relevance, rather than a forceful truncation of the sample. Regardless of how one views the use of politically relevant dyads to study the democratic peace, and whether one uses politically relevant dyads or rejects their relevance, the use of the dyad at the appropriate unit of analysis is unquestioned.

Another branch in the democratic peace vein explores the relationship between a state’s regime type and its “war proneness.” Work of this variety included Small and Singer (1976) and Rummel (1983), who claimed that democratic states are generally more peaceful (that is, less war prone) than non-democracies (a finding that was then questioned by Weede 1984, and others). This monadic analysis of war proneness and democracies began to take a back seat to the dyadic analysis of the democratic peace, largely because the robustness of the monadic peace came into question. Rousseau et al. (1996) considered a “mix” between the monad and the dyad. By “mix,” they meant that democracies are more peaceful than non-democracies toward all nations, but especially toward democracies. Moreover, they found that democracies are less likely to initiate conflicts compared to non-democracies, thereby adding (qualified) support to the monadic peace. Subsequently, Quackenbush and Rudy (2009) used a variety

of measures of democracy to find that while the dyadic democratic peace remained strongly supported, there is little, if any, empirical support for the monadic democratic peace.

But the monad and the dyad were not the only units of analysis to be employed when evaluating the relationship between regime type and conflict. One alternative is to move above the dyad. When doing so, support for the dyadic democratic peace becomes more tenuous. For instance, Cranmer et al. (2015) use the system as the unit of analysis, data from 1948 to 2000, and network analysis tools, to find little evidence for the democratic peace holding at the system level. Specifically, they use a community detection algorithm to find, for each year, a joint democracy layer: that is, whether states are part of the “clique” of all democracies in the system. When taking this system-level approach, they find that democracy itself plays little role in the prevalence of conflict. Instead, the finding appears to be driven by trade connections and joint membership in international organizations. They write that while “in the political science literature, joint democracy enjoys a place of prominence, so much so that the Kantian peace is often reduced to only the democratic peace,” this position of prominence is likely unwarranted, precisely because it cannot “scale up” (Cranmer et al. 2015, 11815):

Our results suggest, at minimum, that the idea of a democratic peace, although generally thought credible at the dyadic level, does not scale up to become a meaningful predictor of system stability. This is problematic for the claim of a democratic peace because states, their relationships, and broader systemic characteristics are all attributes of the same system; a coherent explanation for conflict requires empirical agreement regardless of the resolution with which we measure the process of interest.

Similarly, Gartzke and Weisiger (2013) use a system-level measure: the proportion of the states in the system that are democratic in a given year. Using this unit of analysis, they find that when democracies are scarce across the international system, democracies face a common threat (and, therefore, are likely to band together, rather than seek to enter conflict with one another). But as democracies become more numerous and comprise a larger percentage of states in the system, the threat from autocracies declines and differences among democracies appear more salient. This creates conditions whereby democracies can then enter conflict against one another.⁷

Finally, rather than moving above the dyad, one could move below the monad by considering the democratic leaders themselves. Schafer and Walker (2006) posit that the democratic peace functions through the beliefs of individual leaders. Comparing the behavior of United Kingdom (UK) Prime Minister Tony Blair and United States (US) President Bill Clinton: because Clinton’s leadership style was more pragmatic than Blair’s (who adhered to more dogmatic beliefs), Clinton’s relations toward autocratic regimes was less conflictual (though both viewed democracies as inherently more friendly than autocratic regimes). Keller (2005), conducting a statistical analysis of leaders in 154 foreign policy crises (hence, the unit of analysis is the leader crisis), finds that, within democracies, a leader’s “leadership style” is critical to determining a state’s “war proneness.” Key elements of leaders’ belief systems (for example: Are they distrustful of others? Do they exhibit traits of heightened nationalism?) and interpersonal styles (Do they have a need to control others, and are they highly outcome-oriented?) will shape how aggressively they handle a foreign policy crisis. Democracies led by what he calls “constraint respecters” (that is, leaders who are more trustful, have less nationalistic impulses, and are more cooperatively oriented in their decision-making) are “extraordinarily pacific in their crisis responses” compared to democracies led by “constraint challengers” or by leaders in autocracies of either type.

In short, the choice of unit of analysis appears to matter for claims about one of international relations' "scientific laws": the democratic peace. This points to a fundamental question faced by all scholars near the beginning of any research project: What am I studying? Am I studying how the characteristics of an individual state (for example, its regime type) can lead it to witness a particular type of event (for example, war involvement)? Am I studying how a shared characteristic between two states (for example, a common border) can lead the two states to witness a particular type of event (for example, construct border walls)? Am I even studying states? What about time? Am I only interested in the occurrence of an event, or do I also wish to know when the event did not take place? Ideally, answering such questions would merely require consulting a simple menu or scheme that guides one to choose the appropriate unit of analysis. But this is not possible. Instead, one must simply recognize the importance of answering these questions; and then try to answer them.

To make the process of answering these questions concrete, we now turn to an example of how the unit of analysis was determined in an international relations study. As should be clear, rather than following a process of set steps, the choice of unit of analysis required an inductive and iterative process of theoretical conceptualization combined with data practicalities. Selecting the appropriate unit of analysis (even if one is settled on the state as the unit of observation) requires careful conceptual and theoretical work.

CHOOSING A UNIT OF ANALYSIS: APPLICATION TO SOVEREIGN RECOGNITION

To make concrete the above discussion, we now turn to an application. The goal is to walk the reader through the decisions and choices a scholar must face when settling upon a unit of analysis. Many times, identifying the unit of analysis can happen without much deliberation because there is a clear and simple unit of study applied in the literature. At other times, however, this step requires taking many hard decisions that will inform not just how the theory is interpreted, but also the data sources and methods adopted.

The study under consideration poses the question, "What determines the rhetorical strategy used by third-party states in their decision to recognize a claim for statehood?" International recognition is a core aspect of the current international order. While control over a defined territory is frequently viewed as a core aspect of being a "sovereign" member of the international community, just as important (if not more so) is that this control is acknowledged by other states, typically major powers (Coggins 2011, 2014). Recognition can occur with minimal controversy, but secessionist entities have also faced varying degrees of conflict related to their recognition. For example, Kosovo's recognition of independence from Serbia has not been without controversy. While the major powers must make a decision *to* recognize a new state, often overlooked is how they go about making that recognition. Specifically, how do major powers explain the recognition decision (or non-decision), meaning what rhetoric is used and what is the intent underlying this choice of rhetoric? Their statement or justification makes a judgment about the legitimacy of both the group being recognized and the state from which the group is gaining independence: if the major power identifies one actor's claim to be valid, this implies that the other's claim is not. Also, is the recognition made quickly and assertively, or is it delayed (perhaps well after the decision of other states) and pensive? This can indicate the extent to which the major power views the secessionist group's claims as

legitimate, which can either bolster or undermine the ability of that group to sustain itself as an independent entity. The quintessential example of recognition rhetoric having power is the phrase “self-determination” found in Woodrow Wilson’s 14 points. This phrase was initially interpreted by various colonized groups as a call and encouragement for them to seek out independence from their colonial power. It was only later, during the Versailles negotiations, that the phrase had to be “walked back” to specify that it did not pertain to the colonies of the major Western European powers.

To study recognition rhetoric, what unit of analysis should one adopt? To begin, three important considerations must be asked: (1) Who are the actors? (2) What is the action of interest? and (3) What is the time horizon and iteration? These three questions make up the core of the unit of analysis. While it may seem simple enough in theory, in practice every decision made creates a limitation in the research moving forward. Some trade-offs are unavoidable and can be changed later. Others can have significant implications for the nature of the data one collects, which can entail substantial time and financial investments.

The best way to begin answering these questions is to take a step back and conceptualize the phenomenon under study. As Poast (2016, 370–371) writes, “research questions and theory should guide the decision of research design.” Based on the above description of the phenomenon being studied, it is clear that there are three actors: the recognizer, the secessionist group, and the host state. The recognizer is the main actor in this process; it makes the decision to recognize or not and it communicates that decision in a myriad of ways. The secessionist entity is the group of people within a bound territory that have declared their independence from a host state; which, in turn, is the actor that currently claims sovereignty over the seceding territory. This suggests that our unit of analysis should be triadic: recognizer, secessionist entity, and host.

But knowing that the event is triadic is not sufficient for identifying the actors that will comprise the unit of analysis in this study. One must decide whether only certain actors of a given type are eligible to be in the unit of analysis. As discussed above, this is the rationale for the use of “politically relevant dyads” in some studies of conflict. While it is possible to consider all actors of a given type, this may not be appropriate for theoretical or practical (that is, data collection) reasons. In the present study, which actors are eligible to be included in the study’s notion of secessionist entity is clear: the group must formally declare its independence from a host state. This avoids considerations of “latent secessionist groups,” such as Scotland in the United Kingdom: while there is a clear desire for independence within the entity of Scotland (and, indeed, a referendum on the issue was held in 2014), the entity has not formally declared independence. But what of the recognizer? On the one hand, all states in the system are eligible to be in this category, as each can make its own decision about whether and how to recognize the secessionist entity. On the other hand, prior research clearly shows that the major powers play a critical role in the extent to which recognition of a secessionist entity becomes widespread (Coggins 2014). Indeed, in the Correlates of War dataset, recognition by the key major powers of an era is a core criterion for a state to be counted as a member of the state system (Sarkees and Wayman 2010, 19). When combined with the practical consideration of access and availability of relevant statements by diplomatic officials, it is reasonable to limit the recognizers to the system’s major powers, namely the Permanent Five (P5) members of the United Nations Security Council.

Once the actor is determined, the next step is to understand the action that comprises an event. What action must be taken by a state in order to trigger (or fully constitute) an event?

Here, the overarching action is the international recognition decision regarding statehood. Based on how we conceptualize the process, the action has two components: the decision and the framing. While the decision is a simple dichotomy between whether recognition is granted or denied, framing pertains to how the recognizer explains its decision.

When thinking about the action that must be taken, one is immediately confronted with a potentially nebulous question: Must the event be explicit? While not a question germane to all forms of international interaction (that is, we know if one state fires a missile at another state), in the case of the phenomenon being explored here, recognition, the question is relevant. Here, “explicit” refers to instances where a recognizer actually makes an official statement of recognition or non-recognition. In contrast, “implicit” pertains to the “deafening silence” associated with no statement on recognition. If a recognizer makes no statement on the sovereign status of a secessionist entity, should this be taken as an implied statement of non-recognition by the recognizer? When determining the unit of analysis, it is important to consider whether both or just one is needed to glean the theoretical aspects at play, as well as the realistic implications for data collection. In this instance, this decision does not have to be a permanent one. It is possible to choose to begin with explicit events and build the initial data around those, leaving implicit events for a follow-up to a project. Practical considerations can be dispositive and, indeed, they are for this project. Focusing on explicit events allows one to reasonably acquire information on the event. Moreover, since the focus of the project is the justification used in the recognition decision, it is reasonable to focus on justifications that directly point to the decision. While there are also justifications that need to be given for implicit actions of recognition, these require significant considerations regarding systematic sourcing. An implication of excluding implicit events is that the universe of cases is not fully accounted for, and this needs to be specified when analyzing any results.

The choice of explicit or implicit events takes us to the next critical consideration for devising the unit of analysis: time. More precisely, what is the duration of a distinct event, and should we consider iterations of the same action by the same actor to be distinct events in a dataset? Answering this question has implications for the labor required in data collection and the generalizability of the overall project.

In the case of recognition statements, a statement positively affirming a secessionist entity’s sovereignty might come after a history of explicit non-recognition statements. Similarly, recognition can be revoked after years of explicitly recognizing the entity. Hence, the choice at hand for this project is whether to focus solely on the first statement following a declaration of independence by the secessionist group, or to review all explicit follow-up statements that may update or downgrade previous decisions. The benefit of reviewing updates is that it allows for change over time as well as an understanding of the evolving nature of a secessionist conflict. It can also account for how other states in the international system adapt to the situation over time. However, this approach requires determining whether to study all updates or a random yearly sample. The latter might mean potentially missing updates on a recognition stance, but the former implies much more laborious data collection with minimal changes in policy positions. For the purposes of initiating this study, practical consideration prevailed: the study focuses on the first explicit statement of recognition by a recognizer.

When dealing with the time element of the unit of analysis, it is also important to determine the time frame of study. The unit of analysis is largely constrained by the nature and availability of data at a certain time. Whether the action or event under study is even observable will depend on the historical context, the number of cases, and whether that event would have

even been recorded at a given time. An action could have a different meaning today compared to 200 years ago. To be precise, recognition of statehood likely looked different before the normative establishment of a closed system of membership through the creation of the United Nations. Hence, the author must choose the starting point for their study. While this decision can be changed later, the study must have a start and end date. While there is no set rule for which date must be chosen, a coherent starting point should control for changes in the international system.

This study starts at 1990, following the end of the Cold War, and ends in 2020. By 1990, the establishment of a closed international community in 1945 with explicit notions of peer recognition was a well-entrenched norm. All countries had by then accepted that to be a member of the United Nations (UN) was to nominally be a recognized sovereign member of the international community. Moreover, for data collection purposes, setting the start data at 1990 for an initial data collection offers the possibility of completely covering all recognition statements for all the relevant secessionist entities. A shorter time frame means fewer cases, which allows for a more in-depth and complete review of documents pertaining to the recognition of all secessionist entities that were granted sovereign status following the end of the Cold War.

All of the above decisions create a complete unit of analysis: the first explicit recognition statement by a UN P5 member regarding the secessionist group from a host state from 1990 to 2020. Examples of this triad-event unit of analysis is found in columns (1) through (5) of Table 14.1. Columns (1) through (3) label the recognizing state, the secessionist entity, and the host state, respectively. Column (4) labels the date of the secessionist entity's declaration of independence, while Column (5) lists the date of the recognizing state's first official statement regarding the recognition of the secessionist entity. These five columns constitute the core features of the unit of analysis for this study.

Table 14.1 Key coding in the dataset

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Recognizing country	Secessionist entity	Host state	Date of declaration of independence (dd/mm/yy)	Date of statement on recognition (dd/mm/yy)	Speed (days)	decision (y = 1 / n = 0)
USA	Kosovo	Serbia	17/2/08	18/2/08	1	1
France	Kosovo	Serbia	17/2/08	18/2/08	1	1
UK	Kosovo	Serbia	17/2/08	18/2/08	1	1
USA	Kurdistan	Iraq	25/9/17	25/9/17	0	0
France	Kurdistan	Iraq	25/9/17	2/12/17	68	0
USA	Crimea	Ukraine	11/3/14	14/3/18	1464	0
France	Crimea	Ukraine	11/3/14	25/4/19	1871	0

After forming the unit of analysis, one can now devise and code variables. The triadic-event unit of analysis forms the basis for constructing both the dependent and independent variables. Indeed, once the unit of analysis is established, it is easier to determine the variables to use and the data to code those variables. For instance, since the key action in the unit of analysis is by the recognizer state, data on this action should come from the recognizer state, perhaps by collecting data from a country's foreign ministry. By collecting all the documents pertaining to recognition statements, it is possible to then construct a dataset that captures the various aspects of the study's dependent variable. While the recognition statement directly reveals the recognition decision (that is, to recognize or not recognize), the dates on the statement

can capture the speed of recognition, meaning the number of days from the secessionist entity's declaration of independence and the statement by the recognizing state (column 6). One can also identify whether it was a positive statement, meaning that the recognizing state recognized the secessionist entity's sovereignty, or a negative statement, meaning that the recognizing state did not recognize a secessionist entity's claim (column 7). One can then go further to look at and code other aspects of the recognition statement, such as information on the framing of a recognition decision.

GUIDELINES FOR CHOOSING A UNIT OF OBSERVATION

The heart of any empirical study, whether qualitative or quantitative, is the unit of analysis: What is the object being studied? This question is so basic that the difficulties and complexity of identifying it can often be overlooked. While it might be straightforward to identify a type of event, say war, as the object of interest, boiling it down into constitutive parts that can then be evaluated is not as straightforward.

As could be observed in the example above, the unit of analysis is not always easily defined. It requires understanding the context of an action and what its implications can be at the individual, state, and system levels. Starting from this bird's-eye view is the critical first step to determining the unit of analysis. Making unstated assumptions about the level of analysis limits the exercise of understanding the scope of study from the onset. Taking the time to identify it explicitly, however, allows scholars not just to discern what their object of study is, but also to facilitate the decisions to be made in the data collection process. Armed with a conceptual unit of analysis, it becomes easy to discern what the units of observation will be and how they will be gathered.

This chapter laid out the conceptual difficulties and walked the reader through an example of scholars making the decision of the appropriate unit of analysis. While this showed that much of the process is trial and error, and contingent on the theory being evaluated, some ready rules of thumb can be offered:

- First, what is, in general, the event being studied? Is it the onset of a trade war or a shooting war? The signing of a treaty or the imposition of sanctions?
- Second, who are the actors involved in the given event and, not to be overlooked, how many? If states, is it a multilateral event or a bilateral event? If non-state actors are involved, how many? If non-state actors are involved, what are their relation to states? Are they supported by the states, are they the leaders of the state, are they organizations in which states are members?
- Third, what are the actions taken by the actors and are they directional? Is the action simply "event happens" (that is, war occurs; treaty signed) or are the actions actor-specific (for example, state A attacks state B with support of state C; state A offers trade agreement to states B and C).
- Fourth, given the action being considered, when does the event start and when does it end (if at all)? Does the event occur when any of the actors become involved (for example, war begins when any of the states begin fighting; treaty comes into existence when any of the states sign it), or is it conditional on some threshold (for example, the majority of

states who eventually are part of the war become involved; a plurality of signatories ratify a treaty)?

This is not a comprehensive list of questions. But questions of these types are what scholars must pose to themselves in order to ensure that they have a unit of analysis that makes the creation and coding of explanatory and dependent variables tractable.

NOTES

1. Yurdusev (1993, 81) does identify Waltz's "images" as units of analysis, not levels of analysis.
2. The basic formula for CON computes, for a system of size N in a given year t , each state i 's percentage share of a capability (P), then computes the standard deviation of those shares, and finally divides that standard deviation by the maximum possible standard deviation in a system of size N .
3. He also referred to Bremer (1980), who looked at whether countries with higher levels of national capabilities are more "war prone."
4. Poast suggests that one should include in the dataset the k -ad equal to the largest observed occurrence of the event. To address concerns about excessively large datasets, one can rely on choice-based sampling to create a feasibly sized dataset.
5. The citations to Russett's work with Oneal on the democratic peace continue on into the 2000s and 2010s. It is a truly massive body of work.
6. Studies with useful overviews of the "politically relevant dyad" concept include Benson (2005), Quackenbush (2006), and Lemke and Reed (2001).
7. It should be noted that this article induced a number of responses and rebuttals (Lee Ray 2013; Dafoe et al. 2013; Crescenzi and Kadera 2016; Choi 2016).

REFERENCES

- Allendoerfer, Michelle Giacobbe, Amanda Murdie and Ryan M. Welch. "The path of the boomerang: human rights campaigns, third-party pressure, and human rights." *International Studies Quarterly* 64, no. 1 (2020): 111–119.
- Ausderan, Jacob. "Reassessing the democratic advantage in interstate wars using k-adic datasets." *Conflict Management and Peace Science* 35, no. 5 (2018): 451–473.
- Babst, Dean V. "Elective governments: a force for peace." *The Wisconsin Sociologist* 3, no. 1 (1964): 9–14.
- Baccini, Leonardo, Andreas Dür and Manfred Elsig. "Intra-industry trade, global value chains, and preferential tariff liberalization." *International Studies Quarterly* 62, no. 2 (2018): 329–340.
- Barnett, Michael N., and Kathryn Sikkink. "From international relations to global society." In Christian Reus-Smit and Duncan Snidal (eds), *The Oxford Handbook of International Relations*. New York: Oxford University Press (2008): 62–83.
- Benson, Michelle A. "The relevance of politically relevant dyads in the study of interdependence and dyadic disputes." *Conflict Management and Peace Science* 22, no. 2 (2005): 113–133.
- Braumoeller, Bear F., and Gary Goertz. "The methodology of necessary conditions." *American Journal of Political Science* (2000): 844–858.
- Bremer, S.A. National capabilities and war proneness. In J. David Singer (ed.), *The Correlates of War II: Testing some realpolitik models*, New York, London: Free Press: 57–82.
- Collier Macmillan (1980).
- Bremer, Stuart A. "Dangerous dyads: conditions affecting the likelihood of interstate war, 1816–1965." *Journal of Conflict Resolution* 36, no. 2 (1992): 309–341.
- Bush, Sarah Sunn. "National perspectives and quantitative datasets: a silver lining?" *Journal of Global Security Studies* 4, no. 3 (2019): 372–383.

- Choi, Seung-Whan. "A menace to the democratic peace? Dyadic and systemic difference." *International Studies Quarterly* 60, no. 3 (2016): 573–577.
- Coggins, Bridget. "Friends in high places: international politics and the emergence of states from secessionism." *International Organization* 65, no. 3 (2011): 433–67.
- Coggins, Bridget. *Power Politics and State Formation in the Twentieth Century: The Dynamics of Recognition*. New York: Cambridge University Press (2014).
- Colgan, Jeff D. "American bias in global security studies data." *Journal of Global Security Studies* 4, no. 3 (2019): 358–371.
- Cranmer, Skyler J., and Bruce A. Desmarais. "A critique of dyadic design." *International Studies Quarterly* 60, no. 2 (2016): 355–362.
- Cranmer, Skyler J., Elizabeth J. Menninga, and Peter J. Mucha. "Kantian fractionalization predicts the conflict propensity of the international system." *Proceedings of the National Academy of Sciences* 112, no. 38 (2015): 11812–11816.
- Crescenzi, Mark J.C., and Kelly M. Kadera. "Built to last: understanding the link between democracy and conflict in the international system." *International Studies Quarterly* 60, no. 3 (2016): 565–572.
- Croco, Sarah E., and Tze Kwang Teo. "Assessing the dyadic approach to interstate conflict processes: aka 'dangerous' dyad-years." *Conflict Management and Peace Science* 22, no. 1 (2005): 5–18.
- Dafoe, A., J.R. Oneal and B. Russett. "The democratic peace: weighing the evidence and cautious inference." *International Studies Quarterly*, 57, no. 1 (2013): 201–214.
- Dorussen, Han, Erik A. Gartzke and Oliver Westerwinter. "Networked international politics: complex interdependence and the diffusion of conflict and peace." *Journal of Peace Research* 53, no. 3 (2016): 283–291.
- Fjelde, Hanne, and Desirée Nilsson. "Rebels against rebels: explaining violence between rebel groups." *Journal of Conflict Resolution* 56, no. 4 (2012): 604–628.
- Fordham, Benjamin, and Paul Poast. "All alliances are multilateral: rethinking alliance formation." *Journal of Conflict Resolution* 60, no. 5 (2016): 840–865.
- Gartzke, Erik, and Alex Weisiger. "Permanent friends? Dynamic difference and the democratic peace." *International Studies Quarterly* 57, no. 1 (2013): 171–185.
- Hafner-Burton, Emilie M., Miles Kahler and Alexander H. Montgomery. "Network analysis for international relations." *International Organization* 63, no. 3 (2009): 559–592.
- Keller, Jonathan W. "Leadership style, regime type, and foreign policy crisis behavior: a contingent monadic peace?" *International Studies Quarterly* 49, no. 2 (2005): 205–231.
- Lake, David A. "The state and international relations." In Christian Reus-Smit and Duncan Snidal (eds), *The Oxford Handbook of International Relations*. New York: Oxford University Press (2008): 41–61.
- Lee Ray, James. "War on democratic peace." *International Studies Quarterly* 57, no. 1 (2013): 198–200.
- Lemke, Douglas. "The tyranny of distance: Redefining relevant dyads." *International Interactions* 21 (1995): 23–38.
- Lemke, Douglas, and William Reed. "The relevance of politically relevant dyads." *Journal of Conflict Resolution* 45, no. 1 (2001): 126–144.
- Long, Karen J. "Unit of Analysis." In Michael S. Lewis-Beck, Alan Bryman and Tim Futing Liao (eds), *The SAGE Encyclopedia of Social Science Research Methods*. London: SAGE Publications (2003): 1157–1158.
- Lupu, Yonatan, and Paul Poast. "Team of former rivals: a multilateral theory of non-aggression pacts." *Journal of Peace Research* 53, no. 3 (2016): 344–358.
- Madeira, Mary Anne. "New trade, new politics: intra-industry trade and domestic political coalitions." *Review of International Political Economy* 23, no. 4 (2016): 677–711.
- Mansfield, Edward D., Helen V. Milner and B. Peter Rosendorff. "Why democracies cooperate more: Electoral control and international trade agreements." *International Organization* 56, no. 3 (2002): 477–513.
- Maoz, Zeav, and Bruce Russett. "Normative and structural causes of the decentralized peace, 1946–1986." *American Political Science Review* 89 (1993): 914–924.
- Most, Benjamin A., and Harvey Starr. *Inquiry, Logic, and International Politics: With a New Preface by Harvey Starr*. Columbia, SC: University of South Carolina Press (1989).
- Oneal, John R., and Bruce M. Russett. "The classical liberals were right: democracy, interdependence, and conflict, 1950–1985." *International Studies Quarterly* 41, no. 2 (1997): 267–293.

- Oneal, John R., Frances H. Oneal, Zeev Maoz, and Bruce Russett. "The liberal peace: interdependence, democracy, and international conflict, 1950–85." *Journal of Peace Research* 33, no. 1 (1996): 11–28.
- Oren, Ido. "The subjectivity of the 'democratic' peace: changing US perceptions of imperial Germany." *International Security* 20, no. 2 (1995): 147–184.
- Piazza, James A. "Transnational ethnic diasporas and the survival of terrorist organizations." *Security Studies* 27, no. 4 (2018): 607–632.
- Poast, Paul. "(Mis) Using dyadic data to analyze multilateral events." *Political Analysis* 18, no. 4 (2010): 403–425.
- Poast, Paul. "Dyads are dead, long live dyads! The limits of dyadic designs in international relations research." *International Studies Quarterly* 60, no. 2 (2016): 369–374.
- Quackenbush, Stephen L. "Identifying opportunity for conflict: politically active dyads." *Conflict Management and Peace Science* 23, no. 1 (2006): 37–51.
- Quackenbush, Stephen L., and Michael Rudy. "Evaluating the monadic democratic peace." *Conflict Management and Peace Science* 26, no. 3 (2009): 268–285.
- Rousseau, David L., Christopher Gelpi, Dan Reiter and Paul K. Huth. "Assessing the dyadic nature of the democratic peace, 1918–88." *American Political Science Review* 90, no. 3 (1996): 512–533.
- Rummel, Rudolph J. "Libertarianism and international violence." *Journal of Conflict Resolution* 27, no. 1 (1983): 27–71.
- Russett, Bruce, John R. Oneal, and David R. Davis. "The third leg of the Kantian tripod for peace: International organizations and militarized disputes, 1950–85." *International Organization* 52, no. 3 (1998): 441–467.
- Sarkees, Meredith Reid, and Frank Whelon Wayman. *Resort to War: A Data Guide to Inter-State, Extra-State, Intra-State, and Non-State Wars, 1816–2007*. CQ Press, 2010.
- Schafer, Mark, and Stephen G. Walker. "Democratic leaders and the democratic peace: the operational codes of Tony Blair and Bill Clinton." *International Studies Quarterly* 50, no. 3 (2006): 561–583.
- Singer, J. David. "International conflict: three levels of analysis." *World Politics* 12 (1959): 453–461.
- Singer, J. David, Stuart Bremer, and John Stuckey. "Capability distribution, uncertainty, and major power war, 1820–1965." *Peace, War, and Numbers* 19, no. 48 (1972): 19–48.
- Small, M. and Singer, J.D. "The war-proneness of democratic regimes, 1816–1965." *Jerusalem Journal of International Relations* 1, no. 4 (1976): 50–69.
- Temby, Owen. "What are levels of analysis and what do they contribute to international relations theory?" *Cambridge Review of International Affairs* 28, no. 4 (2015): 721–742.
- Waltz, Kenneth Neal. *Man, the State, and War: A Theoretical Analysis*. New York, NY: Columbia University Press (1956).
- Waltz, Kenneth Neal. *Theory of International Politics*. McGraw-Hill (1979).
- Weede, Erich. "Democracy and war involvement." *Journal of Conflict Resolution* 28, no. 4 (1984): 649–664.
- Xiang, Jun. "Relevance as a latent variable in dyadic analysis of conflict." *Journal of Politics* 72, no. 2 (2010): 484–498.
- Yurdusev, A. Nuri. "'Level of analysis' and 'unit of analysis': a case for distinction." *Millennium* 22, no. 1 (1993): 77–88.