

# .10. *What Is a Foreign Policy Event?\**

CHARLES F. HERMANN

The story goes that late one night a passerby saw a man down on his hands and knees beside a lamppost obviously looking for something. "What're you doing?" asked the curious stroller. "I've lost my keys" came the thick-voiced reply, which suggested the man had passed considerable time in the nearby bar. Since there was no sign of the keys on the pavement, the onlooker asked: "Did you lose them under this streetlight?" After a moment's pause, the drunk replied: "No, but the light is better here."

## *I. BASIC UNITS OF FOREIGN POLICY*

At present, those scholars looking for ways to measure foreign policy are not unlike the drunk searching under the lamppost. We have various alternative units and measures designed to shed a little light in the murky darkness of foreign affairs, but we cannot be sure that the keys to unlock explanations of foreign policy behavior will

\*Although he has not had the opportunity to participate in the drafting of this paper, Stephen A. Salmore of Rutgers University is my co-investigator in the research from which this paper is drawn. The study is supported by a small grant from the Voluntary International Coordination project and is a part of the larger effort (The Inter-University Comparative Foreign Policy Project) under the leadership of James N. Rosenau. I wish to express my appreciation to the graduate students in the joint Princeton-Rutgers seminar on comparative foreign policy—particularly, Linda Brady, Ralph Craft, and Donald Munton—who have been involved in the refinement of the coding rules. I am also grateful to the study's three able coders: Madeline Lonski, Elizabeth Schorske, and Joe Spina.

be found under any of the existing methodological lampposts. This paper examines one of the potential units for measuring foreign policy—the discrete event.

A moment's reflection will confirm the importance that standardized units of comparison have played in other areas where the scientific method has been applied. For example, consider the role of British Thermal Units (BTU's) in certain areas of physics, calories in human nutrition, currency in economics, or even intelligence quotients in psychology. All these units have certain limitations and one or more of them may be replaced at some point by a more satisfactory unit, but each currently makes possible in its respective area comparisons and measurements with a reasonable degree of reliability. Moreover, they play a significant role in theoretical statements.

The selection of a coding unit will depend upon the particular aspects of foreign policy under investigation. What seems quite appropriate for one research effort will be unsatisfactory for another. This diversity appears particularly likely in the early stages of scientific research. Given the infancy of quantitative foreign policy research it is hardly surprising to find "100 flowers blooming." In order to clarify the basis for the assessments of alternative units made in this paper, it is appropriate to specify the criteria drawn from our own particular research interests. Although there seem to be a number of studies with similar concerns, to the extent that others have different research concerns they may find our requirements for a basic unit less relevant.

The first criterion states that the unit must be defined in terms of operational procedures that will permit all users who faithfully apply the specified operations to identify the same number of units in a given source. Second, it must be applicable to all the political entities capable of undertaking foreign policy behavior at a given point in time. Third, it must be comprehensive, that is, it must be relevant to the entire range of foreign policy behavior that an international actor is capable of undertaking. Finally, it must provide a basis for probing into the processes and perceptions that led to the policy occurrence.<sup>1</sup>

<sup>1</sup>Other criteria also are salient to this study for the selection of a basic unit such as theoretical relevance and construct validity. Although we can enumerate them, it is difficult to judge alternative units against these standards at this early stage of quantitative foreign policy research.

The first criterion, coder reliability, requires little elaboration. A unit that fails to yield a considerable degree of stability from one user to another or for the same user across time is like the proverbial rubber yardstick. It serves no useful purpose in comparing one occurrence of foreign policy behavior with another.

Given our interest in the cross-national comparisons of foreign policies, the unit must apply to the foreign policy activities of any actors in the international political system at a given point in time. This second criterion means, for example, that a unit applicable only to the policies of relatively industrialized societies or only to entities which maintain standing armed forces would be inappropriate. For the practical task of data gathering, this criterion also implies that the unit must be constructed in a way that makes possible the acquisition of data on a regular basis for all (or, at least, almost all) actors. A unit would be of little value if it required data that almost never become available for closed societies or from those with minimal governmental statistical services.

The third requirement demands that the unit encompass all possible behaviors of the foreign policy actor. When inquiry is confined to one particular type of foreign policy behavior—for example, formal treaty agreements—this criterion does not come into play. The important empirical and theoretical work that is being conducted in such areas as conflict behavior or perceptions of hostility usually does not pretend to identify the entire range of foreign policy activity in which a nation might engage. If the scheme does allow for the exhaustive scaling or classification of all activity, it includes a large neutral category. (McClelland and Hoggard, for example, refer to the extensive comment activity in the interactions of nations that is neither conflict nor cooperation, but rather what they call the diplomatic overhead necessary to keep the system going.<sup>2</sup>) In such inquiries of limited scope, the basic unit can be tailored to the specific type of activity being analyzed. Yet there are numerous questions about foreign policy that require attention to the entire spectrum of possible behaviors that an actor might undertake. For inquiries of this type—including our own—the unit of comparison must be designed to include widely divergent kinds of behavior.

<sup>2</sup>Charles A. McClelland and Gary D. Hoggard, "Conflict Patterns in the Interactions Among Nations," in *International Politics and Foreign Policy*, ed. James N. Rosenau (rev. ed. New York: Free Press, 1969), pp. 711-24.



The fourth criterion indicates our interest in relating foreign policy outcomes to the structures and processes from which they are derived—what Rosenau calls the policy input clusters.<sup>3</sup> Students of foreign policy are increasingly sensitive to the limitations of an external stimulus-governmental response model of foreign policy and to the pitfalls of treating the state as a single-minded monolith.<sup>4</sup> We seek a unit of foreign policy behavior that could be associated with the policy process. It must be sufficiently discrete, for example, to be associated with the actions of one department or bureau of the government as opposed to another. It must provide a basis for identifying differences in perceptions and reasons for taking action. The implications of all four requirements will be made more evident by examining the kinds of quantitative research on foreign policy currently underway.

Perhaps the most obvious unit for measuring and comparing policy in the study of politics is the vote. Work by Alker and Russett on voting in the United Nations<sup>5</sup> comes to mind as a use of this unit with applications for the comparative study of foreign policy. With the increasing availability of United Nations' voting data in machine-readable form for an extended period of time, we can expect more exciting work based on the voting behavior of states. It should be apparent, however, that despite the access to the data and the good reliability, voting studies provide indicators of only a fraction of the total range of foreign policy activity. And even when the forum is the United Nations General Assembly, some significant national actors are excluded. Moreover, a single vote often suggests a unity within the government that simply may not exist.

Thematic perceptions derived from content analysis provide another unit for the quantitative study of foreign policy. Much of the

<sup>3</sup>James N. Rosenau, "Pre-theories and Theories of Foreign Policy," in *Approaches to Comparative and International Politics*, ed. R. Barry Farrell (Evanston: Northwestern University Press, 1966), pp. 27-92.

<sup>4</sup>For example, see the elaboration of the S-R model by Ole R. Holsti, Richard A. Brody, and Robert C. North, "Measuring Affect and Action in International Reaction Models," *Journal of Peace Research* 1, 3-4 (1964), pp. 170-89; and the bureaucratic model of Graham T. Allison, "Conceptual Models and the Cuban Missile Crisis," this volume, chapter 11.

<sup>5</sup>Hayward R. Alker and Bruce M. Russett, *World Politics in the General Assembly* (New Haven: Yale University Press, 1965).

work by North, Holsti, Zinnes, Brody, and others who have been associated with the Stanford Conflict and Integration Project<sup>6</sup> illustrates this form of foreign policy analysis. Because the foreign policy leaders of virtually all countries continuously make statements that are reported in public documents, cross-national comparisons are possible. Content analysis provides a technique for abstracting thematic units which, for a wide range of themes, can result in acceptable reliabilities. Themes, however, are normally applied to verbal behavior of policy makers and lead to difficulty when applied to nonverbal foreign policy activity.

Transaction data—characterization of nonverbal foreign policy behavior—are described by McClelland as activities between nations that occur with such regularity and in such quantity as to either remain unreported because they are not “newsworthy” or to be reported in aggregate form.<sup>7</sup> Examples include volume of international trade, various communication flows (letters, tourists, magazines, films, etc.), and combat deaths. Among scholars who have worked with such data are Deutsch, Russett, and Brams.<sup>8</sup> This class of measures can be extended to include what might be called quantitative indicators of national behavior. The latter may occur with somewhat less regularity (and, hence, be treated by the press as “news”). Like transaction data, however, each unit of measure is defined in terms of one specific kind of behavior. Thus, we have one unit defined in terms of anti-foreign riots, another defined as abrogation of treaties, and so on. Much of the existing machine-readable data on foreign policy is of this type. Behavioral variables in Rummel’s Dimensionality of Nations (DON) project and Burgess’ Comparative Analysis of Policy Environments (CAPE) are

<sup>6</sup>For example, see Robert C. North, Ole R. Holsti, M. George Zaninovich, and Dina A. Zinnes, *Content Analysis* (Evanston: Northwestern University Press, 1963).

<sup>7</sup>Charles A. McClelland, “International Interaction Analysis: Basic Research and Some Practical Applications,” Technical Report #2, November 1968, Department of International Relations, University of Southern California, pp. 17–22.

<sup>8</sup>Illustrative of their writings are: Karl W. Deutsch et al., *Political Community and the North Atlantic Area* (Princeton: Princeton University Press, 1957); Bruce M. Russett, *Community and Contention* (Cambridge: Massachusetts Institute of Technology Press, 1963); and Stephen J. Brams, “Transaction Flows in the International System,” *American Political Science Review* 60 (December, 1966), pp. 880–99.

examples.<sup>9</sup> Often such variables have been assembled for projects that have not aspired to encompass the entire range of foreign policy activity, but rather have examined one particular type of behavior such as conflict. Both transactional and behavioral units have dealt with physical actions and have seldom been used to record types of verbal behavior such as the substance of diplomatic negotiations. In this respect they represent the opposite of thematic perceptions. One approach encodes activity primarily in the verbal sector, the other in the nonverbal sector. Even in the domain of nonverbal activities, however, categories of transactions or behavior remain separate indicators without any common underlying unit of observation that applies to all nonverbal behavior.

Other kinds of units for quantitative, comparative foreign policy analysis are possible. For example, one might use decisions as the basic unit and record the frequency of various kinds of decisions.<sup>10</sup> Cross-national surveys of policy-makers' attitudes toward policy issues offer another means of treating perceptual data and provide an alternative basic unit for comparing foreign policy. The inaccessibility of policy-makers and their general unwillingness to give fully candid responses undoubtedly contribute to the lack of applications of this technique.<sup>11</sup> Simulations of foreign policy actions can generate perceptions and behaviors, but they leave unresolved the problem of the unit for characterizing the simulated policy unless the range of possible outputs is completely programmed.

Yet another approach to the quantitative study of foreign policy uses the "event" as the basic unit. Interest in event analysis has shown a marked increase in recent years. In 1969 there were con-

<sup>9</sup>The research activities of the project for the Comparative Analysis of Policy Environments currently are described in a series of mimeographed papers written under the direction of Philip M. Burgess, Behavioral Sciences Laboratory, Ohio State University. For an overview of the project directed by R. J. Rummel, see his chapter, "The Dimensionality of Nations Project," in *Comparing Nations*, ed. Richard Merritt and Stein Rokkan (New Haven: Yale University Press, 1966).

<sup>10</sup>For example, see James A. Robinson, *Congress and Foreign Policy-Making* (rev. ed. Homewood, Ill.: Dorsey Press, 1967).

<sup>11</sup>We have cross-national attitudinal surveys of general publics on foreign policy issues and some studies of foreign policy elites—for example, Karl W. Deutsch, Lewis J. Edinger, Roy C. Macridis, and Richard L. Merritt, *France, Germany and the Western Alliance* (New York: Scribners, 1967). Cross-national surveys of policy makers on substantive issues, however, are another matter.



ferences at Michigan State and Southern California on event analysis and a panel at the Midwest Political Science Association was devoted to the topic. The International Relations Archive Advisory Committee has recommended that the Inter-University Consortium for Political Research expand its holdings of this type of data. To the pioneering work of McClelland can now be added the acquisitions by Azar, Burrowes, Corson, and a number of others.<sup>12</sup>

An event can be loosely defined as an action enclosed in some kind of boundaries. In fact, if the designation "event analysis" were not so widespread, we might do well to describe this as "action analysis" thus properly associating this type of unit with work going on elsewhere in the social sciences. With respect to foreign policy behavior it is possible to include any nation-state or other political entity as a generator of events. Obviously not all foreign policy events are a matter of public record, but a large number are. (For instance, although the substantive issues of a conference may remain private, the existence of the meeting and the parties to it are hard for states to conceal for any period of time.) Not only do events characterize the activity of any foreign policy actor, the generic term covers almost all foreign policy behavior in which a state or other actor might engage—perceptions or outputs, verbal or non-verbal.<sup>13</sup> Like all the other possible units, events are not without their limitations as Burrowes has recently noted.<sup>14</sup>

<sup>12</sup>For example, see Edward E. Azar, "Gathering and Measuring Phenotypic Events in the Study of International Integration," and Walter Corson, "Measuring Conflict and Cooperation Intensity in International Relations," both of which were prepared for the Michigan State University's Event Data Conference in February-March 1969. Also see Robert Burrowes with Bert Spector, "Conflict and Cooperation Within and Among Nations," a paper prepared for the annual meeting of the International Studies Association, April 2-4, 1970.

<sup>13</sup>As with most of the other units mentioned, events seldom record certain conscious nondecisions, that is, a choice made by policy-makers to say or do nothing about a given issue.

<sup>14</sup>In his paper for the International Studies Association, cited above, Burrowes makes the following criticisms of event data: (1) They contain distortions and inaccuracies in substance and dating; (2) the emphasis given to different kinds of events varies from source to source and within the same source over time; (3) the often cryptic or ambiguous reporting of an event makes it difficult to evaluate one datum independent of others; (4) coders need knowledge of the countries involved to avoid the exclusion or miscoding of a large number of items; (5) concepts and theoretical goals have to be adjusted to square with the available data and its format; and (6) event data on external affairs are more easy to come by than domestic data.

At this point, the blind men have hold of many different parts of what we hope is the same elephant. As it becomes clearer that we do indeed have different units and categories for dealing with the *same* phenomena of policy, then some fruitful combination of several units and the deletion of others may occur. For the moment, however, it seems encouraging that there is vigorous activity underway with the various alternatives. Given our interests in the criteria described above, we have concluded that the most appropriate basic unit for our purposes is the event. This conclusion forces us to address a fundamental problem—defining an event so that it can be reliably abstracted from multiple data sources.

## II. TWO DEFINITIONS OF EVENT

As of this writing little of the existing work on event analysis has been published. Therefore, attempts to summarize the thinking of those using the event as a unit of analysis is uncertain at best. However, it is our impression that investigators using event data have been more concerned in their writing with the classification and application of event data than with reliable operational procedures for the identification of events. If this impression is correct, the reason for the lack of careful attention to the definition of events is understandable. Some who are described as doing event analysis are not really concerned with event as a concept, but with what we have earlier described as quantitative indicators of specific kinds of behavior. If one is interested in the frequency of riot events or the number of violent conflict events, then the concern is with the definition of a riot or a violent conflict rather than of the more generic term "event." Others, we suspect, see the identification and abstraction of events as a preliminary to the more arduous task of sorting events in a theoretically relevant manner once they are assembled. After all, whatever payoff is to be found from event analysis resides not in the accumulation of large numbers of undifferentiated events but rather in their implications after they are classified or scaled according to some system.

Such explanations may account for the relatively little attention to the problem of event definition and identification in the writing of pioneers conducting event analysis. These explanations may also



account for the infrequent reference to the literature on events and actions in philosophy and sociology.<sup>15</sup> It is difficult to quarrel with the enormity of such questions as the relative richness and reliability of alternative data sources, or techniques for classifying and scaling event data—two problems that have received considerable attention in the mimeographed papers on event analysis currently in circulation.<sup>16</sup> But to ignore the problem of reliable operational procedures for defining the event unit may mean that we end up with impressive techniques for scaling garbage. If one seeks to record and compare the range of foreign policy events generated by nations, then the task of defining the event unit must be faced.

McGowan has dealt with the operational definition of event in his research on the foreign policy behavior of thirty-two African states. He defines a foreign policy event as “a simple declarative sentence about an activity undertaken by a state or its official representative(s) wherein it may be inferred that the actor has undertaken the activities in order to affect the behavior of the external recipient of the act.”<sup>17</sup> More specifically, the sentence format of an event requires:

<sup>15</sup>Examples include the series of essays on “Interaction” in the *International Encyclopedia of the Social Sciences* 7 (New York: Macmillan and Free Press, 1968), 429–71; George Herbert Mead, *The Philosophy of the Act* (Chicago: University of Chicago Press, 1938); Talcott Parsons and Edward A. Shils, eds., *Toward a General Theory of Action* (New York: Harper Torchbooks, 1962); Nicholas Rescher, ed., *The Logic of Decision and Action* (Pittsburgh: University of Pittsburgh Press, 1967); Alfred Schuetz, “Choosing Among Projects of Action,” *Philosophy and Phenomenological Research* 12 (December, 1951), 161–84; and William H. Riker, “Events and Situations,” *Journal of Philosophy* 54 (January, 1957), pp. 57–70.

<sup>16</sup>In addition to items previously cited, this literature is represented by such papers as: Edward E. Azar, Thomas O. Jukam, and James M. McCormick with Stanley H. Cohen, “A Quantitative Comparison of Source Coverage for Event Data,” Michigan State University (no date); Robert A. Young and Wayne R. Martin, “A Review of International Event/Interaction Category and Scaling Methods,” University of Southern California, January 1968; Daniel G. Youra, “Multi-Dimensional Scaling and its Application to the Study of International Politics and Foreign Policy,” Project CAPE Working Paper #31, Ohio State University (no date); Dennis K. Benson, Edward H. Seidler, and Richard H. Sinnreich, “Foreign Policy Behavior,” Project CAPE Working Paper #33, Ohio State University, Winter 1970; Gary Hoggard is currently preparing a paper that compares several data sources; already published in L. E. Moses et al., “Scaling Data on International Action,” *Science* 156 (May 26, 1967), pp. 1054–59.

<sup>17</sup>Patrick J. McGowan, “The Unit-of-Analysis Problem in the Comparative Study of Foreign Policy,” a paper prepared for the Events Data Measurement Conference, Michigan State University, April 15–16, 1970, p. 20.

- i. A singular, plural or compound subject.
- ii. One transitive verb that states something rather than asks a question or gives a command.
- iii. A singular, plural, or compound direct object of the verb.
- iv. And frequently an indirect object of the verb or of a prepositional phrase.<sup>18</sup>

This system has the extremely important asset of providing operational procedures that can be easily applied by coders with a minimum of training. Intercoder reliabilities should be quite high. Yet it is totally dependent upon the grammatical structure and the amount of description devoted to a foreign activity by the constructor of the data source. Several problems result. The groups that control the content of the data source will demand vast quantities of simple declarative sentences on phenomena they regard as important. Consider the number of sentences written in any major American newspaper on the daily activities in Vietnam or the weekly sessions of the Paris peace talks. In brief, this procedure will confront the coders with a task of enormous proportions if the data are abstracted from any source that provides more than extremely short telegraphic accounts of all reported events. If project sources are substantial, however, and if the focus of the project remains limited, both in period of time covered and number of actors included, then this obstacle can be surmounted. A more serious problem is that no two untrained observers are likely to report the same occurrence with the same number of simple declarative sentences. Indeed the same observer would not be expected to replicate his own account if his editor requested a longer or shorter story on the subsequent trials. In other words, the trained coders who abstract sentences from the data source may be highly reliable, but the reporters or editors who originally encoded the material—and upon whom this system is totally dependent for its construction of an event—are likely to be extremely unreliable and to provide low construct validity. The event compiler has no reason to use the sentence as a standardized unit for recording acts of equivalent levels of substance or generality.

McClelland and his associates in the World Event/Interaction Survey (WEIS) also have wrestled with the problem of defining an

<sup>18</sup>*Ibid.*, p. 28.

event. Like McGowan, they require that an event have a national actor, some behavior (an act), and a target. Unlike McGowan, they do not require that all these components appear in a single sentence or other grammatical unit. Instead they have several pages of rules (plus, I understand, a considerable number of unwritten norms) indicating the kinds of actors that may initiate behavior and the specifications for determining if the action is single, discrete, international, and interactional (as opposed to transactional). Extensive attention is given to the 63 categories into which actions are classified. According to one of their reports, "the current data collection represents an 83 percent level of agreement among ten research assistants."<sup>19</sup>

The users of the WEIS system have experimented with a number of data sources and Hoggard has done a study combining events obtained from several sources. However, the primary data set has been derived from one source, the daily *New York Times*. McClelland clearly states the argument for this procedure:

The question of whose coding system matches correctly to reality is virtually unanswerable. When everybody has a bias, what test could there be for objectivity? . . . The solution that is available is to set aside the problem of unbiased truth and to analyze over time the reports that pass through *one* bias-filter of a national press. The best choice that can be made of a single source is the *New York Times*. This paper must be assumed to filter the news with an American bias but it is an exceptionally rich source of reports of international political events.<sup>20</sup>

McClelland contends that academic investigators are not direct observers of international events. They must depend on other sources—each of which has some nonrandom biases. In order to avoid confounding those biases in a totally uninterpretable way, the scholar must adhere for the present to one data source. Thus it is the world as viewed through the *New York Times*, *Le Monde*, *British White*

<sup>19</sup>Gary D. Hoggard, "The World Event/Interaction Survey Data Collection: A Status Report," Technical Report #4, February 1969, Department of International Relations, University of Southern California, p. 3. I understand from Hoggard that this reliability represents an overall rating including the assignment into categories, rather than simply the reliability on the first step of abstracting the events from the data source.

<sup>20</sup>Charles A. McClelland, "International Interaction Analysis; pp. 46-48 (emphasis in the original).



*Papers, Deadline Data on World Affairs* or some other source. At a future point after considerable experience has been acquired with various sources, McClelland believes that some effort can be made to identify biases and take corrective measures.

This stance certainly avoids some of the problems noted earlier. We would make one caveat, however, about this solution. All data sources, and certainly the *New York Times*, are the product of numerous individuals each with their own perceptual screens for interpreting reality. The owners and readership of a data source may insist that the source maintain a certain general posture. Within those broad boundaries there is considerable variation in the interpretation of events as reporters and editors change or are re-assigned. A number of people, for example, contend that David Halberstam's reporting on the Vietnam War was quite different from that of other *New York Times* reporters. In brief, even a single data source is the integration (conglomeration?) of many different interpretations of events. Moreover, if one wishes to learn something of the internal processes by which an actor chose to initiate an action, then no one data source is adequate. National or regional data sources often provide information for one or two countries on such matters as the channel by which an action is announced, the publicly stated reason for an action, or the relevant prior and associated events. But if the study is limited to one comprehensive data source, then this type of information will be missing for most actors.

McClelland and McGowan both recognize two inescapable facts about the acquisition of foreign policy event data. First, we are dependent upon others for the description of international events. Second, sources vary greatly in their detail and interpretation of events. In response to these problems, each scholar has advanced an alternative solution. McGowan seeks to overcome the great variation in detail and interpretation by defining an event in terms of a grammatical structure found in almost all English language sources. In this way a certain uniformity is established. McClelland, on the other hand, increases uniformity by limiting himself to one source. Despite their quite different solutions, both seem to make their definition of an event dependent upon the data source.<sup>21</sup>

<sup>21</sup>The dependency on the source to define an event may be less evident in McClelland than in McGowan's use of sentence structure. However, in addition to stipulating the components necessary for the existence of an event, McClelland implicitly adds

### III. AN ALTERNATIVE DEFINITION OF EVENT

We propose an alternative approach for constructing an operational definition of event. We acknowledge with McClelland and McGowan that knowledge of foreign policy activity is almost exclusively from secondary sources, but we seek to avoid letting the source define what comprises an event. The definition of an event can be thought of as analogous to a box or some other container. The data sources provide the stream of filtered information that serves as the content of the containers. We depend upon the data sources to determine what we place in the boxes, but we need not be dependent upon them to determine the nature of our conceptual containers.

The basis for this argument is that an event is an analytical concept imposed on a seamless web of "reality" by those who have reason to break out certain aspects or apparent cycles of that phenomena. As Riker has noted:

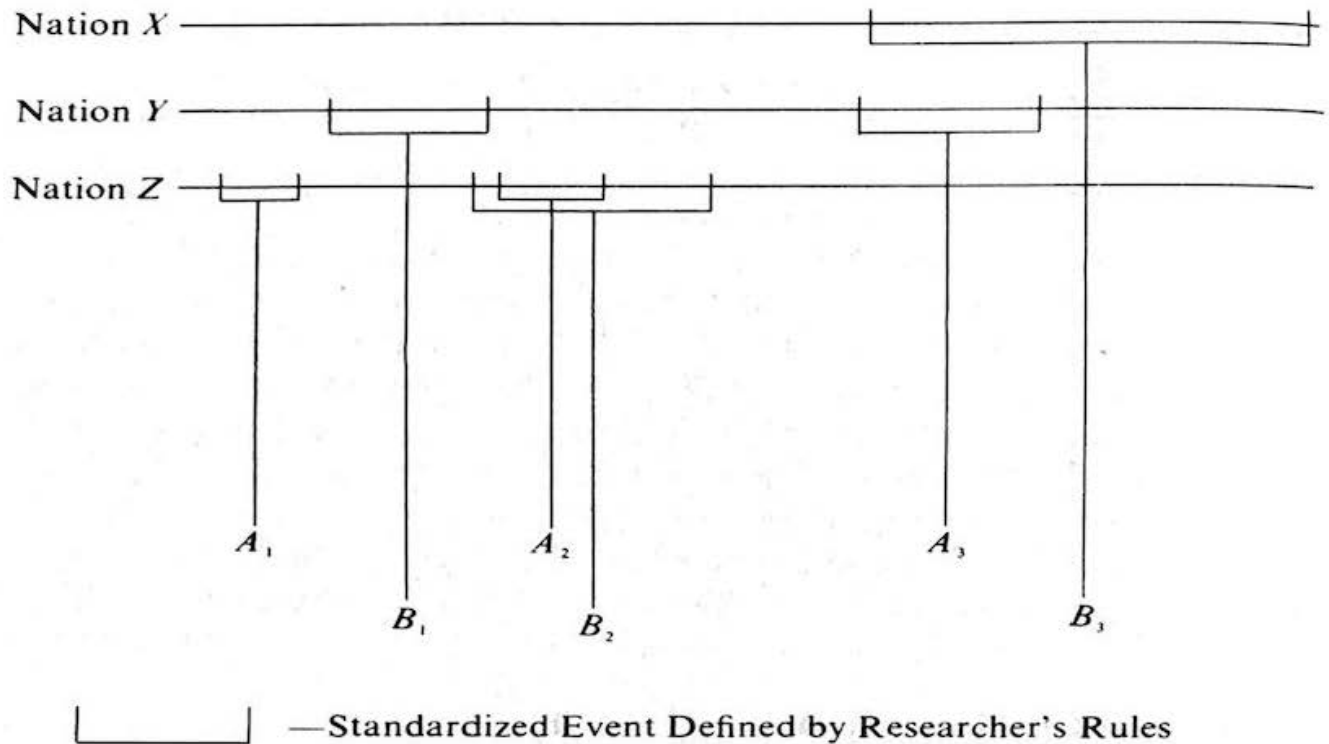
Although a continuous reality cannot, by definition, consist of discrete motions and actions, we imagine starts and stops. What lies between the starts and stops we call events. Events are motion and action separated out of continuous reality by the verbal imposition of boundaries. So accustomed are we to separating out events by verbal processes that we often lose sight of the subjective character of the separation.<sup>22</sup>

Riker's description of events can be represented as we have done in Figure 1. Each of the lines at the top of the diagram depicts the continuous stream of activity for one set of actors whom we call a nation. The bracketed segments of these lines that are reproduced below as  $A_1$ ,  $A_2$ ,  $A_3$ , and  $B_1$ ,  $B_2$ ,  $B_3$  represent six reports of "continuous reality" by two different data sources ( $A$  and  $B$ ). It will be noted that the distance between boundaries—that is, the size of the events—differs not only between  $A$  and  $B$  but between different reports of the same source. This represents the idea that most data sources do not impose a uniform standard on what they report as an

---

another—that the event be reported in the *New York Times*. By using a data source as part of the definition of an event, McClelland avoids many detailed specifications about his other stipulated components of that concept which would be necessary if multiple sources were used.

<sup>22</sup>William H. Riker, "Events and Situations," p. 59.



*Figure 1.* A diagram illustrating how two different data sources (*A* and *B*) abstract events from the continuous stream of activity for three nations (*X*, *Y*, and *Z*)—represented by the three horizontal lines). In the figure each data source is shown as having abstracted three events ( $A_1$ ,  $A_2$ ,  $A_3$ , and  $B_1$ ,  $B_2$ ,  $B_3$ ). The two hypothetical data sources do not overlap in their reports of the activities, except for  $A_2$  and  $B_2$ . Nor do they give equal coverage to each of the three nations' activities. The width of the brackets which intersect the horizontal lines represents the arbitrary boundaries, or definitions of events, imposed by the data sources. No two brackets (i.e., events) are of the same width. It is suggested in the text that a standardized definition of event can be imposed—not directly on the continuous streams of national activity but on the events of varying boundary size observed and reported by the data sources.

event. In the diagram, the scholar is limited in his knowledge of the three nations' activities to the reports of the combined data sources— $A_1$ ,  $A_2$ ,  $A_3$ ,  $B_1$ ,  $B_2$ , and  $B_3$ . However, he does not need to accept their various boundaries in his definition of an event. He may impose his own as illustrated by the bracket in the lower left corner of Figure 1. With his own uniform definition of an event, the scholar may do a secondary analysis of the activities reported by the data sources. He may discard some reports (like  $A_2$  in the diagram) as being less than his definition of an event, while breaking



other single reports into multiple events according to his standard (like  $B_3$  in the diagram).

An event, we contend, can be defined by anyone even if he is not the direct observer of that phenomenon. Although we certainly cannot code what the observers of international phenomena do not report, we need not adhere to the units or boundaries that are implicit or explicit in their reporting. In other words, we propose that the definition of an event should not be dependent upon a data source for its boundaries, level of generality, or components. Does a source fail to report all the components necessary for the stipulated definition of an event? Then we have a missing data problem that may be resolved by the use of other sources. Do two sources differ in the extensiveness of detail with which some activity is described? Then we must have rules available to determine which aspects of the reported activity will be treated as descriptions of the same event and which will constitute separate events. Do sources contradict one another as to the nature of the event? Then we must have rules to indicate how the problem is to be resolved in accordance with our research interests.<sup>23</sup>

The insistence on defining events so they are not dependent upon a single data source or upon the grammatical structure of English language data sources requires the specification of a number of rules for the operations necessary to identify events. At present we propose thirty-six coding rules for event identification. These rules can be conceived as attempting to establish a standardized set of boundaries or set of units for events. The first rule establishes our broad definition of an event, which does not differ substantially from that of McClelland, McGowan, or others who have defined the term.<sup>24</sup> All the rules are reproduced below with a brief justification

<sup>23</sup>Rules for coping with contradictions can be devised in part from data quality control techniques such as those advanced by Raoul Naroll, *Data Quality Control: A New Research Technique* (New York: Free Press, 1962). Thus, for example, in deciding which of two conflicting sources one would use for identifying an event, a check could be made to determine whether either source had an observer on the location when the activity occurred. Our present coding rules are not complete with respect to data source contradictions.

<sup>24</sup>For example, see Nicholas Rescher, "Aspects of Action," in *The Logic of Decision and Action*, ed. Nicholas Rescher (Pittsburgh: University of Pittsburgh Press, 1967), pp. 215-19.

or explanation for each. Further information on the operational procedures for each rule appears in a separate coding manual.

*Rule 1: Definition of Event.* TO BE CODED, AN EVENT MUST HAVE AN ACTOR, AN ACTION, AND AT LEAST ONE DIRECT TARGET; IT MAY ALSO HAVE ONE OR MORE SEPARATE INDIRECT OBJECTS. The identification of an actor, action, and target is common in analyses of actions and events. The addition of an indirect object as a fourth component results from our assumption that governments engage in foreign policy activity to influence other external entities. Often the objects of influence include entities other than the immediate target to whom the behavior is addressed. Whenever possible we wish to identify the other objects of an actor's influence attempts, while maintaining the distinction between the addressee (direct target) and others (indirect object).

*Rule 2: Governments vs. Individuals.* CODE NATIONAL GOVERNMENTS AS ACTORS, BUT RECORD NAME AND ROLE OF INDIVIDUAL REPRESENTATIVE(S) WHO IS THE INITIATOR, IF THIS INFORMATION IS PROVIDED. Actions are always taken by individuals, but data sources often report only the nation that the unidentified individuals represent. When given by the source, the name and role of the individuals involved can provide clues about the nature and importance of the decision.

*Rule 3: States as Actors.* ONLY THE OFFICIALS OF THE EXECUTIVE IN A NATION-STATE CONSTITUTE ACTORS. Our current research interest is restricted to the foreign activity of the executive of a government and his representatives. (Private groups, legislative bodies, international organization leaders, and so on are not included.)

*Rule 4: One Actor per Event.* WHEN MORE THAN ONE NATION PARTICIPATES IN INITIATING AN ACTIVITY, REPEAT THE ACTIVITY AS A SEPARATE EVENT FOR EACH NATION EXCEPT IN CASES OF UNANIMOUS INTERGOVERNMENTAL BEHAVIOR. We wish to reconstruct all coded actions in which a given nation participated. Even in collaborative actions, we assume that each government originally made a separate decision to participate. However, when an international organization with well-defined national membership takes action on behalf of its members and without dissent, the name of the organization can be used as a shorthand for all the individual actors.

*Rule 5: Contested States.* IN CONTESTED AREAS, CODE THE PRIOR

GOVERNMENT AS AN ACTOR AS LONG AS IT CONTINUOUSLY CONTESTS ITS OPPONENTS BY FORCE, AND CODE OPPONENTS AS AN ACTOR WHEN THEY CONTROL TERRITORY FOR A YEAR AND HAVE DIPLOMATIC RECOGNITION FROM TWO OF THE SPECIFIED POWERS (given in manual). We arbitrarily specify that a government becomes a relevant actor in foreign policy if it sustains control over people and territory and is acknowledged by more than one power in the international system.

*Rule 6: Events Reported by Third Parties.* ONLY NONGOVERNMENTAL SOURCES CAN BE USED TO IDENTIFY ANY NATIONAL ACTOR OTHER THAN THEIR OWN GOVERNMENT. Official sources (including a state-controlled press) that report on activities initiated by other states often prove to be unreliable, and therefore, are excluded.

*Rule 7: Official Newspapers as Actors.* ONLY NEWSPAPERS CONTROLLED BY THE RULERS OF THE STATE ARE ACTORS. Newspapers operated on behalf of the rulers of the state can be assumed to serve as channels for official announcements of actions.

*Rule 8: Unidentified Sources.* SOURCES IDENTIFIED NEITHER BY NAME NOR BUREAUCRATIC POSITION ARE NOT ACTORS, BUT CAN BE SOURCES OF INFORMATION ABOUT NAMED ACTORS IF ACCEPTED AS RELIABLE BY THE DATA SOURCE. Officials who provide information to the press but withhold their identity may be engaging in bureaucratic maneuver or the release of trial balloons. Under these conditions they may not be acting in ways that commit any departments or agencies of the government to an action, and therefore, should not be treated as actors.

*Rule 9: Who Is a Direct Target?* THE DIRECT TARGET CAN BE AN INDIVIDUAL OR A PUBLIC OR PRIVATE COLLECTIVE ENTITY. Officials of states direct their foreign policy behavior not only at other states but at individuals and various collective entities. Because we are interested in the entire range of foreign policy activities, all of these potential addressees are included.

*Rule 10: Multiple Direct Targets.* ONE EVENT CAN HAVE MORE THAN ONE TARGET. Officials acting for a state often direct the same foreign policy behavior toward multiple direct targets. They are included in the same event because it is unlikely that a separate decision process was associated with each target.

*Rule 11: Specificity of Direct Targets.* A DESIGNATED DIRECT TARGET MUST HAVE A DEFINITE "REAL WORLD" REFERENT. We are



interested in a state's attempt to influence foreign entities. Although some actions with an indefinite or unspecified target may be attempts to influence other entities, many such actions are part of the diplomatic rhetoric which is without direct implications for anyone. We assume that if influence is intended, eventually the actor will clarify the target; until then the activity is not coded.

*Rule 12: Organizations Can Be Own Direct Target.* WHEN THE EVENT IS A MEETING BETWEEN NATIONS WITHOUT DISSENT TO THEIR ANNOUNCED DECISIONS, THE ACTOR AND THE DIRECT TARGET CAN BOTH BE THE SAME INTERNATIONAL ORGANIZATION OR ALLIANCE. In any meeting between representatives of different states, the immediate targets of each participating state's actions are the other conferees. Therefore, when the name of the international organization can appropriately be used as a shorthand for the actor (Rule 4), it can be used in the same way for the target. This coding can always be decomposed to mean each member is directing action toward all the other participants.

*Rule 13: Direct Target Mentioned in Context.* EVEN THOUGH A DIRECT TARGET IS NOT REPEATED BY A DATA SOURCE FOR EACH COMBINATION OF ACTOR AND ACTION, AN EVENT MAY BE FORMED PROVIDED THAT THE RELEVANT DIRECT TARGET APPEARS IN THE IMMEDIATE CONTEXT. This rule about direct targets is a specific application of our meta-rule that the components necessary for an event do not have to appear within a given grammatical unit such as a clause, sentence, paragraph, or fixed number of words.

*Rule 14: Official Announcements and Press Releases.* IN PUBLIC ANNOUNCEMENTS, PRESS RELEASES, AND OTHER CONTACTS BETWEEN THE PRESS AND GOVERNMENT, THE PRESS IS NEVER THE DIRECT TARGET (EXCEPT IN ADDRESSES TO PRESS CLUBS); INSTEAD THE DIRECT TARGET MUST BE A FOREIGN ENTITY FOR WHICH THE DATA SOURCE PROVIDES A SPECIFIC INFLUENCE RELATIONSHIP WITH THE ACTOR. Except in an extremely small number of cases, the press is not the government's intended target or object of influence. The media serve primarily as a channel for the communication of governmental actions.

*Rule 15: Direct Target of a Speech.* THE DIRECT TARGET OF A SPEECH, WITH THE EXCEPTION OF A PRESS CONFERENCE, IS THE AUDIENCE. Although government officials seldom confine the target of

their actions to the audience that is physically present, we regard the immediate audience of an address as the first-level target and one which should be recorded separately in the coding scheme. This rule parallels the one that designates targets of actions at meetings as the other conferees (Rule 12).

*Rule 16: Target of Official Newspaper Actions.* WHEN AN OFFICIAL NEWSPAPER IS THE ACTOR, THE "DOMESTIC READERSHIP" IS ALWAYS THE DIRECT TARGET. We consider it an important clue as to whether a government elects to transmit a foreign policy action through a channel that it knows will be scrutinized by domestic elites as well as by foreign political entities.

*Rule 17: Separation of Indirect Objects from Direct Targets.* NEVER CODE THE SAME ENTITY AS BOTH INDIRECT OBJECT AND DIRECT TARGET. Rejection of double coding maintains the distinction between the direct channeling of an action to another party (i.e., the direct target) and the more indirect, tacit approach (i.e., the indirect object) even though both components may be subjects of attempted influence.

*Rule 18: Multiple Indirect Objects.* ONE EVENT CAN HAVE MORE THAN ONE INDIRECT OBJECT. The argument for multiple indirect objects parallels that made for multiple direct targets (see Rule 10).

*Rule 19: Inferring the Indirect Object.* AN INDIRECT OBJECT MUST BE EXPLICITLY STATED IN THE DATA SOURCE; ALTHOUGH IN A REPORT THAT YIELDS SEVERAL EVENTS, THE INDIRECT OBJECT NEED ONLY APPEAR ONCE TO BE APPLIED TO ALL EVENTS. An individual with substantial knowledge about a country and its foreign relations will occasionally recognize indirect objects of stated actions even when the objects are not explicitly stated. However, because this coding system is designed for coders with no substantive political expertise, permitting such inferences would lead to considerable miscoding.

*Rule 20: Is Influence of Indirect Object Intended?* WHEN THE ACTOR OR DATA SOURCE MENTIONS A HUMAN ENTITY OTHER THAN THE DIRECT TARGET AS SOMEHOW INVOLVED IN, OR AFFECTED BY, THE OUTCOME OF THE ACTOR'S ACTION, THEN THAT ENTITY SHOULD BE CODED AS AN INDIRECT OBJECT. When an entity other than the direct target is explicitly mentioned in the data source, the coder should err on the side of overreporting potential indirect objects even though the key question of whether the actor intends to influence

that entity remains uncertain. The question of influence can be checked in later screenings of other sources.

*Rule 21: When is Action Foreign?* TO BE CODED AS A FOREIGN POLICY EVENT, EITHER THE DIRECT TARGET OR THE INDIRECT OBJECT MUST BE AN ENTITY EXTERNAL TO THE NATIONAL TERRITORY OF THE ACTOR OR NOT COMPRISED OF CITIZENS OF THE ACTING NATION. This arbitrary distinction between domestic and foreign policy is based on the assumption that one of the explicitly mentioned entities toward which influence is attempted must be outside the acknowledged boundaries of the nation or be comprised of non-nationals.

*Rule 22: Level of Description.* WITH THE EXCEPTIONS NOTED IN THESE RULES, CODE THE MORE SPECIFIC CHARACTERIZATION OF THE EVENT AND NOT THE MORE GENERAL. Data sources vary drastically in the level of generality with which actions are reported. This meta-rule establishes our preference for the more discrete actions on the grounds that they can always be aggregated into macro-actions later. However, many of the subsequent rules are designed to place limits on this process of reduction with regard to certain kinds of action (e.g., travel, wars, and conferences).

*Rule 23: Differentiating Actions.* WHEN THE ACTIVITIES OF AN ACTOR TOWARD AN EXTERNAL DIRECT TARGET OR INDIRECT OBJECT CAN BE DIFFERENTIATED BY COMMITMENT, TIME, OR RESOURCE-AREA, THEN EACH SUCH ACTIVITY IS A SEPARATE ACTION AND IS THE BASIS FOR A SEPARATE EVENT. A data source that provides a lengthy account of some activity or a series of itemized proposals raises the problem of separating that activity into a standardized number of actions. The three criteria in Rule 23 are based upon several meta-rules. First, actions will be treated as separate if it can be reasonably inferred that the actions result from partially or completely different decision processes. Second, actions will be treated as separate if it can be reasonably inferred that the actions were differentiated by the actors in the course of a common decision process. Third, actions will be treated as separate if it can be reasonably inferred that the initiating actors expect the actions will influence different specified nations or other entities in different ways. The first criterion refers to differences in commitment, which we infer to be intended differences in the effect of the action on different entities. If the action begins at a different time from other activity, the second criterion, then sep-



arate executing and monitoring processes are required in the initiating government. The third criterion of resource-areas is intended to identify different decision processes.

*Rule 24: Travel as Action.* TRAVEL IS CODED AS A SEPARATE ACTION ONLY WHEN THE ACTIVITIES FOR WHICH THE TRAVEL WAS UNDERTAKEN ARE NOT REPORTED. Foreign travel by governmental officials is assumed to be undertaken for some purpose and is not of interest in and of itself except when the purpose is not publicly known.

*Rule 25: Procedural and Social Activities.* PROCEDURAL ACTIVITIES ARE NOT CODED AS SEPARATE EVENTS UNLESS EXPLICIT REFERENCE IS MADE TO THEIR SUBSTANTIVE (POLICY) SIGNIFICANCE; SOCIAL AND PROTOCOL ACTIVITIES ARE SEPARATE EVENTS ONLY IF THEY ARE ATTENDED BY OFFICIALS OF TWO OR MORE COUNTRIES WHO ARE AFFORDED AN OPPORTUNITY TO TALK PRIVATELY WITH ONE ANOTHER. Procedural and social activities involving a foreign direct target or indirect object can have substantive significance, but in this coding scheme they are not recorded as events unless the opportunity for substantive interaction is demonstrated.

*Rule 26: Military Actions.* EVERY 24 HOURS THAT MILITARY CONFLICT CONTINUES, EACH CONTENDING STATE WILL BE DESIGNATED AS THE INITIATOR OF ONE EVENT FOR EACH OF THE OTHER STATES INVOLVED IN THE CONFLICT ON THAT DAY. EACH SUMMARIZING CONFLICT EVENT WILL INDICATE ANY CHANGE IN THE INTENSITY OF FIGHTING SINCE THE PREVIOUS DAY. Military conflicts often receive quite uneven coverage in data sources, but on some occasions they are reported in extensive detail. Hence, a check on the meta-rule of selecting the most discrete level of reporting is necessary to avoid confronting coders with voluminous accounts of combat activities and to avoid overweighting such actions. The rule provides a means of recording the duration of the conflict and a crude indicator of changes in intensity. Weighting devices can be added subsequently to such events if an expression of national commitment is desired. Coders are supplied with a list of continuous military conflicts for the time period they are examining.

*Rule 27: Comments by Combatants.* DURING A CONTINUING MILITARY CONFLICT, CODE COMMENTS MADE BY A COMBATANT ON THE MILITARY SITUATION AS SEPARATE ACTIONS ONLY IF THE DIRECT TARGET

IS NEITHER A PARTY TO THE CONFLICT NOR AN ENTITY WITHIN A STATE ENGAGED IN THE CONFLICT. Almost as numerous as reports of combat activities are reports of verbal charges, countercharges, denials, and so on made by parties to the conflict. We assume these add little to the intensity or significance of the military action recorded in the summary of physical encounters. Hence, they are not considered as separate events unless their direct target is not a party to the conflict. Designating noncombatants as direct targets indicates influence attempts directed at other actors in the international system, and therefore such actions are considered as separate behavior.

*Rule 28: Reiterations of Position.* WITHOUT SOME NEW DEVELOPMENT (I.E. SHIFT IN A RELATED POSITION, AN EXTERNAL INQUIRY, RELEVANT CHANGE IN THE SITUATION) REPETITION OF PREVIOUSLY ESTABLISHED POSITIONS SHOULD NOT BE CODED AS NEW ACTION. Diplomats, particularly when exposed to repeated encounters with the press, engage in considerable reiteration of positions taken previously by their government. If there is no explicitly mentioned and relevant intervening activity between repetitions, it seems unlikely that the actor's government has conducted any reexamination of the position since the last statement. Under these circumstances the reiteration is not treated as a new action.

*Rule 29: Statements about Past or Future Actions.* PREVIOUSLY UNCODED EVENTS IDENTIFIED BY MONTH AND YEAR AS FALLING WITHIN THE TIME PERIOD OF THE STUDY SHOULD BE CODED. ACTOR ANNOUNCEMENTS OF FUTURE ACTIONS CAN BE EVENTS, BUT NOT SPECULATIONS ABOUT THE FUTURE BY NONACTORS. We seek to obtain as complete a record of foreign policy events as our data sources will permit. Therefore, references to past events not previously recorded are coded if the specification of the time when the event occurred is exact enough to determine that it happened within the period under examination. Announcements by actors of intended future events can influence foreign entities, and therefore, are coded even if the announced event subsequently fails to materialize.

*Rule 30: Minimum Events in Meeting.* EVERY PARTICIPATING GOVERNMENT IN AN INTERNATIONAL MEETING IS AN ACTOR IN WHAT WILL BE ONE OR MORE EVENTS. If a government sends a representative to a meeting we assume that such behavior results from a conscious government decision, which means it will be recorded as having initiated at least one action.

*Rule 31: Short vs. Extended Meetings.* IN THE ABSENCE OF INFORMATION ABOUT THE SUBSTANCE OF AN INTERNATIONAL MEETING, A SHORT, CONTINUOUS MEETING WILL BE ONE EVENT FOR EACH PARTICIPANT WITH ALL OTHER PARTICIPANTS AS THE DIRECT TARGETS. UNDER THE SAME CONDITIONS, EACH SESSION OF AN EXTENDED, PERIODIC MEETING COMPRISES ONE EVENT FOR EACH PARTICIPANT WITH THE OTHERS AS THE DIRECT TARGETS. We assume that in any extended meeting with days or weeks between sessions the delegates consult their home governments for new instructions, thus forcing each government to review its position before every session. This consultation can happen in shorter meetings but cannot be automatically assumed.

*Rule 32: Substance in Meeting without Agreement.* WHEN SOME OF THE SUBSTANCE IS KNOWN OF A MEETING WHERE NO AGREEMENT IS REACHED, THE NUMBER OF ACTIONS DEPENDS ON EACH PARTICIPATING NATION'S POSITION ON EACH KNOWN PROPOSAL OR TOPIC. PROPOSALS ARE DIFFERENTIATED ON THE BASIS OF COMMITMENTS, TIME, AND RESOURCE-AREA (SEE RULE 23). Each known proposal on which the representatives of a government take a position at an international meeting is assumed to reflect some internal decision process within the representative's government.

*Rule 33: Incomplete Information in Meeting without Agreement.* WHEN A MEETING REACHES NO AGREEMENT AND ON A GIVEN ISSUE THE POSITION OF ONLY ONE PARTICIPANT IS KNOWN, THE ACTION FOR EACH OF THE OTHER PARTICIPANTS IS: "CONSIDERED COUNTRY X'S POSITION." This rule allows for the recording of as much substance of a meeting and the position of the participants as is known.

*Rule 34: Meetings with Known Substantive Agreements.* EACH ACCEPTED PROPOSAL IN A MEETING IS AN ACTION FOR EACH OF THE PARTICIPATING GOVERNMENTS. It is assumed that the acceptance of an agreement represents a decision by each of the participating governments.

*Rule 35: Verbal Statements with Physical Deeds.* AN ACTOR'S VERBAL STATEMENT DESCRIBING OR EXPLAINING ONE OF HIS PHYSICAL DEEDS IS NOT CODED SEPARATELY FROM THE PHYSICAL ACTION IN THE ABSENCE OF RELEVANT BEHAVIOR BETWEEN THE DEED AND THE STATEMENT BY THE DIRECT TARGET OR THE INDIRECT OBJECT. Verbal statements about physical deeds are assumed to have resulted from the same decision process unless the direct target or indirect object



takes action between the physical action and verbal behavior which can be inferred to have forced the original actor to reconsider or justify some prior physical deed with verbal behavior.

*Rule 36: Continuing Actions.* ALTHOUGH A PHYSICAL ACTION (SUCH AS THE EXECUTION OF A TRADE AGREEMENT) MAY REQUIRE AN EXTENDED PERIOD OF TIME TO COMPLETE, IT IS CODED ONLY ONCE UNLESS THE ORIGINAL EXPECTATIONS ARE BROKEN OR REVISED. Data sources occasionally report that a given activity established at some prior point is continuing. By itself such a report is assumed to be an observation by the data source that reflects no new decision activity by the initiating government.

#### IV. AN INITIAL TEST OF THE CODING RULES

The proof of the pudding, the saying goes, is in the eating. We will be the first to acknowledge that this coding scheme is not a banquet—certainly not at present anyway. There are two immediate tests the scheme must meet. First, it must guide coders to abstract events with a high degree of reliability both with each other and with the investigators' judgment of how the material should be coded. Second, it must be applicable to any data source.

In an effort to meet the first requirement we have conducted a simple reliability test. Three coders were trained for approximately four hours during which they read an earlier version of the above rules and practiced coding. Then each received a test booklet containing ten selections of approximately 100 words each from *Deadline Data on World Affairs*. Working independently, the coders abstracted events and their components from each selection. Their results were compared first with the ratings made by the two principal investigators and then with each other. All comparisons were done with a statistical measure of comparison proposed by W. S. Robinson. It should be noted that Robinson's measure of agreement is a far more stringent test of reliability than the two more commonly used tests—percentage of judgments for which there is agreement or the Pearson correlation between ratings. Like the Pearson correlation the values for the measure of agreement can range from -1.00 (complete disagreement) to 1.00 (complete agreement). Unlike the Pearson correlation, "agreement requires that paired values be identical, while

correlation requires only that paired values be linked by a linear relationship, or, if one defines correlation more broadly, that the paired values be linked according to some mathematical function. Perfect agreement has but one form,  $X_1 = X_2$ , whereas correlation may variously be written  $X_1 = a + bX_2$ ,  $X_1 = a + bX_2 + cX_2^2$ ,  $X_1 = \log X_2$ , etc. Thus agreement is a special case of correlation, since two variables that agree must be correlated, but variables which are correlated do not necessarily agree."<sup>25</sup>

The two research directors or investigators found 37 foreign policy events in the ten selections. The average measure of agreement between us and the three coders on exactly those 37 events was .92. The reliability on events among the three coders themselves was .79. The average agreement between the investigators and the coders on actors was .80 and among the coders it was .88. Between investigators and coders an average correlation of .17 was obtained for the direct target—the lowest level of average agreement computed. Among the coders themselves, the agreement on exactly the identical targets was .72. These two values suggest that the instructions for identifying the direct target meant much the same thing to all three coders, but it was not what we had intended. Clearly, this level of agreement was unacceptable and the rules for identifying direct targets were revised. The revised rules, which are those reported in the previous section, appear to have improved the reliability scores as is noted below. The average correlation of agreement for the indirect object was .63 between each coder and the investigators and was .65 among the coders themselves. Finally, on the specification of the action, the average level of agreement between coders and investigators was .86. The agreement among coders on action was .87. Table 1 reports all these values in columns 4 and 5. Furthermore, it provides the measure of agreement between each individual coder and the researchers in columns 1, 2, and 3. Further reliability tests are planned including some at another university where the coders will have no contact with the original investigators except through the revised written coding rules.

How can we determine if the operational procedures are applicable to various data sources? We have experienced difficulty in

<sup>25</sup>W. S. Robinson, "The Statistical Measure of Agreement," *American Sociological Review* 22 (February 1957), p. 19.

Table 1. Measures of Agreement for Two Data Sources

	<i>Deadline Data</i>					<i>Times of India</i>				
	Reliabilities between Coders and Investigators				Inter- Coder Reliability	Reliabilities between Coders and Investigators				Inter- Coder Reliability
	Coder 1	Coder 2	Coder 3	Average		Coder 1	Coder 2	Coder 3	Average	
Events	.96	.88	.91	.92	.79	.76	.87	.74	.79	.83
Actors	.83	.82	.95	.80	.88	.68	.94	.61	.74	.76
Direct										
Targets	.27	-.04	.29	.17	.72	.68	.92	.61	.74	.75
Indirect										
Objects	.88	.78	.22	.63	.65	.68	.61	.61	.63	.97
Actions	.81	.82	.95	.86	.87	.68	.94	.52	.71	.70



devising a satisfactory technique for exploring this second requirement. Given the variation in sources, we would not expect them to report the same events nor even the same number of events. Instead, as was noted earlier, we can expect contradictions and differences in level of event reporting between different data sources. As an initial test, we assume that the reliability scores for the same coder should not vary significantly between one data source and another. The relative stability of the reliability scores would suggest that the procedures operate equally well regardless of data source.

As a pilot test of this procedure, the same three coders were given a sample of the newspaper, *Times of India*. With the revised coding rules, but without any prior introduction to that particular data source, they coded the number of events in each front page story. The results are displayed in columns 6–10 of Table 1. As can be seen by comparing the average agreements between coders and investigators for *Deadline Data* (column 4) with those for the *Times of India* (column 9), there is considerable variation between each pair of scores. In general, coders 1 and 3 had lower scores for the *Times of India* and coder 2 had higher scores. Thus, on this initial test, the coding rules did not lead to very stable reliability scores. However, the coders were not alerted to several special problems associated with the second data source before the test. Some prior exposure to the source with an opportunity to ask questions before coding begins (as was done with *Deadline Data*) could be expected to make the pairs of scores more similar. One encouraging note is that the revised coding rules with respect to the direct target did improve the reliability scores between coders and investigators on that component.

In conclusion, it is well to recall the story with which we began about the drunk looking for his keys under the lamppost. It is too soon to tell if we will find any keys around the conceptual and operational lamppost reported in this chapter. Given the subsequent task of classifying and scaling events, perhaps the story should have a sequel about the drunk with keys in hand searching for the right door to unlock. Fortunately, other scholars have already taken some first steps on the classification and scaling problem. Therefore, the reader can be spared a second story.

*Comparative  
Foreign Policy*  
**THEORETICAL ESSAYS**

*Edited by*

**WOLFRAM F. HANRIEDER**

University of California at  
Santa Barbara

**DAVID MCKAY COMPANY, INC.**  
NEW YORK

# Contents

	INTRODUCTION .....	1
1.	HENRY A. KISSINGER, <i>"Domestic Structure and Foreign Policy"</i> .....	22
2.	J. P. NETTL, <i>"The State as a Conceptual Variable"</i> .....	51
3.	KENNETH E. BOULDING, <i>"National Images and International Systems"</i> .....	90
4.	WOLFRAM F. HANRIEDER, <i>"Actor Objectives and International Systems"</i> .....	108
5.	RUDOLPH J. RUMMEL, <i>"Some Dimensions in the Foreign Behavior of Nations"</i> .....	132
6.	DINA A. ZINNES and JONATHAN WILKENFELD, <i>"An Analysis of Foreign Conflict Behavior of Nations"</i> .....	167
7.	ROGER D. MASTERS, <i>"World Politics as a Primitive Political System"</i> .....	214
8.	WOLFRAM F. HANRIEDER, <i>"Compatibility and Consensus: A Proposal for the Conceptual Linkage of External and Internal Dimensions of Foreign Policy"</i> .....	242
9.	EUGENE J. MEEHAN, <i>"The Concept 'Foreign Policy'"</i> .....	265
10.	CHARLES F. HERMANN, <i>"What is a Foreign Policy Event?"</i> .....	295
11.	GRAHAM T. ALLISON, <i>"Conceptual Models and the Cuban Missile Crisis"</i> .....	322
12.	WAYNE WILCOX, <i>"Forecasting Models and Foreign Policy"</i> .....	385