Data Development for International Research (DDIR)
DDIR II: Event Data Research

by Rchard L. Merrit [1]
and Dina A. Zinnes
University of Illinois at Urbana-Champaign

Numerous scholars of international relations have recently sought to improve the empirical quality of their research. They feel that quantitative approaches, properly designed and applied, can significantly enhance our ability to understand international events and interactions among nation-states. One result has been a plethora of analytic techniques that rely on mathematical bases. Global modeling is an example of this direction. Another result is the generation of new data sources. This article focuses on the latter tack: data development for international research.

Growing emphasis on quantitative data has not been without problems. For one thing, some researchers flat out reject their usefulness or validity. Such intransigence obfuscates a central fact: Our growing comprehension of social scientific knowledge is linked inextricably to the computer-based information revolution. Whether we like it or not, whether we comprehend it or not, we cannot avoid their implications for political analysis. Both developments—new analytic techniques and data sources—demand greater sensitivity. Another problem is weaknesses in early data collections. We cannot deny the fact that some important datasets were flawed, just as we cannot ignore criticisms about some analytic methods researchers have used. Such weaknesses have contributed to misunderstandings, skepticism, and even occasional hostility.

This article describes a particular research project undertaken in the field of international and cross-national relations by a community of U.S.-based social scientists. The Data Development for International Research (DDIR) project seeks to maintain, extend, and develop new data banks for the study and analysis of cross-national and international political phenomena. It was the outgrowth of three years of discussion, correspondence, and seminars involving both data collectors and data users. Funding for 1986-89 by the National Science Foundation enabled the project's first phase (DDIR I) to focus on four tasks: datasets in the areas of national attributes and interstate disputes, data planning, research organization, and international broadening. New NSF funding for 1991-93 permits a second phase (DDIR II) to concentrate on the area of event data.

The article describes how DDIR began, what it has done, and where it is heading. It seeks neither to assay the often sterile debate on the usefulness of quantitative approaches, nor to offer a definitive answer to the question of what analytic techniques and data sources are appropriate for what purposes. Its concern is rather how the DDIR community envisages the status of quantitative research in international and cross-national relations. It summarizes DDIR's organizational background, philosophic orientation, and goals.

Origins: Need for Quantitative Data
Four trends in the social sciences are particularly relevant for understanding the need to develop data for international research:

* An explosion in the scientific study of national development and processes of interstate interaction has characterized the last five decades.

Questions concerning the relationship between national attributes and the domestic and foreign policy behavior of nations, the evolving structure of the international system, causes and consequences of international crises and war, and the dynamics of interstate interaction both conflictual and cooperative have come under careful and systematic scrutiny. Many of the cherished maxims of international behavior have been shown to be false; and new insights into causes and consequences of national and international processes have been observed.

* The awareness has grown that datasets are crucial within the context of the entire research process, and integral in the continuing feedback relationship between theory and research.

Contradicting the often trite argument that we allow our data to shape our questions, having large datasets that researchers know exist—and which continue to be maintained—opens up the range of research questions and continues development of theory in international relations and comparative politics.

* Funding for data development has been at best sporadic. This has meant an inability to mount a concentrated and coordinated attack on fundamental problems facing the field.

Currently existing datasets are largely the work of a few dedicated researchers scattered throughout the country, who have been dependent on the vicissitudes of changing national funding strategies. There is no guarantee that these data collections will be continued and certainly no clear opportunity for extending and further developing them in response to the evolving needs of the research community. Furthermore, while data collectors are generally aware of one another, there is no overarching mechanism to integrate and compare their results. This has led to unfortunate duplications of effort, differences in definitions, and differences in usage of sources.

• The data movement of the past several decades has enhanced the methodological expertise for the extraction of data from public sources, development of indicators for basic concepts, and quality control through reliability checks. This, together with the extensive technological advances of recent years in computer technology, makes feasible the future development of considerably more valid and reliable datasets.

These facts—the research record; recognition of the need for systematic datasets; the currently scattered, *ad hoc* nature of data collection activities; and the available methodological/technological expertise—point to the desirability of a large-scale, integrated effort that can contain, extend, and further develop the data resources available to the research community of international relations scholars.

Such perspectives on the state of the art in international and cross-national relations generated an interest in taking action to improve the field's quality. A series of informal meetings, piggybacked on to professional conferences, and workshops at the University of Illinois at Urbana-Champaign and elsewhere led to a remarkable degree of consensus among several dozen researchers. These meetings and workshops eventually honed in on a basic decision. If those interested in using quantitative data did not take action, the participants argued, then opportunities to have such data would atrophy. Accordingly, an effort to organize the relevant community of scholars was warranted and, indeed, long overdue. The researchers then focused on the overall strategy that such an organization, Data Development for International Research, should pursue: Should DDIR serve solely as an interest group, or should it encourage and seek funding for research activities? And, if the latter, which kinds of relevant research should have DDIR's initial attention?

The organizational task was easily resolved provided that some colleagues were willing to devote some of their time and energy. The point of departure for DDIR was in a sense the National Election Study project. As a large-scale, long-term data collection project for the enhancement of social science research, the NES clearly stands as a model. In another sense, however, important differences distinguish, on the one hand, the theoretical framework, goals, and structure of the NES and, on the other, the needs of the research community studying international and cross-national phenomena. The two research communities are diverse in the questions they ask, degree of consensus regarding fundamental methodological issues, and sheer number of researchers currently relying on the data collections.

This diversity suggested the need for more decentralization in the data-collection efforts and communications framework than has been needed in the NES. DDIR thus supports not a single, massive project, but rather individual researchers at different universities carrying out separate—though clearly related— projects. The diversity should be seen as a major strength of DDIR I; and it is this orientation that guides DDIR II since it also points to a multiplicity of research agendas. While the pitfalls of decentralization are apparent, these dangers do not obviate possibilities for successful coordination and integration.

The task of choosing areas for research focus proved to be more difficult simply because the potential areas are many and the competition for needed funding and other scarce resources is even greater. Some of the principle supporters, none of them with any immediate claim for DDIR-related resources, distributed questionnaires and carried out other research to ascertain how members of the potential community evaluated data priorities (McGowan *et al.*, 1988). A solicitation of research ideas, further consultation in meetings and workshops, and much telephoning eventually produced substantial if not complete agreement on a particular strategy.

The informal consensus saw three activities: First of all, DDIR would seek funding to carry out a discrete number of projects in two research areas, national attributes and interstate war, that its growing number of members considered most relevant and likely to be carried out. Second, those administering DDIR, in conjunction with an advisory committee, would assess the prospects for similar research projects in two other areas, event data and international political economy (IPE) data, which DDIR might wish to initiate later. Third, DDIR would also try to improve communications among scientists interested in quantitative research in international and cross-national relations. This meant on the one hand setting up a regular newsletter, *DDIR-Update*, and, on the other, scheduling at professional conferences both research sessions and organizational meetings

DDIR I: National Attributes and Interstate War
DDIR's first task, aimed at improving the quality of data

on national attributes and interstate war, proceeded from a rich background. Significant international and cross-national data collections were developed well before World War II (Merritt, 1990). Not until the late 1950s and early 1960s, however, did the large-scale data movement begin. As part of the general behavioral movement in political science away from assessments based on intuition or folk wisdom and toward more rigorous, systematic analyses, scholars became sensitive to their need for adequate data bases to study key questions. Is inequality in the world at large becoming more or less intense? Do alliance configurations and power distributions enhance or decrease the probability of war? Do internal domestic problems have specific effects on foreign policy behavior? To move beyond simple, impressionistic answers to such questions, it was necessary to begin collecting data on the attributes of nations and events characterizing their interactive behavior.

It is possible, in retrospect, to trace three broad data collection efforts that sought to provide the evidence necessary to facilitate the scientific study of international processes: data collections that focused on the quantitative and qualitative characteristics of (1) national attributes, (2) major conflicts and wars, and (3) interactive events within and between nations. Intriguingly, these projects all began within a year or two of one another and spread rapidly across the scientific geography of the United States and even abroad.

DDIR I: National Attributes Dimension
Historical background.

In the late 1950s and early 1960s Karl W. Deutsch at Yale University was arguing for the use of data to confirm or disconfirm hypotheses about international and cross-national politics. He had demonstrated the practicability of the search for such data and their analytic value in his studies of *Nationalism and Social Communication* (Deutsch, 1953) and *Political Community and the North Atlantic Area* (Deutsch *et al.*, 1957), and in a series of articles (most notably Deutsch, 1960, 1961) that showed how important questions were not being addressed because of the absence of valid, comparable indicators based on reliable (that is, replicable), impersonal, and quantitative data.

The year 1963 was a watershed for these innovative ideas. At the Yale Data Conference (Merritt and Rokkan, 1966) held in September international scientific researchers gathered to discuss systematic means to compare nation-states, outline organizational efforts to further such research, and learn at first hand of three major data-collection activities reaching fruition in the United States.

• **Russett et al. (1964): Yale Political Data Program.**

With financial support from the National Science Foundation, Deutsch and Harold Lasswell created the Yale Political Data Program, which in 1962, under the direction of Bruce M. Russett, had begun the cross-national collection of political, social, and economic data (see Deutsch *et al.*, 1966). Its immediate result was the *World Handbook of Political and Social Indicators*, known as "World Handbook I"; and in later years World Handbooks II and III appeared (Taylor and Hudson, 1972; Taylor and Jodice, 1983).

• **Banks and Textor (1963): Cross-Polity Survey.** Arthur S. Banks, a political scientist, and Robert B. Textor, an anthropologist, had combined forces to classify 115 polities according to 57 sets of carefully operationalized criteria.

• **Rummel (1964): Dimensionality of Nations.** At Northwestern University, initially as a component of Harold Guetzkow's Inter-Nation Simulation (INS), Rudolph J. Rummel had compiled data characterizing nation-states (see Rummel, 1979). (He also—and we shall return to this later—systematically searched the *New York Times Index* and other sources to record domestic-political and foreign-conflict events.)

Years subsequent to this burst of creativity saw three important developments. The first was the growing use of data already collected to examine theoretically interesting propositions. Second, the efforts of the 1960s were continued and expanded during the 1970s. Particularly important here were (1) Taylor *et al.*'s World Handbooks II and III, (2) Gurr's (1974, 1978) research on polities and segmental groups, and (3) the national characteristics assembled by the Correlates of War project directed by J. David Singer (see Singer and Small, 1972; Small and Singer, 1982). Third, researchers became more sophisticated in both measurement and data-collection techniques. They also met increasingly frequently to discuss their work, as well as to exchange preprints and sometimes tables of data. A "community" of quantitative international relations (QIP) scientists was emerging.

These years of an emerging QIP community were heady ones for scientific advances. Data-generators and users did not simply rest on the intellectual platforms given them in the 1950s by Karl Deutsch, Harold Guetzkow, Harold Lasswell, and others, but used them to push understanding forward. The analyses themselves were not always elegant, but nevertheless established clearly at least two points. First, "the activities of nations," in Rummel's (1966:205) words, "are highly patterned behavior," and, second, it is both possible and intellectually profitable to establish data banks to help ascertain what these behaviors are, how they are structured, and what impact this structured behavior has on such vital

issues as war and peace.

By the mid-1980s the need for aggregate data was growing but so was the reality that past data sources were aging and not being kept alive. It is thus not surprising that research scientists saw in DDIR the opportunity to resuscitate and significantly improve this field. A coordinated effort by both data generators and data users assembled a research design on national attribute data and, through DDIR, submitted it to the National Science Foundation. NSF support enabled DDIR I to carry out the following projects:

DDIR I-1. Correlates of War civil war datasets. Principal investigator: J. David Singer, University of Michigan. Updating and revalidating the Correlates of War (COW) dataset on civil wars, 1816-1988.

DDIR I-2. Correlates of War national capabilities dataset. Principal investigator: Ted Robert Gurr, then University of Colorado, Boulder, and now University of Maryland at College Park. Cooperation with J. David Singer, University of Michigan, to produce an integrated dataset, 1816-1988, on the COW project's national capability variables (population, urban population, iron/steel production, energy consumption, military expenditures, and military personnel) plus government revenues and expenditures for all states at one time members of the central-state system and, insofar as possible, peripheral systems.

DDIR I-3. Correlates of War dyadic relationships dataset. Principal investigator: J. David Singer, University of Michigan. Collaboration with Michael Wallace, then University of British Columbia and now Simon Fraser University, to update and revalidate the COW dyadic relationship dataset, 1816-1988, on shared membership in foreign alliances, diplomatic representation, and shared membership in international bodies.

DDIR I-4. Political structures dataset. Principal investigator: Ted Robert Gurr, then University of Colorado, Boulder, and now University of Maryland at College Park. For all international-system members, 1816-1985, development of a complete and updated dataset on régime characteristics, collaboration with Mark Lichbach, University of Illinois at Chicago, to transform into time-series data coding on each authority dimension.

DDIR I-5. World Handbook national attributes dataset. Principal investigator: Charles Lewis Taylor, Virginia Polytechnic Institute and State University. Expanding and filling in yearly data, 1950-1985, on readily accessible economic,

demographic, and related variables.

These projects are now for the most part complete, reports included in DDIR's newsletter, *DDIR-Update*, and the datasets sent to the Inter-University Consortium for Political and Social Research (ICPSR) in Ann Arbor, Michigan, for access to the scientific community.

DDIR I: International Conflict Dimension
Historical background.

Just as the national-development datasets had their predecessors in the many efforts of individual scholars, government agencies, and international bodies, so, too, current efforts to generate data on international conflicts can look back on a tradition of earlier projects. Though flawed, these early studies pioneered the path that modern researchers, with their richer sources and com-puter-based operations, continue to treat. Among these are four deserving particular attention:

• **Woods and Baltzley (1915): Is War Diminishing?** Frederick Adams Woods, an eminent biologist, and a young political scientist, Alexander Baltzley, provided a list of wars and their participants for most of the major European states for 1450-1900 (and back to 1100 in the cases of England and France). Chapters on each of eleven countries indicated the years of initiation and termination of national war, and hence the duration of war, and statistical graphs showed national percentages of interstate, imperial, and civil wars.

• **Sorokin (1937): Social and Cultural Dynamics.** Pitirim A. Sorokin identified "almost all the known wars" for the major European states from antiquity to 1925—including internal disturbances as well as interstate, civil, and imperial wars. He gathered the dates of initiation and termination, the war's duration for each major state, and estimates of the average army size, percentage of casualties, and total number of casualties for each state.

• **Wright (1942): A Study of War.** For 1480-1936 Quincy Wright listed balance of power, civil, imperial, and defensive wars involving each major or minor state. The dataset includes for each war its initiation and termination dates, identity of participants, their individual day of entry, and number of important battles. Wright also assembled data on the frequency and types of battles, casualties, and internal systemic disturbances.

• **Richardson (1960): Statistics of Deadly Quarrels.** Lewis Fry Richardson's compilation of conflict data includes all deadly quarrels—imperial wars, civil wars, and other forms of domestic conflict—between 1820 and 1949 that caused death of

humans. The war is the unit of analysis, and wars are organized by magnitude and then chronologically within magnitudes. The data on each war include the magnitude, dates of initiation and termination, participants, identity of the initiator, and ostensible cause or issue at stake.

Each of these data collections, of course, had problems. That by Woods and Baltzly was remarkably incomplete, especially by modern standards. Sorokin gave no explicit operational criteria for interstate, civil, and imperial wars. Wright's use of legalistic criteria for including and excluding wars was questionable. Richardson's research raises serious questions about the reliability of the data and validity of the categorical codings. But each was a significant beginning. Each, in a sense, set the groundwork for the subsequent ones (although the lateness of Richardson's publication makes it easy to ignore the fact that his research efforts were contemporaneous with those of Wright).

More to the point, however, each has been superseded by modern datasets that have not only initiated major data collections but also spawned the use of those datasets to conduct significant, empirical research on questions of war and peace. Some recent international conflict datasets are:

Singer: Correlates of War (COW). J. David Singer's intellectual focus was on the correlates of war, those factors that seemed to covary and thus be associated with the occurrence, duration, and magnitude of wars. His COW project, which began formally in 1963 under the auspices of the National Science Foundation, initially concentrated on obtaining information on the attributes of the international system that theorists had argued were the "cause" of war. COW researchers systematically culled historical texts to obtain as complete a listing as possible of all wars since 1815, together with major identifying characteristics, such as the number of participants, battle deaths, and durations (Singer and Small, 1972; Small and Singer, 1983). Subsequent data collections expanded COW's horizons—to such independent variables as population, iron and steel production, energy consumption, military expenditures, and military personnel, and, with several colleagues, to forms of international behavior, such as crises.

Siverson and Tennefoss (1982, 1984): Interstate Conflict. Unaware of COW's shift in focus and data gathering, Randolph Siverson and Michael Tennefoss independently developed a dataset on major international crises since 1815. They classified into three types the implicit/explicit level of war: threats to use force, unilateral uses of force, and reciprocated

military interactions.

Levy (1983): Great Power War. Jack Levy's dataset overlaps that of COW but extends the latter back to 1495 for the great powers. Concentrating on interstate, great-power wars (excluding civil and imperial wars) which had more than 1,000 battle deaths, it provides data on their magnitude, severity, and intensity.

Overlapping but not identical with these efforts were several other projects. Robert Butterworth (1976, 1980), Michael Brecher and Jonathan Wilkenfeld (1982), and Hayward R. Alker, Jr. and Frank L. Sherman (1982; cf. Sherman, in progress) collected data on significant attributes of international crises since World War II. Along a somewhat different but related dimension, Frederic S. Pearson (e.g., 1974) collected data on international interventions in the post-World War II period.

As was the case along the national development dimension, participants in the DDIR enterprise saw a clear need to update and expand data on the international conflict dimension. Major insights had come from the analyses based on the earlier datasets. But, as was true with respect to the national development dimension, inadequate coordination had led to duplication and incomparability. Members of the emergent DDIR community responded to the need by preparing research proposals that eventually formed a component part of DDIR's main task: research to develop adequate measures of international conflict and systematically to collect relevant data. NSF funding supported the following projects:

DDIR I–6. Great-power war dataset. Principal investigator: Jack S. Levy, then University of Texas and now Rutgers University. Collaboration with T. Clifton Morgan to revalidate and fill in missing data in Levy's dataset on participation, casualties, and initiation/termination dates for all wars among great powers, 1495-1815.

DDIR I–7. International crisis behavior dataset. Principal investigator: Jon Wilkenfeld, University of Maryland at College Park. Revalidating the International Crisis Behavior (ICB) dataset, 1929-79, and updating it through 1987.

DDIR I–8. Interstate war catalog. Principal investigator: Claudio Cioffi-Revilla, then University of Illinois at Urbana-Champaign and now University of Colorado, Boulder. Completion of a master catalog comparing (with reliability indicators) existing datasets on interstate wars (see Cioffi-Revilla, 1990).

DDIR I–9. Interstate war dataset. Principal

investigator: J. David Singer, University of Michigan. Updating for 1980-88 the COW dataset on the initiation of interstate wars, participation (in nation-months), and casualties; defining and coding additional variables for 1816-1988 on interventions by third parties, war phases, monthly casualty rates, and characteristics of war terminations.

DDIR I–10. Interventions dataset. Principal investigator: Frederic S. Pearson, University of Missouri-St. Louis. Filling in the dataset on unilateral, multilateral, and international organization interventions for 1816-1988 on interventions by third parties, war phases, monthly casualty rates, and characteristics of war terminations.

These projects are now for the most part complete, reports on most included in the DDIR's newsletter, *DDIR-Update* (and one, Cioffi-Revilla's [1990] interstate war catalog, published), and datasets sent to the ICPSR for access to the scientific community.

DDIR II: Event Data

A second, and equally important, DDIR I activity was planning future data-gathering activities on two dimensions: interstate events and international political economy (IPE). For the field of international relations to keep up with and anticipate data needs deriving from new theoretic growth requires imaginative and sustained attention to such matters as conceptualization, indicator validity, and collection procedures. DDIR's organizational goal was to hold separate sets of conferences on the two dimensions, at which active scholars would discuss needs, priorities, and procedures. The long-term hope was that conferences would produce specific research programs which could be developed for future funding.

Accordingly, with respect to the event-data dimension, planning conferences took place in May 1987 in Columbus, Ohio (Hermann, 1987), November 1987 in Cambridge, Massachusetts (Alker, 1988), and March 1990 in Chicago, Illinois. What emerged was a two-year proposal to the National Science Foundation that included researchers at seven different academic institutions who will carry out distinct but generally integrated research projects. NSF funding, awarded in January 1991, permits the realization of DDIR II. And, as in the past, the Merriam Laboratory for Analytic Political Research, located at the University of Illinois at Urbana-Champaign, serves as DDIR II's administrative umbrella.

From Political Arithmetic to Event-Data Research
Narratively oriented diplomatic historians generally view the course of international relations as a series of events—démarches, protests, treaties, crises, wars, conferences, and the like. An event in this sense is an

occurrence that stands out against the gray background of everyday living. In principle an event is a discrete unit of action, with its own beginning and ending points. In practice we often view events as nested sequences of yet smaller events. Thus an historian may view the Franco-Prussian war of 1870-71 in the light of *inter alia* Bismarck's wars against Denmark and Austria, the Ems dispatch, declaration of war, military hostilities, siege of Paris, conclusion of a peace treaty, and such consequences as indemnification, territorial transfer, and formation of the German empire; and each of these in turn comprises a congeries of lesser events. Is there another, more systematic, way to look at international events?

Analysts have devised various ways to study the events they define as important in our individual and social lives. Indeed, modern statistics finds one of its main roots in the "political arithmetic" used in the 17th century by John Graunt and William Petty to examine mortality tables. Sickness and death are individual events. And yet knowledge of how many of a society's members suffer from particular illnesses and die of particular causes tells us something about the society itself, and enables us to predict the need for medical services and the proper price for insurance. Similar considerations led Petty and other social philosophers to argue for the collection of criminal statistics (Walker, 1971; Collmann, 1973), and the occasional monarch or cabinet minister undertook a survey from time to time.

Such studies had individuals as their unit of analysis. Not until the late 19th century, with the flowering of labor unions throughout the industrialized West, did government agencies begin to gather data on social events. The target was the strike or lock-out, industrial disputes leading to stoppage of work in some firm or branch of industry. Nor is it surprising, given the general attitude then prevailing toward labor unions as a whole, that data on strikes took on the character of criminal statistics (International Labour Office, 1926). In the United States, the Department of Labor's Bureau of Labor Statistics combed newspapers and other sources to identify work stoppages, sent questionnaires to key participants to ascertain the dimensions of these events, and reported on the number of strikes, workers involved, duration, days idle, and so forth (see U.S. Department of Labor, 1976: 195-202).

The 1930s saw three major social scientific efforts to collect data on social events. The first, described earlier, focused on aspects of wars. A second was Harold D. Lasswell's (1936) intentionalist/instrumental view of politics in terms of "who gets what, when, and how." The third was Lasswell and Blumenstock's (1939) study of social unrest and world revolutionary propaganda in Chicago from 1919 to 1934. They recorded the number

and characteristics of communist-sponsored meetings, demonstrations, parades, and other social gatherings; strikes; group and individual complaints about violations of civil rights; and evictions, foreclosures, and arrests of "radicals." Lasswell and Blumenstock concluded among other things that communist propaganda was most successful during tough economic times and when it incorporated American symbolism instead of harping on Soviet accomplishments. But at the same time, by giving hardstrapped citizens an outlet to vent their frustrations and business a scapegoat to blame for the country's economic woes, communist agitation worked ultimately to deflect any truly revolutionary spirit and hence to strengthen the capitalist system.

Two decades later scholars of international relations renewed their interest in systematically studying events. One starting point was growing concern with processes of political development and the place of violence in them. Cross-national studies using data for a single year (that is, synchronic) aimed at discovering the correlates of unrest and violence; longitudinal (diachronic) studies traced patterns over time among some more limited set of countries. The nation-state was the unit of analysis. Researchers tabulated such events occurring within a state's boundaries as demonstrations, *coups d'état*, and revolutions.

Another starting point for event analysis centered on foreign-policy decision-making. Scientists conducting simulations of international processes—whether using people only, computers only, or some combination of the two—discovered they needed hard data both to feed into the simulation itself and/or to check the realism of their findings. Eventually the focus shifted from the nation-state as the unit of analysis to interactions between pairs of nation-states: ongoing processes such as trade and diplomatic exchanges as well as more or less distinct occurrences such as a threat or militarized intervention. From there it was a short step to taking seriously the new emphasis on the international system *qua* system (Kaplan, 1957) and tabulating the attributes of that system as a whole and the events taking place within it.

Still a third and doubtless the most important starting point was a growing concern with international crises and war. In the nuclear age, the possibility of war cannot be taken lightly. If analysts had had the correct tools, scientists asked, could they have recognized the probable outcome of the sequence of events in mid-1914 or in the 1930s early enough to have prevented the outbreak of war? Is there some means to ascertain when international crises are reaching the boiling point? What steps can governments take to de-escalate crises? Answers to such questions seemed to require detailed information on the course of events occurring in the global arena.

Progress in Developing Event Datasets
Initial efforts to assemble data about the events of nation-states electrified the discipline of international politics. They were, broadly speaking, of two types. First, global studies defined events of interest, specified coding rules, and, in such universal sources as the *New York Times* or *Facts On File*, coded every single occurrence of such events. (Regional studies pursued the same procedures but focused primarily on regional issues and sources.) Second, event-specific studies proceeded from the opposite direction. That is, they identified critical events of interest, such as the Suez crisis of 1956, and searched a wide variety of newspapers and historical treatises to describe, in detail, their characteristics and the chronology that preceded the key event.

Not only did these event studies set the standards that subsequent researchers would use and contend with, but they resulted in empirical studies that opened scientists' minds to new modes of research. As the data movement captured the field of international politics a series of datasets were compiled by different researchers. Limitations of one dataset for a new research question being posed led to the development of new datasets. A glance at the history of this evolution suggests at least seven major compilations.

• **Dimensionality of Nations (DON)**. Rummel, as we saw earlier, generated one of the original collections of national-attribute data. He also focused his research on interactions within and among states. He used five sources to assemble data for 1955-57 on the domestic-politics and foreign-conflict behavior of 77 nation-states (Rummel, 1964, 1967, 1972). Among other things DON tabulated the presence or absence of guerrilla warfare, number of assassinations, and seven other domestic conflict events. Rummel's thirteen foreign-conflict variables were, besides the presence or absence of military action, the number of anti-foreign demonstrations, negative sanctions, protests, countries with which diplomatic relations were severed, ambassadors expelled or recalled, diplomatic officials of less than ambassador's rank expelled or recalled, threats, wars, troop movements, mobilizations, accusations, and people killed in all forms of foreign conflict behavior.

• **World Event Interaction Survey (WEIS)**. At roughly the same time Charles A. McClelland initiated at the University of Southern California an unrelated data enterprise. This collection focused on the events, or interactions, that took place over time between pairs of countries (and in this sense was not dissimilar to the foreign conflict events coded by Rummel). WEIS consisted of a very detailed set of coding categories (63 mutually exclusive and exhaustive categories) designed to capture the type of

hostile or cooperative action that one country directed toward another, but not the intensity of hostile or cooperative behavior. Relying on reports published in the *New York Times*, McClelland and his colleagues (1971) recorded such acts in terms of initiator, target, type of act, and date of occurrence, covering the period after 1946. The extensive historical chronicle of interstate interactions that resulted made it possible to observe patterns in the activities of states and to determine the degree to which special patterns preceded major crises or wars. That such a monitoring system might facilitate forecasting of the onset of future crises was an integral part of McClelland's overall research design.

• **Conflict and Peace Databank (COPDAB).**
Edward E. Azar's particular interest in recurring Middle Eastern conflicts led him to develop a new and somewhat differently focused event dataset (see Azar, 1970, 1980a, 1980b; Azar and Sloan, 1975; Azar and Havener, 1976; Azar and Lerner, 1981). Building on earlier work by Robert C. North, Lincoln E. Moses, and their collaborators (Moses *et al.*, 1967; Choucri and North, 1975), Azar defined events as occurrences between or within nation-states that were sufficiently distinct from the constant flow of "transactions" (such as trade or mail flow) to stand out as reportable or newsworthy against this background. The coding categories were very similar to those of McClelland (see Howell, 1983; McClelland, 1983; Vincent, 1983), but the sources Azar used for coding the events went far beyond the *New York Times* to include a variety of international as well as local reporting sources.

• **Comparative Research on the Events of Nations (CREON).** Yet another important event dataset, developed by Charles F. Hermann and his colleagues (1973) primarily at The Ohio State University, sought to examine the correlates of foreign policy behavior. It focused on events that characterized different foreign policy positions of states. The coding categories were therefore somewhat different from those developed for WEIS or COPDAB. Further, since the central question concerned the relationship between certain attributes of states and types of foreign policies, extensive and costly time-series were not necessary. CREON rather provided snapshots at various points in time of the foreign policy behaviors of states.

• **World Handbook of Political and Social Indicators (1983).** In the late 1970s, Charles Lewis Taylor and David A. Jodice (1983) significantly expanded the data-gathering approaches originally developed, as noted above, at the Yale Political Data Program by Russett *et al.* (1964) and Taylor and Hudson (1972). World Handbook III provided daily event data for domestic political events only, for 136 nation-states for 1948-77. The event categories include political unrest (e.g., protests, riots), state coercive behavior (e.g., government sanctions, political executions), and governmental change (e.g., elections, executive transfers). The number of deaths from events involving domestic violence is also recorded, and additional codings for event duration, intensity, scale, and impact are included for events from 1968. World Handbook III also separately compiles for each state statistical indicators of political, economic, and social change, thus helping to define the broader context within which coded events occur.

These five event datasets, despite their apparent differences, share two important similarities. First, the definition and coding of an event are in terms of actors (national or subnational) and actions; and events are classified into a set of predetermined categories which provide descriptors of the event. Second, they pursue *global* coverage, that is, they are concerned with the entire international system.

These were not, of course, the only event datasets to emerge after the 1950s. For the *Political Instability Data Bank*, Ivo K. and Rosalind L. Feierabend (1966a, 1966b) codified 28 types of events occurring for 1955-61 in 84 countries. In his *Comparative Study of Civil Strife*, Ted Robert Gurr searched standard sources for the occurrence in 1961-68 of civil violence in 119 polities; this data collection, which he analyzed in various forms and made available to the scholarly community, provided the empirical basis for Gurr's important, prize-winning theoretic work, *Why Men Rebel* (Gurr, 1970).

Two other datasets are *event specific* and thus differ from the others in significant ways. In effect, two levels of "events" characterize these datasets. One is the identification of a key event, for example, an international crisis. The other is a minute examination in considerable detail of all preceding events, where "event" in this second instance is considerably more fine grained.

• **Behavioral Correlates of War (BCOW).** The BCOW dataset, developed by Russell J. Leng as an offshoot of the Correlates of War project, starts with Leng's earlier data on militarized interstate disputes (MID)—defined in terms of disputes in which parties on both sides threaten, display, or use military force— but focuses only on a subset of more intense disputes, called *militarized crises* (Leng and Singer, 1988). It then provides for the time period prior to each militarized crisis a fine-screened description of all events. Unique features of the BCOW coding scheme (beyond the core coding of who does or says what to whom and when) include: location of each event;

duration and variations in intensity of multi-day events; assignment of physical events to one of 103 categories of military, diplomatic, economic, or unofficial behaviors; and detailed analysis of sequential verbal interactions (allowing identification of bargaining strategies). This fine-grained coding of verbal actions allows for a detailed analysis of interstate bargaining and the development of an "hierarchical choice tree."

• **SHERFACS.** Using criteria to select and merge conflict cases from the FACS dataset (Farris, Alker, Carley, and Sherman, 1980) and nearly 40 other studies, Frank L. Sherman's SHERFACS produced a combined file of 730 international disputes and 980 domestic quarrels that provide data on, among other things, the identification of conflict phases, means of referrals to management agents, and nature of actions taken by all parties (see Alker and Sherman, 1982; Sherman, 1987a, 1987b, in progress). Sherman then developed a phase structure for domestic quarrels similar to the CASCON structure for international conflicts.

Some related datasets were mentioned earlier: Butterworth (1976), Brecher and Wilkenfeld (1982), and Pearson (1974). Then, too, empirical studies of conflict management, such as SHERFACS, have a rich tradition: Ernst B. Haas's (1968) disputes referred to the United Nations for management, Joseph S. Nye's (1971) added conflicts referring similarly to regional international organizations, the joint effort by Haas, Robert Butterworth, and Nye (1972) added to the existing set three new types of conflicts—interstate disputes in which some kind of international organization, e.g., the United Nations Security Council, sought involvement; civil strife in which one side of the dispute enjoyed the support of another government; and "non-managed" interstate conflicts in which fatalities occurred—and the CASCON phase structure developed by Bloomfield and Leiss (1969).

These event-data projects saw enormous use by scholars. This was particularly the case with Azar's COPDAB, which continued until 1979 to collect data, and, like other event-data collections, was made generally available to users. But these projects—and hence the fundamental idea underlying event datasets—also came under fire by critics, both friendly and hostile. Complaints ranged from the usefulness of particular sources, such as the *New York Times*, to the modes of categorizing the data. The level of hostility had multiple effects. It diminished funding and shifted intellectual concerns. It discouraged previous and emerging event-data researchers from either undertaking new collections or updating the old ones. The scientific progress of the 1960s soon began to languish. But, at the same time, challenging the past

value and uses of event data encouraged researchers to spend time thinking through various dimensions of previous projects, exploring new ideas, and, particularly, adapting their research plans to take advantage of modern computational capabilities.

DDIR II: Developing New Event-Data Research
DDIR's three event-data conferences sought first of all to assess the state of the art, then to review new data priorities, and finally to develop an effective research strategy. Several considerations shaped a decision to pursue a mixed strategy: the need to (1) generate a rich and general, core dataset; (2) improve the capabilities of key specialized event datasets that already exist; (3) enhance software so as to minimize the time and cost of expanding datasets in the future; and (4) explore the possibilities for new styles of event-data research.

Enhancing existing and generating new event datasets.
If we are to enhance the quality and quantity of some existing datasets, which ones should they be? Our survey of the literature (McGowan *et al.*, 1988) together with a study of each event dataset's time-span and comprehensibility across a wide range of theoretically interesting issues strongly suggested a central focus on the COPDAB file. Not the least reason for this is the fact that, of the five global event datasets—DON, WEIS, COPDAB, CREON, and World Handbook III—COPDAB best met the combined criteria of past scientific usage, availability over a long time series, and attention to a broad range of new styles of computer-aided, event-data research (Starr, 1987). Other factors included COPDAB's compatibility with case-oriented datasets (most notably BCOW and SHERFACS), the needs of those initiating regional event datasets, COPDAB's appropriateness for testing new software, and, by no means least significantly, the fact that the Center for International Development and Conflict Management (CIDCM) at the University of Maryland at College Park was planning to update and expand the COPDAB dataset.

Thus the Global Event-Data System (GEDS) project at Maryland became the natural focal point for organizing DDIR II's core data-generation part. The CIDCM's research team will establish GEDS for computer-assisted identification, abstracting, and coding of daily international and domestic events, as reported primarily in comprehensive, on-line news sources such as the Reuters news service. GEDS thus aims at developing a core event-data stream from 1979 forward. It will include:

• the actions *vis-à-vis* each other of (1) nation-states, (2) major nonstate communities, and (3) international organizations,

• detailed event summaries and coding, including

direct quotations and cross-referencing, and

• information allowing users to access those full-text source articles which are available on-line.

GEDS software will permit partially automated, continuous updating after 1990 of the core event-data stream. In the discussion that follows, the term GEDS refers to the event-data stream generated by using Maryland's computer-assisted coding procedures on on-line news sources. Each of the projects described below produces a specialized dataset based on GEDS.

DDIR II–1. University of Maryland: Updating and Extending Existing Datasets. As part of its larger GEDS effort, the Maryland team—John L. Davies, Ted Robert Gurr, and Chad K. McDaniel—will update to 1990+ the existing COPDAB dataset, and incorporate updated WEIS and, as they become available, World Handbook III (and BCOW and SHERFACS) event data. The updated dataset will be compatible with each of these previously-coded datasets, but expanded to include new foci (e.g., inclusion of nonstate actors) and new sources made available through computer-assisted coding.

DDIR II–2. American University: Foreign Policy Behaviors of Southeast Asian States (SAS). Llewellyn D. Howell will use the GEDS computer-assisted procedures on regional sources to produce a data bank on 10 Southeast Asia states. The SAS event-data stream, to be added to the Maryland core event-data stream, will thus enrich the latter and provide a check on the comparability of global sources vs. regional sources.

DDIR II–3. University of Kansas: Kansas Event-Data Sources (KEDS) for Central Europe and the Middle East. Philip A. Schrodt, Ronald A. Francisco, and Deborah J. Gerner have two tasks. First, they will extend their existing software for automated coding. Using the GEDS files as inputs, the current software automatically generates WEIS-coded data. Resources permitting, the software can be expanded to produce COPDAB-coded data. Second, the Kansas team will assemble a high-density, international, event dataset for Central Europe and the Middle East. It uses specialized journals and government publications around the world to increase regional coverage without the time and expense involved in working with regional journalistic sources such as newspapers. Like Howell's SAS project, the use of regional sources will provide the basis for comparing alternative, global vs. regional sources of events.

DDIR II–4. Middlebury College: Behavioral

Correlates of War (BCOW). For 40-55 militarized crises occurring in 1979-90, and starting with the core data provided by GEDS, Russell J. Leng will apply BCOW data-collection procedures to produce a fine-screened dataset. The BCOW coding manual specifies as many as 103 descriptors of each action (such as alert, mobilization, or evacuation) that could take place during a militarized crisis. Each such event action is categorized according to the date of occurrence, actor, target, location, whether the actor was acting unilaterally or with another state, and "tempo" of the action.

DDIR II–5. Miami University: Nonstate Actors in Interstate Conflicts (SHERFACS). Frank L. Sherman at Miami University of Ohio will enhance and bring up to date the SHERFACS dataset, which comprises fine-screened accounts of several kinds of episodic conflict situations. Inclusion is global, but limited to international conflicts and domestic quarrels, especially those involving collective management (e.g., UN mediation) and nonstate actors. The expanded event summaries generated by GEDS will increase the number of international conflicts and domestic quarrels that will be coded using the SHERFACS template. And, like the BCOW dataset, the SHERFACS dataset will augment the analytic capabilities inherent in the expanded COPDAB dataset to be developed by CIDCM at the University of Maryland.

DDIR II–6. Massachusetts Institute of Technology: Data Development for Interpretive Analysis. Hayward R. Alker, Jr., at MIT, will develop methods for the interpretive analysis of detailed event summaries by adding narrative depth and varieties of interpretive perspectives for specific conflict episodes in the GEDS dataset. The three data components to be studied are (1) explicitly coded WEIS/COPDAB/BCOW/ SHERFACS event data, (2) humanly constructed narrative summaries of each event, and (3) quotations attributed to principal actors/ interactors of the event being described. In addition, original and secondary source stories will be made conveniently accessible, possibly as part of each record, for the purposes of detailed textual and interpretive analysis of both quantitative and qualitative, political data.

These various data-collecting activities can significantly improve the quality of research in the field of quantitative and textual international politics. First, they will bring up to date and expand the more important event datasets identified by publications and by quantitative and textual scientists. Second, they will provide procedures for routinizing future such event-data collections. This will sharply reduce the need to turn to funding

agencies every five years or so in the search for new support to update the datasets. Third, they aim at achieving an integrated event dataset. Interaction among the principal investigators through DDIR's ægis can ensure that interchangeable datasets are in the public domain. Fourth, the coordinative thrust nevertheless permits maximum flexibility among these principal investigators to carry out their individual research strategies.

Software Developments to Aid Data Collection and Analysis.

Recognizing the need for a core data-collection effort such as GEDS was only one step. Researchers in recent years also began to appreciate the important role that computerized methods could play. With major international news sources, such as Reuters, Associated Press, and United Press International, as well as local news reports (as translated, for instance, by the *FBIS Reports*) either now or soon to be accessible on-line, the retrieval of source stories begs for automation. Moreover, the enormously expanded storage capacity, processing speed, and programming flexibility at the microcomputer level now makes it possible to develop an event-coding system which sacrifices neither the comprehensiveness of global coding efforts nor the depth and diversity of coverage of the episodic coding projects.

DDIR II proceeds from the conviction that the development of computerized methods for the *collection of data* is not only a desirable but a necessary innovation. It includes several projects in this area:

DDIR II–7. University of Maryland: Computer-Assisted and Partially-Automated Coding in GEDS. With a grant from DDIR and backing from their institution, the Maryland team has developed and tested a preliminary version of software for computer-assisted entry, coding, and editing of Reuters on-line source stories to produce GEDS event records. As a significant product of its software development, the team will set in place at CIDCM a process for continuously coding GEDS records.

DDIR II–8. University of Kansas: Partially-Automated Procedures for the KEDS Machine Coding Systems. The KEDS machine-coding systems will be enhanced to permit continued development of event-data generating software, which will use inexpensive, machine-readable data sources and personal computers. The KEDS-X rule-based coding system will (1) add a practical English parser to handle grammatical tasks associated with text analysis, (2) experiment with non-English source text, and (3) implement a parallel processing network for increased coding speed. Schemes for coding time-dependent datasets, such as BCOW, will also be

explored. The software developed at Kansas will provide inexpensive, up-to-date, and easily customized datasets on international and domestic conflict and cooperation, and will also aid in developing the partially automated coding software being written by the Maryland team.

In addition, machine-assisted coding procedures will be implemented by two other projects. Howell's SAS project will make extensive use of the computer-assisted (and ultimately partially automated) methods that the Maryland team will develop. Some of these methods are even now in use in the SAS project. In addition, Leng's BCOW project will use machine-assisted coding software recently developed as a part of that project. This software is specifically designed to use as input for the detailed data records produced by the GEDS project.

The software component of DDIR II also focuses on *software for data analysis.* Included are four projects at the participating institutions as well as an evaluation to be carried out in Illinois:

DDIR II–9. Massachusetts Institute of Technology: Computerized Textual and Interpretive Analysis of Conflict Episodes. Alker is exploring software development for the interpretive analysis of event histories. This will allow subsequent validity- and reliability-oriented comparisons of original sources, GEDS codings, human narrative summaries, speech fragments, and such computational interpretations as would be produced. Central to redefining available software routines for computational text analysis in the Schank-Abelson tradition are developing and implementing an "event description framework" motivated by Lasswell's work on interactions, and a translation scheme for "filling in" this framework using, in particular, SHERFACS data. The interpretive routines would then operate on this framework to produce event interpretations computationally.

DDIR II–10. Middlebury College: Extension of Computerized Procedures for the Analysis of BCOW Data. Leng is modifying and enhancing two currently existing software packages developed for analyzing BCOW data. Because of the richness of BCOW coding categories, software is the only efficient way for aggregating the data for subsequent analyses. One program, CRISIS, permits users to select, count, and scale events along various dimensions. Another, INFLUENCE, is designed specifically for analyzing crisis bargaining. Both programs currently exist only in the environment of a (VAX) mini-computer, and the goal is to increase their functionality and availability by converting them to microcomputer environments.

DDIR II–11. Miami University: Computerized Preparation of SHERFACS Data for Interpretive Analysis. Sherman will also explore means to fit the SHERFACS coding schema into Lasswellian frames, which Alker proposes to use for interpretively describing conflict episodes. Computer-assisted or partially automated coding sequences are needed to transform into Lasswellian categories the SHERFACS information (and, by extension, the associated event summaries and event categories of GEDS). The software will be compatible with the GEDS data-collection system.

DDIR II–12. University of Maryland: GEDS User Software. The Maryland team will develop GEDS end-user software for browsing, data selection, temporal and spatial aggregation, graphic display, and to interface with related databases with full-text sources as well as statistical and interpretive software packages.

The Merriam Lab is considering the possibility of enhancing the utility of software developed by the various projects. For example, its numerous computer language compilers (e.g., C, Pascal, Lisp) for several different operating system environments (e.g., IBM, Macintosh, UNIX) are available for coordination tasks; and it can develop simple macros designed to link processing across the different executables so as to reduce the amount of time needed by users to perform multiple research tasks.

Though focusing primarily on data collection, DDIR II can creatively enhance software facilities that expand the usage of such data. To some measure it banks on enhanced hardware and software technologies. An ideal and very "futuristic" automated system for handling unstructured data would provide multiple interpretations of one unstructured data stream—just as ordinary citizens, political activists, and scientists working within different research traditions while looking at the same ordinary language texts might draw different interpretations. While several experimental parsers already exist, more basic research is needed before they can become reliable components of a data development infrastructure. Although a multiple-interpretive parser for ordinary language text will probably not be available for some time, we recognize the need to anticipate future technological advances in the more modest coordination outlined here. Future technological developments undertaken by other researchers will eventually permit some further extensions such as semi-automated technologies for processing unstructured, that is, ordinary-language, text.

DDIR II itself can also contribute to enhancing the hardware and software technologies that are needed. It is also essential, however, to look more closely at the degree to which coding judgments stray from case-study level understandings. The Merriam Lab will thus include some general comparisons across the basic event datasets (COPDAB, WEIS, BCOW, and SHERFACS) to assess their relative validity against original source texts regarding, say, the crisis leading up to the Persian Gulf war. The point is *not* that these datasets are invalid, but rather that their quality will reflect coders' perceptions, and that, therefore, independent analysts would have to take this fact into account in using the data for their own research.

Toward the Future: DDIR III on International Political Economy Data

About a dozen years ago, international relations scholars rediscovered the importance of international political economics (IPE). It had of course remained alive and well in some quarters, particularly in Great Britain where the field of political economy was nurtured some two centuries ago. But it tended to interest economists, not political scientists, just as such issues as social change in developing countries tend to interest sociologists. Political scientists, even those concentrating their studies on international relations, by and large treated economic considerations as peripheral to the main struggle for national power and global order. Especially in recent decades the main thrust of their scholarship and instruction had been power politics, with its emphasis on military security, East-West confrontations, and guiding the political development of new nation-states. The long-standing tradition of political economy paled in the perspectives of all but a few of those who were shaping the post-1945 directions of international political research.

The renewed interest in IPE caught empirical researchers in a state of acute embarrassment. As we have seen, QIP scientists had focused on national characteristics, conflict, events, and a wide variety of other topics. By the end of the 1970s, when they looked in the larder of systematically evaluated IPE data, they found the cupboard bare.

A curious sequence of events then took place. The availability of IPE-related data from the United Nations and other agencies posed a delicious dilemma. On the one hand, a wide variety of such data sources existed but, on the other, they were of mixed quality for the type of analysis conducted by QIP scientists. The data were not always compatible, nor did they address some of the key questions relating to the broad domain of IPE research. This led to dismay in some circles. Perhaps scientists had become too accustomed to readily available, reliable, and paradigmatically similar data from such agencies as the ICPSR, the European Consortium for Political Research (ECPR), and the Zentralarchiv für empirische

Sozialforschung at the University of Cologne. The absence of any comparable storehouse of IPE data may have led these scientists to ignore the fact that such rich data sources were not the product of a single day's labor.

The virtual lack of appropriate data had different consequences in other circles. Some researchers—possibly following Admiral David Farragut's injunction during the American Civil War to "Damn the torpedoes: Full speed ahead!"—wrote treatises based on existing data sources, however disparate they may have been. The predictable result was sharp criticism from their colleagues, and especially from those who were fundamentally disposed to favor data-based research. (Those opposed to the basic idea of such research merely found their predilections confirmed!) Viewed from a more distant perspective, such studies could be described as courageous but flawed efforts to make sense out of a complicated field. Still other researchers explored the means to generate new, more sophisticated IPE data (Bornschier and Heintz, 1979; Groenick, 1988; Müller, 1988). What they quickly discovered is that such reliable data, especially those encompassing long time series, are as scarce as the proverbial hen's teeth. This discouraged the faint of heart. The result was that, though many researchers called for better IPE data, few proved willing, in the words of the famous American challenge, to put their money were their mouths were.

The DDIR community held two workshops—in October 1987 in New Haven, Connecticut (Russett, 1988), and April 1988 in Tempe, Arizona (McGowan, 1988; Pollins, 1988)—to address three questions about data important for studying political dimensions of international economic transactions.

• What is the current status of IPE data?

Of particular importance are their availability and quality, and differences among datasets generated by national and the international institutions, commercial firms, and university research institutes. The concern is a very pragmatic one: To what extent can QIP researchers interested in a broad range of IPE issues actually use existing datasets?

• What IPE data do active researchers need?

Two issues are problematic here. First, given unlimited resources, including funding, computational facilities, and qualified research assistance, which datasets are the most significant in terms of probable intellectual or scientific payoff? Second, given the fact that such resources are not unlimited, how can we prioritize among competitive claims of significance?

• How can we enhance IPE data development?

The assumption sometimes seems to be that desired datasets will drop from the clear blue sky. To the contrary, they must be developed. The question thus focuses on two issues—especially in an international framework. One is, How can we enhance institutional arrangements to facilitate data development? The other is, How can we support or persuade leading IPE researchers to take on leadership roles in these endeavors?

DDIR's plan to initiate a third research phase on IPE data remains in its pre-planning stage. An earlier effort to organize a team of researchers interested in generating data programs proved to be premature. The reason for this may have been simply that scientists invited to participate were too involved in other projects to undertake new, time-consuming ones. It may also be that the most active IPE researchers view their own roles as chiefs rather than braves, as theoreticians willing to recommend and eventually to use improved data files rather than practitioners willing to dig out the data. But, whatever the cause, the result is that any DDIR effort to encourage IPE data collections will require renewed vigor. In the meantime, word of mouth and conversations at professional meetings have revealed a number of younger and perhaps less well-known scientists with a keen interest in assembling new data collections so that they can use them for their own research. This suggests a revised DDIR strategy. It should doubtless solicit requests for proposals for IPE data programs, first to ascertain the extent to which the community of IPE scientists is interested in undertaking data-gathering activities and, second, if this proves to be the case, to work out joint procedures to coordinate these activities and seek funding.

A key element of a projected DDIR III will be the internationalization of any joint data-gathering activities. DDIR I and II have been directly related to datasets generated and carried out predominantly in the United States. It thus made sense to seek initial funding from the U.S. National Science Foundation. In the future, of course, given the international response to data on national capabilities, interstate conflict, and international events, we may expect more data-gathering activities to emerge in other countries. Accordingly, it will make sense to enhance international collaboration and seek international funding. These conditions already exist in the field of IPE, for both data-producers and data-users; and, indeed, the most significant IPE datasets to be created in recent years came from West Europe (Groenink, 1988; Müller, 1988). Going it alone, either for individual researchers or those at a single country's academic institutions, may continue to be feasible but is not the best research strategy.

Clearly, international collaboration is needed. In April 1989 a study group on QIP data was established within the framework of the International Political Science Association (IPSA). IPSA's 15th World Congress, to be held in Buenos Aires, Argentina, in July 1991, provided an opportunity for the study group to hold sessions on IPE data development and data uses. Letters sent to several dozen U.S. and foreign scientists, however, found virtually no response—and only one expression of interest in participating in such a session (and this a U.S. scientist). Establishing the basis for better international cooperation appears to be something yet in the future.

The scientific field of international political economy is clearly in an exciting state of flux. While it is burgeoning in an intellectual sense, its data needs continue to be substantial. Governmental and nongovernmental agencies create many datasets, of course, but, for theoretic research carried out at academic institutions, these clearly need assessment to ascertain their value and sometimes much reworking to ensure consistency across time and space. An increasing number of scientists working in the field has recognized the need for IPE data to carry out their research activities. Also important is the fact that some of these scientists express interest in improving existing datasets and/or generating new ones. Multi-institutional and multinational organizations can facilitate such research activities. If DDIR's current organizational efforts can be carried out—or modified so that they function more effectively—the prospect is for a new era of data-based research on IPE that can significantly address important human issues.

References

Alker, Hayward R., Jr. (1988) "Second (MIT-CIS/) DDIR conference on event data." *DDIR-Update* 2:2 (January), 2-5.

—— and Frank L. Sherman (1982) "Collective security-seeking practices since 1945," pp. 113-145 in *Managing International Crises*, ed. Daniel Frei. Beverly Hills, Calif., London, and New Delhi: Sage Publications, Inc.

Azar, Edward E. (1970) "Analysis of international events." *Peace Research Reviews* 4,1 (November), 1-113.

—— (1980a) *Codebook of the Conflict and Peace Data Bank (COPDAB)*. Ann Arbor, Mich.: Inter-University Consortium for Political and Social Research.

—— (1980b) "The Conflict and Peace Data Bank (COPDAB) project." *The Journal of Conflict Resolution* 24:1 (March), 143-152.

—— and Thomas N. Havener (1976) "Discontinuities in the symbolic environment: a problem in scaling." *International Interactions* 2:4, 231-246.

—— and Steve Lerner (1981) "The use of semantic dimensions in the scaling of international events." *International Interactions* 7:4, 361-378.

—— and T. J. Sloan (1975) "Dimensions of interaction: a sourcebook for the study of the behavior of 31 nations from 1948 through 1973." International Studies Association, Occasional paper #8.

Banks, Arthur S. (1971) *Cross-Polity Time-Series Data*. Cambridge, Mass., and London: The M.I.T. Press.

—— and Robert B. Textor (1963) *A Cross-Polity Survey*. Cambridge, Mass.: The M.I.T. Press.

Bloomfield, Lincoln P. and Amelia C. Leiss (1969) *Controlling Small Wars: A Strategy for the 1970's*. New York: Alfred A. Knopf.

Bornschier, Volker and Peter Heintz (reworked and enlarged by Thanh-Huyen Ballmer-Cao and Jürg Scheidegger) (1979) *Compendium of Data for World-System Analysis: A Sourcebook of Data Based on the Study of MNCs, Economic Policy and National Development*. Zürich: University of Zürich, Sociological Institute.

Brecher, Michael and Jonathan Wilkenfeld (1982) "Crises in world politics." *World Politics* 34:3 (April), 380-417.

Butterworth, Robert Lyle (1980) "Managing interstate conflict, 1945-79: data with synopses." Final Report, February.

—— with Margaret E. Scranton (1976) *Managing Interstate Conflict, 1945-74: Data with Synopses*. Pittsburgh, Pa.: University of Pittsburgh, University Center for International Studies.

Choucri, Nazli and Robert C. North (1975) *Nations in Conflict: National Growth and International Violence*. San Francisco, Calif.: W. H. Freeman.

Cioffi-Revilla, Claudio (1990) *The Scientific Measurement of International Conflict: Handbook of Datasets on Crises and Wars, 1495-1988 A.D.* Boulder, Colo., and London: Lynne Rienner Publishers.

Collmann, Hans-Jürgen (1973) *Internationale Kriminalstatistik: Geschichtliche Entwicklung und*

gegenwärtiger Stand. Stuttgart: Ferdinand Enke Verlag.

Deutsch, Karl W. (1953) *Nationalism and Social Communication: An Inquiry into the Foundations of Nationality.* Cambridge, Mass: The Technology Press of The Massachusetts Institute of Technology; and New York: John Wiley & Son, Inc.

—— (1960) "Toward an inventory of basic trends and patterns in comparative and international politics." *The American Political Science Review* 54:1 (March), 34-57.

—— (1961) "Social mobilization and political development." *The American Political Science Review* 55:3 (September), 493-514.

——, Sidney A. Burrell, Robert A. Kann, Maurice Lee, Jr., Martin Lichterman, Raymond E. Lindgren, Francis L. Lowenheim, and Richard W. Van Wagenen (1957) *Political Community and the North Atlantic Area: International Organization in the Light of Historical Experience.* Princeton, N.J.: Princeton University Press.

——, Harold D. Lasswell, Richard L. Merritt, and Bruce M. Russett (1966) "The Yale Political Data Program," pp. 81-94 in *Comparing Nations: The Use of Quantitative Data in Cross-National Research,* ed. Richard L. Merritt and Stein Rokkan. New Haven, Conn., and London: Yale University Press.

Farris, Lee, Hayward R. Alker, Jr., Kathleen Carley and Frank L. Sherman (1980) "Phase/Actor Disaggregated Butterworth-Scranton Codebook." Cambridge, Mass.: The Massachusetts Institute of Technology, Center for International Studies, working paper.

Feierabend, Ivo K. and Rosalind L. Feierabend (1966a) "Aggressive behaviors within polities, 1948-1962." *The Journal of Conflict Resolution* 10:3 (September), 249-271.

—— and —— (1966b) "The relationship of systemic frustration, political coercion, international tension and political instability: A cross-national study." Paper prepared for delivery at the Annual Meeting of the American Psychological Association, New York City, 2-6 September.

Groenick, Ronald J (ed.) (1988) *Data on Europe, 1945-1980.* Bilthoven: Prime Press.

Gurr, Ted Robert (1970) *Why Men Rebel.* Princeton,

N.J.: Princeton University Press.

—— (1974) "Persistence and change in political systems, 1800-1971." The American Political Science Review 68:4 (December), 1482-1504.

—— and Associates (1978) *Comparative Studies of Political Conflict and Change: Cross-National Dataset.* Ann Arbor, Mich.: Inter-University Consortium for Political and Social Research.

Haas, Ernst B. (1968) *Collective Security and the Future International System.* Denver, Colo.: University of Denver, Monograph Series in World Affairs, vol. 5, no. 1.

——, Robert L. Butterworth and Joseph S. Nye (1972) *Conflict Management by International Organizations.* Morristown, N.J.: General Learning Press.

Hermann, Charles F. (1987) "First DDIR conference on event data." *DDIR-Update* 1:5 (July), 4-5.

——, Maurice A. East, Margaret G. Hermann, Barbara G. Salmore and Stephen A. Salmore (1973) *CREON: A Foreign Events Data Set.* Beverly Hills, Calif., and London: Sage Publications, Inc., Sage Professional Papers in International Studies Series, No. 02-024.

Howell, Llewellyn D. (1983) "A comparative study of the WEIS and COPDAB data sets." *International Studies Quarterly* 27:2 (June): 149-159.

International Labour Office (1926) *Methods of Compiling Statistics of Industrial Disputes.* Geneva: International Labour Office, Studies and Reports, Series N (Statistics), No. 10.

Kaplan, Morton A. (1957) *System and Process in International Politics.* New York: John Wiley & Sons.

Lasswell, Harold D. (1936) *Politics: Who Gets What, When, How.* New York and London: McGraw-Hill Book Company, Inc., and Whittlesey House.

—— and Dorothy Blumenstock (1939) *World Revolutionary Propaganda: A Chicago Study.* New York and London: Alfred A. Knopf, Inc.

Leng, Russell J. and J. David Singer (1988) "Militarized interactive crises: the BCOW typology and its applications." *International Studies Quarterly* 32:2 (June), 155-173.

Levy, Jack S. (1983) *War in the Modern Great Power System, 1945-1975.* Lexington: University of Kentucky Press.

McClelland, Charles A. (1961) "The acute international crisis." *World Politics* 14:1 (October), 182-204.

—— (1983) "Let the user beware." *International Studies Quarterly* 27:2 (June), 169-177.

——, Rodney G. Tomlinson, Ronald G. Sherwin, Gary A. Hill, Herbert L. Calhous, Peter H. Fenn and J. David Martin (1971) *The Management and Analysis of International Event Data: A Computerized System for Monitoring and Projecting Event Flows.* Los Angeles, Calif.: University of Southern California, School of International Relations (September), mimeographed.

McGowan, Patrick J. (1988) "Second DDIR international political economy data conference." *DDIR-Update* 2:3 (April), 4-8

——, Harvey Starr, Gretchen Hower, Richard L. Merritt and Dina A Zinnes (1988) "International data as a national resource." *International Interactions* 14:2, 101-113.

Merritt, Richard L. (1990) "Data in international research: Confluence of interest and possibility." *DDIR-Update* 4:3 (April), 1-11.

—— and Stein Rokkan (1966) *Comparing Nations: The Use of Quantitative Data in Cross-National Research.* New Haven, Conn., and London: Yale University Press.

Moses, Lincoln E., Richard A. Brody, Ole R. Holsti, Joseph B. Kadane and Jeffrey S. Milstein (1967) "Scaling data on inter-nation action: A standard scale is developed for comparing international conflict in a variety of situations." *Science* 156:3778 (26 May), 1054-1059.

Müller, Georg P., with the collaboration of Volker Bornschier (1988) *Comparative World Data: A Statistical Handbook for Social Science.* Baltimore and London: The Johns Hopkins University Press.

Nye, Joseph S. (1971) *Peace in Parts: Integration and Conflict in Regional Organization.* Boston, Mass.: Little, Brown and Company.

Pearson, Frederic S. (1974) "Geographic proximity and foreign military intervention." *The Journal of Conflict Resolution* 18:3 (September), 432-460.

Pollins, Brian M. (1988) "IPE data: a preliminary survey." *DDIR-Update* 2:3 (April), 8-12.

Richardson, Lewis Fry (1960) *Statistics of Deadly Quarrels*, ed. Quincy Wright and C. C. Lineau. Pittsburgh, Pa.: Boxwood Press; and Chicago, Ill.: Quadrangle Books.

Rummel, Rudolph J. (1964) "Dimensions of conflict behavior within and between nations," pp. 1-50 in *General Systems: Yearbook of the Society for General Systems Research*, vol. 8, ed. Ludwig von Bertalanffy and Anatol Rapoport. Ann Arbor, Mich.: Society for General Systems Research.

—— (1966) "A foreign conflict behavior code sheet." *World Politics* 18:2 (January), 283-296.

—— (1967) "Dimensions of dyadic war, 1820-1952." *The Journal of Conflict Resolution* 11:2 (June), 176-183.

—— (1972) *The Dimensions of Nations.* Beverly Hills., Calif., and London: Sage Publications, Inc.

—— (1976) "The roots of faith," pp. 10-30 in *In Search of Global Patterns*, ed. James N. Rosenau. New York: The Free Press.

—— (1979) *National Attributes and Behavior: Data, Dimensions, Linkages and Groups, 1950-1965.* Beverly Hills, Calif., and London: Sage Publications, Inc.

—— and Hayward R. Alker, Jr., Karl W. Deutsch, and Harold D. Lasswell *et al.* (1964) *World Handbook of Political and Social Indicators.* New Haven, Conn., and London: Yale University Press.

Russett, Bruce M. *et al.* (1988) "First DDIR international political economy data conference." *DDIR-Update* 2:3 (April), 1-4.

Sherman, Frank L. (1987a) "Four major traditions of historical events research: a brief comparison." Paper presented at the Second DDIR Event Data Conference, Boston, Mass., The Massachusetts Institute of Technology, 13-15 November.

—— (1987b) *Partway to Peace: The United Nations and the Road to Nowhere.* State College, Pa.: Ph.D. dissertation, The Pennsylvania State University.

—— (in progress) *Recognizing and Responding to*

International Disputes.

Singer, J. David and Melvin Small (1972) *The Wages of War, 1816-1965: A Statistical Handbook.* New York: John Wiley & Sons.

Siverson, Randolph M. and Michael R. Tennefoss (1982) "Interstate conflicts, 1815-1965." *International Interactions* 9:2, 147-178.

—— and —— (1984) "Power, alliance, and the escalation of international conflict, 1815-1965." *The American Political Science Review* 78:4 (December), 1057-1069.

Small, Melvin and J. David Singer (1982) *Resort to Arms: International and Civil War, 1816-1980.* Beverly Hills, Calif., London, and New Delhi: Sage Publications, Inc.

Sorokin, Pitirim A. (1937) *Social and Cultural Dynamics,* vol. 3: *Fluctuation of Social Relationships, War, and Revolution.* New York: American Book Company.

Starr, Harvey (1987) "What is a national resource in international data?" *DDIR-Update* 1:3, app. (February), 5-6.

Taylor, Charles Lewis and Michael C. Hudson (1972) *World Handbook of Political and Social Indicators* (2d ed.) New Haven, Conn., and London: Yale University Press.

Taylor, Charles Lewis and David A. Jodice (1983) *World Handbook of Political and Social Indicators* (3d ed.). New Haven, Conn., and London: Yale University Press.

United States, Department of Labor (1976) *BLS Handbook of Methods.* Washington, D.C.: U.S. Department of Labor, Bureau of Labor Statistics, Bulletin No. 1910.

Vincent, Jack E. (1983) "WEIS vs. COPDAB: correspondence problems." *International Studies Quarterly* 27:2 (June), 160-168.

Walker, Nigel (1971) *Crimes, Courts and Figures: An Introduction to Criminal Statistics.* Harmondsworth, Middlesex: Penguin Books Ltd.

Woods, Frederick Adams and Alexander Baltzly (1915) *Is War Diminishing? Study of the Prevalence of War in Europe from 1450 to the Present Day.* Boston, Mass., and New York: Houghton Mifflin Company.

Wright, Quincy (1942) *A Study of War.* Chicago, Ill.: University of Chicago Press.

[1] Richard L. Merritt is Professor of Political Science and Research Professor in Communications, and Dina A. Zinnes is the Charles and Ethel Merriam Professor of Political Science, both at the University of Illinois at Urbana-Champaign. DDIR is housed at: Merriam Laboratory for Analytic Political Research, University of Illinois at Urbana-Champaign, 512 East Chalmers Street, Champaign, Illinois 61820-3696, U.S.A. (Tel: 217-244-0739; FAX: 217-333-4369). Its newsletter, *DDIR-Update*, is currently distributed without cost; as its subscribers expand in number, however, we can expect a nominal fee to be charged. Subscriptions are available from the Merriam Lab.