Theory and Practice of Events Research

Studies in Inter-Nation Actions and Interactions

Edited by

EDWARD E. AZAR

and

JOSEPH D. BEN-DAK

with an introduction by

PHILIP M. BURGESS

—

1975

Testing International Theory:

Methods and Data in a Situational Analysis of International Politics

RONALD GOODMAN, JEFFREY HART and RICHARD ROSECRANCE

Situational Analysis Project, Cornell University

Using as sources general diplomatic histories for 1870-90, 987 events (2,046 interactions) were coded, with inter-coder reliability about 80 per cent. Histories were used because they have filtered out all trivial events and because they are more manageable than primary sources. The authors feel that the events of 1870-90 have a bearing on contemporary international politics; the notion of "balance of power," for example, was as important then as now. Having chosen the directed dyad as the unit best suited for analyzing the period, categories were rank-ordered and scaled according to Stevens' technique. The result was the Corkeley (Cornell-Berkeley) Scale. The authors found that historical and newspaper sources differed significantly (the London *Times*, for example, was consistently biased in a cooperative direction). They anticipate only minor revisions in the scale for use in scaling events of other historical periods and feel that many international relations theories can be tested against Situational Analysis Project data.

INTRODUCTION

The theoretical study of international politics has benefited enormously in the past ten years from the development of quantitative data against which the generalizations of international relations theory could be tested. Rather than summarizing this important work, let us merely point out that such investigators as Deutsch, Russett, Singer and Small, McClelland, Rummel and Tanter, North, Brody and Holsti, M. Haas, and others have furnished us with quantitative measures that have helped to challenge some of the main propositions in the field.[1] The next few years will undoubtedly witness the further refinement of such work, looking toward the development of even more politically sensitive indicators. ·

Researchers have become increasingly interested in scaling or categorizing international events. (See among others Azar, 1970a; Corson. 1070; McClelland, 1969b; and North and Choucri, 1968.) In each case, investigators seek to establish a reliable listing of international outcomes and then to scale or categorize the events in question to test for patterns of participation or cooperation conflict. With the exception of a study by Robert North and Nazli Choucri, these efforts have largely been confined to the contemporary period. The data used has been derived from contemporary newspaper or journal sources.

NEWSPAPERS AS DATA SOURCES

The first comparison of newspaper and other sources was made by Edward Azar. He found that in reporting Middle East events for 1955-58, the *New York Times* and the *Middle East Journal* overlapped only 9.7 percent of the time. Azar (1970a, pp. 35-36) writes:

Given that the two sources mutually reported only 9.7 percent of the events during the entire period. we therefore decided to determine if this overlap were related to event violence. Intuitively one might expect that the concentration of duplicated reporting would be for highly cooperative and/or highly conflictual events—the extreme values of the scale. Such events would seem to be most "newsworthy" and "dramatic." Our research disconfirms this notion. Instead, the most predominant overlap was for events classified at the intermediate levels of violence—communicative behaviors and domestic political unrest. Examining this overlap of reporting by actor. we found more duplicate reporting of events in the high violence region for acts initiated by Israel than by Egypt, while both sources reported more Egyptian-initiated events classified in the low violence (cooperative) region.

These findings on the question of source coverage serve as a warning against using a single events data source without a comparison of sources by frequency as well as content. Our work has been directed at the whole spectrum of inter-nation behaviors regarding source coverage. but this research warning also is applicable to the study of a particular inter-nation behavior (i.e., coup occurences, border conflicts, etc.). One may be able to use a single source for a particular behavior *only* after a systematic comparison of sources.

It is quite apparent from our investigation that reliance on a single source might generate strikingly dissimilar conclusions

about inter-nation behavior. For example, in terms of the violence dimension, the *New York Times Index* tended to report more events in the intermediate categories of violence (6) in contrast to the *Middle East Journal* which reported relatively more events in the low violence cooperative regions (1-5). Since events would be the basic units-of-analysis in the quantification of inter-nation interactions, the cooperative-conflictual character of a nation-state might be greatly determined by a source.

DIPLOMATIC HISTORIES AS DATA SOURCES

These cautionary remarks lead one to inquire whether additional source materials need to be developed to check the findings based on newspaper data collections. One source of additional data is that given by the diplomatic historian. For a number of diplomatic eras, the historian has provided a reliable guide to the important events, those on which the international system turned. Against the strictly diplomatic collections of governments, historians have been able to balance biographical and monographic accounts. They have been able to evaluate critically the claims of statesmen in memoirs written well after the event. The diplomatic historian, in short, not only provides a guide to the sources; he also represents a critic of the sources.

In certain periods, moreover, documentary sources have been available for some time, and diplomatic historical surveys of this material have reached a relatively stable plateau of scholarship. The general diplomatic histories for the period 1870-90 are regarded as exemplars of the historian's craft. It seems unlikely at this point that new documentary or archival collections will be unearthed that will substantially alter our understanding of the period. Even in 1950, W. L. Langer, writing in the preface to the second edition of *European Alliances and Alignments, 1871-1890,* was able to claim that the new literature since the original edition (1931) had not changed in any substantial way the account he would offer of these years. Partly because the systematic origins of World War I were deemed to have had their roots in the 1870-90 period, historians have devoted a great deal of attention to the events of the Bismarckian system.

The importance of diplomatic history as a guide to the theoretical analysis of international relations is coming to be recognized among scholars. Recently, Frank Denton surveyed histories of war and diplomacy for the period 1750-1960, emphasizing the reasons historians gave for war or military conflicts (Denton, 1969). It appears likely that historical sources will be plumbed more and more for events which would validate international relations propositions. It is therefore possible, though not certain, that the data offered in diplomatic histories could be used to check international theories and perhaps to generate new ones.

The present project seeks to use the works of general diplomatic histories for the period 1870-90 (though the project will eventually be extended to 1914) to develop a list of significant diplomatic actions. These actions in turn would be used as dependent variables in an attempt to validate or disconfirm existing international theoretical propositions. For the period 1870-81, 987 significant events have actually been coded, with inter-coder reliability levels substantially above 80 percent.

PRESUMPTIONS IN THE USE OF DIPLOMATIC HISTORICAL DATA

Such an attempt clearly rests upon certain presumptions about the nature of the historian's craft. In the first place, the list of events is based upon the historian's culling of the significant from the insignificant events. It should be noted that differences of historical interpretation do not necessarily bias such a listing. Historians might disagree greatly about the interpretation of events while agreeing generally about which events were significant. Their description and listing of events would be similar, while different interpretive conclusions were drawn from the events recorded. It is also, of course, possible that in addition to disagreements over interpretation, historians might disagree over which events were significant and which insignificant. Under certain circumstances one could anticipate different event lists from individual historians on the basis of different historical interpretations. If so, such differences could be checked and corrected by including the works of historians embracing different historical interpretations. Sectarian principles of inclusion could in this way be overcome.

Secondly, and more important than the scope of events, would be the description of individual events. In some cases historians, while surveying the same events, have given varying accounts of the events in question. This is not a chronic problem in historical research; or, to put the matter in relative

litary con-
kely that
and more
ernational
possible,
offered in
check in-
erate new

: works of
od 1870-90
xtended to
omatic ac-
e used as
validate or
heoretical
)-81, 987
oded, with
y above 80

SE   OF

ı presump-
s craft. In
l upon the
ım the in-
that dif-
ı do not
ans might
ı of events
vents were
: of events
'tive con-
orded. It is
ldition to
ians might
icant and
ımstances
; from in-
different
lifferences
uding the
historical
inclusion

the scope
ndividual
eving the
ts of the
roblem in
ı relative

terms, historians are much less likely to disagree on their description of events than on the explanation of those events. Still, in the cases where general diplomatic histories do disagree about the description of an historical event, one must proceed to the monographic or primary sources. Differences in descriptions of events among historians usually relate to partial emphases. One historian's description may emphasize one aspect of an event, another, another. More complete or exhaustive treatments reconcile such differences by putting the partial emphases into a more general and detailed context.

Thirdly, and perhaps most significantly, our study presumes that historical data for the period 1870-90 can apply to international theory in the present epoch. If one takes the position that history is a "seamless web," with each event totally unique in all its manifestations, then the past offers no data which can help validate generalizations about the contemporary period. We might, however, view each event as representing a unique combination of causal factors, without conceding that the causes themselves are unique. Military technology, for example, has changed greatly since the nineteenth century. Military size, strength, effectiveness, and so on, however, have a bearing on diplomacy at all times and places.

In contemporary international politics, actions occur which are both more cooperative and more conflictive than events which occurred in the nineteenth century. The amplitude of international politics has increased. It is still interesting to note, however, which independent variable factors are associated with increases of conflict and which with increases in cooperation. In addition, the nature of the modern state does not differ much from that of the late nineteenth-century state. Structurally, the problems of adjusting relationships among constituent state units and accommodation of interests should be similar for the period since the French Revolution, when states began to emerge as the instruments of nationalized publics. If, for instance, concepts of "polarity" affected patterns of conflict and cooperation in previous European diplomacy, they should have some application to the present system. Balances of power, though now perhaps measured in nuclear terms, are probably as important today as they were a century ago.

More fundamentally, however, in basic respects human behavior must be similar across historical epochs. If so, then the results of behavior will also show similarities. The problem then becomes that

of discovering these similarities—a feasible task if comparisons are generic rather than specific. The greater the historical distance between behaviors to be compared, the more generalized and aggregative must be the principles of comparison. The same principle holds with comparisons of different social and political institutions in different geographic locations in a given historical epoch. The more universal the political behavior one wants to describe and analyze, the less system-specific must the terminology be. Thus the typical traditional political categories used in the analysis of Western democratic government (such as executive, legislative, and judicial) could not be used in the analysis of non-Western societies and polities. More general terminology (such as "rule-making," "interest articulation, interest aggregation," "political socialization," and "political communication and political culture") had to be employed to compare political systems across geographic and developmental lines.

Thus, in principle, cross-temporal comparisons or relation of the behavior of the nineteenth-century diplomatic system to the twentieth-century diplomatic system present no problems. The difficulty is a practical one; on what conceptual bases do we make the linkages? Data concerning the Bismarckian system can apply to current international problems and theories so long as they are couched in variables which are not system-specific or system-restricted. In essence, therefore, only the analytic and conceptual use of the data presents problems. The data themselves may have great value.

None of this of course should be interpreted as an argument against studies of the nuclear epoch or for making direct extrapolations from nineteenth- or early twentieth-century behavior to the present era. As additional diplomatic collections and archives are opened to historical inspections, and as historians are able to survey them, their data will be added to that of the present project. In addition, many other scholars are focussing primarily on events since 1945 (McClelland, Corson, Azar, North, and others); there is no paucity of work in this area, and the generalizations derived from it may then be compared with those emanating from the earlier period.

DIPLOMATIC HISTORY AND GOVERN-MENTAL COLLECTIONS AND ARCHIVES

Given the relevance of historical materials to con-

temporary international events, one may be tempted to ask. Why survey only the diplomatic histories? Why not plumb the primary collections of governmental documents themselves? There are two reasons for not doing so. First, there is the sheer impossibility of surveying all the relevant sources in order to unearth each happening of diplomatic significance. Robert North and his associates plumbed more than 3,000 documents in the compilation of data for the six weeks' period June 28-August 4, 1914. At this level of intensity the study could not have been expanded to cover even the three-year period ranging from Agadir to the Balkan Wars to World War I. Such techniques are very useful in studying short crisis periods (like Bosnia, Fashoda, or Cuba), but they could not possibly be used for 30- or 40-year intervals.

The second reason is that even access to all the sources in bulk form, assuming they could be sub-divided into separate diplomatic events, does not solve the problem. Even if it were possible to study the entire output of diplomatic documents for a particular period, a mere listing would not distinguish the important from the unimportant events. The total diplomatic record would provide, in addition to the significant acts, masses of irrelevant detail. It would offer events so relatively routine or trivial that even the diplomatic participants themselves would not attend to them. Today, an American secretary of state may specifically note and act upon 6 to 20 cables per day, between .1 and .3 percent of the incoming communications (Rosecrance, 1968). While desk officers take account of some portion of the rest, most tend to fall beneath the threshold of diplomatic significance. Thus the question crops up again. How do we screen the important from the relatively unimportant events? Who should be our guide to such a task? Once more, the diplomatic historian serves as an indispensable filter. On the basis of his expert knowledge, he can distinguish between significant and unimportant happenings. Presumably, those acts and events which diplomatic historians have seen fit to record in their works are those that the most highly trained and knowledgeable scholars regard as most important to an understanding of the period.

DATA-MAKING PROCEDURES IN THE SITUATIONAL ANALYSIS PROJECT

The Situational Analysis Project, at this time, has completed five phases of data-making and analysis:

1) Selection of historical sources
2) Compilation of events from each source
3) Compilation of a master list of events
4) Creation of a cooperation-conflict scale
5) Scaling of specific events

Certain primitive terms will be defined prior to discussion of these phases. An *event* is any overt input and/or output of the type. Who does or says what to and/or with whom and when? which is communicated and which affects the behavior of an international actor or actors. For coding purposes, an event must be recorded in one of the sources which meets the coding rules.

This definition differs slightly from the one used by McClelland and Azar. McClelland (1961) would exclude

constant, routine exchange of goods, services, concerns, and interests between nations which tend to bring them together in a condition of interdependence

(such as transaction flows) from his list of event/interactions. Both Azar and the current project would recognize that changes in transactions could well become diplomatically important. In short, the rule is to record the transaction when a source mentions it. Our definition of an event differs from that proposed by Azar in that it requires *communication* of the action in question to some other state. Consider the following case: in 1878 the British Cabinet took a secret decision to move Indian troops to Malta in the Russo-Turkish crisis. Until those troops were actually in process of being moved or the movement was announced publicly, the event would *not* be coded. Nations, in short, must not be talking to themselves. An event only becomes international when it is communicated to some other actor or when overt action takes place. For the same reason, *perceptions* of statesmen or governments which are *entirely* confined to national decision-making arenas and are not communicated are *not* coded.

Events or actions are performed *only* by official representatives of a political entity. Antiforeign riots would not be coded. Actions by members of an opposition party would not be coded. Changes of government, even though involving official actions, would not be coded unless the source consulted mentioned the possible impact of the change. Events must be *discontinuous* and capable

ıking and

ource
nts
scale

ıed prior to
any overt in-
does or says
n? which is
behavior of
coding pur-
one of the

the one used
(1961) would

concerns, and
em together in a

t of event/in-
rent project
actions could
In short, the
ıen a source
: differs from
·quires *com-
o some other
in 1878 the
to move In-
ırkish crisis.
icess of being
:ed publicly,
ns, in short,
n event only
nunicated to
takes place.
statesmen or
confined to
nd are not

ly by official
Antiforeign
members of
led. Changes
g official ac-
source con-
pact of the
and capable

of being given a date. *Continuous* events like wars or on-going negotiations are coded *only* when their inception and conclusion are mentioned, or when special changes or departures take place in the ongoing process. The general rule is to note such departures only when the source in question mentions them. Since the focus of the international system, 1870-90, was on Europe, it was decided to restrict diplomatic/interaction events for this period to those in which at least one *European* state was a participant as either actor or target. For later periods, this coding rule will be modified as a more global system develops.

International *actors* are either sovereign states or political movements which have attained some degree of international recognition (such as that accorded the Bosnia insurgents after their revolts against Turkish rule in 1875). A *sovereign state* is a

clearly defined territory and an associated self-identified social and political system characterized by a governmental structure with externally recognized authority. (Alker and Bock, 1968)[2]

An *initiator* of an event is an actor who is responsible for originating an international event. A *target* is an actor who is the object of an event, one to whom something is said or done. There are two kinds of targets: (1) a *direct target*—one explicitly mentioned in the account of the event in the original source; and (2) an *indirect target*—one implicit in the account of the event. Take for example the following account of an event—"Germany tells Austria that she opposes France." In this event, Germany is the initiator and Austria is the direct target. The indirect target is France. A *dyad* is a pair of actors, and a *directed dyad* is an ordered pair of actors in the form (initiator, target). *Cooperation-conflict* is a continuum which can be used to describe the consequences of an event from the relations between a specific directed dyad. Therefore, since an event may involve one or more directed dyads, an event consists of a set of directed dyadic interactions (or simply *interactions*). An *international system* is simply a set of international actors and a set of cooperation-conflict levels for each directed dyad. The Situational Analysis Project has been restricted to the international system in which, during 1870-90, at least one of the actors in each directed dyad was a European power.

I Selection of Historical Sources

The sources used for the coding of data, 1870-81,

were: Albertini (1952), Fay (1930), Hinsley (1962), Langer (1950 and 1952), Schmitt (1934), Sontag (1933), and Taylor (1954). We selected primarily sources included in the American Historical Association's *Guide to Historical Literature* for the period in question which covered at least a 12-year period of international diplomacy. The specification of a 12-year rule prevents inclusion of essentially monographic treatments of very short time periods. It also rules out accounts that deal solely with a functional area of interest (such as colonial or naval policy). The reason for making certain that general treatments and surveys will form the basis for the events list is that reliance upon monographic treatments would alter our criteria for considering only significant events. The general treatments aim at a roughly common level of abstraction and therefore at a roughly similar standard of significance. To have included monographic accounts of the Near Eastern Crisis (1875-78) or the negotiation of the Dual Alliance (1878-79) would have skewed the list toward two specific episodes: it would not have provided relatively "even" coverage of the entire period. Moreover, since detailed treatments do not exist for each microscopic time period in the monographic literature, there is no way in which an evenness of abstraction and significance could have been attained had the monographic accounts been included in the compilation of the events list.

There were other sources on the A. H. A. list which met the coding rules and which also could have been plumbed for events. These include works by Hauser, Mansergh, Schmitt, Roubaud, Meyer, Renouvin, and Hayes. In a second-stage investigation these works will be covered by our coding procedure. However, even using the eight sources which were noted in the first stage of the project, we found that each additional source added fewer and fewer events to our master list. Coding additional sources would have limited usefulness. Nonetheless, to make sure that no events of significance are omitted, the second stage of the project will proceed to cover the additional works and any others that subsequently appear and that fall within our coding categories.

II Compilation of Events from Each Source

The principal investigator, along with two graduate students of history, two graduate students of political science, and an undergraduate student of political science, systematically selected events for

the period 1870-81 from each historical source. Each coder was given instructions specifying what an event is and how to record it (see above). Coders were instructed to list the date of the event, the nature of the event (in language as close as possible to the original wording of the historian), and the initiator and the target of the event on 3″ × 5″ index cards along with the page and paragraph of the source in which the event was found. If the exact date of the event were missing, the event was still recorded—but events without relatively specific descriptions or without specific initiators or targets were not recorded.

An initial check on the reliability of event selection was made. Several coders selected events from the same passage of several sources. The average overlap of events recorded fell in the vicinity of 85 percent—that is, less than 15 percent of the events listed by any given coder were not listed by other coders. This seemed to be an acceptable level of reliability for the initial stages of event selection, and reliability improved in later stages. After the initial cross-coder reliability check was completed, each coder went on to code all the events for the required period in only one or two of the historical sources. Events were coded in the order in which they appeared in the text. Coders were instructed to arrange the index cards in rough chronological order prior to the third phase of data-making.

III   Compilation of a Master List of Events

Events were recorded on index cards for all the historical sources. At this point, a committee was formed consisting of all the coders. The committee reviewed the selected events in chronological order to decide which events could be placed on a master list and to establish correspondences between events mentioned by different historians. The criteria for forming the master list were as follows:

1) If only one historian mentioned a given event, the event was included in the master list

2) If more than one historian mentioned a given event and one historian's version was clearly more specific, detailed, and accurate than the others', that historian's version was included in the master list

3) If more than one historian mentioned a given event and the content of the different versions was not contradictory, a composite event including all the information contained in different versions was included in the master list

4) If none of the versions of the event containe[d] a specific date or if different versions were con[-] tradictory, outside sources were consulted

The number of events selected for the master li[st] was 987. The compilation of the master li[st] resulted in a higher level of reliability of even[t] selection for each individual source, since com[-] mittee members frequently had to check the sourc[es] to determine whether a given event was comparabl[e] to an event in another source. The procedure wa[s] rather time-consuming but worthwhile in the lon[g] run. For other historical periods, we will probabl[y] try to teach one person how to compile master list[s] reliably.

Two versions of the master list were recorded[.] The first was a hand-written copy with the date o[f] the event, the content of the event, and the pag[e] and paragraph of each source in which the even[t] was mentioned. This version will be used in late[r] analyses of the overlap and/or bias of individua[l] historians. The second version was a deck of IB[M] computer cards with the date and content of th[e] event. IBM cards were used because of the[ir] flexibility—later information on dates or o[n] additional events may be added (and cards re[-] arranged) without difficulty. Copies of the maste[r] list can be made inexpensively on IBM car[d] printing machines.

IV   Creation of a Cooperation-Conflict Scale

The next phase of data-making was the creation o[f] a scale for measuring the amount of cooperation o[r] conflict in each directed dyadic interaction. Th[e] first step was to devise a relatively exhaustive ty[-] pology of events. The master-list committe[e] developed a typology with 42 categories of directe[d] dyadic interactions by synthesizing an[d] generalizing three previously developed typologie[s] and by adding or revising categories when o[ld] categories did not work (see Table I). The typolo[gy] was checked for exhaustiveness, generality, an[d] reliability by applying it to Langer's Encycloped[ia] of World History for a variety of historical period[s]. Although no formal statistical checks were made [at] that time, the committee judged it to be sufficien[tly] exhaustive, general, and reliable.

The next step was an attempt to rank-order t[he] categories with respect to the degree [of] cooperation-conflict. Three graduate student[s at] the University of California and three professor[s at] Cornell ranked the full list of types in Table I. T[he]

it contained
s were con-
ed

e master list
master list
ity of event
since com-
k the source
comparable
ocedure was
e in the long
vill probably
e master lists

re recorded.
h the date of
nd the page
ch the event
used in later
of individual
deck of IBM
ntent of the
se of their
ates or on
d cards re-
f the master
IBM card-

t Scale

e creation of
operation or
action. The
haustive ty-
committee
s of directed
izing and
l typologies[3]
s when old
The typology
erality, and
ncyclopedia
ical periods.
ere made at
sufficiently

k-order the
degree of
students at
rofessors at
Table I. The

TABLE I
The event typology

1.	Inquiry for information from X—Favourable
2.	Warning to X—Military
4.	Start of Negotiations between A and X
6.	Limited Use of Nuclear Weapons (nucs) against X
10.	Trade Ban
11.	Unlimited Use of Nuclear Weapons against X
13.	Comment Unfavourably on Statement of X
15.	Withdraw from Military Alliance which was Directed against X
22.	Withdraw Proposal toward X—Proposal was Unfavourable
23.	Acceptance of Proposal by X—Diplomatic
24.	Pledge to Improve Relations with X
34.	Rejection/refusal of Proposal made by X—Military
35.	Withdrawal from Military Alliance with X
38.	Declaration of War on X
39.	Arms Control Agreement with X
40.	Request Positive toward X
45.	Military Occupation of X
48.	Informal Agreement with X—Diplomatic
49.	Increase Trade with X
50.	Blockade/Siege of X
51.	Statement of Policy Unfavourable to X
52.	General Observation that Relations with X are Deteriorating
53.	Arms Reduction or Reduction of Military Budget—Previously thought to be Directed against X
55.	Warning to X—Diplomatic
62.	End Major Attack on X
64.	Full Mobilization against X
65.	Ultimatum to X
67.	End Economic Sanction against X
70.	Sign Military Alliance with other Power against X
71.	Major Attack on/Invasion of X
80.	Severance of Diplomatic Relations with X
82.	Internal Change in A—Unfavourable to X
83.	Internal Action Favourable to X
86.	Breach of Arms Control Agreement with X
87.	Assurance to X—Military
88.	End Military Aid to X
90.	End Informal Agreement with X—Diplomatic
94.	Agreement with X—Formal Diplomatic
99.	Military Alliance with X
100.	Establish Economic Community with X
101.	Supply Military Aid to X
102.	Establish Political Federation with X

level of agreement on these rankings was quite high—rank-order correlations were all above .85—but it was decided that several categories were very ambiguous, several were too similar with respect to cooperation-conflict to be distinguishable, and that a higher level of measurement would be both possible and desirable. Therefore, the number of categories was reduced to 32 by eliminating or replacing the ambiguous ones and by collapsing the nearly equivalent ones (see Table II), and a technique for obtaining a ratio or interval scale was

TABLE II
Reduced typology in rank order
(According to the average rankings of the six individuals)

most cooperative		
	32.	Establish Political Federation with X
	31.	Establish Economic Community with X
	30.	Military Alliance with X
	29.	Supply Military Aid to X
	28.	Arms Control Agreement with X
	27.	Agreement with X—Formal Diplomatic
	26.	Assurance to X—Military
	25.	Informal Agreement with X—Diplomatic
	24.	Acceptance of Proposal by X—Diplomatic
	23.	Increase Trade with X
	22.	Arms Reduction or Reduction of Military Budget Previously Thought to be Directed against X
	21.	Pledge to Improve Relations with X
	20.	Request Positive Toward X
	19.	Start of Negotiations between A and X
	18.	Inquiry for Information from X—Favourable
	17.	Comment Unfavourable on Statement of X
	16.	General Observation that Relations with X are Deteriorating
	15.	Statement of Policy Unfavourable to X
	14.	Rejection/Refusal of Proposal made by X—Military
	13.	Withdrawal from Military Alliance with X
	12.	Warning to X—Diplomatic
	11.	Trade Ban
	10.	Sign Military Alliance with Other Power against X
	9.	Warning to X—Military
	8.	Ultimatum to X
	7.	Full Mobilization against X
	6.	Declaration of War on X
	5.	Blockade/Siege of X
	4.	Major Attack on/Invasion of X
	3.	Military Occupation of X
	2.	Limited Use of Nuclear Weapons against X
least cooperative	1.	Unlimited Use of Nuclear Weapons against X

devised.

The technique used was an adaptation of the procedure for obtaining ratio scales in psychophysics. Eight scalers, all experts on international politics, were selected. Each scaler was given the following verbal instructions:

You will be presented with a series of interaction categories in irregular order. Your task is to tell how cooperative they seem by assigning numbers to them. Call the first category any number that seems to you appropriate. Then assign successive numbers in such a way that they reflect your subjective impression. For example, if a category seems 20 times as cooperative as another, assign a number 20 times as large as the first. If it seems one-

fifth as cooperative, assign a number one-fifth as large, and so forth. Use fractions, whole numbers, or decimals, but make each assignment proportional to the intensity of cooperation as

TABLE III

Scaling form used for obtaining a modified "ratio scale"

	Scale Score	Geometric Mean
Agreement with X—Formal Diplomatic		
Pledge to Improve Relations With X		
Unlimited Use of Nuclear Weapons against X		
Withdrawal from Military Alliance with X		
Ultimatum to X		
Establish Political Federation with X		
Blockade/Siege of X		
Assurance to X—Military		
Arms Reduction or Reduction of Military Budget Previously Thought to be Directed against X		
Request Positive toward X		
Inquiry for Information from X Favourable		
Military Occupation of X		
General Observation that Relations with X are deteriorating		
Warning to X—Diplomatic		
Full Mobilization against X		
Establish Economic Community with X		
Military Alliance with X		
Arms Control Agreement with X		
Increase Trade with X		
Start Negotiations between A and X		
Informal Agreement with X—Diplomatic		
Comment Unfavorably on Statement of X		
Rejection/Refusal of Proposal Made by X—Military		
Limited Use of Nuclear Weapons Against X		
Declaration of War on X		
Major Attack on/Invasion of X		
Supply Military Aid to X		
Acceptance of Proposal by X—Diplomatic		
Trade Ban		
Warning to X—Military		
Statement of Policy Unfavorable to X		
Sign Military Alliance with other Power against X		

you perceive it. Consider the number 1 to be the lowest possible number (corresponding to minimal cooperation) and the number 100 to be the highest possible number. The number 50 should be used for those categories which seem to you to be neither cooperative nor uncooperative.

Prior to scaling interaction categories, scalers were given a practice session in which they used the instructions to judge the lengths of straight-line segments.[5] Usually one practice session was needed to introduce the concept of proportionality. After the practice session, the scalers were presented with the 32 categories of interactions listed in Table II in a random order on a sheet of paper like that illustrated in Table III.

With this technique, scale scores of high inter-scaler reliability were obtained (see Table IV). No inter-scaler correlation was lower than .927. The high product-moment correlations suggest that an interval scale was obtained. It would be incorrect, however, to claim that a ratio scale was obtained since the way in which the Stevens technique was adapted for use here—e.g., the addition of maximal and minimal scores and a mid-point—violates some of the requirements for a true ratio scale. Final scale scores were obtained by computing the geometric mean of the eight scalers' scores for each category. (Stevens suggests the use of the geometric mean rather than the arithmetic mean). The final result is illustrated in Table V and will henceforth be called the "Corkeley" Scale to commemorate the collaboration of Cornell and Berkeley in its formation.

At this point, a scholar trained in diplomatic history went through the entire master list of events and identified all the directed dyadic interactions in each event. This was considered to be more desirable than to ask each coder to individually identify interactions, since the possibility of disagreement on indirect targets was great.

V Scaling of Specific Events

Coders were instructed to assign cooperation-conflict scores to specific interactions using the Corkeley Scale. They were told to find a category in the Corkeley Scale which most closely approximated the nature of the interaction in question. If they thought that the interaction was either more or less cooperative than the Corkeley Scale score suggested, they could assign the interaction a score anywhere within the limits created by the next highest and the next lowest category in the Corkeley Scale

TABLE IV
Pearson correlation matrix

	GQ	NU	MW	EF	WP	RG	CC	RS	BG	AA	AS	JH	RM	DL
GQ		0.9762	0.9638	0.9770	0.9655	0.9744	0.9637	0.9744	0.9605	0.9594	0.9589	0.9602	0.9819	0.9591
NU			0.9591	0.9675	0.9539	0.9826	0.9801	0.9796	0.9658	0.9740	0.9722	0.9633	0.9645	0.9660
MW				0.9553	0.9666	0.9774	0.9739	0.9758	0.9699	0.9655	0.9607	0.9267	0.9734	0.9753
EF					0.9777	0.9726	0.9746	0.9776	0.9780	0.9739	0.9757	0.9490	0.9792	0.9708
WP						0.9692	0.9739	0.9736	0.9825	0.9769	0.9780	0.9397	0.9781	0.9721
RG							0.9935	0.9911	0.9826	0.9885	0.9861	0.9599	0.9775	0.9809
CC								0.9896	0.9921	0.9939	0.9928	0.9515	0.9796	0.9803
RS									0.9834	0.9795	0.9776	0.9445	0.9798	0.9843
BG										0.9935	0.9909	0.9385	0.9788	0.9737
AA											0.9965	0.9526	0.9717	0.9729
AS												0.9602	0.9755	0.9737
JH													0.9530	0.9439
RM														0.9729
DL														

Two kinds of reliability tests were made of the Corkeley coding of interactions; the first was a simple intercoder reliability test, and the second was a test of the stability of each coder's application of the Corkeley Scale over time. Inter-coder reliabilities for the first 262 interactions in the master list were checked as well as the reliabilities of three subsets of these interactions. The results of this test are given in Tables VI through IX. Compared with previous cooperation-conflict studies (see Moses et al., 1967), very high levels of inter-coder agreement were obtained. There appears to be a slight decrease in agreement in the second 100 and final 62 interactions—a reversal of the finding by the Stanford group that reliability improved over time. But this may be attributed to the greater familiarity of the coders with the first 100 interactions, since the first 100 were used in training sessions. The stability of scoring was tested by having the group of coders code the same hundred interactions at two different times. An interval of two weeks between codings assured that short-term memory effects would be minimal. The results of the stability test are given in Table X. All the correlations are above .83. Thus, the Corkeley scaling technique provides the a way of obtaining reliable, interval-level measurement of the degree of cooperation-conflict in directed dyadic interactions.

It is necessary to note that coding was always done in chronological order, so that the context of events and interactions would be clear to coders. It would be possible and desirable to see if similar scores could be obtained if the order of events was randomized before coding. Judging from the reactions of coders, however, this does not appear very likely. Thus, the context may be very important for the judgment of cooperation-conflict levels.

After the initial reliability testing, coders were given, on the average, 100 interactions per week to code. At the end of the week, meetings were held for the purpose of discussing the historical context and checking for general procedural snags, typographical errors and the like. The entire 1870-81 period was coded in this manner (see the coding example in Table XI). A new set of data cards was punched containing the following information: the identification number of the event in question, the date of the event, the identification number of the interaction, code numbers for the initiator and the target, the cooperation-conflict score (the geometric mean of all the coders' scores), and whether the target was direct or indirect. These data cards are being used in Phase VI of the Situational Analysis Project, Preliminary Data Analysis. They allow for a high degree of flexibility in the analysis of the cooperation-conflict data in the following respects:

TABLE V
The Corkeley scale

Category	Scale Score
Unlimited Use of Nuclear Weapons	1.01
Limited Use of Nuclear Weapons	3.27
Military Occupation	4.51
Major Attack on/Invasion of	7.99
Blockade/Siege	10.6
Declaration of War	11.7
Full Mobilization against X	16.0
Ultimatum to X	16.3
Sign Military Alliance with Other Power against X	24.7
Warning to X—Military	25.2
Trade Ban	27.6
Withdrawal from Military Alliance with X	28.6
Warning to X—Diplomatic	32.5
Rejection/Refusal of Proposal Made by X—Military	34.9
Statement of Policy Unfavorable to X	42.1
General Observation that Relations with X Are Deteriorating	44.5
Comment Unfavorably on Statement of X	45.8
Inquiry for Information from X—Favourable	52.0
Request Positive toward X	54.0
Pledge to Improve Relations with X	58.2
Start Negotiations between A and X	60.3
Acceptance of Proposal by X—Diplomatic	64.8
Arms Reduction or Reduction of Military Budget Previously Thought to be Directed against X	65.5
Increase Trade with X	65.8
Informal Agreement with X—Diplomatic	69.1
Assurance to X—Military	73.7
Agreement with X—Formal Diplomatic	76.3
Supply Military Aid to X	81.2
Arms Control Agreement with X	82.9
Military Alliance with X	88.0
Establish Economic Community with X	94.0
Establish Political Federation with X	99.6

1) Analysts may choose the time periods they wish to investigate simply by aggregating scores over time; they are not limited to calendar or fiscal years or even to months but only by days and by the requirements of their analysis

2) Analysts may choose among the following units of analysis—directed dyads, dyads, actors, subsets of actors or systems

3) Analysts may choose between synchronic (or cross-sectional) and diachronic (or time-series) analysis—or they may combine both approaches.

4) Analysts may control for the possibility of bias introduced by the identification of indirect targets or the disaggregation of events into interactions.

TABLE VI
Inter-coder reliability for the Corkeley scaling of interactions: correlations between coders for 262 interactions

	BS	CC	AA	BG	AS	RG	GM
BS	—	88	87	88	87	87	87
CC	92	—	88	89	89	89	89
AA	91	92	—	91	90	90	91
BG	91	93	94	—	88	91	91
AS	91	92	92	92	—	87	91
RG	90	93	94	93	90	—	92
GM	91	92	91	91	89	89	—

Key: Correlations are given without decimals. ("92" means that the correlation was .92.) Spearman correlation coefficients are given in the upper-right portion of the matrix; Pearson correlations are given in the lower-left.

[a] GM stands for geometric mean.
BS, CC, AA, BG, AS, RG are the initials of the Cornell students who coded interactions.

TABLE VII
Intercoder reliability for Corkeley scaling: the first 100 interactions (nos. 1-100)

	BS	CC	AA	BG	AS	RG	GM
BS	—	91	90	90	90	90	90
CC	95	—	94	94	93	94	94
AA	95	97	—	94	94	94	93
BG	95	97	97	—	93	93	93
AS	94	97	97	96	—	93	93
RG	93	96	96	95	95	—	93
GM	94	96	95	94	93	93	—

Key: See Key for Table VI.

METHODOLOGICAL JUSTIFICATION AND RESEARCH RELEVANCE

I Monadic, Dyadic and Directed Dyadic Events
The event is used in the above data-making procedures because it enables the researcher to use and compare a variety of historical sources in systemic fashion (Moses et al., 1967; Holsti et al. 1968; Corson, 1970; McClelland, 1969a and 1969; Azar, 1970.) The directed dyad is less frequently used, however, as the unit of analysis. Thus, some justification for this choice of basic unit—besides

of interactions:
ractions

	RG	GM [a]
	87	87
	89	89
	90	91
	91	91
	87	91
	—	92
	89	—

"92" means that
coefficients are
matrix: Pearson

Cornell students

aling:
(60)

	RG	GM
	90	90
	94	94
	94	95
	93	95
	93	95
	—	95
	93	—

TION AND

adic Events
data-making
earcher to use
sources in a
Holsti *et al.*,
9a and 1969b:
s frequently
. Thus, some
unit—besides

TABLE VIII
Inter-coder reliability for Corkeley coding:
the second hundred interactions (nos. 101-200)

	BS	CC	AA	BG	AS	RG	GM
BS	—	89	88	86	89	83	85
CC	91	—	88	87	88	87	83
AA	90	91	—	91	88	88	90
BG	87	90	93	—	85	90	89
AS	89	90	87	84	—	83	87
RG	88	91	93	94	84	—	90
GM	89	88	90	88	87	86	—

Key: See Key for Table VI.

TABLE IX
Inter-coder reliability for Corkeley coding:
the last sixty-two interactions (nos. 201-262)

	BS	CC	AA	BG	AS	RG	GM
BS	—	84	80	88	81	86	84
CC	89	—	77	88	81	82	83
AA	82	79	—	82	83	85	82
BG	88	85	86	—	83	87	86
AS	84	83	83	88	—	82	84
RG	87	85	89	89	84	—	87
GM	86	83	83	87	82	84	—

Key: See Key for Table VI

TABLE X
Stability of coding over time

Coder	Correlations between Different Codings Pearson	Spearman
BS	88	88
CC	84	88
AA	86	83
BG	96	95
AS	90	91
RG	94	92
GM	90	90

the added flexibility it allows the analyst—must be given.

Other studies of cooperation-conflict have used a variety of different units—the dyad (Klingber, 1961), the monad (Tanger, 1966), the pole or alliance subset (Corson, 1970; Singer and Small, 1965), and the system (Denton, 1969). The assumptions involved in using these units of analysis are not fully appropriate for the study of international politics in the 1870s for several reasons. First, the use of the dyad presumes a high degree of dyadic symmetry. It is assumed that if Nation A is cooperating with Nation B, then Nation B is cooperating with Nation A. Thus there can be no asymmetrical relationships. Some of our data analysis has convinced us, however, that relationships are not fully symmetrical for the period 1870-81 (Hart, 1972). Second, the use of the monad involves an assumption that such a thing as "generalized conflict-cooperation" exists independent of target—that is, that some actors uniformly cooperate or clash with all or most other actors. It thus makes sense to look at the attributes of "aggressive" and "pacific" actors. One should use monadic events, therefore, only when actors appear to indiscriminately cooperate or clash with other actors. This is clearly not the case in the Europe of the 1870s. In this period nations definitely did discriminate in their behavior toward others: a conflict with one power often amounted to cooperation with another. Thirdly, the use of the "pole" or alliance as a unit of analysis assumes unanimity on the part of "pole" members. That is, if Nation C and Nation D are aligned against E and F, it is assumed that Nations C and D are successfully coordinating their politics toward E and F—that no difference exists between them. Certainly in the 1870s this was not true. Finally, the use of the "system" as the unit of analysis is always appropriate where time-series analysis is possible—even though there may be problems involved in devising indices for systemic cooperation-conflict. Nevertheless, if measurement begins at the level of the system, no information about the contribution of individual actors, dyads, or directed dyads can be obtained without going back to the original data sources—a key restriction.

Thus, the use of the directed dyad—which expresses cooperation or conflict going in one direction—allows one to test for more detailed and sensitive relationships—for symmetry, for example. Further, the directed dyad can be aggregated.

TABLE XI

An example of SAP cooperation-conflict coding

(FM = Foreign Minister: AMB = Ambassador: EMP = Emperor: CHANC = Chancellor)

September 6, 1872

06091872 002

RUSSIAN FM EXPRESSES CONCERN OVER AUSTRIAN INTRIGUES IN BOSNIA AND HERZEGOVINA

Initiator	Target	Type of Target	Cooperation-Conflict Score
Russia	Austria	Direct	45.79

September 6, 1872

06091872 003

AUSTRIAN FM TELLS RUSSIAN FM THAT 1) AUSTRIAN POLICIES IN BOS-HERZ ARE DEFENSIVE, 2) AUSTRIA WISHES GOOD RELATIONS WITH SERBIA, 3) AUSTRIA WILL NOT ALLOW EXTENSION OF SERBIAN TERRITORY, AND 4) AUSTRIA DESIRES STATUS QUO IN TURKEY.

Initiator	Target	Type of Target	Cooperation-Conflict Score
Austria	Russia	Direct	48.72
Austria	Serbia	Indirect	44.04
Austria	Turkey	Indirect	55.28

September 6, 1872

06091872 004

RUSSIAN FM TELLS AUSTRIAN FM THAT RUSSIA IS NOT CONNECTED WITH SERBIAN AGITATION AND THAT RUSSIA IS SATISFIED WITH THE STATUS QUO IN THE NEAR EAST.

Initiator	Target	Type of Target	Cooperation-Conflict Score
Russia	Austria	Direct	57.03
Russia	Serbia	Indirect	45.12
Russia	Turkey	Indirect	56.55

September 6, 1872

06091872 005

AUSTRIAN AND RUSSIAN FMS AGREE TO POLICY OF NONINTERVENTION IN THE NEAR EAST AND TO FOLLOW POLICY OF STATUS QUO.

Initiator	Target	Type of Target	Cooperation-Conflict Score
Austria	Russia	Direct	71.22
Russia	Austria	Direct	71.22

September 6, 1872

06091872 006

RUSSIAN FM IN PRESENCE OF FRENCH AMB AND GERMAN EMP MENTIONS AGREEMENT WITH GERMAN CHANC AT THREE EMPERORS CONFERENCE.

Initiator	Target	Type of Target	Cooperation-Conflict Score
Russia	France	Direct	45.26
Russia	Germany	Direct	55.23

while it is not possible to disaggregate the dyad, or the monad. Hence all the advantages of dyadic, monadic, pole-oriented, or systemic analysis can be obtained without their ensuing complications. There is, of course, a price to pay for fewer assumptions and greater generality, and that is slightly more intricate data format that requires more sophisticated mathematic and statistical techniques than those necessitated by other units of analysis.

II Discontinuous vs. Continuous Events

Some justification must also be given for the decision to use only discontinuous events rather than both continuous and discontinuous events as some studies have done (Corson, 1970). Continuous events, such as wars, negotiations, and trade bans do have an effect on international actors, but is it really a continuous one? That is, even though a raging war influences the cooperation-conflict behavior of some actors, probably the greatest effect of the war is felt at its inception and at its end, or at major turning points. The alternatives to this assumption are extremely problematical. One can assume, for example, that continuous events have an effect described by some function of their time span—e.g., the effect of a war intensifies as time goes on. But there is no obvious way to assign such functions—a priori. Corson, for example, assumes that continuous events have constant effects on the cooperation-conflict levels of the actors involved. This assumption might apply to events of short, intense duration, such as the Cuban missile crisis, but not to a long drawn-out trade ban.

In addition, the use of continuous events presents conceptual difficulties. Does a technological innovation qualify as a continuous event? Must a continuous event be communicated like a discontinuous event? Do continuous events have initiators and targets? Finally, while the Situational Analysis Project study does not use continuous events, it would be possible to include the effects of continuous events like wars or negotiations to see what the impact would be on the pattern of dyadic interaction. It would be possible to add conflict or cooperation scores to the entire range of dyadic events between the inception and termination of a continuous event. Thus the benefits of both a continuous and noncontinuous procedure could be obtained.

For most continuous events, however, the present procedure will likely be most satisfactory.

Our coding rules require that we mention a continuous event again when any change in its course is recorded by the historian. In short, we take account of the event when anything occurs that appears historically significant. If a crisis is very intense, historians will refer to its ongoing course again and again and measure each minor development. If any major battles are fought, they will be listed separately. Thus the intensity of the crisis is likely to be mirrored in the intensity of coverage by the historian. Relying upon his judgment to distinguish the significant from the insignificant should help assuage the problem of treating only discontinuous events in our master list.

III Disaggregation and Targeting

The 987 events coded from historical sources for the period 1870-81 have been subdivided into 2,046 separate interactions. Thus, it was hypothesized that we should examine each event's effects on specific actors rather than the rest of the international system (monadically). When states engage in actions, they usually have specific targets in mind. In some cases, there is an indirect as well as a direct target. For example, in 1870 on the eve of the Franco-Prussian War, Prince Bismarck asked for Russian help against Austria-Hungary in the event that Austria supported France. Russia, however, refused to give such support. In this instance, the direct target of the Russian refusal was Prussia, but the indirect targets were Austria and France. While Russian action harmed the direct target, Prussia, it helped the indirect targets Austria and France.

In theory, of course, the number of indirect targets is indefinite, and scalers cannot be expected to know which nations to include and which to exclude as indirect targets. The solution adopted was to ask a scholar trained in diplomatic history of the period to specify which targets should be scaled in each case. Since the distinction between direct and indirect targets has been maintained on the data cards, however, it is possible to exclude the indirect targets for purposes of data manipulation.

In some cases, of course, multiple interactions occur. When Austria-Hungary and the other powers made demands upon Turkey in 1875-76, some actions were taken collectively which could also be accurately represented as individual dyadic interactions. We have disaggregated them accordingly. On the other hand, if multiple interactions are sanctioned, average scale scores will

be weighted in terms of multiple interaction events. Frequencies of individual action will also be distorted by the inclusion of large numbers of such events. Azar (1970. p. 17) discusses this problem also. In the period 1870-90, however, few multiple interactions—none apparently biased in one direction—occurred. Multilateral interactions take place in cooperation as well as conflict.

IV Historical and Newspaper Sources

Obviously historical and newspaper sources contain different accounts of a specific time period. As we have noted, Azar's comparison of *New York Times* and *Middle East Journal* listings of events for 1955-58 in the Middle East displayed an overlap of only 9.7 percent. In order to test differences and similarities between newspaper and historical treatments in the nineteenth century, we compared our historical source master list for the period January to June, 1875, with the accounts of the contemporary London *Times*. In this instance there was only a 3.3 percent overlap (see Figure 1). Further, there appeared to be some systematic differences in the recording of events in the two sources. The *Times* included many more cooperative events than noncooperative ones; the diplomatic histories did not (see Table XII). An inspection of the data reveals that the *Times* missed some events of great diplomatic significance. It did not systematically record, for example, the distorting influence of the "war scare" between Germany and France, which was the highlight of the diplomatic treatments. Perhaps surprisingly, however, the scale scores for newspaper events between France and Germany taken singly followed a path not unlike, though they lagged behind the scale scores for the historical data events between the two powers. The *Times* certainly noted the tension between France and Germany, though it did not seize upon the gravity of the crisis either systemically or bilaterally.

There were other important differences. Historical accounts were also focused largely on the actions of the great powers, which performed 91.6 percent of all initiations and served as target of 87.8 percent of the actions (Gray, 1971). Thus, very great differences in the frequency of action for individual powers is found between the *Times* and the historical sources. This, of course, is not surprising. Much of what the *Times* regards as "news" does not pass the historian's filter of diplomatic significance. Again, similarly, the actions of many
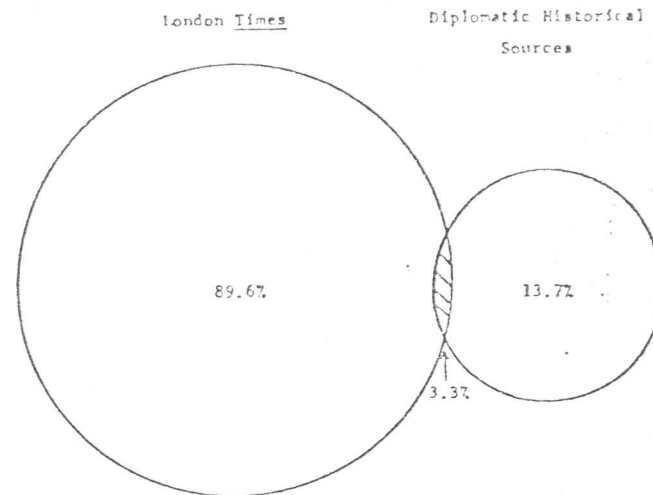

FIGURE 1 Overlap of events in newspaper and historical sources

of the smaller powers are devalued in the historical accounts as less important for the international system.

As we have seen, the *Times* record is biased in a cooperative direction relative to the historical accounts. This is also perhaps not surprising. The war scare was not fully understood nor its gravity fully recognized in contemporary accounts. The conflict largely took place at the diplomatic and hence covert level. If the tension had been more open, possibly newspaper sources might have recorded about the same levels of cooperation as those found later in the historical treatments. To observe whether this is so, other periods of greater overt conflict should be surveyed to note similarities and differences among newspaper and historical accounts. Perhaps newspapers miss the significant but covert types of conflict and cooperation. When secret alliances are signed, they may miss some of the important cooperation taking place. When covert diplomatic crises occur, newspapers may underestimate the seriousness of the conflict. The latter seems to have been the case in the period January to June, 1875.

It is difficult to reach an unequivocal conclusion on newspaper sources. They include many interactions that historians with the benefit of hindsight and full access to the sources will regard as trivial. They probably slight those activities of governments which can be concealed, or whose real gravity can be masked. Still, newspapers rank as an indispensable source. In a second stage of the Situational Analysis Project, we hope to survey newspaper data for the period 1870-80 and compare scalings and patternings with those of the

TABLE XII
Dyad frequencies by scale values of cooperation conflict for source

| | CONFLICT | | | COOPERATION |
	% of dyads which fall between scale scores 21-40	% of dyads which fall between scale scores 41-60	% of dyads which fall between scale scores 61-80	Total Dyads
London Times Data	N=63 %=7.7	N=599 %=73.5	N=153 %=18.8	815
Historical Source Data	N=18 %=11.6	N=119 %=76.8	N=18 %=11.6	155

historical accounts.

V   Revision of the Corkeley Scale?

How generally applicable, in fact, is the Corkeley scale? Efforts were made to assure that the categories used in the scale would not be historically specific (by testing them with events from other periods taken from Langer's *Encyclopedia of World History*). Some categories like those for nuclear attacks were included so that the categories could be applied to the contemporary era. But since the scale scores were assigned by experts of international relations of the nineteenth-century (although some twentieth-century experts were also used), and since the main concern was to create a scale valid for the period 1870-90, one might question its historical generality. For example, is an ultimatum more conflictive in the contemporary era than a declaration of war (see Table V)? Stated in more familiar terms, the question is whether the diplomatic "rules of the game" have changed since 1890. Another aspect of this question is the extent to which the rules are universally accepted by members of the system. Thus, it is possible to have *simple* rule changes in which all members of the system accept a new set of rules or *complex* rule changes in which only some members accept the new rules. For example, some authors have proposed a separate set of rules for the international relations of communist nations in the contemporary system (Alker and Bock, 1968, p. 124). The "rules of the game" such as they existed during 1870-90, were probably almost universally accepted. This raises for the Situational Analysis Project the problem of determining when rule changes occur and for which actors they occur. It is hoped that, if they oc-

cur, the changes will be reflected in lower interscaler reliability scores and in counter-intuitive results. If this be the case, then the Corkeley scale must and will be revised to reflect the new rules. We optimistically anticipate only minor revisions, however.

FUTURE RESEARCH AND RELEVANCE TO THE WORK OF OTHER INVESTIGATORS

Up to the present time, the quantitative study of international politics has tended to concentrate on crises (North, Koch, and Zinnes, 1961; Holsti, Brody, and North, 1965; McClelland, 1969a; Hermann and Hermann, 1969) and wars (Richardson, 1960a, 1960b; Singer and Small, 1970; and Denton, 1969). The impact of alliance cooperation
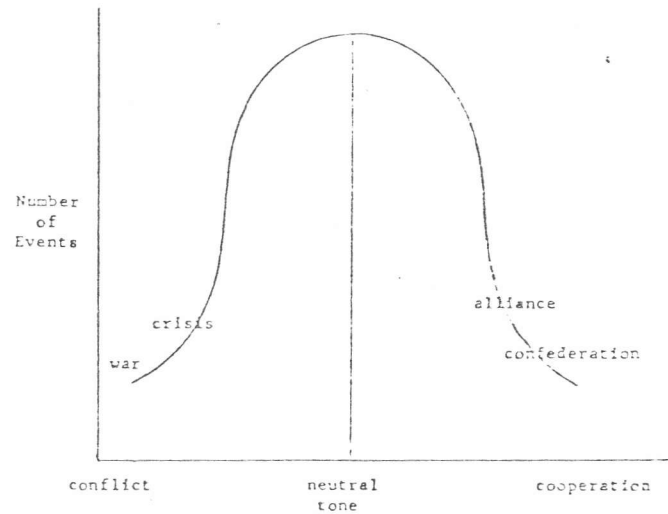
FIGURE 2   "Normal" curve of conflict and cooperation in international politics.

(Singer and Small, 1968) has also been surveyed. Crises do not occur frequently in international politics, despite their importance; nor do they represent typical behavior. Wars, though even more important, are rare. Finally, extreme evidences of cooperation like alliance, economic integration, or confederation are equally rare and are typical on the other side of the scale. The advantage of events interaction approaches is that they take up the rest of the normal curve, the vast preponderance of international acts (see Figure 2).

The work of the Situational Analysis Project will in the future be devoted to the refinement of hypotheses and testing of propositions generated in a range of studies. Though we have emphasized the scaling of events on a conflict-cooperation continuum, the events can be categorized along lines of the WEIS project under McClelland and thus permit some comparisons between the international systems of the nineteenth and late twentieth centuries. Crisis propositions of McClelland, Hermann, Brody, and North could be tested by a time-series analysis of the "conflict spiral" in several of the nineteenth-century crises. Crisis stages might even be approached through Guttmann scale techniques, prescribing necessary escalatory and deescalatory steps in a sequence of conflictive or cooperative events. Many arguments of other investigators can be carried further.[6] Singer and Small hypothesize that alliances reduce interaction opportunities and thereby limit flexibility in the system, making for greater conflict. By seeking to elucidate the event flow between alliances and subsequent patterns of cooperation or conflict, one may observe the mechanism by which the Singer effect may operate (or not operate). In more general terms, investigators have hypothesized that

international conflict would be accentuated by either bipolarity or multipolarity; by the presence or absence of a balance of power; by certain thresholds in transaction indices; by certain types of military threats; by certain degrees of social and/or geographic distance between states; by certain types of ideological or economic orientation; by the presence of internal conflict, and so on. In several of these cases, new measures of independent variables will have to be provided. They may then be compared, however, to our dependent measures of cooperation and conflict. The central postulates of international relations theory, in short, may be tested against the data generated by the Situational Analysis Project. In the years that follow we hope to make a start on actually testing these variables against diplomatic outcomes.

NOTES

1. In the course of these inquiries, propositions linking domestic conflict with international conflict, successful deterrence with capability perceptions, alliance aggregation, disaggregation and war, bipolarity and multipolarity, and conflict have been subjected to searching examination, in several cases, suggesting new hypotheses.
2. Contrast this definition with that given in Alker and Bock (1968, p.106).
3. The three typologies were McClelland's World Event Interaction Survey (WEIS) typology, Corson's modified version of the WEIS typology, and the categories used by Moses, Brody, Holsti, Kadane, and Milstein in their Q-sort scaling technique. See Corson (1970) and Moses et al. (1967).
4. A full explanation of the differences between rank-order, internal, and ratio scales appears in Torgerson (1958, p.16).
5. This technique was adapted from that suggested in Stevens (1966) for obtaining ratio scales.
6. See for example Healy and Stein (1971) and Hart (1972).