

AD-A014 616

DECISIONS, EVENTS AND PERCEPTIONS IN
INTERNATIONAL CRISES. VOLUME I.
MEASURING PERCEPTIONS TO PREDICT
INTERNATIONAL CONFLICT

David M. McCormick

First Ann Arbor Corporation

Prepared for:

Office of Naval Research
Advanced Research Projects Agency

July 1975

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE

262108

FAAC

FIRST ANN ARBOR CORPORATION

209 East Washington St. • Ann Arbor, Michigan 48106 • 313-761-5836

AD A914616

FAAC REPORT FR267U/2537

**DECISIONS, EVENTS AND PERCEPTIONS
IN
INTERNATIONAL CRISES**

VOLUME I

**MEASURING PERCEPTIONS
TO PREDICT INTERNATIONAL CONFLICT**

Reproduced by
**NATIONAL TECHNICAL
INFORMATION SERVICE**

US Department of Commerce
Springfield, VA. 22151

July 1975

DDC
RECEIVED
AUG 29 1975
RECEIVED
B

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

Sponsored by

Advanced Research Projects Agency
ARPA Order No. 2873

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) First Ann Arbor Corporation 209 East Washington St. Ann Arbor, Michigan 48108		2a. REPORT SECURITY CLASSIFICATION Unclassified	
		2b. GROUP	
3. REPORT TITLE Decisions, Events and Perceptions in International Crises - Volume I			
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Final Report			
5. AUTHOR(S) (First name, middle initial, last name) David M. McCormick			
6. REPORT DATE July 1975	7a. TOTAL NO. OF PAGES 139 + iii 153	7b. NO. OF REFS 41	
8a. CONTRACT OR GRANT NO. N00014-75-C-0328	9a. ORIGINATOR'S REPORT NUMBER(S) FR267U/2537		
b. PROJECT NO. c. 5W10	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)		
d.			
10. DISTRIBUTION STATEMENT <div style="border: 1px solid black; padding: 5px; width: fit-content; margin: auto;">DISTRIBUTION STATEMENT A Approved for public release Distribution Unlimited</div>			
11. SUPPLEMENTARY NOTES		12. SPONSORING MILITARY ACTIVITY Advanced Research Projects Agency Human Resources Research Office Arlington, Virginia 22209	
13. ABSTRACT This Final Report describes research on Israeli perceptions of the Arabs during the 1967 and 1973 conflicts. A content analysis scheme was devised to measure perception of threat and time pressure from publicly available documents. The results of this manual content analysis were compared with an automated content analysis of the same documents, and threat perception indices were used in conjunction with event data to statistically predict the outbreak of violent conflict. It was found that the manual content analysis index of threat perception was a good predictor of conflict, increasing sharply several days before both wars. The automated content analysis index was not as good a predictor, and the actions each side directed at the other did not predict conflict escalation with a measurable time lag. The measure of time pressure was not sufficiently sensitive to be useful. Thus, the conclusions are that a reasonable index of threat perception can be derived from the public statements of at least Israeli decision makers, and this index apparently predicts conflict much better than events alone.			

FAAC

FR267U/2537

DECISIONS, EVENTS AND PERCEPTIONS
IN
INTERNATIONAL CRISES

VOLUME I

MEASURING PERCEPTIONS
TO PREDICT INTERNATIONAL CONFLICT

July 1975

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Advanced Research Projects Agency or the U.S. Government.

TABLE OF CONTENTS

	Page
Executive Summary	i
Chapter I - Introduction	1
Chapter II - Development of Data Measures	4
Background	4
Approach	7
Pre-Test Development	8
Results of Pre-Test	14
Coding the 1967 and 1973 Crises	25
Chapter III - Index Construction	29
Manual Content Analysis	30
Inquirer II	39
Missing Data	43
Event Data	44
Validity of the Measures	48
Chapter IV - Analysis	63
Hypotheses	64
Hypothesis Testing	67
Chapter V - Conclusions	92
Suggestions for Further Research	102
Appendix A - Coding Rules	105
Appendix B - Computer Content Analysis	125
Appendix C - List of Documents	135
Bibliography	137

LIST OF FIGURES

Figure Number	Title	Page
1	Expected Time Pressures	10
2a	Time Pressure During June 1967 Crisis	34
2b	Time Pressure During 1973 Crisis	35
3	Polynomial vs. Straight Line Fit (1967 Crisis)	45
4a	Manual Threat Perception - 1967 Crisis	50
4b	Inquirer Threat Perception - 1967 Crisis	51
4c	Manual vs. Inquirer Threat Perception - 1967 Crisis	52
4d	Manual Threat Perception - 1973 Crisis	53
4e	Inquirer Threat Perception - 1973 Crisis	54
4f	Manual vs. Inquirer Threat Perception - 1973 Crisis	55
5a	Predicted Values from Model Including First Day of War	89
5b	Model Without First Day of War	89
6	Conflict vs. Predicted Conflict	101

LIST OF TABLES

<u>Table Number</u>	<u>Title</u>	<u>Page</u>
1	Sample of Crises for the Pre Test	8
2a	Bivariate Correlations - 1967 War	57
2b	Bivariate Correlations - 1973 War	57
3a	Arab Physical Conflict (t-1) → Israeli Physical Conflict	69
3b	Israeli Physical Conflict (t-1) → Arab Physical Conflict	69
4a	Arab Verbal Conflict (t-1) → Israeli Physical Conflict	71
4b	Israeli Verbal Conflict (t-1) → Arab Physical Conflict	71
5a	Lagged Arab Verbal Conflict → Perceptions of Active/Negative	72
5b	Lagged Arab Physical Conflict → Perceptions of Active/Negative	72
6a	Correlation Between Arab Verbal Conflict and Threat Perception	73
6b	Correlation Between Arab Physical Conflict and Threat Perception	74
7	Correlations Between Lagged Inquirer Threat Index and Manual Threat Index	77
8a	Regressions Predicting Israeli Conflict at t_0	81
8b	Regression of Israeli Physical Conflict on Threat Perception	82
8c	Rank Order Correlations of Israeli Physical Conflict with Threat Perception (First Day of War Included)	82
8d	Manual Threat Index (-1) → Israeli Physical Conflict	83
8e	Residual Threat Index (-1) → Israeli Conflict	83

LIST OF TABLES (Cont'd)

<u>Table Number</u>	<u>Title</u>	<u>Page</u>
8f	Manual Threat (-5) + Israeli Physical Conflict	84
8g	Residual Threat (-5) + Israeli Physical Conflict	84
8h	Smoothed Manual Index (-5) + Arab Physical Conflict (-1) + Israeli Physical Conflict	85
8i	Residual Threat Index (-5) + Arab Physical Conflict (-1) + Israeli Physical Conflict	86
8j	Residual Threat Index (-1) + Arab Physical Conflict (-1) + Israeli Physical Conflict	87
A-1	Examples of Value Threatened	120

EXECUTIVE SUMMARY

In times of crises, U.S. decision makers face difficulties anticipating leadership decisions of other nations, particularly those leading to violent conflict. Events move rapidly and the ability to reliably predict behavior by even a few days would be extremely valuable. Although crises may sometimes appear to break out spontaneously with no more than a few hours to act, it was our expectation that the escalation of serious crises (that is, those which result in violent conflict) develops over a matter of days. While alerts and mobilizations may be ordered very rapidly, we anticipated that at least one party to serious crises will betray some characteristic stress a few days before the actual outbreak of violence. This project was intended to determine if there are such early warnings of crises and what characterizes them.

Existing techniques used to make predictions of leadership behavior include the study of military movements, captured plans, other explicit indicators of intended behavior, and the subjective examination of public documents. These techniques are not particularly successful at predicting the outbreak of violent conflict.

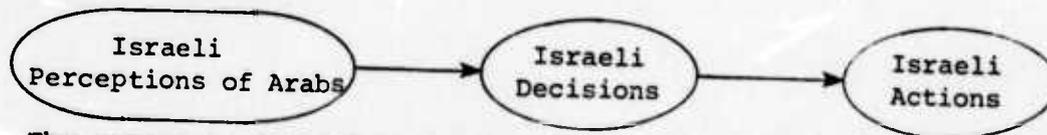
We believed a viable alternative would be the systematic measurement of national leaders' perceptions via content analysis of public documents combined with the linkages between perceptions and decisions and actions. Therefore, we developed a methodology to measure perceptions and examined two potential linkages to decisions and actions. In

summary, we found that perceptions can be measured and that they did predict the outbreak of violent conflict in the cases studied.

Based on these findings, we can envision, in the long run, a real time system to receive written or spoken statements, automatically content analyze them, display the results and provide estimates of the likelihood of the outbreak of violent conflict. DoD intelligence analysts could then combine this with other information, interpret it in light of their knowledge of the circumstances and develop their total short range forecasts of the outbreak of conflict.

We examined twelve major international crises from the outbreak of World War I through the Cuban Missile Crisis to develop the procedures for measuring national leaders' perceptions. Subsequently we successfully tested the measures' validity on the Arab-Israeli crises of 1967 and 1973. In addition, we tested with positive results, the linkages between perceptions and actions.

The study was originally based on the assumption that perceptions of the external world led to decisions which then led to subsequent actions as shown below:



The report is divided into two volumes, the first examines the link between perceptions and actions statistically, bypassing the actual decision, assuming only that the relationship is systematic. The methodology made no assumptions about the wisdom of the decision or the goals of the decision maker.

The second volume investigates the linkage between perceptions and decisions employing the use of decision trees. This methodology examines

the extent to which decision makers reach decisions which are value maximizing given their perceptions of the situation. Thus, if we could measure their perceptions in historical situations and find their decisions to be consistently value maximizing, we could predict their actions in future situations.

Using the 1967 and 1973 Arab-Israeli wars as test crises, we arrived at the following results:

- Threat perception predicts the outbreak of violent conflict quite well (explained variance = 82 percent) approximately five days in advance. (See pages 79-91.)
- The perception of threat can be measured with content analysis. (See pages 14-29.)
- Decision trees have limited ability to explain or predict decisions from perceptions, depending on the characteristics of the decision. (See Volume II.)
- The use of content analysis of decision makers' statements as inputs to decision trees is marginal. (See Volume II.)

The first two findings are the most important. They show that there exists a strong, useful linkage between threat perception and the outbreak of violent conflict in the crisis escalation period. We showed that we were able to bypass the actual decision process and explain actions directly from perceptions even though we had only inconsistent ability to explain the actual decision using decision trees.

Because of these positive results, we decided to extend the tests back one further step to determine if we could identify the linkages between Arab actions and Israeli perceptions or Israeli reactions. We concluded from this that the perceptual mechanism is sufficiently complex that it cannot easily be predicted or bypassed. The analyses indicate that there are clear advantages to measuring the perceptions and predicting actions directly from them rather than from previous events.

Now the study must be expanded to determine if leaders in other recent crises react similarly. We must also test the sensitivity of the predictor to assure minimum false alarms. Additionally, while the manual content analysis we employed is adequate for research purposes, it must be computerized for operational use. A computer would efficiently and reliably process the continuous stream of inputs in a real time system.

Finally, a methodology must be developed for computerizing the evaluation of coded content analysis data. For example, we would want the computer to compare current patterns of observations with the patterns exhibited in previous crises. The system should be able to display to the analyst previously observed patterns of attitudes most similar to those currently observed, and the estimated types of behavior most likely to occur in the near future.

Thus, we see the following work as most essential to continued progress of this system.

- Expand the study to include other crises.
- Check false alarm rate by examining non-crisis periods.
- Computerize the manual content analysis.
- Develop computerized pattern identification and conflict estimation techniques.

CHAPTER I

INTRODUCTION

It has become obvious in the post-World War II era that short-intense international crises are one of the realities of life. This study is one of a series of ARPA-sponsored research projects designed to understand and, thereby, to better anticipate and manage these crises.

One difficulty scholars and decision makers alike have had is a lack of simple measurement instruments they can use to monitor and study crisis dynamics. The purpose of this study was to use content analysis to measure three of the concepts developed by Hermann (1972), namely; threat perception, time pressure, and surprise. Additionally, we employed these measures in conjunction with measures of the sequence of events to empirically examine two Arab-Israeli crisis periods that included war: May-June 1967 and September-December 1973. The argument generally presented in favor of measuring perceptions is that they are a necessary intervening variable between a set of input behaviors serving as a stimulus, and the set of output behaviors which are the response. It was our intention to empirically determine how much additional explanatory/predictive power is gained by examining the linkage between perceptions and behavior.

There are, broadly speaking, three sets of research which jointly underlie this project;

Hermann's work with the concepts time pressure, threat and surprise in crises.

The "Stanford Studies" of crisis escalation.

The event-interaction studies of McClelland, Azar and others.

Hermann's man-machine simulations demonstrated that the concepts time pressure, threat and (to a lesser extent) surprise are useful predictors of decision makers' behavior during times of crisis. Because of his findings, we sought to develop measures of decision makers' perceptions of these concepts during times of crisis. We expected that decision makers in actual crises would also act predictably on their perceptions of time pressure, threat and surprise.

To test the predictive power of these measures, we used a modification of the mediated Stimulus-Response (S-R) model as employed by North, Holsti, Brody, Zinnes, Choucri and others.* They developed and tested in several experiments, a model proceeding from stimulus events through two perceptual processes to reaction events. Overall, these studies have found support for the linkages between events, perceptions, and reactions in that temporal sequence.

Finally, a number of scholars working with event data alone have attempted to predict events directly from events completely bypassing the role of the decision maker.** While these studies have consistently shown an ability to predict a behavioral response to a stimulus in the same time frame as the stimulus, they have been much less successful at predicting a systematic reaction at any time after the initial stimulus.

The current study developed measures of the Hermann concepts and employed them as the perceptual variables in a mediated S-R model. As

* Holsti (1972), Holsti, North and Brody (1956, 68), North (1967), North, Holsti and Brody (1964), North and Choucri (1968), Zinnes (1972)

** Azar (1970, 1974), McClelland (1961), McClelland and Hoggard (1969), Tanter (1974), James McCormick (1974).

an alternative, we examined the event-interaction model to determine how much predictive power is added by measuring perceptions and including them as an intermediate step between actions and reactions.

CHAPTER II

DEVELOPMENT OF DATA MEASURES

BACKGROUND

The purposes of this portion of the research are twofold: (1) to develop measures of decision makers' perceptions during crises and, (2) identify relationships between those perceptions and events characterizing the Arab-Israeli conflicts of 1967 and 1973.

In its broadest sense, the research is driven by the mediated stimulus-response model as used in the series of Stanford studies of conflict and integration. As we are using it, the model assumes events occur, are observed by decision makers, are translated into a revised or reinforced image of reality and are then again translated into a behavioral response to this image. The initial perception of the stimulus is more abstract than the revised/reinforced image. This distinction is different from the Stanford distinction of "image of reality vs. intention to act" (r vs. s). We have simply broken the image of reality into two components. The intention to act is the subject of Volume II of this report.

Political scientists have spent an extraordinary amount of time and effort developing means of measuring events and, subsequently, collecting event data from various open sources. Because of this, we decided we would use existing methods and/or data to measure events characterizing the crises. We chose to concentrate our data gathering efforts in the area of perceptions of international crises.

From the standpoint of a decision maker, crises can be characterized in innumerable different ways. They may, for example, differ in terms of the

relative characteristics of the participants, their relationships to the existing international system, potential impact on the total international system, total military/political power of the primary participants, etc. Charles Hermann (1972) has worked, with some success, with the concepts of time, threat, and surprise in simulated crises.

Despite his emphasis on these concepts, he has done relatively little to develop any good measures useable for analytic purposes. Since we are specifically concerned about the decision makers' perceptions of the crises, we chose to emphasize information which most directly taps that aspect. One of the measurement systems addressing this type of question is content analysis.

There are two types of content analysis which can be employed when attempting to tap decision makers' perceptions of crises: analysis of manifest content, and latent content. The former examines the thoughts the decision maker more or less consciously attempts to convey to the readers or listeners. The latter attempts to measure underlying attitudes the decision maker has regardless of the communication's explicit content. Manifest content analysis is often performed manually with coders using explicit coding rules to guide them in reading and interpreting decision makers' statements. It has the advantage of being able to make relatively subtle distinctions in meaning, hence it may be less sensitive to the size of the sample. It has the disadvantage of being subject to variations in human coding, and in being relatively susceptible to intentional deceit by the decision maker. If large volumes of information are being processed, computer algorithms can be developed to perform manifest content analysis.

Latent content analysis normally examines word choices as indicative of the decision makers' beliefs regardless of manifest content. Because it

is usually little more than a complex word count mechanism, latent content analysis has very rigorous coding rules which are frequently computerized. As a consequence, reliability is very high.

Additionally, the speaker or writer presumably will find it more difficult to hide his true feelings from this type of measurement device than from a manifest measure. On the other hand, the latent measure suffers from its simplicity. In many instances it is difficult for it to distinguish major differences in meaning. For example, given one standard approach,* the sentence, "The Arabs have issued a ceaseless torrent of threats against our independence and territorial integrity" would be coded as barely different from the sentence, "The Arabs have consistently issued statements supporting our independence and territorial integrity." Both sentences would have the words "issued," "independence," and "integrity" in common. The words "consistently" and "ceaseless" are coded similarly. The two sentences would, therefore, have four words with nearly identical codes. The words "stream" and "statements" do not carry any codes. The only differences arise from "supporting" versus "threats" and "against." And these differences are only additive. Codes for the two sentences are shown below.

	Sentence 1 (...threats against...)	Sentence 2 (...supporting...)
Weak 1	1	0
Strong 1	2	2
Strong 2	2	1
Strong 3	0	1
Negative 2	1	0
Negative 3	1	0
Positive 2	1	2
Positive 3	1	1
Active 2	0	1
Passive 3	1	0

* The Inquirer II using an unmodified Stanford Dictionary is described in Appendix B.

We can see that although the intent of the two sentences are completely reversed, the coding is only slightly different. Content analysis of this type depends on carefully considered dictionaries and retrieval systems to eliminate as much distortion as possible. It further depends on large samples to minimize competing effects of sentences such as those shown above.

APPROACH

In this research, we employed manifest content analysis in an attempt to measure Herman's dimensions: time, threat, and surprise. We also used the Inquirer II with the Stanford dictionary to measure dimensions of active/passive, strong/weak, and positive/negative. Since these dimensions have been used (Mogdis, 1970) to index threat, we were able to compare the Inquirer measures versus manual measures of threat.

Since no content analysis schemes explicitly measure time, threat, and surprise, it was necessary to develop a set of coding rules to extract this type of information from verbatim public statements of national decision makers.

Our research strategy was to semi-inductively examine twelve crises looking for patterns of words reasonably consistent with our theoretical ideas. After identifying these patterns during pre-test we tightened our hypotheses and coding rules and applied them more deductively to a set of documents covering the Suez crisis of 1956. This analysis permitted us to check the extent categories held up, and also to check inter-coder reliability.

We used these latter results to make the final alterations to the coding scheme before finally applying it to the 1967 Six-Day War and the 1973 Yom Kippur War.

PRE-TEST DEVELOPMENT

For the pre-test we selected the twelve crises listed in Table 1 and instructed coders to follow a set of loosely formulated coding rules.

TABLE 1

SAMPLE OF CRISES FOR THE PRE-TEST

German invasion of Belgium (1914)	Soviet invasion of Hungary
German invasion of Russia (WW II)	Quemoy - Matsu
Pearl Harbor	Soviet invasion of Czechoslovakia
Berlin crisis of June 1948	Cuban - Bay of Pigs
Communist take-over of China	Cuban missile crisis
Korea	Gulf of Tonkin

Preliminary coding rules were developed from the following initial ideas concerning the concepts of time, threat, and surprise.

Surprise Element

We expected "surprise" would be expressed by individual words appearing in many of a large number of different contexts. These words, we believed, would include some directly referencing surprise such as the words "surprise," "unexpected," and "startling." We also believed that words reflecting secrecy on the opponent's part would evidence surprise. Included would be words such as "secretly," "unprecedented," and "unprovoked." Yet another category of words we thought might indicate surprise are those describing the opponent's actions as rapid. Because of this, we expected words describing the opponent as having acted precipitously, suddenly, abruptly, etc., as indicative of surprise.

Time Element

Initially, we thought any individual reference to time would be codeable according to both intensity and type. We hypothesized that perceptions of limited time would be reflected either as time related demands one nation makes upon another, or as the direct perception an actor nation has only a limited amount of time within which to perform some action.

We directed the coders to look for statements translatable into:

"A" demands that "B" perform "X" action in some limited time "Y".

or

"A" perceives that "A" must perform "X" action in some limited time "Y".

While the latter is a more direct measure of time pressures we argued that time related demands do indicate the extent time is prominent in the decision maker's mind, and very well may be a projection onto the opponent of his or her own internal time pressures. While there are some real questions of the extent time related demands reflect the sort of time pressure Hermann spoke of, it does appear reasonable to argue it reflects time pressures characterizing the total crisis. Because of this, we believed at these early inductive stages, time related demands were worth including.

We had expected time pressures to be non-linear and continuous and described by the graph of Figure 1.

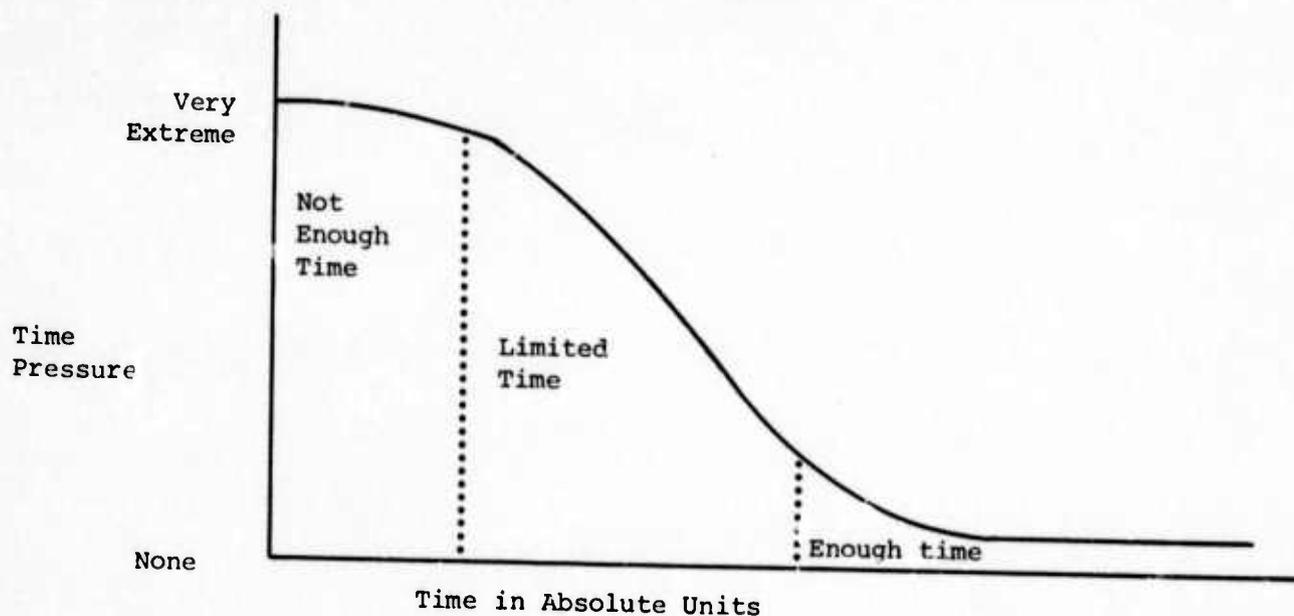


FIGURE 1
EXPECTED TIME PRESSURES

For any given required action, there are two critical points from an analyst's viewpoint. The first is that point where there is insufficient time to act. This will vary depending on the required action. Nonetheless, once that point is reached, further reductions in absolute time available for action does little to increase time pressure (unless it forecloses other potential actions).

The other point, although less definitive, identifies the point at which time pressure is minimal, or non-existent. At this point, the actor has time to proceed in a methodical manner. Additional time in this case does little to reduce time pressure.

Between these points, we believed variations in time pressure would be reflected in different intensities of time references such as "instantly" versus "with all due speed." Thus, we had hoped to categorize each reference to time as either pressure on the actor or as a demand on an object nation.

Additionally, we hoped to develop a dictionary to measure variations within the range of limited time.

Threat

We thought threat would be contained in three distinctly different types of statements. The first is a general description of the situation. In this type of statement we expected decision makers to make blanket references to the situation according to some adjective such as grave, better, excellent, disastrous, etc. This is, simply stated, the most direct evaluation of the threat level. One of the expected weaknesses with the measure derives from its simplicity. It is obvious the decision makers may intentionally distort their own perceptions of the threat either to arouse or calm the public. Thus, our initial concerns were that while most certainly useful as a measure of propaganda, there must be some reservations about it as a measure of the decision maker's own perceptions of the crisis. Even short of such intentional deception, there is a question of the interpretation of such broad adjectives. For example, consider the situation in which a credible threat is being made to something having great value, but the decision maker truly believes with an all-out effort he will allay the threat. It is not easy to predict whether he would describe the situation as good, optimistic, poor, serious or something else. Indeed, blanket references can mean so many different things that, even assuming complete candor and honesty, it is not certain we could evaluate the general descriptions. Despite these very serious potential problems, the measure is so obvious we believed the data should be collected, if only for the pre-test, in order to verify these difficulties.

Value Threatened

The second method of measuring threat involves use of the decision maker's own statements about the object(s) (in the loose sense of the word) being threatened. Thus, for example, if a decision maker believes national existence is threatened, this would be a considerably more serious threat than economic ties. The codeable statements for this category are essentially the publicly declared reasons for participating in the crisis. Here also, the decision maker has some latitude for intentionally attempting to direct public opinion one way or another; by making the value threatened seem particularly severe or particularly innocuous. Nonetheless, because he is tied to reality, it was our expectation that there would be useful information identifying the value threatened during a crisis. For example, we expected vague statements of potential loss such as "world peace" or "victories of the revolution" would indicate a lower threat than more specific statements such as territory, national sovereignty, etc. We expected that when concrete values are threatened, they will be mentioned. Predominant references to platitudes, we expected, would be linked to situations in which the real threat is not particularly severe.

Level of Effort

The third measure of perceived severity of the threat would be measured by the level of effort a nation perceived it would have to exert to subdue the enemy. Initially we believed that as the threat increases, so would the level of effort required to allay it. As a consequence, we decided to examine this as one potential element of threat.

With these initial ideas in mind we formulated a set of coding rules for use in analyzing verbatim speeches of major decision makers in each of

the 12 pre-test crises mentioned above. The documents were obtained primarily from the New York Times. Since we did not intend to make any substantive inferences about these crises, we did not make a serious effort to examine either the universe or a systematic sample of the documents. The coding rules given the coders are listed below.

SURPRISE ELEMENT

Words to check for surprise.

Sudden, surprise, precipitous, unexpected, unforeseen, abrupt, hasty, unprovoked, amazing, astounding, astonishing, shocking, stunning, startling, unprecedented

TIME ELEMENT

Translate sentences into logical statements.

A: Demands

"A" demands that "B" perform "X" action (time reference)

e.g. Israel demands that Egypt return prisoners immediately.

B: Time Pressures

"A" perceives that "A" must perform "X" action (time reference)

e.g. Israel perceives that it must mobilize reserves eventually.

THREAT ELEMENT

Translate sentences into logical statements.

A: General description of the situation

"A" perceives the situation to be (threat)

e.g. Israel perceives the situation to be grave.

B: Potential value threatened

"A" perceives that _____ is threatened.

e.g. Israel perceives that its existence is threatened.

C: Effort Required

"A" must expend "X" effort in order to "Y"

e.g. Israel must expend an enormous effort in order to repel the attacking Egyptians.

RESULTS OF PRE-TEST

Surprise

Our inductive analyses had significant impacts on the subsequent coding schemes. Generally, we found that the expression of surprise or even oblique references to surprise are so inconsistent it is doubtful whether there is any hope of measuring it in the statements of leaders during or after crises.

It appears as though there are conflicting values associated with the element of surprise seriously affecting the decision maker's public references to its occurrence. For the international audience, decision makers seem to find it advantageous to argue that their enemy did something "surprising" in the sense of an extraordinary violation of international law or something done secretly or covertly. They will occasionally denounce such activities as "cowardly." On the other hand, few politicians want to admit that they were taken by surprise, meaning "unprepared." The domestic repercussions of such an admission can be quite severe. It (lack of preparedness) normally results in accusations, hearings, and ultimately in punishment of those persons responsible for the poor state of preparedness. Even Pearl Harbor, for example, prompted both House and Senate investigations of the surprise attack. Understandably, decision makers want to keep that type of accusation to a minimum. The easiest way to do it is to deny being surprised, or even keep the question from being raised. Because of

these two strong norms, it is quite difficult to infer any validity to statements referring to the surprise element.

As examples, consider the following:

In the Cuban missile crisis, Kennedy referred in his speech of October 22 to the placement of missiles twice as "a surprise," once as done under a "cloak of secrecy," and once as done "secretly." While Kennedy did not know months in advance that the missiles were being installed in Cuba, the crisis was not a surprise in the conventional sense of the word. Rumors had circulated for some time of the missile build up, and Kennedy himself appears to have known of the missiles for about a week prior to deciding what to do. Thus, using our measure of surprise we would have observed it well after the surprise had passed.

Similarly in the Cuban Bay of Pigs crisis, which was hardly a surprise, Fidel Castro accused the United States of acting "secretly and cowardly."

Interestingly, the Czech leaders claimed that the Russian invasion of Czechoslovakia was a surprise while the Hungarians argued that the Russian invasion of Hungary was not. If anything, we would have expected the reverse.

In the outset of World War I, Britain and Belgium both expressed surprise that Germany violated Belgian territorial rights while France explicitly stated it was not surprised. The Germans, while accusing the Russians of having instigated the crisis, said they were not surprised. Knowing what we do now, there seems little reason to believe that any of the participants were surprised by the course of events. The examples could run on, but they basically all show the same thing; expressions of surprise do not fit very well with what we currently know about the decision makers' knowledge of the situation during these incidents.

An additional characteristic of the concept of surprise became apparent after the pre-test, and that is it does not have the temporal characteristics of threat or time pressure. Surprise does not increase or decrease as the crisis develops. It normally occurs only once when there is an unexpected outbreak of violence. Thus, it is but a single spike in the time series plot of the crisis. As a consequence, even if it were mentioned in public documents we would not be able to rely on the laws of large numbers which, to some extent, can be relied on for the more continuous concepts. Because of these reasons, we concluded that surprise is not measurable from content analysis.

Time Element

Decision makers make occasional references to time easily divisible into demands versus internal pressures. It appears as though we can easily categorize individual statements into references of "not enough time," "limited time," and "enough time." Indeed, the overwhelming number of references to time pressures fall in the middle category. Decision makers only rarely will admit to not having enough time to act even if it is the case. If they have plenty of time, they will occasionally make references to that fact, but most frequently will not bother to mention the time element at all.

In those instances when time is a matter of concern, we found it difficult to identify the time pressure severity. The tendency of decision makers toward imprecision produces numerous situations wherein it is obvious there is significant time pressure even though it is nearly impossible to judge how severe it is.

The time references we identified are listed below:

Immediately	Quickly
Without delay	Speedily
Forthwith	Rapidly
Instantly	Promptly
Urgently	Soon
Rush	Without wasting time
Within 24 hours	In a brief span of time
Within 48 hours	

There are two characteristics of this list worth mentioning. First, the word "immediately" is by far the most widely used of all of these words, accounting for nearly 50% of all time references. "Immediately" is so widely used, it is quite apparent it conveys less urgency than would normally be associated with its dictionary definition. Normally, it seems to imply recognition of physical or bureaucratic obstacles to action. Some activity will or should begin as soon as possible.

The second point is that with the exception of the word "soon," the words are so close to being interchangeable there seems little point in developing weightings for the time references. Attaching different weights would impose a precision on the measurement system greater than precision of the words' usage. Thus, we treated all references to short time equally. As a consequence, the coding rules require only that time pressure be recorded as "not enough time," "short time" and "enough time."

The distinction between time-related demands and the perception that a nation has only limited time to act appeared clear and were retained. One additional category, however, was identified and was added to the final coding rules. Decision makers occasionally make references to the rapidity

with which actions are occurring. While they may not make direct references to either the time constraints acting on them or on the demands they make on their opponents, the inference of rapidly moving events and the need for quick action is unmistakable. For example, a decision maker refers to "this deadly drama evolved quickly." Because of these kinds of statements, we included a category to record references to rapidly moving events. Each such reference is counted as the equivalent of a reference to "short time."

Value Threatened

Decision makers seem to speak freely about values being threatened during a crisis. Fortunately, they appear to speak relatively truthfully. Where there is distortion, rather than exaggeration, we find a tendency toward platitudes. References to world peace, freedom of all peoples, etc., tend to appear frequently. Our examination showed references to the following values threatened with the associated frequencies.

Value Threatened	Percent of Total Values
Peace	24.0
Freedom	13.5
Security	10.5
Independence	6.5
Specified Military Assets	5.5
Existence	4.5
Status Quo	3.5
Honor	3.5
International Agreements	3.5
Obligations	3.0
National Interest	3.0
Domestic Political Values	3.0
International Law	2.0
Life of Citizens	2.0
Cultural Values	2.0
Integrity	2.0
Rights	1.0
Order	1.0
International Relationships	1.0
Personal Losses	1.0
Democracy	1.0
Dignity	1.0

We were able to classify these values into the following categories.*

Military-Security
Economic-Development
Cultural Status
Political (Domestic)
Political (International)

We used Brecher's definitions except for the two political categories. The definitions are:

Military-Security (MS) comprises all values focusing on questions pertaining to violence, including military alliances and weaponry, and those perceived by the foreign policy elite as constituting a security threat.

Economic-Development (ED) comprises all values which involve the acquisition and allocation of resources, such as trade, aid and foreign investment.

Cultural-Status (CS) consists of those values involving cultural, educational, and scientific matters as well as international status issues as perceived by the decision makers (e.g., honor, dignity).

Political (Domestic) (D) includes values associated with the basic philosophy underlying the domestic political system, e.g., democracy - freedom - Communism.

Political (International) (I) includes values associated with the relationships between states, or the methods by which such relationships are governed.

* These categories are virtually identical to Michael Brecher's "issue areas" (Brecher 1973). It should be noted that we used Brecher's "issue areas" rather than value categories. The latter proved to be a poor taxonomy for empirically observed values threatened.

It can be seen that the value threatened category creates a record of the cognitive identification of a threat to a value rather than a measure of the level of the value threatened. That is, we record the existence of one of a number of different types of values without attempting to measure which is greater than, or less than, others. This measure is then combined interactively with a measure level of threat using the level of effort categories described below.

Level of Effort

One significant theoretical problem with using the value threatened concept as a measure of the level of threat to that value, is that value threatened does not contain, within itself, a good measure of threat credibility. For example, both the Soviet Union and Ecuador can, if they choose, attempt to raise a threat to the U.S. military security. Credibility of the threat may be measured by the relative level of effort required by the U.S. military to counter the threat.

Implications of this point are critical to our discussion. If a value threatened will certainly be lost in the absence of any resistance, we can infer that it may not be lost, given enough resistance. Admittedly, the recipient of the threat may not be able to mount sufficient resistance, but the point is that it is always possible to define a level of resistance to allay the threat. We can also assume that as the level of resistance increases, so will the probability the value threatened will be saved. However, the threatening nation may increase its efforts, thus altering the level of threat and required resistance. On the surface, it appears all that would be necessary to measure the threat is to keep a running account of the probability of loss under different force mixes. Unfortunately, this position suffers from the fact that in documents designed for public consumption, decision

makers only rarely make reference to a finite probability of losing some valued asset. It is a universal dictum that decision makers never acknowledge the possibility of defeat until it has happened (and if possible, not even then).

The fact that decision makers do not speak in probabilistic terms, we believe, is not simply a matter of speech patterns. Rather, we argue that probabilities in the conventional sense are not even a part of the decision process. The entire concept of expected utility or expected loss was developed by economists for situations in which the businessman is interacting with a partially unpredictable market over which he has no control. Because the businessman must make decisions knowing only there are certain probabilities the market or competitors will act in certain ways, the expected utility concept is reasonable for him.

In the case of international political crises, however, the situation is considerably different. The decision maker is not playing a game against nature. Rather, he has considerable control over the probabilities. That is, when faced with a threat to some value, the decision maker is normally in a position of being able to choose some set of behaviors which will minimize the probability the value will be lost. His primary task is to determine how much of what kind of effort is needed to drive that probability of loss as close to zero as he possibly can. It is our belief that the decision maker recognizes this and acts accordingly. His mental calculus thus rarely explicitly addresses the question of the probability of losing the threatened value, but rather what level of effort is needed to save it.

The answer to the latter question can usually be interpreted as a measure of threat severity. One exception to this rule occurs when the object being threatened has a low value. In that case the maximum level of effort which the decision maker is willing to expend to save the object is equal to the decision maker's perception of the severity of the threat.

At the outset of this project, we believed there was an important distinction between the level of effort a decision maker expects to have to commit, the level he/she is willing to commit if need be, and the level ultimately committed.

We had expected that the level a decision maker expects to commit would be his/her estimate of the severity of the crisis at that moment of time. Secondly, the statement of willingness to exert greater effort would define the perceived crisis upper boundary. In practice we found such clear distinctions are not made for public consumption.

The last element, the level of effort actually committed posed some conceptual difficulties which were ultimately resolved. The act of committing resources is an event comprising part of the dependent variable for one phase of the study. While, on the surface, the perceived level of effort and the actual commitments would appear to be measures of the same phenomena, it was our working assumption that they, in fact, are not. Statements such as "we are fully prepared to defend our security" and, "it is our duty to adopt all necessary steps to meet any possible development" reflect a perceived recognition of some threat and do not simply present a plan of action.

Actual behaviors in a crisis may differ significantly from the statements of expectation of willingness for a number of reasons, not the least of which is the impact of the statements on opponents' behavior. In the Cuban missile

crisis, for example, no shots were fired, and no Soviet ships were stopped or sunk. Kennedy's statement that we would do that was sufficient to prevent such drastic action. We are the first to admit that there is a danger of circularity in our development. However, the reader will note as we proceed through the analysis that we go to rather extensive pains to avoid that trap.

In examining the twelve crises, we found that level of effort can reasonably be divided into four major categories. There are: 1) military, 2) social/domestic, 3) diplomatic, and 4) economic sanctions. These categories are relatively self-evident except for social/domestic, which taps the dimensions of effort in which decision makers call for some sacrifice by the civilians to allay the threat. Statements imploring people to show "courage" and "determination" are examples of this dimension.

Finally, as a result of the pre-test we decided to eliminate the general description category. We found it was extremely difficult to refine into sufficiently rigorous coding rules to prevent its serving as a "catch-all" category. Additionally, we were able to observe no systematic variance in the data we collected. Aside from this, we refined the remaining concepts and subsequently formulated the coding rules as presented in Appendix A.

The basic output of a document coded under these rules will contain the following variables:

Time

Not enough

Short

Enough

Value Threatened

Military Security

Economic Development

Cultural Status

Political Domestic

Political International

Other Domestic

Level of Effort

Military Security

a. All out

b. Major

c. Moderate

d. Minor

Social Domestic

a. Total

b. Significant

c. Some

Diplomatic

a. Bilateral

b. International Organization

c. Miscellaneous

Economic Sanctions

For each document coded, the raw data comprise the frequency of references to each of the above categories. For example, we might find five references to all out military level-of-effort, four to moderate military effort, etc.

Before the coders began coding the 1973 and 1967 crises they were instructed to code identical documents for the 1956 Suez crisis. Our purpose in doing this was twofold: (1) to obtain measures of inter-coder reliability and (2) to resolve any final problems with the coding rules.

To test for inter-coder reliability, we correlated the coded raw frequencies for each document and the percentages in each category as a function of the total number of coded items. We found little agreement between coders on raw frequencies but strong (r^2 .90) correlations between them on the percentages. That is, they differed occasionally by as much as 50% in their decisions of how much material should be coded. They did, however, consistently code the same proportions of each item even at the most severe test - the document level. Since the percentage distribution of perceptions was our primary concern, we felt confident to proceed with the coding rules as they were presented above.*

CODING THE 1967 AND 1973 CRISES

For the 1967 and 1973 Arab-Israeli conflicts we content analyzed as many of the verbatim speeches of high level Israeli decision makers or spokesmen as we could find during and immediately surrounding the crises. We defined decision makers to be cabinet level, and spokesmen including U.N. delegates. Some of the latter clearly were not high level persons, but we assumed the content of their official U.N. speeches could be taken as their nation's position.

* The problem of determining what should be coded was never totally resolved although, as the coding progressed, we did identify some of the major sources of variance.

For the 1967 Six Day War we gathered and analyzed 42 documents from May 1 to June 30. For the Yom Kippur War of 1973, we gathered and analyzed 38 documents from September 1 to December 31.* Occasional cross checks were made between the coders generally showing reliability patterns similar to those we uncovered in the pre-test. We did, however, notice a tendency for coders to become confused in terms of when to include references to the recent past. We had originally expected decision makers' references to values which had been threatened in the past, and which continue to be threatened, would indicate one aspect of the threat today. Similarly, we expected references to ongoing levels-of-effort would partially indicate threat perception. Unfortunately these decisions required more subjective interpretation by the coders than we wanted. One of the coders was never able to grasp the subtle distinctions between references to historic and current threat. Much of his work had to be recoded. As the project progressed, it became apparent that, for reasons of reproducibility, the current activity references would have to be dropped altogether. Thus, for methodological, logical, and substantive reasons, subsequent analyses using this coding system would be greatly facilitated by eliminating all references of this type.

The coders did not evidence serious problems in identifying the proper category for any codable statement. Some minor problems arose in determining the particular level of threat for the Military-Security and Social-Domestic categories. We found, however, that where differences occurred, they generally varied by only one scale value (e.g., from moderate to minor military

* A list of the documents is presented in Appendix C.

actions). These problems tended to cancel each other out resulting in the observed strong inter-coder reliability. Additionally, the elements of these dimensions were eventually combined into a weighted scale thus further reducing the effects of any incorrect codings.

Another of the potential difficulties we encountered resulted from verbal statements occasionally having characteristics in common with more than one category. This was particularly true of the value threatened category in which a threat to one phenomena (such as peace) in fact has significant meaning to at least two value categories (Military-Security and International-Political). We required the coders to make a decision concerning the single category most affected. This requirement was made to avoid problems associated with the fact that, to some extent, all values are probably threatened by virtually any threatening act. In retrospect, the problems arising from this decision were not serious.

Despite the fact that, for this contract, the problems associated with multiple categories being referenced were not severe, they remain a potential source of error. Subsequent to completing coding, we became acquainted with the recent mathematical developments in fuzzy sets (Zadeh 1965). His developments are, in part, intended to make more practical the assignment of different weights (membership numbers) to phenomena simultaneously associated with different sets. Although a full scale investigation of fuzzy sets was beyond the scope of this contract, our initial impressions are that it may have some significant contribution to data making efforts in research of this type.

In summarizing the coding, we found it went reasonably smoothly. However, as previously noted, subsequent content analyses would be well advised

to severely restrict the coding of references to current activities. Not only would that resolve some coding difficulties but, as we will show later, little useful information is lost in the process. Additionally, there exists a potential for difficulty in the area of references to multiple categories. While this did not empirically create any serious problems, it was obvious it made coding difficult, and certainly had the potential of introducing distortions in the results. The solution to this problem is not clear although we believe the use of fuzzy sets is a reasonable consideration.

CHAPTER III

INDEX CONSTRUCTION

The project deals with three distinct data sets: manual (manifest) content analysis, machine (latent) content analysis, and event data. For each of these sets, we assumed the actors were Israelis versus Arabs. We defined the Arabs as a single actor comprised of Egypt, Jordan, and Syria. We performed this aggregation for three reasons. First, we found that in the public statements used in the content analysis, the Israelis did not make sufficient distinction between the different Arab states to warrant disaggregating their perceptions by nation. Since the event data were used in conjunction with the content data, there would have been little point in keeping the events disaggregated. Secondly, the Israelis appeared to treat the Arabs as something of a unitary bloc during the crises under examination. Certainly their public statements bear this out. Also, Egypt and Syria had formal defense pacts during the entire period of 1967 examined, and Jordan put its forces under an overall Arab command on May 31. In 1973, there was even more coordination between Egypt and Syria, although Jordan was more of an independent actor. Third, aggregating the Arab events together minimized the statistical problems associated with sparse data matrices, of particular importance when working at such a low level of aggregation as the one-day period. We will discuss the index construction for these sets systematically.

MANUAL CONTENT ANALYSIS

As with all content analysis strategies, we were faced with the dilemma surrounding the fact decision makers make statements of varying length, and this length appears inconsistently related to importance. Since content analysis normally involves counting numbers of references to one phenomenon or another, the raw frequency counts obviously increase with length. To compare these raw frequencies across documents hopelessly confuses the variance in document length with the variance in emphasis within the document.

Rather than weight documents by length, we chose to compare the relative emphasis on concepts within documents. Additionally, we chose to compute our indices by comparing the emphasis on various subcategories within the major concepts of time, level of effort, and value threatened. For example, we compared the frequency of major military effort to the total references to all level-of-effort categories. Similarly, we wanted to compare the number of references to military-security values threatened with the references to political international values threatened. Therefore, for each document, we computed the references to each subcategory as a percentage of the total references to the major category.

That is,

$$X_{ij} = \frac{F_{ij}}{\sum_{j=1}^K F_{ij}}$$

where

i = the major categories; time, level-of-effort, and value perceived

j = subcategories within major categories

K = total number of subcategories within a given category

X_{ij} = score for a document of ij

F_{ij} = raw frequency of ij

If there were no references in a document to one of the major categories, the X_{ij} for all j in i were set equal to zero.

For those days on which we had more than one document, we computed the mean value of each X_{ij} across the documents. As a result of these aggregations, we had two data sets of 21 variables for the manual content analysis. For the 1967 and 1973 crises we had respectively 21 and 25 days with measured perceptions.

Time

As mentioned above, manual content analysis was designed, in part, to measure time pressure under each of three situations, namely; 1) Israeli perceptions of time available to carry out action, 2) perceptions of the speed of current events, and 3) Israeli demands on the Arabs reflecting time pressure on the Israelis. Time pressure in each case was measured as "insufficient time to act," "short time," and "enough time." In summary, results of the analysis revealed:

- a) In all three categories there were no references to "insufficient time to act"
- b) In category (3), the Israeli demands on the Arabs generally failed to convey the impression of a time pressure on the Israelis.
- c) In all cases, Israeli decision makers made infrequent reference to time pressures.

The absence of any reference to "insufficient time to act" clearly required we delete this category of pressure from further consideration. Our examination of the data showed that the third category was coded most

frequently during latter portions of the crises. Primarily, these statements were demands for the immediate return of prisoners, the immediate cessation of ceasefire violations, etc. These types of statements seem only marginally related to time pressures acting on the Israeli leaders. Therefore, the time demand category was dropped from the analysis.

The data show that Israeli decision makers do not make frequent references to time pressures. On those days when time pressures are mentioned, there are rarely more than three or four such references. Because of this sparseness, the index construction mentioned above (page 30) is too sensitive to variations in references to the "enough time" category. Therefore we chose to standardize perception of time as a function of document length. To partially control for variations in irrelevant content, document length is defined as the number of words identified by Inquirer II (see Appendix B).

If we examine the plots of the resulting measure of short time pressure index, we see that it has reasonable face validity. The major point supporting the measure is the fact that in pre-war 1967 documents we observe considerable time pressure, whereas in 1973 no mention of time pressure appears until the first day of war. Since in 1967 we know there was more or less traditional escalation, and the Israelis were actually aware of the speed events were moving, we would expect to have observed these references.

The 1973 crisis, at least from the Israeli standpoint, was considerably different from the 1967 crisis. While they were aware of Egyptian and Syrian maneuvers on their borders, they appear to have believed war would not break out.* While they were clearly uneasy about the Arab activities,

* Commission of Inquiry - Yom Kippur War, Partial Report, April 1974.

they did not seem to be as concerned with the tempo of events as they were in 1967. The data reflect this.

Also supporting the measure is the fact the time pressure measure is high for the first few days after the onset of war, and remains near zero thereafter. This does not seem unreasonable.

The major difficulty with the measure, as seen in Figures 2a and 2b, is its sparseness is not suitable for standard time series analysis. During the two crises we examined, it appeared suddenly and disappeared almost as rapidly. It does not exhibit the fairly continuous form that characterize our threat measures. Nonetheless, the peaks do appear at periods generally considered to be points of most severe time pressure. (c.f., Brecher 1974.)

Because the measure is extremely coarse, it fails to pick up minor gradations in time pressure which almost certainly existed in some of those periods. The measure can be used only as a categorical variable and not as a continuous time series variable. For example, one could use the observations of time pressure in pre war 1967 with other variables to categorize that situation as one exhibiting more or less classic crisis characteristics. The absence of such observations in 1973 indicates that if any "crisis" situation existed, it did not create severe time pressures.

Threat

Although time pressure was not mentioned frequently, the threat indicators yielded considerable data. Recall that the threat measures were: value threatened and level-of-effort. The coding rules specify four major subcategories

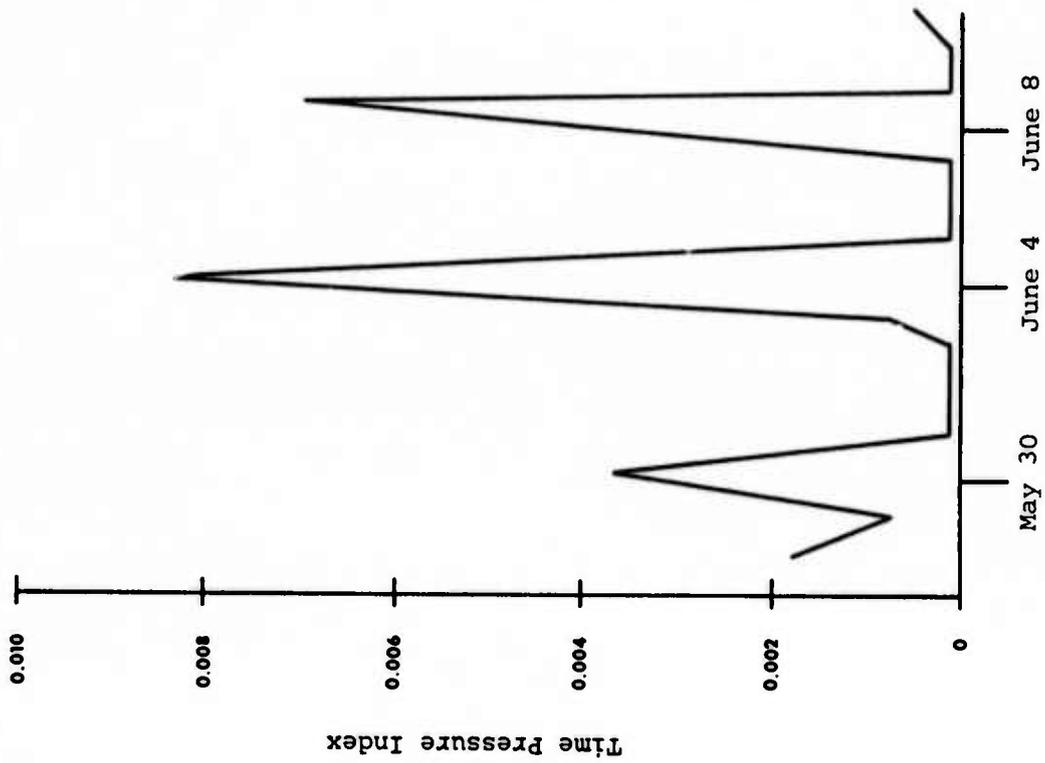


Figure 2a

Time Pressure During June 1967 Crisis

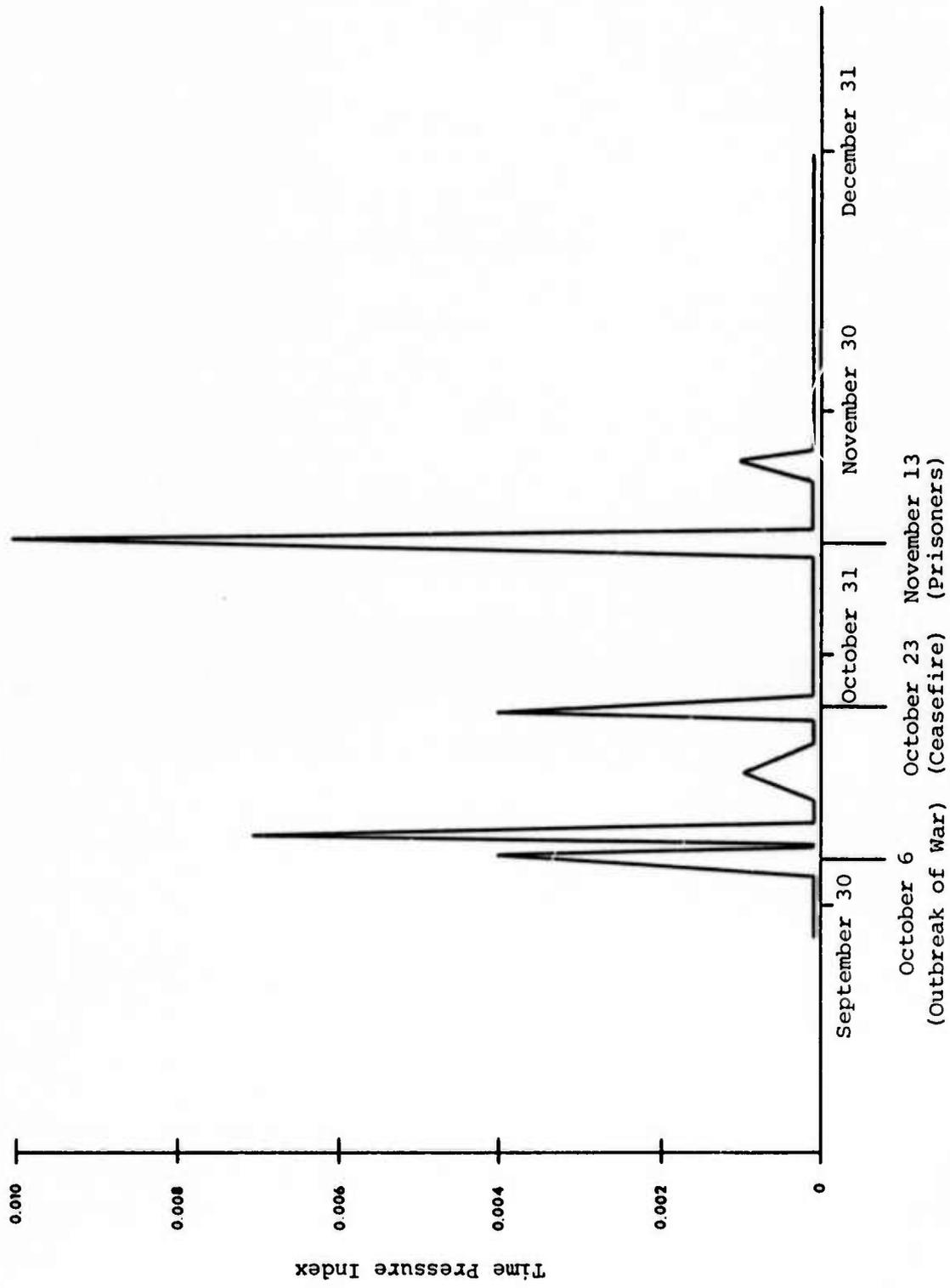


FIGURE 2b

TIME PRESSURE DURING 1973 CRISIS

within level-of-effort: Military, Social-Domestic, Diplomatic, and Economic. The economic level-of-effort category was never mentioned by the Israelis during either crisis, so it was eliminated. In addition, statements of diplomatic level-of-effort are quite distinct from Social-Domestic and Military levels-of-effort. Diplomatic effort requires voluntary cooperation of other parties rather than relying on one's own capabilities. The contention that diplomatic level-of-effort taps a different underlying concept is supported by a -.55 correlation between diplomatic and military level-of-effort.

We argue that emphasis on diplomatic level-of-effort indicates such low level threat it should either be left out of the calculation of threat or, possibly, should be subtracted from it. We chose a method which does the latter indirectly. It is not included in the weighted sums of other elements, but implicitly entered the equation as an element in the formation of internal percentages. Recall that the percentage (X_{ij}) of a variable within a category is:

$$X_{ij} = \frac{F_{ij}}{\sum_{j=1}^K F_{ij}}$$

Since diplomatic level-of-effort is one of the F_{ij} terms, increases in its frequency will increase the denominator thus reducing the values of other X_{ij} .

After examining the types of statements coded as Military-Security and Social Domestic, we concluded that they indeed measure our original conception of threat perception. For example, in the following four statements, the first two fall in the category of a Military-Security level-of-effort, while the latter two are Social-Domestic.

"We shall not permit our borders to be opened to attack."

"If they will try to sow unrest on our border - unrest will come to theirs."*

"We will undertake every sacrifice to open the straits."

"Israel will make unlimited effort to open the straits."**

Therefore, we decided they should be parts of the underlying concept of threat.

Recall Military-Security level-of-effort has four categories ranging from "minor" to "all out war," and Social-Domestic has three categories ranging from "some effort" to "total national effort." "Minor" military effort refers to calling up reserves, small alerts, etc. Based on the figures presented in Kanovsky (1970, p. 94-95) we concluded the cost of "all out war" is at least ten times the cost of a "minor" level of effort. Using these figures and assuming a linear relationship within the categories*** we formed the basic component of threat index as a weighted sum of the Military-Security and Social-Domestic levels of effort. We used the following weights:****

MS A = 10	SD A = 10
MS B = 7	SD B = 5
MS C = 3	SD C = 1
MS D = 1	

* Levi Eshkol Broadcast on May 13, 1967.

** Abba Eban Press Conference on May 30, 1967.

*** By employing a conservative estimate of the range between "minor" and "all out war" categories, we hoped to reduce the tendency toward exaggeration.

**** Intuitive "interval" weighting as opposed to ordinal ranking is advocated by Abelson and Tukey (1959) and Tufte (1969).

Once the weighted sums of the level-of-effort categories were taken at the one day level of aggregation, they were multiplied by the percentage of Military-Security values threatened. This was done to control for levels-of-effort to overcome threats to non Military-Security. It also ensured the content analysis measured more than the Israelis' statements of their intentions. As it turned out, this multiplicative index had more face validity than either of its components alone.

This measure of threat, therefore, is computed from the following formula:

$$T = \sum_{j=1}^7 \alpha_j E_j \times V_1$$

where:

T = Threat

E_j = The percent of level-of-effort subcategory j of the total level-of-effort category

α_j = The weight assigned to E_j

V_1 = The percent Military-Security value threatened of the total value threatened category

We have mentioned our concern about possible circularity in our measure of level-of-effort. Recall we had difficulties in the coding rules regarding references to current activities as measures of threat. We did review the coded documents to eliminate some of the more obvious instances of this potential coding problem. Additionally, we have multiplied the level-of-effort by the percent of references to Military-Security values threatened in a further attempt to reduce this potential problem. As yet

another attempt to assure we are not simply measuring events in a different manner, we decided to create an additional measure of threat which should eliminate all questions of this type of circularity.

Since it was our desire to anticipate behaviors in the future from a measure of threat, we were primarily concerned with lagged relationships between threat perception and Israeli conflict. Because of this, the simultaneous occurrence of threat and conflict is not of interest. Therefore any common variance at time zero, whether "real" or spurious can be eliminated. Recognizing this we regressed Israel's conflict at time, t_0 , on the threat index with missing data estimated. At time t_0 , we then used the residuals as a threat index. Presumably it controls for a description of events occurring simultaneously with the observation of the threat.

INQUIRER II

Our purpose in employing the Inquirer II was to obtain a measure of threat perception in addition to the manifest measure obtained from manual content analysis. Our index of threat is the interaction of Israeli perceptions of the Arabs as strong, negative and active. This is the index used by Mogdis (1970) in an analysis of Sino-Soviet interaction, and is consistent with J. David Singer's (1958) formulation of threat as the interaction of capabilities and intention.

The analysis of latent content attempts to measure underlying attitudes held by decision makers independent of the thoughts they are overtly trying to convey. One of the most widely used algorithms in international relations has been the Inquirer II (and subsequent revisions) computer package combined with the Stanford dictionary. The dictionary incorporates a set of concepts

and related words which are then associated with textual input supplied by the user. The Stanford dictionary is basically an operationalization of Osgood's (1957) semantic differential. That is, it defines three scales: positive-negative, strong-weak, and active-passive. Using the results of extensive testing, Osgood and others have been able to identify the location of roughly 3000 words along these scales. The Stanford dictionary assigns a six point scale for each of these with an implicit seventh point at zero. For example, there are three levels of positiveness, three of negativeness and, by default, a seventh which is neither positive nor negative. Each word in the dictionary is given "tags" associated with its location along these dimensions. The word "fight," for example, is coded negative 2, strong 3, and active 3. "Security," on the other hand, is positive 3, strong 2, and neutral on the activeness dimension.

The program scans the input verbal text* and attaches the proper tags to the words it reads. Finally, the user can supply retrieval commands and receive the aggregated references to each tag.

For example, in this study we instructed the program to total the tags for all Israeli references to the Arabs for each day we had at least one document.

Data that has been tagged can be tabulated according to various retrieval specifications. The tabulation lists all the categories of interest and provides raw frequency of occurrence for each concept in the tabulated text.

* While the Inquirer is logically capable of extraordinarily complex retrieval tasks, it was far more practical for our purposes to permit human intervention to prepare the data for analysis. There are three basic purposes to the pre-processing: in breaking up complex sentences, identification of object, and elimination of extraneous material. For further explanation of Inquirer II, see Appendix B.

The tabulate option also provides the total number of units in the document (e.g., words, sentences, paragraphs, etc.). If the investigator is interested in only a few concepts and not the total contained in the dictionary, he may specify which concepts are to be tabulated and which are to be suppressed. In addition, the tabulate program provides index scores which are obtained by the division of the various frequency scores. The index scores produced are:

$$\text{WORD INDEX} = \frac{\text{total assignments of a given concept}}{\text{total words in the entire document}} \times 100$$

$$\text{CONCEPT INDEX} = \frac{\text{total assignments of a given concept}}{\text{total number of concepts assigned in the entire document}} \times 100$$

$$\text{WORD CONCEPT INDEX} = \frac{\text{total assignments of a given concept}}{\text{total words assigned any concept in the entire document}} \times 100$$

$$\text{SENTENCE INDEX} = \frac{\text{total sentences containing a given concept}}{\text{total sentences in the entire document}} \times 100$$

$$\text{SENTENCE CONCEPT INDEX} = \frac{\text{total sentences containing a given concept}}{\text{total sentences assigned any concept in the entire document}} \times 100$$

Obviously, each of these indices provide different measures of the relative importance of references to the concepts. We chose to use the concept index for the following reasons:

1. It provides an indicator of the importance of each concept relative to all of the concepts,
2. The total of the concept indices is always 1.0, thus treating all documents equally despite variation in speech patterns (such as excessive reference to irrelevant material) or value laden words (which carry multiple tags).

Using the concept index, Inquirer II provides us with one measure of the relative frequencies of the elements of the three perceptual dimensions, however neither the dictionary nor the program actually computes values for the dimensions. To do that, we applied weights ranging from +3 to -3 for each dimension. For example, a tag of strong 3 was assigned the weight of +3, and weak 3 was assigned the weight of -3. The value for a given day for that dimension was, therefore, the weighted sum of the elements along the dimension.

Because we intended to use the Inquirer II measure to get another threat index, we multiplied the values of the three dimensions. Thus, threat = strong x negative* x active. This index of threat perception is consistent with Singer's (1958) formulation of threat. He defined threat as intentions times capabilities. In using the Inquirer, intentions can be considered to be measured along the positive-negative and active-passive dimensions, and capabilities along the strong-weak dimension.

* In order to maintain the proper sign of the threat index, we computed the affect dimension such that negative 3 = +3 and positive 3 = -3.

Thus, our measure of threat using the Inquirer II is the following:

$$T_{lc} = \prod_{J=1}^3 \sum_{i=1}^6 \alpha_i D_{ij}$$

where

T_{lc} = Threat measured from latent content

α_i = Weights for each element of a dimension. These range from -3 to +3

D_{ij} = Concept index measure for each element i of a dimension j

j = Dimensions of active/passive, strong/weak, and negative/positive

MISSING DATA

Unfortunately, decision makers do not make public statements every day to inform us of their perceptions. Obviously, their perceptions do not go to zero on those days they make no public statements. Therefore, we initially treated days with no documents as missing data.

There are, of course, two options available to handle missing data: 1) simply exclude the observations from analysis, or 2) estimate the missing values. The dilemma is a familiar one and can be reduced to a tradeoff between small (and thus possible biased) samples with observed measurements against a superior sample of cases with some data of questionable quality.

Because time series analysis permits the use of some powerful assumptions about the ordering of reality, we used polynomial regressions in time to fit the observed data and then to evaluate the missing periods. That is, we regressed the observed values of threat perception against a polynomial in time, and then evaluated the polynomial of those time periods with missing data.

This approach has an advantage over straight line interpolations because it minimizes the discontinuity at each of the observation points. This point is illustrated by the following portion of the 1967 crisis shown in Figure 3. Because this method also treats the set of points as an entity rather than individually, it tends to depress apparently random fluctuations. Thus, it serves simultaneously to estimate missing data and smooth the time series function. It should be pointed out that since this is little more than a curvilinear interpolation between points, the normal tests of statistical significance are irrelevant. However, this creates no difficulties because we make no pretense the polynomial equations have any applicability outside the time range included within our data base. The real question is how much of the apparently random variance the user is willing to "smooth" away. This is admittedly a judgmental decision.

EVENT DATA

There has been considerable effort expended by others gathering data measuring international events. In the course of this contract, we engaged in a preliminary investigation of the spacial and temporal domains of the more prominent data sets available. Included in our survey were the CREON, DON, MEIS, WEIS and COPDAB data sets. Of these only WEIS and COPDAB covered both the 1967 and 1973 Arab-Israeli wars. While Azar's COPDAB collection focuses on the Middle East, WEIS is global. Azar was not prepared to release his data at the daily aggregation level. WEIS, on the other hand, was easily available at the daily aggregation for both wars. Additionally, despite criticisms occasionally leveled at WEIS for its heavy reliance on the New York Times, the data set enjoys one of the best reputations of the various event collections.

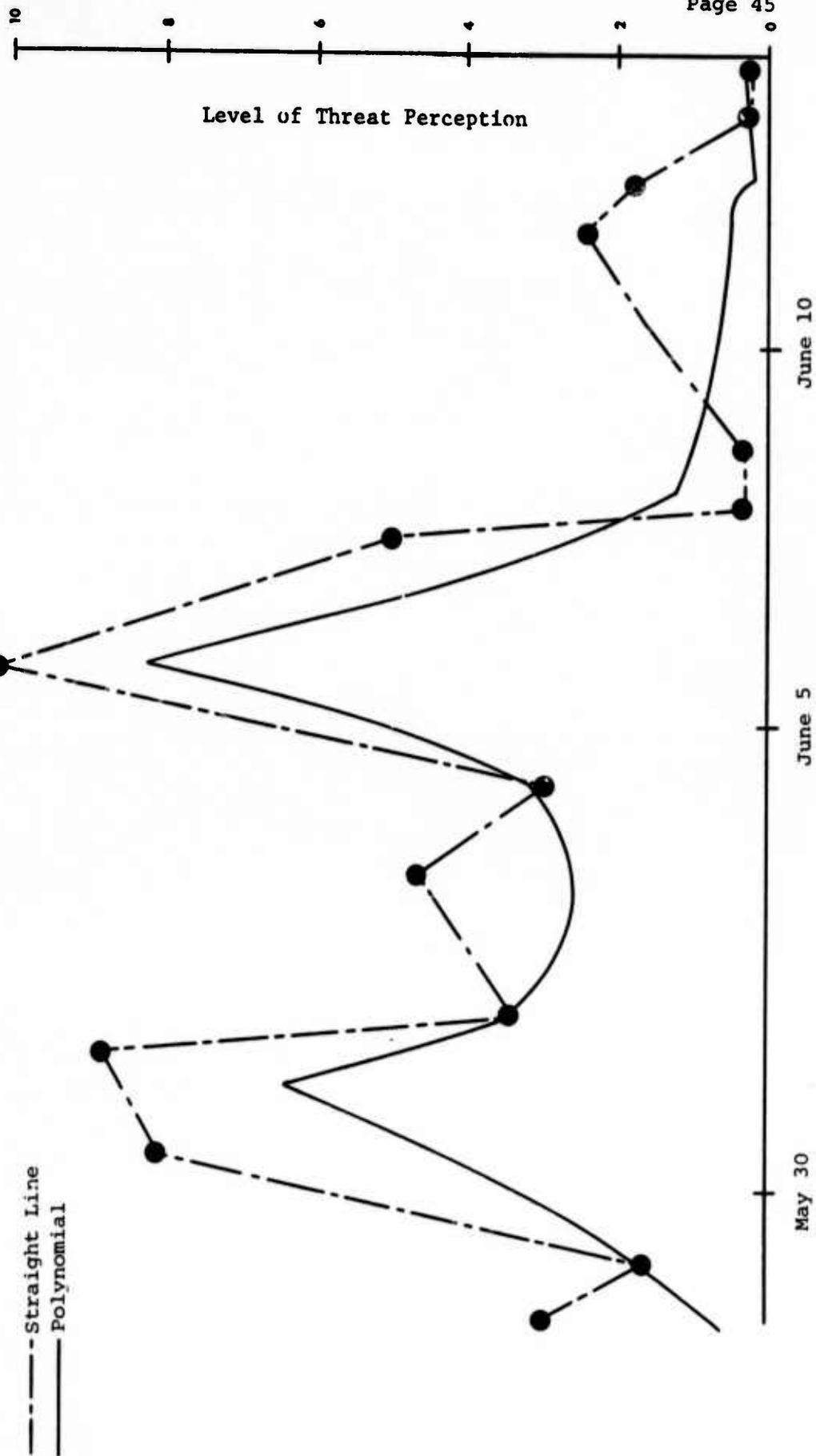


Figure 3
Polynomial vs. Straight Line Fit

Upon receipt of the data, however, we found the 1967 collection was quite incomplete. For example, we found from May 1 to August 1, 1967 only 138 events were recorded with such major events as the May 22 blockade of the Gulf of Aqaba missing. After several cross checks using Menahem Mansoor's (1972) event chronology as a source, we concluded we would have to collect data for the 1967 crisis. To do this, we used Mansoor's (1972) chronology, coding the data with the WEIS codes.

The resulting data set contained records of 349 events for the May 1 to July 1 period. Of particular interest is that rather than the 22 pre-war events contained in the WEIS data we identified 145 events in this period. We used Mansoor as the data base because we were convinced he had performed the most thorough reconstruction available.

Whereas the 1967 WEIS data were quite unsatisfactory, we found records for the 1973 war acceptable. The data we received contained over 300 recorded events for the September 1 to November 1 1973 period, and nearly 500 events for the four month period from September through December. Additionally, an examination of the chronology seemed to indicate most of the events with which we were familiar had been recorded.

As a final check on the 1973 data, we compared it with the Middle East Journal Chronology and with a clipping file taken from the major Israeli newspapers. Where we found references not contained in the WEIS data we added them to the data set. Interestingly, this check did not add more than thirty (30) events.

Because the 22 basic WEIS codes are far too numerous to analyze, we aggregated them into the following categories:

Seven Major Categories	Twenty-Two General Variables
Verbal Cooperation:	Approve, promise, agree, request, propose
Cooperative Action:	Yield, grant, reward
Participation:	Comment, consult
Verbal Conflict-Defensive:	Reject, protest, deny
Verbal Conflict-Offensive:	Accuse, demand, warn, threaten
Conflict Action:	Demonstrate, reduce relationship, expel
Military Conflict:	Seize, force

The WEIS coding rules, of course, only permit identification of occurrence of the discrete WEIS categories such as demonstrations or accusations. There is not an explicit weighting either between or within the categories even though intuitively it appears many of the codeable items should fit some form of scale. For example, the demonstration of offensive weapons judgmentally appears less serious than the deployment of these weapons which, in turn, seem less serious than a mobilization, etc. This problem has not gone unnoticed and is the source of considerable debate among event data analysts. Walter Corson (1970) and Edward Azar (1970) are among those who have developed scaling techniques. Corson's system permits the coder to assign interval measures from 1 to 515 along four dimensions of physical and verbal conflict and cooperation. Azar permits ordinal ratings from 1 to 7 along conflict and cooperation dimensions.

On the surface at least, Corson's scale has considerably more to commend it than does Azar's. Corson's is more plausibly an interval scale than Azar's is; Corson's weightings were developed through a systematic study of State Department FSO's whereas Azar's was based on the understanding

he and his colleagues had of the Middle East; and Corson's scaling recognizes the findings by numerous scholars that verbal and physical conflict are apparently not related.* Even though we believed Corson's scale was more valid than Azar's, we decided to invest the relatively small amount of time necessary to code both scales and to do some simple comparisons between them.

For all event variables, an absence of information was treated as zero interactions rather than missing data. This can be easily justified for the 1967 data because the Mansoor chronology is very complete. For 1973, however, the assumption is less warranted because there has not been a similar historical reconstruction of the war. Nonetheless, we feel confident that even the weaker 1973 data present a reasonable time series data base of the events.

Summarizing the collection of event data, we had records of 549 events from May 1 to July 1, 1967, and 534 events from September 1, 1973 to January 1, 1974. These events were coded using WEIS categories and aggregated into the eight categories mentioned above. Additionally, we scaled the data by verbal and physical conflict and cooperation using Walter Corson's weightings, and by cooperation and conflict using Edward Azar's scale.

VALIDITY OF THE MEASURES

In general, there are three ways the validity of social science indicators can be assessed. One is the "construct validity" of the measure, i.e., the soundness of the reasoning which justifies the transformations.

* Rummel (1966, 1969), McClelland and Hoggard (1969)

Another is "face validity," or the degree variables resemble what is generally known about the phenomena under investigation. The third is "convergent/discriminant validity," the extent multiple indicators of one underlying concept are intercorrelated while being uncorrelated with indicators of unrelated concepts. The construct validity of our indices has been defended above; this section addresses both face validity and convergent/discriminant validity through plots and correlation matrices of the variables.

The concept we are most concerned with measuring is perception of threat by Israeli decision makers. For an index of threat perception to have face validity, it should be high in periods we are reasonably sure the Israelis felt threatened, and vice versa. For example, one would strongly expect the Israelis to have felt quite threatened before the 1967 war, and during the first few days of the 1973 war, but not feel threatened immediately after either war, especially in 1967.

Time series plots of the variables in Figures 4a-f indicate the threat indices from both the manual and machine content analysis have considerable face validity across both wars. The variables in Figures 4a & 4f rise before the wars and drop off before the end of the fighting, and remain at a fairly low level afterward. However, the Inquirer threat index rises and falls somewhat more slowly than the manual content analysis index in 1967. In 1973 both measures rise and fall at about the same rate. The face validity of both measures on the whole looks rather good. One possible exception is a dip in the threat indices from both types of content analysis between June 1 and June 5, 1967. One might interpret this as evidence against the measure's validity, or of a substantively interesting phenomena. These conflicting

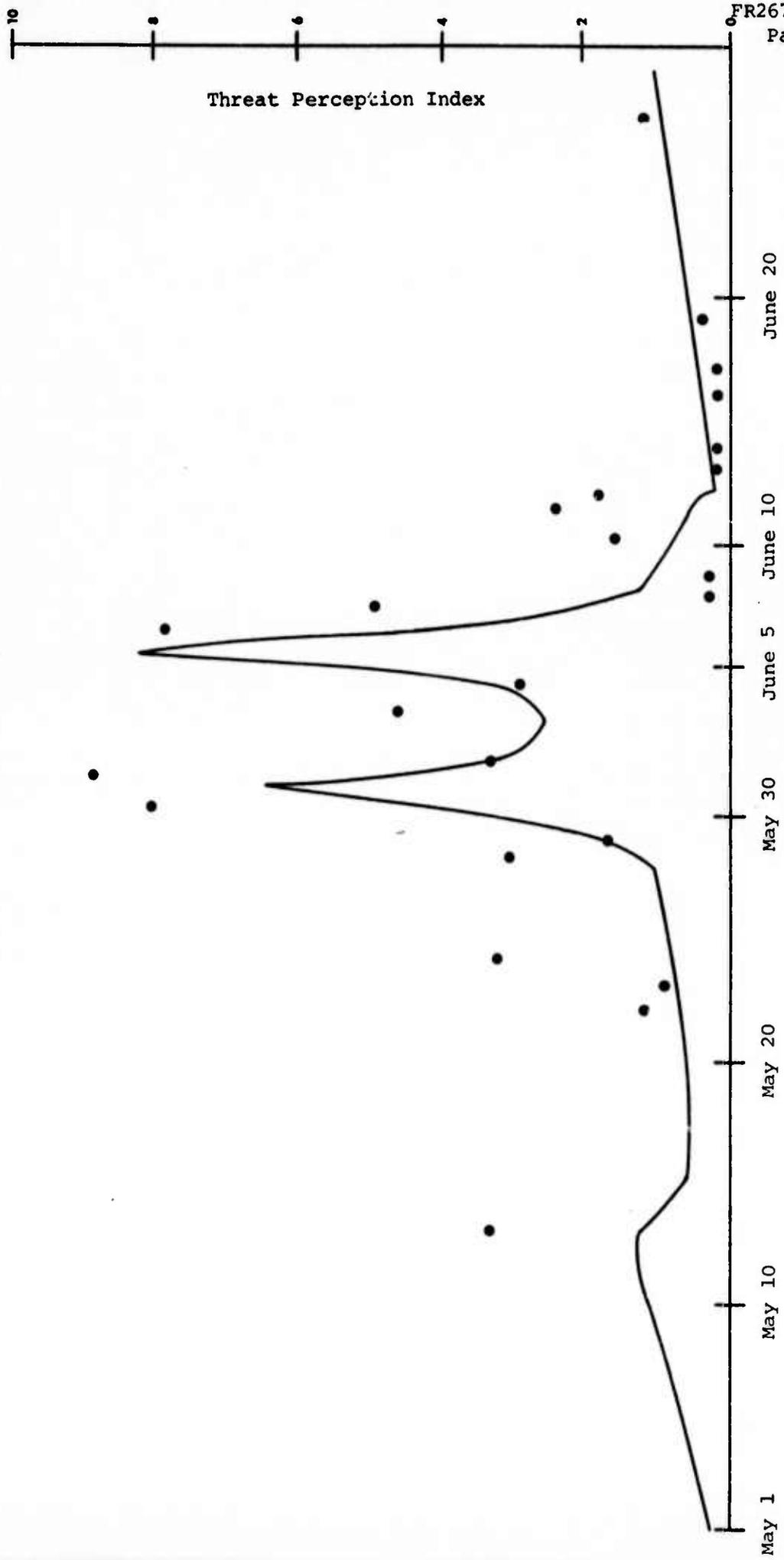


Figure 4a

Manual Threat Perception - 1967 Crisis

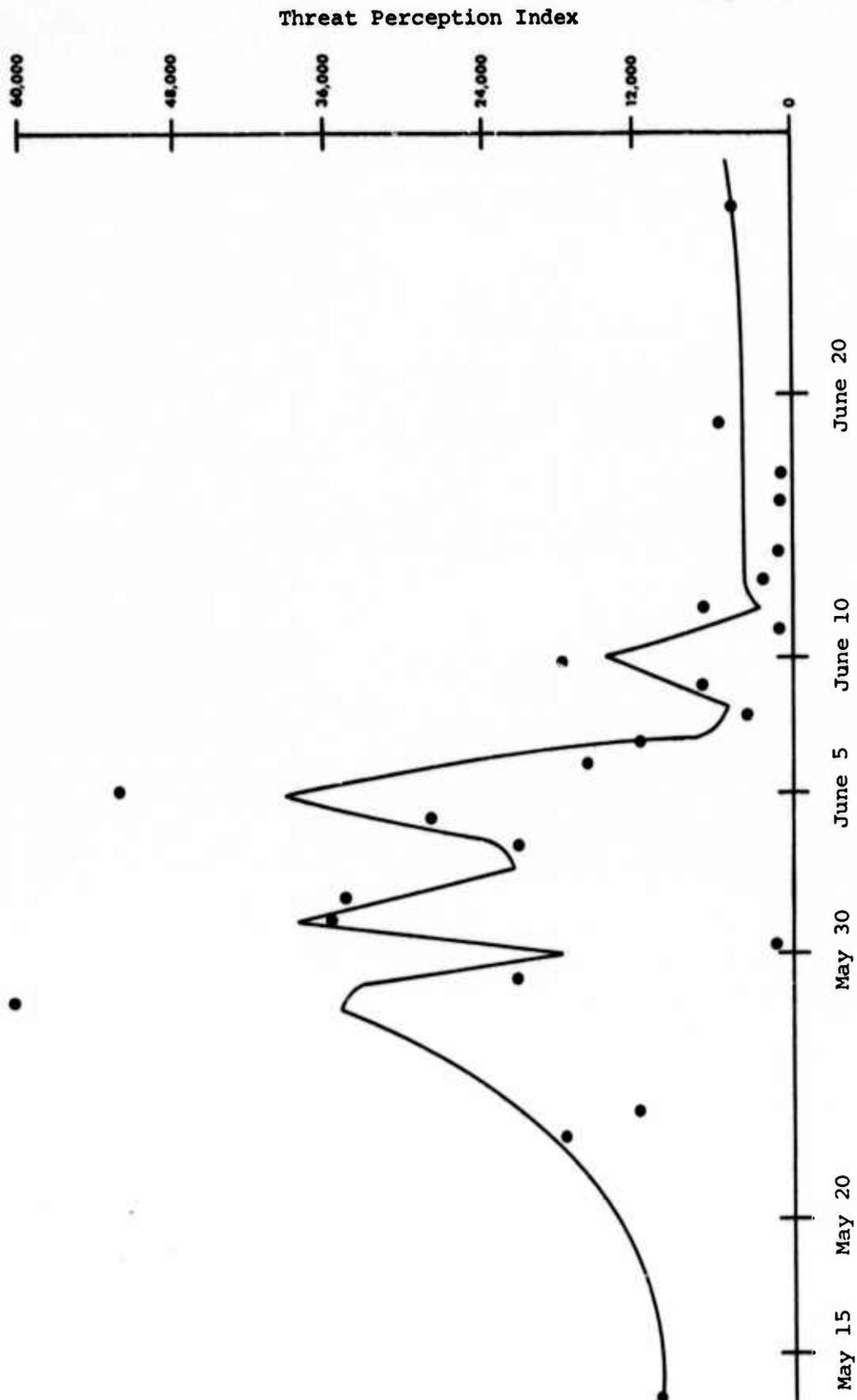


Figure 4b
Inquirer Threat Perception - 1967 Crisis

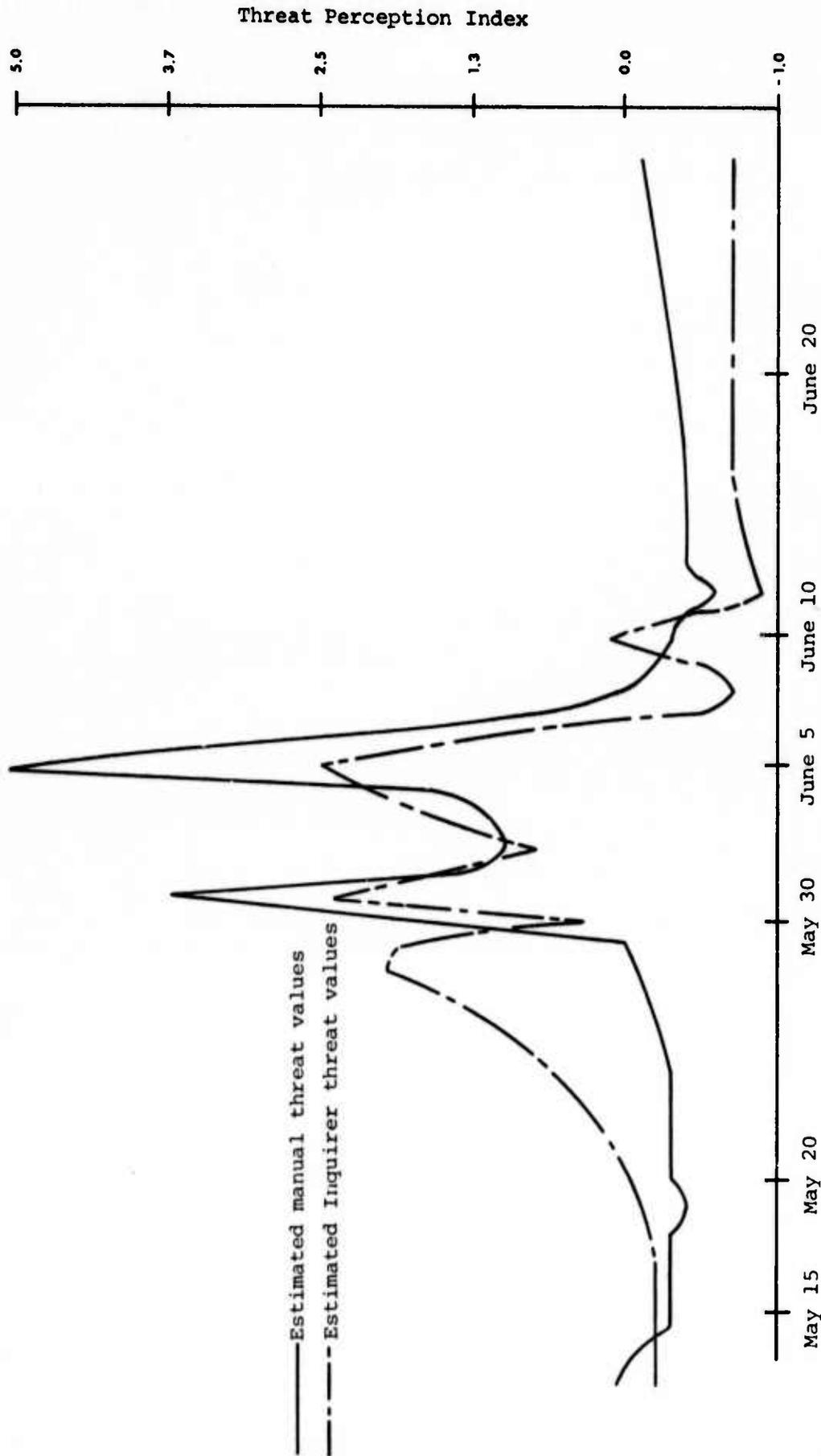


Figure 4c
Manual vs. Inquirer Threat Perception - 1967 Crisis

Figure 4d
Manual Threat Perception - 1973 Crisis

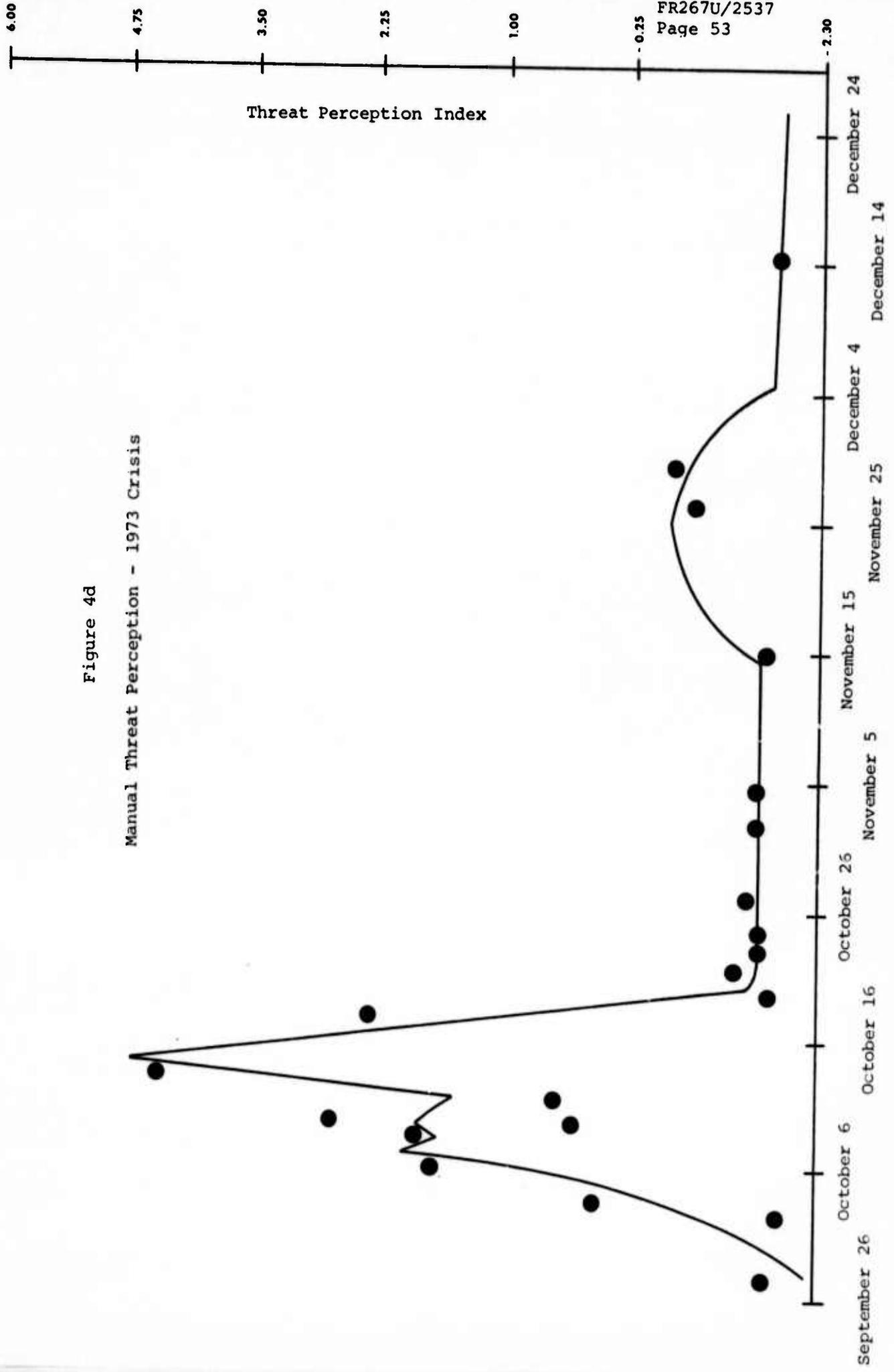
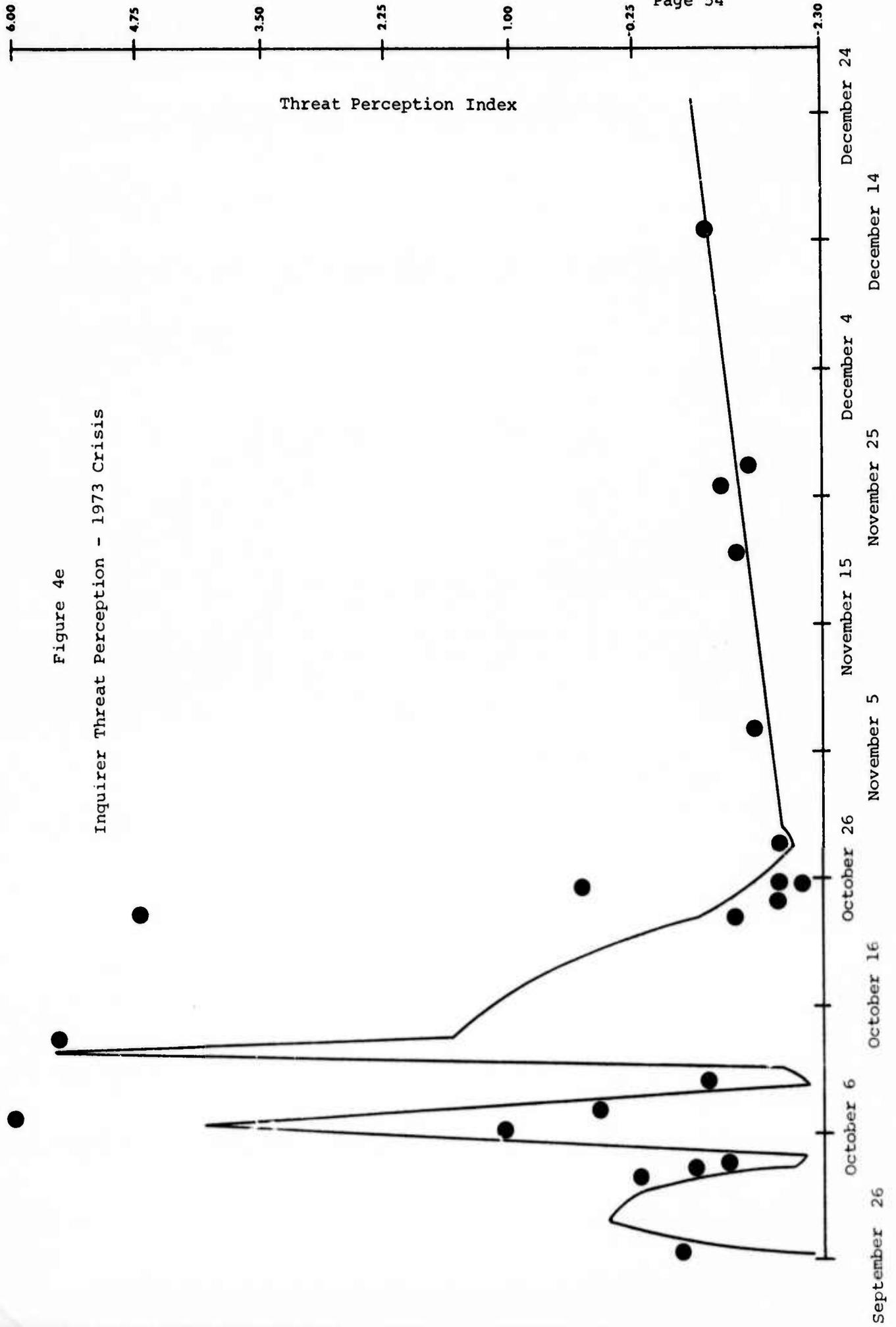


Figure 4e
Inquirer Threat Perception - 1973 Crisis



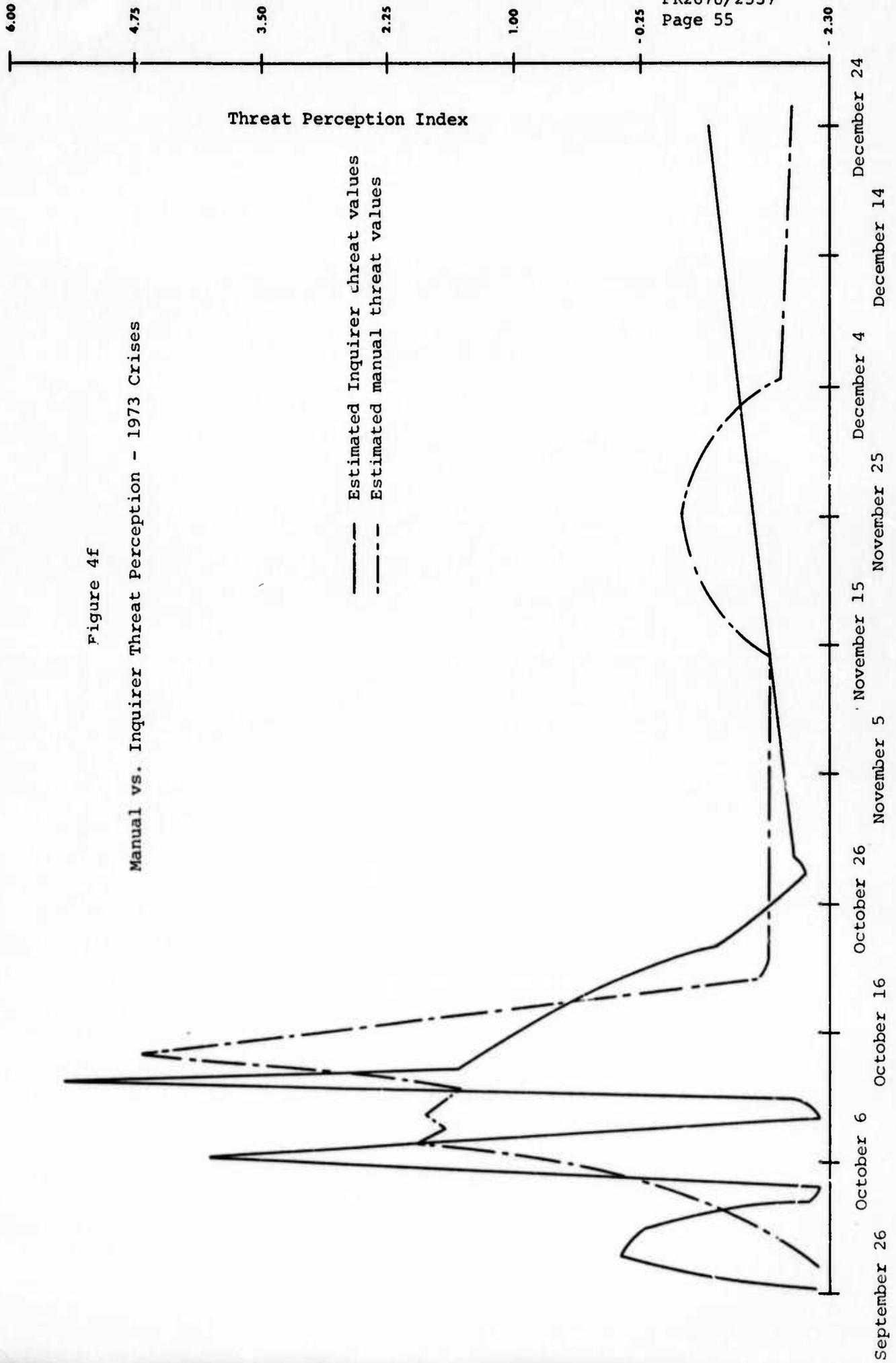


Figure 4f
Manual vs. Inquirer Threat Perception - 1973 Crises

Threat Perception Index

— Estimated Inquirer threat values
- - - Estimated manual threat values

interpretations are discussed below.

Tables 2a and 2b show correlations among the variables used to test the hypotheses* giving further indication of the measure validity. In this regard, observe the correlations among the perceptual indicators as compared with the relationships with other, non-threat, variables.

Except for one pair of variables, the threat indicators are reasonably well intercorrelated, ranging between .43 - .93. The weakest relationships, in the .4 - .5 range, are between the raw manual index and the raw General Inquirer index. Those indicators without missing data are more highly correlated. This is probably a result of the removal of random variance. Both types of threat measures show good correlation with the physical conflict of both the Arabs and Israelis, but are not consistently correlated with other indicators of different concepts. Since we originally expected the concepts of "threat" and "conflict" to be related, we are fairly confident in the convergent and discriminant validity of the threat variables.

Turning to the indicators of the "conflict" concept, we had three basic types of measures. The WEIS coding scheme is based on frequency of events across seven aggregated dimensions as listed above. The COPDAB scheme scales events from 1 to 7 on two dimensions, conflict and cooperation. The Corson scheme scales conflict from 1-515 on 4 dimensions: verbal cooperation, physical cooperation, verbal conflict, and physical conflict.

We originally hypothesized differences in assumptions would have a significant impact on the hypothesis testing. The correlations in Tables 2a and 2b however, indicate this is not so. The Azar conflict, Corson physical conflict

* Presented in Chapter IV.

TABLES 2a and 2b

DESCRIPTIONS OF VARIABLE LABELS

Variable	LABEL
Arab Cooperation (Weis)	A COOP
Arab Offensive Verbal Conflict (Weis)	A VCON O
Arab Military Conflict (Weis)	A MJLCON
Arab Conflict (Azar)	A AZCON
Arab Verbal Conflict (Corson)	A COVCON
Arab Physical Conflict (Corson)	A COPCON
Israeli Cooperation (Weis)	I COOP
Israel Offensive Verbal Conflict (Weis)	I VCON O
Israeli Military Conflict (Weis)	I MILCON
Israeli Conflict (Azar)	I AZCON
Israeli Verbal Conflict (Corson)	I COVCON
Israeli Physical Conflict (Corson)	I COPCON
Israeli Perception of Time Pressure	TIMEPRES
Manual Threat Index	RAWTHRET
Residual Threat Index (with estimates)	THRESID
Manual Threat Index with Estimated Values	THRETEST
Inquirer Threat Index with Estimated Values	GI EST
Inquirer Threat Index	RAW GI

TABLE 2a
BIVARIATE CORRELATIONS - 1967 WAR

VARIABLE						
A COOP	1.0000					
A VCON O	.3334 (60)	1.0000				
A MILCON	.3760 (60)	.3759 (60)	1.0000			
A AZCON	.4031 (60)	.5278 (60)	.9298 (60)	1.0000		
A COVCON	.7000 (60)	.7189 (60)	.5736 (60)	.6843 (60)	1.0000	
A COPCON	.3882 (60)	.3835 (60)	.9834 (60)	.9500 (60)	.5711 (60)	1.0000
I COOP	.7655 (60)	.1956 (60)	.2190 (60)	.2384 (60)	.4546 (60)	.2463 (60)
A VCON O	-.0135 (60)	.4248 (60)	.4022 (60)	.4308 (60)	.4718 (60)	.3890 (60)
I MILCON	.6404 (60)	.3914 (60)	.9237 (60)	.8782 (60)	.6951 (60)	.9313 (60)
I AZCON	.5683 (60)	.4852 (60)	.9471 (60)	.9295 (60)	.7351 (60)	.9578 (60)
I COVCON	.0154 (60)	.4579 (60)	.5038 (60)	.5286 (60)	.5314 (60)	.4888 (60)
I COPCON	.5693 (60)	.4216 (60)	.9344 (60)	.8986 (60)	.6670 (60)	.9483 (60)
TIMEPRES	.4016 (25)	-.1759 (25)	.3140 (25)	.2831 (25)	.1015 (25)	.3257 (25)
RAWTHRET	-.1222 (25)	.2140 (25)	.5603 (25)	.5688 (25)	.1906 (25)	.5451 (25)
THRESID	-.2706 (60)	.0255 (60)	.1870 (60)	.1778 (60)	-.0032 (60)	.1590 (60)
THRETEST	.0646 (60)	.2409 (60)	.6453 (60)	.6189 (60)	.3438 (60)	.6286 (60)
GI EST	-.1426 (48)	.0048 (48)	.3588 (48)	.3898 (48)	.0687 (48)	.3411 (48)
PAW GI	-.3222 (21)	-.0126 (21)	.2168 (21)	.2033 (21)	-.0377 (21)	.1907 (21)
	A COOP	A VCON O	A MILCON	A AZCON	A COVCON	A COPCON

TABLE 2a (Cont'd)

VARIABLE						
I COOP	1.0000					
A VCON O	-.0292 (60)	1.0000				
I MILCON	.4064 (60)	.2529 (60)	1.0000			
I AZCON	.3620 (60)	.3963 (60)	.9743 (60)	1.0000		
I COVCON	.0093 (60)	.9074 (60)	.3331 (60)	.4765 (60)	1.0000	
I COPCON	.3460 (60)	.2567 (60)	.9860 (60)	.9822 (60)	.3311 (60)	1.0000
TIMEPRES	.3757 (25)	-.2002 (25)	.4382 (25)	.3423 (25)	-.2454 (25)	.3962 (25)
RAWTHRET	-.1748 (25)	.2993 (25)	.4034 (25)	.4627 (25)	.4012 (25)	.4381 (25)
THRESID	-.1863 (60)	.3359 (60)	.0059 (60)	.0685 (60)	.4362 (60)	-.0000 (60)
THRETEST	.0207 (60)	.4203 (60)	.5174 (60)	.5688 (60)	.5448 (60)	.5196 (60)
GI EST	-.1675 (48)	.2936 (48)	.2034 (48)	.2300 (48)	.3104 (48)	.1957 (48)
RAW GI	-.3725 (21)	.3595 (21)	.0551 (21)	.0848 (21)	.4036 (21)	.0343 (21)
	I COOP	A VCON O	I MILCON	I AZCON	I COVCON	I COPCON

VARIABLE						
TIMEPRES	1.0000					
RAWTHRET	.1907 (25)	1.0000				
THRESID	-.0579 (25)	.7488 (25)	1.0000			
THRETEST	.1615 (25)	.9045 (25)	.8544 (60)	1.0000		
GI EST	.2584 (25)	.6526 (25)	.6794 (48)	.6874 (48)	1.0000	
RAW GI	.0617 (21)	.4385 (21)	.5935 (21)	.5627 (21)	.8337 (21)	1.0000
	TIMEPRES	RAWTHRET	THRESID	THRETEST	GI EST	RAW GI

TABLE 2b
BIVARIATE CORRELATIONS - 1973 WAR

VARIABLE						
A COOP	1.0000					
A VCON O	.1215 (119)	1.0000				
A MILCCN	.7037 (119)	.5718 (119)	1.0000			
A AZCON	.0339 (119)	.6798 (119)	.9785 (119)	1.0000		
A COVCON	.0045 (119)	.7822 (119)	.7301 (119)	.7617 (119)	1.0000	
A COPCON	.7323 (119)	.5793 (119)	.9767 (119)	.9824 (119)	.6932 (119)	1.0000
I COOP	.7912 (119)	.1215 (119)	.0256 (119)	.0521 (119)	-.0015 (119)	.0423 (119)
A VCON O	.1578 (119)	.4753 (119)	.5803 (119)	.6066 (119)	.4824 (119)	.5805 (119)
I MILCON	.0785 (119)	.5803 (119)	.9326 (119)	.9340 (119)	.6892 (119)	.9390 (119)
I AZCON	.0467 (119)	.6050 (119)	.9501 (119)	.9573 (119)	.7313 (119)	.9615 (119)
I COVCCN	.0022 (119)	.4418 (119)	.5628 (119)	.5724 (119)	.6173 (119)	.5281 (119)
I COPCON	.0345 (119)	.5894 (119)	.9300 (119)	.9427 (119)	.7043 (119)	.9597 (119)
TIMEPRES	.1779 (28)	-.0667 (28)	-.0865 (28)	-.0905 (28)	-.0288 (28)	-.0986 (28)
RAWTHRE ^T	-.1952 (28)	.0510 (28)	.4453 (28)	.3859 (28)	.2475 (28)	.4208 (28)
THRESID	-.1299 (95)	-.0610 (95)	.3500 (95)	.2911 (95)	.2743 (95)	.3269 (95)
THRETEST	-.0734 (95)	.1805 (95)	.6059 (95)	.5848 (95)	.4269 (95)	.6354 (95)
GI EST	-.0913 (92)	.1190 (92)	.4645 (92)	.4539 (92)	.2298 (92)	.5033 (92)
RAW GI	-.1289 (23)	-.1293 (23)	.2353 (23)	.2369 (23)	-.0280 (23)	.3165 (23)
A COOP	A VCON O	A MILCCN	A AZCON	A COVCON	A COPCON	

TABLE 2b (Cont'd)

VARIABLE						
I COOP	1.0000					
A VCON O	.0784 (119)	1.0000				
I MILCON	.1018 (119)	.5767 (119)	1.0000			
I AZCON	.0439 (119)	.6935 (119)	.9631 (119)	1.0000		
I COVCON	-.0172 (119)	.6027 (119)	.4231 (119)	.5547 (119)	1.0000	
I COPCON	.0456 (119)	.5583 (119)	.9716 (119)	.9753 (119)	.4305 (119)	1.0000
TIMEPRES	.3772 (28)	.0539 (28)	.0156 (28)	-.0629 (28)	-.0658 (28)	-.0476 (28)
RAWTHRET	-.1236 (28)	.1932 (28)	.4296 (28)	.4355 (28)	.2109 (28)	.4788 (28)
THRESID	-.1269 (95)	.0550 (95)	.2658 (95)	.3158 (95)	.3279 (95)	.3199 (95)
THRETEST	-.0711 (95)	.2153 (95)	.6144 (95)	.6258 (95)	.2897 (95)	.6794 (95)
GI EST	-.0816 (92)	.2617 (92)	.4673 (92)	.4568 (92)	.1389 (92)	.4985 (92)
RAW GI	-.1289 (23)	.0951 (23)	.1633 (23)	.1655 (23)	.1306 (23)	.1894 (23)
	I COOP	A VCON O	I MILCON	I AZCCN	I COVCON	I COPCON

VARIABLE						
TIMEPRES	1.0000					
RAWTHRET	.1659 (28)	1.0000				
THRESID	-.0462 (28)	.6551 (28)	1.0000			
THRETEST	.0820 (28)	.9254 (28)	.7842 (96)	1.0000		
GI EST	.2254 (28)	.5330 (28)	.2890 (92)	.6123 (92)	1.0000	
RAW GI	.1470 (19)	.4998 (19)	.2333 (23)	.5358 (23)	.8261 (23)	1.0000
	TIMEPRES	RAWTHRET	THRESID	THRETEST	GI EST	RAW GI

and WEIS military conflict measures are very highly correlated (> 0.90) for both actors in both crises.

Given these findings we decided the measures were, for all practical purposes, totally substitutable. We chose Corson's scale for hypothesis testing primarily because of its superior construct validity. We did, however, compare Azar's scale with Corson's for one of the hypotheses and found the results were inconsistent. For reasons which we will discuss later, we concluded Azar's scale introduces some rather unusual problems as a result of its quasi-interval nature.

CHAPTER IV

ANALYSIS

Data were collected with the intention of testing some aspects of the response model. For this study, we employed a modification of the Stanford model used by North, Holsti, Zinnes and others. We posit that the mental process can be described as follows:

Some stimuli in the real world occur which impact on the perceptual sensors of an observer. The immediate response is an abstract impression of the stimuli. The response is then formulated into a clearer action-oriented perception of the stimuli which becomes the immediate stimulus to the resulting behavior.

This model is closer to that presented by Coutu (1949) than to Stanford studies. The major difference is that Coutu argues both intermediate mental processes are treated as perceptions, whereas the Stanford studies describe the second intermediate process as the actual plan for action.

There are serious questions whether any of these steps can be bypassed. The traditional action-reaction model, for example, presumes a consistent and reasonably non-complex linkage between all of these elements. If this assumption is made, the intermediate mental processes of the decision maker can be treated as a linear black box. If this is true, we should find event-to-event predictions are more or less sufficient by themselves, or that little is added by knowledge of the perceptual mechanism.

On the other hand, the work of cognitive map modelers (Steinbruner (1968), Shapiro and Bonham (1973)) argue the mechanism transforming stimuli into

perceptions is extremely complex, and basically dooms the event → event model to failure. Similarly, research into selective perception suggests there are higher level core beliefs which act complexly on stimuli to create revised mental images (Newcombe, Turner and Converse, 1965). These models suggest the mental process cannot simply be bypassed in our attempts to predict behaviors. These alternative models will be tested through the following set of hypotheses.

HYPOTHESES

- #1. Increases in Arab physical conflict with Israel will lead to increases in Israeli physical conflict with the Arabs, and vice versa.
- #2. Increases in Arab verbal conflict with Israel will lead to increases in Israeli physical conflict with the Arabs and vice versa during the escalation periods.

These first two hypotheses assume the most simplistic role for decision makers. Basically, they assume decision makers either consciously or unconsciously operate in a simple action-reaction mode and have only simplistic procedures for processing information or pursuing goals. Thus, it should be possible to treat the decision maker as a linear black box.

During the pre-war period hypothesis 1 should predict escalation. During the post-war it should predict de-escalation. Hypothesis 2 is derived from the traditional escalation arguments which posit verbal conflict is low-level physical conflict and, in escalations, is the forerunner of physical conflict.

- #3. Increases in Arab physical or verbal conflict will lead to increases in Israeli perceptions that the Arabs are negative and active.
- #4. Increases in Arab physical or verbal conflict with Israel will lead to increases in Israeli perceptions of Arabs as threatening.

- #5. Israeli perceptions of Arabs as strong, negative and active will covary with Israeli perceptions of Arabs as threatening.
- #6. Increases in Israeli perceptions of Arabs as threatening will lead to increases in Israeli physical conflict with the Arabs.

Hypotheses 3-6 in toto reflect a conventional S-O-R model of decision makers that argues decision makers interpret information in a relatively consistent and non-complex manner, and then react linearly to their images of the world.

Hypothesis #3 reflects our belief the Inquirer II measures of negativeness and activeness are the first mental images of threat to be influenced by events. We argue that negative behavior by the Arabs would not necessarily alter the Israeli perception of strength. That should be governed by the Israeli's ongoing mental image and those aspects of the events, such as numbers of troops, quality of armaments, and relative success of the activities, none of which are measured by standard event data.

Hypothesis #4 assumes the mental process is only a one step process, and that events are stimuli which evoke a well formulated perceptual response. The latter we hypothesize to be the immediate predecessor to behavior.

Hypothesis #5 provides linkage between the abstract latent measure of threat and the more well formulated action manifest measure of threat. That is, we hypothesize that some stimuli (events) create a rather abstract feeling of threat which becomes reformulated into an explicit recognition of threat. This hypothesis posits the linkage between these portions of the mental process.

Hypothesis #6 presents linkage between the explicit perception of threat and behavior. As we discussed in the published study plan for this project we expected the threat perception → behavior sequence to exhibit characteristics

of the psychological phenomena of approach-avoidance. Because of this, we further hypothesized in a non-surprise crisis there should be a peak followed by a reduction in threat perception just prior to the crisis. We suggest that the peak prior to the crisis is that point at which the participants recognize they may go to war (even though a formal decision may not occur until later). The act of making this decision will temporally reduce feelings of threat until actual war activities are put in motion.

Further, we argued that in surprise crises the conscious attempts to deceive the enemy may make this pattern less pronounced. Nonetheless, this pattern should still be observable in the aggressor in surprise crises.

- #7. Increases in Israeli perceptions of Arabs as threatening plus increases in Arab physical conflict toward Israel will lead to increases in Israeli physical conflict with the Arabs.

This hypothesis simply reflects a recognition of the fact there may be information contained in event measures which are more or less directly passed on in reaction to events and are not measured by the perceptual data. This phenomena is most likely to occur when events are moving quite rapidly.

The final hypothesis requires the following assumption: the crises examined here can be divided into three stages: escalation, war and post-war.

- #8. Parameters which define the relationships between variables will vary by crisis stage.

In this hypothesis, we are arguing that we cannot expect a general model to apply to all stages of a crisis.

The escalation phase is characterized by high level political decisions made under tension. These decisions, whether or not to go to war, are roughly

within the areas normally addressed by political scientists. The actual fighting, on the other hand, while not devoid of policy implications, tends to adopt a life of its own centering on military strategy. While the scope of the war probably rests, at least partially, in the hands of political leaders, its actual conduct would seem to reflect different motivational factors than were dominant in the escalation. For example, troop movements often reflect tactical factors during a war whereas they would reflect political factors during escalation.

Additionally, event data present some severe difficulties during a war. More traditional operational reporting statistics are far better indicators of the ongoing progress of a war than frequencies of physical and verbal conflict events. When we consider both the data and theoretical problems, we feel the likelihood is quite slim that the same parameters should describe these two phases of war.

The post-war periods are characterized by political decisions, but it is quite difficult to consider this clearly within the crisis period. For example, we know the levels of time pressure, and threat, are not very high during this period. On the assumption that crisis periods are different from non-crises, we believe there is a reasonable expectation this hypothesis will be supported.

HYPOTHESIS TESTING

Hypothesis #1

The first hypothesis postulates an actor will reciprocate conflict the other actor directs at it. Note in Tables 2a and 2b correlations between Arab and Israeli conflict at time t_0 are very high for the crises. These correlations do not necessarily support the hypothesis for a number of reasons. First, it

is impossible to sort out the direction of causality since one side may be reacting to the other, or vice versa, or the causation may be reciprocal. Second, the coding rules make it virtually certain that a force event will be "double coded," that is, a border clash would be coded once with Israel as the actor and an Arab as the target, and once with an Arab as the actor and Israel as the target. This means that the high correlation at time t_0 only supports the model in the sense that if one side shoots, the other side will shoot back.

A more conventional way to test the hypothesis is to examine the lagged relationship between one side's conflict and that of the other. The results of analyses with a one-day lag are presented in Table 3a and 3b. The hypothesis is supported in only one non-war period: the prediction of Arab conflict at time t_0 from Israeli conflict at time $t-1$ for the 1967 pre-war.*

It should be noted that for an escalation period, regression coefficients should be greater than 1.0, indicating conflict $A \rightarrow B$ at time t_0 will be reciprocated by more severe conflict at time $t + 1$. The pre-war 1967 results meet this criterion only for Israeli \rightarrow Arab conflict.

As an additional test we ran one day lags for the pre-war period, excluding the first day of war. Eliminating the outlier of the first day of war decreases the correlation to .04.

During war, the model does well only for Arab \rightarrow Israeli events in 1967. In fact, there are significant negative correlations for both sides during 1973. We did not attempt a substantive interpretation of this observation.

* "Pre-war" includes the first day of war for the dependent variable when the independent variables are lagged, unless otherwise noted.

TABLE 3a #

ARAB PHYSICAL CONFLICT (t-1) →
ISRAELI PHYSICAL CONFLICT

<u>Period</u>	<u>b</u>	<u>SE of b</u>	<u>Partial r</u>	<u>R²</u>
Pre-War 1967	.82	.77	.18	.03
War 1967	1.25**	.38	.85	.73
Post-War 1967	-.13	.39	-.08	.007
Pre-War 1973	-.47	1.14	-.05	.004
War 1973	-.42*	.22	-.43	.18
Post-War 1973	-.009	.10	-.01	.000

TABLE 3b

ISRAELI PHYSICAL CONFLICT (t-1) →
ARAB PHYSICAL CONFLICT

<u>Period</u>	<u>b</u>	<u>SE of b</u>	<u>Partial r</u>	<u>R²</u>
Pre-War 1967	2.69**	1.26	.35	.12
War 1967	.34	.37	.42	.18
Post-War 1967	-.06	.16	-.09	.009
Pre-War 1973	-.80	2.73	-.05	.002
War 1973	-.33*	.19	-.40	.16
Post-War 1973	-.03	.06	-.06	.004

For this table and subsequent tables, we use the following terminology:

- b = regression coefficient
- SE of b = standard error of the regression coefficient
- r = bivariate or partial correlation
- R² = squared multiple correlation = explained variance
- DW = Durbin-Watson Statistic
- * < .10
- ** < .05
- *** < .01

The most likely cause is the data problem during ongoing wars as described above.

Since the results for 1967 pre-war Arab → Israeli events were surprisingly low, the hypothesis was further explored for that period. A multiple regression was run with Arab conflict from $t - 1$ to $t - 5$ as separate predictors. The multiple R was only 0.3. As an additional test, both variables were aggregated to the 2 and 4 day levels; neither produced significant coefficients.

As a final test of the hypothesis, we used Azar's measures of conflict to determine if a relationship could be supported. Generally, we still found no confirmation for the hypothesis, except for a three day time lag for Arab → Israel in the 1967 pre-war period excluding the first day of war. Including the first day reduces the relationship.

Additional examination showed the model was predicting mobilizations but not outbreaks of war. This is discussed in more detail in the concluding chapter of this report.

We conclude there is very little support for this hypothesis. Most results were not significant and those which were failed to stand up under subsequent testing.

Hypothesis #2

Hypothesis #2 argues that one sides' verbal conflict will lead to physical conflict by the other side. This hypothesis deals with escalation spirals and presumes verbal conflict is the low level forerunner of physical conflict. As such, a test of the hypothesis makes sense only during the pre-war periods. The tests of this hypothesis with verbal conflict at $t-1$ (see Tables 4a and 4b).

showed no significant relationships at all. Subsequent examinations of longer time lags similarly failed to produce significant results. We concluded, therefore, Hypothesis #2 is not supported.

TABLE 4a[#]
 HYPOTHESIS #2
 ARAB VERBAL CONFLICT (t-1) → ISRAELI PHYSICAL CONFLICT

<u>Period</u>	<u>b</u>	<u>SE</u>	<u>r</u>	<u>R²</u>
Pre-War 1967	-2.00	2.84	-.12	.01
Pre-War 1973	-1.30	6.27	-.03	.001

TABLE 4b
 HYPOTHESIS #2
 ISRAELI VERBAL CONFLICT (t-1) → ARAB PHYSICAL CONFLICT

<u>Period</u>	<u>b</u>	<u>SE</u>	<u>r</u>	<u>R²</u>
Pre-War 1967	-2.28	2.81	-.14	.02
Pre-War 1973	-3.27	10.58	-.05	.002

Hypothesis #3

Failure of the event/interaction model as a predictor of events may be partially explained by the results of our investigation of Hypothesis #3 as illustrated in Table 5a, and b. None of the correlations were significant at the .05 level, and only three are significant at the .10 level. With over thirty observed correlations, this is no better than we would expect by chance alone. A closer examination of the best of these results (post-war 1967) indicates there was very little variance in active/negative perception during that period except for a slight increase in late June which co-occurred with some minor cease-fire violations.

See footnote on page 69 for symbol code.

TABLE 5a#
HYPOTHESIS #3
LAGGED ARAB VERBAL CONFLICT → PERCEPTIONS OF ACTIVE/NEGATIVE

<u>Period</u>	<u>Lagged Correlations</u>			(N) ¹
	<u>-1</u>	<u>-2</u>	<u>-3</u>	
Pre-War 1967	.08	.03	-.02	(15)
War ² 1967	.14	-.33	-.34	(6)
Post-War 1967	.40	.38	.18	(17)
Pre-War 1973	-.02	.003	.03	(12)
War 1973	-.16	-.26	-.02	(23)
Post-War 1973	.03	.07	.04	(61)

TABLE 5b
HYPOTHESIS #3
LAGGED ARAB PHYSICAL CONFLICT → PERCEPTIONS OF ACTIVE/NEGATIVE

<u>Period</u>	<u>Lagged Correlations</u>			(N)
	<u>-1</u>	<u>-2</u>	<u>-3</u>	
Pre-War 1967	.02	-.20	.13	(15)
War 1967	-.10	-.45	-.45	(6)
Post War 1967	.34	.30	.15	(16)
Pre-War 1973	.14	.10	.14	(12)
War 1973	.26	.11	.11	(23)
Post-War 1973	.25	.14	.04	(61)

We conclude Hypothesis #3 is not confirmed. If, indeed, the S-O-R model discussed above does reflect the process of transforming one set of events to another, the negative results of Hypothesis #3 partially explain the failure

¹ None of these correlations are significant at the .05 level.

² Non-estimated values used in the war period because we had complete data

See footnote on page 69 for symbol code

of Hypotheses #1 and #2. Since we cannot explain the mechanism by which events are transformed into even the most basic perceptions, we clearly would expect to have difficulty treating the event → perception link as a linear black box.

Hypothesis #4

Hypothesis #4 argues that Arab verbal and physical conflict will lead directly to explicit threat perception rather than operating through the elements of abstract threat perceptions by Hypothesis #3.

Tables 6a, and b give the impression our ability to explain the manifest measures of threat perception is slightly better than our ability to explain the active/negative measure.

TABLE 6a #
HYPOTHESIS #4
CORRELATION BETWEEN ARAB VERBAL CONFLICT
AND THREAT PERCEPTION

<u>Period</u>	<u>Lag</u>			(N)
	<u>-1</u>	<u>-2</u>	<u>-3</u>	
Pre-War 1967	-.11	-.22	-.21	(15)
War 1967	.09	.01	.24	(6)
Post-War 1967	-.52**	-.48**	-.55**	(17)
Pre-War 1973	-.13	-.07	-.06	(12)
War 1973	-.01	.26	.14	(23)
Post-War	-.27**	-.31**	-.29**	(61)

See footnote on page 69 for symbol code

TABLE 6b #
HYPOTHESIS #4
CORRELATION BETWEEN ARAB PHYSICAL CONFLICT
AND THREAT PERCEPTION

Period	Lag			(N)
	-1	-2	-3	
Pre-War 1967	.08	.08	.13	(15)
War 1967	-.38	-.82**	-.55	(6)
Post-War 1967	.24	.29	-.37*	(17)
Pre-War 1973	.23	.78***	.79***	(12)
War 1973	.01	.44**	.06	(23)
Post-War 1973	-.55***	-.38***	-.37***	(61)

The tables indicate highly significant positive correlations between threat perception at time t_0 and Arab physical conflict at $t - 2$ and $t - 3$ in the 1973 pre-war analysis. Additionally, we observed moderately significant correlations during the 1973 war period for conflict lagged at $t - 2$.

Further examination, however, showed these findings to be rather fragile. The conflict at $t - 3$ is apparently driven by the extreme threat perception on the first day of the war. This is supported by the observation that the correlation diminishes to insignificance when we remove the first day of war. Removing the auto-correlation further depresses correlation to .11. The apparent relationship between threat perception and Arab conflict at $t - 2$ is almost entirely a function of auto-correlation. After controlling for the auto-correlation, the .78 correlation dropped to .13. The only other significant correlation (.44 during the 1973 war) is a function of an outlier and fell from significance using Kendall's Tau.

See footnote on page 69 for symbol code

The tests also highlight a difficulty with our measure of threat during periods of major diplomatic activity. Tables 6a, and b show consistently strong negative correlations between Arab conflict directed at the Israelis and the latter's perceptions of threat. Since these findings are counter-intuitive, we closely examined our raw data. We discovered in the post-war periods, Israelis reacted to ceasefire violations with pleas for diplomatic solutions. Since increases in perceived diplomatic levels of effort tend to decrease the threat measure, results of the post-war analyses indicate the Israelis' preferences for diplomatic responses.

While it is interesting to note this can be well predicted, it certainly is not closely tied to threat perception. This difficulty calls into question the wisdom of our earlier decision to include diplomatic activity negatively in the index construction. Because including it introduces ambiguity when perceptions of diplomatic effort are high, it might be desirable to eliminate or reduce its effect on the threat index. However, time did not permit this analysis. Therefore, we would not find any support for Hypothesis #4 given our analyses.

Hypothesis #5

Hypothesis #5 argues that the more abstract measure of threat from the Inquirer II should covary with the measure of threat from manual content analysis. Within the general model we employed, the Inquirer II threat is the preliminary abstract perception of threat, while manual threat perception is explicit and action oriented. The latter can be considered the immediate stimulus to further events. Because of this we would expect the Inquirer II threat measure to precede the manual measure. However, because these are internal thought processes, it is not certain the time lag can be observed.

The conclusions shown in Table 7 are inconsistent, but generally supportive. For the 1967 pre-war, we found a significant (.01) correlation between the manual threat at t_0 and the Inquirer II threat at $t - 3$. During the war period, the lag decreased to one day. Because observed correlations for the other hypotheses proved quite fragile, we decided to subject the findings for Hypothesis #5 to severe testing.

Excluding the first day of war from the escalation period did not affect the correlation. Furthermore, rank order statistics to control for unusual distributions still gave significant correlations. The Durbin-Watson statistic was 1.97 indicating the results were not a function of auto-correlation. Thus, these relationships held up despite repeated attempts to destroy them.*

We are, however, faced with the minor problem of why the time relationships are different for the pre-war and war periods. While it is pure speculation at this point, an explanation might be that the Israelis began acting more decisively after the war began. We would not want to make too much of the differences in the time lag. The 1967 war obviously covered only six days and we, therefore, cannot place great confidence in that lag.

We did not find significant relationships in the post-war period. However, this is not surprising given the problems with manual content analysis in the same period.

In comparison with the positive findings of the 1967 war, we found no significant correlations for any phases of the 1973 crisis. As the small number of cases make the positive results in 1967 questionable, the small

* We must be aware that the very small numbers of cases makes the results more tentative than they would otherwise be.

TABLE 7 #
CORRELATIONS BETWEEN LAGGED INQUIRER THREAT INDEX
AND MANUAL THREAT INDEX +

Period	Lag	N	r	rho	tau
Pre-War 1967	0*	10	.14		
	-1	10	-.14		
	-2	9	.13		
	-3	8	.83***	.76	.57
War 1967	0	6	.51	.43	.43
	-1	6	.92***	.71	.52
	-2	6	.63		
	-3	6	.19		
Post-War 1967	0	4	-.20		
	-1	5	.20		
	-2	6	-.30		
	-3	6	-.31		
Entire 1967	0	21	.56***		
	-1	21	.35		
	-2	21	.42*		
	-3	20	.67***		
Pre-War 1973	0	5	too few		
	-1	4	cases to		
	-2	4	analyze		
	-3	3			
War 1973	0	11	.45		
	-1	11	.42		
	-2	10	.21		
	-3	10	-.28		
Post-War 1973	0	7	.48		
	-1	8	.49		
	-2	9	.43		
	-3	10	.39		
Entire 1973	0	23	.54***		
	-1	23	.54***		
	-2	23	.41**		
	-3	23	.23		

+ The General Inquirer measures did not include missing data estimates. The manual threat measure did. Tests could not be run with missing data estimated for both because time is a common element to all of our missing data estimates.

See footnote on page 69 for symbol code

TABLE 7 (Cont'd)

LAGGED CORRELATIONS BETWEEN MANUAL THREAT
AND GENERAL INQUIRER THREAT

Period	Lag	N	r
Entire 1967	-1	21	.30
	-2	21	.05
	-3	21	-.07
Entire 1973	-1	23	.51***
	-2	22	.40*
	-3	22	.43***

number of cases in 1973 caused us to question the absence of significant correlations. As a check, we made the assumption mental processes are less likely than actual behaviors to be influenced by phases of the war. Given that assumption, we examined the entire 1967 and 1973 wars. The results, shown in Table 7, indicate that, in both crises, the two threat measures are positively correlated across all time lags. Six of the eight correlations are significant.

Because the correlations were significant across most time lags, we reversed the test. Here we attempted to explain the variance in Inquirer II threat at time t_0 from the manual analysis at $t - 1$, $- 2$, and $- 3$. For 1967, we found the relationships were insignificant. In 1973 however, the correlations were all positive and significant.

As a conclusion, we can feel certain the null hypothesis is rejected. There clearly is some relationship between the two measures of threat. We cannot, however, state with any confidence that Inquirer II measures an abstract predecessor of the explicit threat from the manual content. Whether they are different measures of the same phenomena, or different phenomenon linked to each other in a feedback manner, is not clear. Given the data base available to us, these questions will have to await further research.

Hypotheses #6 and #7

These two hypotheses argue that Israelis' explicit perceptions of threat either alone or when coupled with hostile Arab events will lead to increased Israeli conflict directed at the Arabs. Conceptually it is based on the belief that explicit threat perception is the final mental step prior to action.

However, hypothesis #7 recognizes there is a possibility that including Arab conflict may add to the explained variance.

From Tables 8a - c we see that for the pre-war periods, these hypotheses are generally supported.* The strongest findings apply to the pre-war 1967 period with a lag of 5 days.** An examination of Figure 4a shows that, as anticipated, there is a peak in threat perception in late May. This peak is primarily responsible for the relationship. As with all potentially important high correlations, we made a serious attempt to determine if it was a result of a statistical artifact (see Tables 8a, b & c). Because we are trying to anticipate the outbreak of war, we included the first day of war in our original analyses. At the bivariate level, we observed our highest correlation of .91 with a five day lag. These findings were supported for the direct measure of threat (threates) and the residual threat measure ((thresid) see page 39).

Even though the substantively most reasonable model should include the first day of the war, the conflict measures are so high on that day, the results could be driven by it as an outlier. To test for this possibility, we tried the analysis excluding the first day of war. The results, although decreased, were still significant for both manual measures of threat. As an additional test, we examined rank correlations to reduce the effects of outliers. Finally, we tested conflict against the raw threat perceptions without the missing data estimates. All of the above results were significant.

* We performed the analyses on all crises periods and results are presented at the end of this section in Tables 8d - 8j. However, because the event data are questionable during the war periods, and threat measures have problems in the post-war, we concentrated our attention on the pre-war periods. We believe that substantively this is the most useful application of the model.

** Although not mentioned in the findings, all lags from -1 to -7 were tested and -5 was the peak.

TABLE 8a #
REGRESSIONS PREDICTING ISRAELI CONFLICT AT t_0
Pre-War 1967 (Including First Day)

Variable	Lag	b	SE	r	Multiple R^2	D.W.
Arab Physical Conflict	-1	.19	.38	.09		
Threatest	-5	492***	44	.9		1.9
Regression Statistics					.82***	
Arab Physical Conflict	-1	.16	.39	.08		
Thresid	-5	480***	45	.90		2.1
Regression Statistics					.82***	
Arab Physical Conflict	-1	1	.74	.23		
Thresid	-1	146*	.79	.31		
Regression Statistics					.13	

Pre-War 1967 (Excluding First Day)

Arab Physical Conflict	-1	.02	.10	.05		
Thresid	-5	60**	25	.42		
Regression Statistics					.18**	2.0
Arab Physical Conflict	-1	.03	.1	.05		
Threatest	-5	61**	25	.42		
Regression Statistics					.18**	2.0

Pre-War 1973 (Including First Day)

Arab Physical Conflict	-1	-18*	7	-.78		
Threatest	-5	2243	750	.83		
Regression Statistics					.73*	2.6
Arab Physical Conflict	-1	-17*	7	-.77		
Thresid	-5	2189**	752	.82		
Regression Statistics					.71*	2.7
Arab Physical Conflict	-1	-11	5	-.59		
Thresid	-1	842***	259	.77		
Regression Statistics					.62**	1.6

Pre-War 1973 (Excluding First Day)

Arab Physical Conflict	-1	.44	.29	.51		
Thresid	-1	-15	.18	-.30		
Regression Statistics					.26	
Arab Physical Conflict	-1	.46	.29	.51		
Threatest	-1	-16	18	-.32		
Regression Statistics					.36	

See footnote on page 69 for symbol code

TABLE 8b#

REGRESSIONS OF ISRAELI PHYSICAL CONFLICT ON THREAT PERCEPTION

Lag	Pre-war 1967			Pre-war 1973			First Day Included?		
	b	SE	r	R ²	b	SE		r	R ²
Threatest -1	151***	77	.32	.10	736***	274	.67	.44	Yes
Thresid -1	146***	78	.31	.10	742***	278	.64	.41	Yes
Threatest -5	495***	43	.91	.82	1348	939	.54	.29	Yes
Thresid -5	493***	43	.91	.82	1340	929	.54	.29	Yes
Thresid -1	8	11	.13	.02	2	15	.05	.00	No
Thresid -5	60***	25	.42	.16	32	37	.40	.16	No
Threatest -1	8	11	.13	.02	2	15	.05	.00	No
Threatest -5	61**	24	.42	.18	31.5	37	.39	.15	No

TABLE 8c

RANK ORDER CORRELATIONS OF ISRAELI PHYSICAL CONFLICT WITH THREAT PERCEPTION (FIRST DAY OF WAR INCLUDED)

	Pre-war 1967		Pre-war 1973	
	Rho	Tau	Rho	Tau
Threatest -1	.31*	.23	Threatest -1	.31
Thresid -1	.26	.20	Thresid -1	.31
Threatest -5	.46***	.37**	Threatest -5	.67**
Thresid -5	.46***	.36**	Thresid -5	.67**
Threat (raw data)	.88**	.77**		

See footnote on page 69 for symbol code.

TABLE 8d #

MANUAL THREAT INDEX (-1) → ISRAELI PHYSICAL CONFLICT

<u>Period</u>	<u>b</u>	<u>SE</u>	<u>r</u>	<u>R²</u>
Pre-War 1967	151.27*	77.75	.32	.10
War 1967	844.23***	100.29	.97	.95
Post-War 1967	111.25	76.51	.34	.12
Pre-War 1973	736.3	274.7	.67	.44
War 1973	141.44**	77.0	.41	.17
Post-War 1973	-60.2***	20.0	-.35	.13

TABLE 8e

RESIDUAL THREAT INDEX (-1) → ISRAELI CONFLICT

<u>Period</u>	<u>b</u>	<u>SE</u>	<u>r</u>	<u>R²</u>
Pre-War 1967	146.14*	78.4	.31	.10
War 1967	785.9***	309.8	.79	.62
Post-War 1967	138.2*	80.1	.40	.16
Pre-War 1973	724.2**	287.8	.64	.41
War 1973	176.23**	75.7	.49	.24
Post-War 1973	-39.7	17.5	-.28	.08

See footnote on page 69 for symbol code

TABLE 8f #

MANUAL THREAT (-5) → ISRAELI PHYSICAL CONFLICT

<u>Period</u>	<u>b</u>	<u>SE</u>	<u>r</u>	<u>R²</u>
Pre-War 1967	495.3***	43.0	.91	.82
War 1967	-751.4	460.4	-.63	.40
Post-War 1967	- 5.9	36.4	-.04	.001
Pre-War 1973	1348.0	939.2	.54	.29
War 1973	65.25	97.4	.16	.03
Post-War 1973	-42.4**	20.9	-.24	.06

TABLE 8g

RESIDUAL THREAT (-5) → ISRAELI PHYSICAL CONFLICT

<u>Period</u>	<u>b</u>	<u>SE</u>	<u>r</u>	<u>R²</u>
Pre-War 1967	493.2***	42.9	.91	.82
War 1967	-62.9	706.3	-.04	.002
Post-War 1967	47.5	42.8	.27	.07
Pre-War 1973	1339.8	928.5	.54	.29
War 1973	96.7	100.9	.22	.05
Post-War 1973	-34.7	14.6	-.28	.08

See footnote on page 69 for symbol code

TABLE 8h[#]

SMOOTHED MANUAL INDEX (-5) + ARAB PHYSICAL CONFLICT (-1) →
ISRAELI PHYSICAL CONFLICT

<u>Period</u>	<u>Variable</u>	<u>b</u>	<u>SE</u>	<u>r</u>	<u>Multiple R²</u>
Pre-War 1967					.82
	Arab Conflict -1	.19	.38	.09	
	Threatest -5	491.8***	44.1	.90	
War 1967					.76
	Arab Conflict -1	1.07*	.5	.78	
	Threatest -5	-250.9	410.0	-.33	
Post-War 1967					.01
	Arab Conflict -1	.13	.40	-.08	
	Threatest -1	-6.46	37.5	-.04	
Pre-War 1973					.72
	Arab Conflict -1	-17.7&	7.2	-.78	
	Threatest -1	2243.0**	750.4	.83	
War 1973					.24
	Arab Conflict -1	-.48*	.22	-.47	
	Threatest -5	105.9	90.3	.28	
Post War 1973					.07
	Arab Conflict -1	.06	.10	-.07	
	Threatest -5	-45.5**	21.6	-.25	

See footnote on page 69 for symbol code

TABLE 8i[#]RESIDUAL THREAT INDEX (-5) + ARAB PHYSICAL CONFLICT (-1) +
ISRAELI PHYSICAL CONFLICT

<u>Period</u>	<u>Variable</u>	<u>b</u>	<u>SE</u>	<u>r</u>	<u>R²</u>
Pre-War 1967					.81
	Arab Conflict -1	.16	.39	.08	
	Thresid -5	480.0***	44.7	.90	
War 1967					.73
	Arab Conflict -1	1.25**	.38	.86	
	Thresid -5	-85.3	420.9	-.12	
Post-War 1967					.09
	Arab Conflict -1	-.21	.31	-.13	
	Thresid -5	51.4	44.3	.29	
Pre-War 1973					.71
	Arab Conflict -1	-17.4	7.2	-.77	
	Thresid -5	2189.4	751.5	.82	
War 1973					.30
	Arab Conflict	-.51**	.21	-.51	
	Thresid -5	150.0	92.3	.31	
Post War 1973					.09
	Arab Conflict	-.08	.10	.10	
	Thresid -5	-38.6***	15.4	-.30	

See footnote on page 69 for symbol code

TABLE 8; #

RESIDUAL THREAT INDEX (-1) + ARAB PHYSICAL CONFLICT (-1) +
ISRAELI PHYSICAL CONFLICT

<u>Period</u>	<u>Variable</u>	<u>b</u>	<u>SE</u>	<u>r</u>	<u>R²</u>
Pre-War 1967					.13
	Thresid -1	145.8*	79.1	.31	
	Arab Conflict -1	1.00	.74	.23	
War 1967					.80
	Thresid -1	1264.4	1229	.51	
	Arab Conflict -1	1.01	.44	.79	
Post-War 1967					.17
	Thresid -1	142.6	84.8	.41	
	Arab Conflict -1	-.19	.37	-.14	
Pre-War 1973					.62
	Thresid -1	892.4***	259.4	.77	
	Arab Conflict -1	-10.8*	5.2	-.59	
War 1973					.25
	Thresid -1	1993**	95.2	.46	
	Arab Conflict -1	-.52*	.28	-.42	
Post-War 1973					.10
	Thresid -1	-47.4***	18.4	-.31	
	Arab Conflict -1	.002	.12	.002	

See footnote on page 69 for symbol code

Hypothesis #7 gives results virtually identical to Hypothesis #6 for the 1967 war. As a further check of model validity, the Durbin-Watson (DW) statistic was run on the residuals from the regression equations with and without the first day of war included; the DW statistic indicates no serial correlation in the residuals.

Finally, we extrapolated predictions of the regression analysis without the first day of war included. The model predicted the outbreak of high level conflict on June 4. (See Figures 5a and b.)

Analysis of the 1973 war shows the results to be considerably more fragile. There are highly significant results when we include the first day of war with a lag of one day. These relationships, however, appear to be at least partially an artifact of the outlier on the first day of the war. When we attempted to control for the extreme conflict on the first day of war by rank correlations, we found a five day lag became significant while strength of the one day lag disappeared. All relationships became negligible when the first day of war was excluded from the analysis.

When we include Arab conflict at $t - 1$, we found roughly the same pattern. Particularly interesting is the fact that including the first day of war, the five day time lag is the strongest. Nonetheless, when we remove the first day of war, the $t - 1$ and $t - 5$ results all fall well below significance.

In addition to our conclusions concerning the viability of our measure of threat perception as a predictor of Israeli conflict, we can make some additional inferences about the effectiveness of conflict events as predictors of reciprocal conflict events. Hypothesis #6 posits threat perception alone as the predictor of conflict. Using it, we are able to explain 82 percent of the variance in pre-war 1967 results. Addition of the conflict

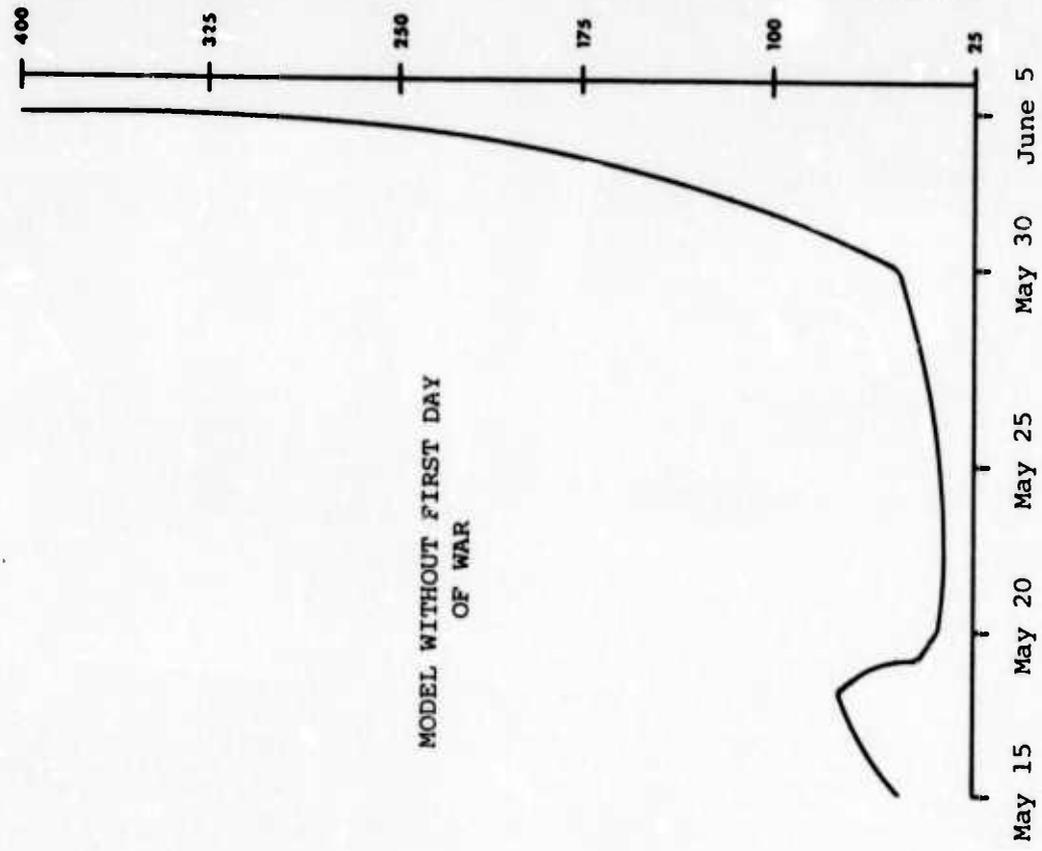


FIGURE 5b

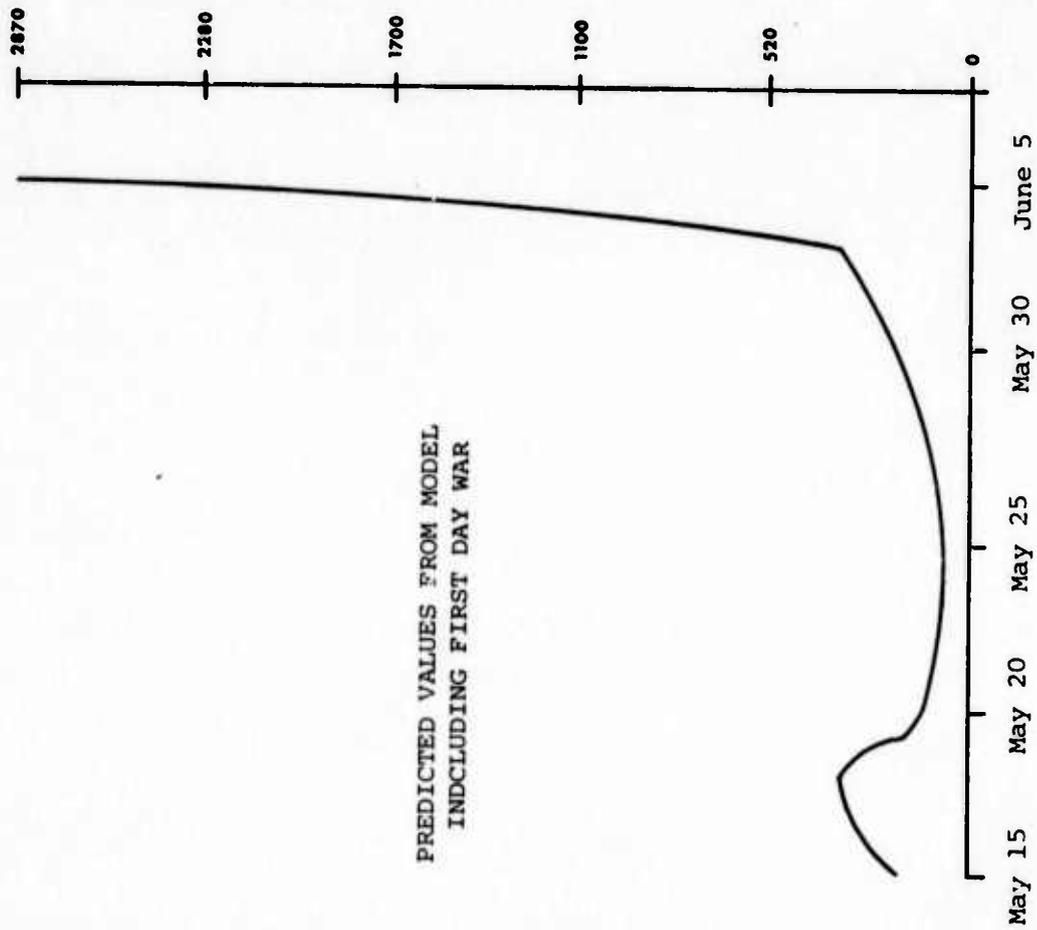


FIGURE 5a

variables fails to increase the explained variance by any notable margin.

The pre-war 1973 results, on the other hand, demonstrate that, in combination with threat perception, Arab physical conflict is able to add appreciably to the explained variance when the first day of the war is included. However, eliminating the first day of war reduces these results to insignificance. We concluded that, in the pre-war 1967 case, we can strongly support Hypothesis #6. The addition of lagged Arab conflict, however, does not improve the results over lagged threat perception alone.

While there are some statistically significant results in 1973, they do not stand up to even moderately severe testing. Using conventional standards then, we would conclude the hypothesis is ambiguously supported for the 1973 pre-war period.

As a final point, although it is outside the hypothesis, we briefly examined the relationship between Inquirer II measures of threat and Israeli conflict. For the 1967 war we found the correlation reaches a peak of 0.51 with a five day time lag when we include the first day of war. This is significant at the 0.05 level. The correlation drops to 0.05 when we exclude the first day of war. Attempts to control for the outlier using rank correlations proved similarly unsuccessful with $Rho = 0.31$ and $Tau = 0.27$. Both are non-significant.

For the 1973 war, the Inquirer II threat does even worse. There are no significant correlations between Inquirer II threat and Israeli conflict for any time lag, with or without the first day of war.

We concluded that while there is some linkage between the Inquirer II and manual threat measures (see Hypothesis #5) this relationship is not strongly passed on as a prediction of behavior.

HYPOTHESIS #8

The final hypothesis argues it is necessary to consider the three phases of war separately. We realized at the outset there are both substantive and methodological reasons for the separate analyses. Substantively, we have no reason to believe nations act similarly in escalation, war, and post-war periods. Their goals are different, the mix of activities is different, etc. Methodologically, we anticipated difficulties measuring events during wars. We did not anticipate difficulty measuring threat in the post-war periods.

To test this hypothesis, we examined all significant findings except those known to be a function of outliers or autocorrelation. We then compared the regression coefficients with coefficients of the same tests for different phases of the war. If parameters of a test from our time period were within two standard deviation units of an identical, but significant, test of a different phase, we concluded they came from the same sample and could have been considered together. If not, we concluded we were correct in separating the phases.*

Of the fifty-four significant pairs** of findings fitting our criteria, thirteen could have come from the same universe, and forty-one appeared to come from different universes. This is significantly different from a $D = 0.50$ at the 0.001 significance level. We, therefore, concluded Hypothesis #8 is confirmed and we decided to analyze the crises by phases.

* Where only bivariate correlations were run, we considered the correlations to have come from different samples if they have different signs and at least one is significantly greater than zero.

** Since there are three phases to both crises, each significant finding is compared to two other phases. Thus, the 54 pairs of findings are generated from 27 significant regressions or correlations.

CHAPTER V
CONCLUSIONS

In a most basic sense, objectives of this portion of the research project were:

1. To examine the feasibility of measuring perceptions of time pressure, threat, and surprise during crises, and to gather data on those evaluated as feasible.
2. To test a series of hypotheses, within the general mediated stimulus-response framework, on the relationships between event/interaction and perceptual processes in the crises surrounding the 1967 and 1973 Arab-Israeli wars.

We treated conflict events as the initial stimuli, threat perceptions as an intervening step, and subsequent events as the response. Thus, we had originally hypothesized that conflict events lead to perceptions of threat which, in turn, lead to conflict.

Summarizing the results of our analyses, we found the following:

- A. Content analysis of the public documents of Israeli decision makers and spokesmen can be used to generate an index we believe to be a reasonably valid measure of threat perception. We were unsuccessful in measuring perception of surprise. Our measure of time pressure using content analysis has some validity but seems to measure only the peaks of "real" time pressure. Computerized content analysis of the same documents produces another index of threat perception that is correlated fairly well with the manually coded index. The machine-coded variable, however, contains more apparently random variation than the manual index. There is some evidence from the 1967 crisis that the two indices tap a somewhat different underlying phenomena because a lagged relationship between the two variables produces a stronger fit than the simultaneous relationship. It was our expectation that as a more abstract measure of threat, we could treat the computer measure of threat perception as the initial emotive perceptual reaction to events.
- B. In both crises, the conflict one side directed at the other was only weakly related to the conflict directed back except when the relationship was measured without any time lag. Even when observed

at the same time point, although the two actors' conflict is highly correlated, there is no evidence conflict alone leads to a simple "conflict spiral."

- C. Perception of threat (with variable time lag) is a strong predictor of Israeli conflict in the pre-war periods of 1967 and 1973. This relationship is strongest when threat is measured with manual content analysis. It is weaker when the automated content analysis measure is used, although there is some evidence that relationship holds in 1967. Combining event data and perception data does not enhance the predictive power to any significant extent in the 1967 escalation. There is, however, some inconsistent evidence of improvement in the pre-war 1973 period.
- D. We were unable to use measures of Arab events to predict threat perception at all well. While logic dictates Israeli perception of the Arabs as threatening should be a result of Arab actions, we were unable to demonstrate this empirically. We conclude the step of the mediated S-R model linking events → perceptions is apparently quite complex.
- E. The parameters linking events and perception to conflict are not invariant across the phases of a particular conflict, or between similar phases of two conflicts. Parameter estimates from regression analysis and the optimal time lags vary considerably within and between the two crises studied. These shifts most likely result from interactions with other unmeasured perceptual phenomena (e.g., time pressure, and surprise) as well as with other environmental characteristics of the situation.

One purpose of the present project was to evaluate some empirical tools for crises research. The analyses were performed with data about Arab and Israeli international events coded from public sources, and estimates of Israeli perception of threat and time pressure, coded from public documents. We have been able to use a manual content analysis of the documents to come up with a variable that predicts reasonably well to Israeli conflict. This variable was constructed to measure threat perception, and we believe that it is a valid measure. It is based on the assumption anticipated levels of effort required to allay a perceived threat to some value is a measure of threat severity. With one exception, manual coding went well. Inter-coder reliability was high, and the measure stood up under tests for face validity

and convergent/discriminant validity. The one major difficulty was the coding rule permitting coding of events which occurred in the immediate past. A suggested revision in the coding rules would strictly forbid any references which are announcements of actions which have already occurred.

Whatever the indicator measures (it may be simply an indicator of what the Israelis are planning to do), the rather impressive predictive power of the index for the 1967 war certainly makes it worthy of further investigation.

The computer measure of threat is comprised of the interaction of Osgood's measures of strength, activeness, and negative affect. The results were something of a disappointment. While the Inquirer II threat index correlated with the manual index at a sufficiently high level to assure us of the latter's validity, it was a considerably weaker predictor of conflict in both 1967 and 1973. This is unfortunate because the manual content analysis is quite tedious and inter-coder reliability can be less than optimal. It would be beneficial, especially for use of the methodology in the policy community, if the computer could replace the human coders.

Machine content analyses often depend on large samples for validity; the manual schemes are less sensitive to sample size. In preparation of the text prior to machine processing (see Appendix B), we attempted to control some sources of extraneous variance. Despite these efforts, the index was overly sensitive to apparently random fluctuations in observations. This results from the index's multiplicative nature. We found, however, the three dimensions measured by Inquirer II do, apparently, interact with each other to produce a systematic variation with conflict. An attempted additive index produced less useful results.

The results obtained here were sufficiently encouraging to warrant further exploration of some form of computer content analysis as a means of generating perceptual data for international relations research. It is quite possible that obtaining larger samples of documents for each day, or making the observations at a higher level of aggregation will minimize the random variance we observed in the Inquirer II analyses. Also, there may be a systematic relationship between the words deleted in pre-processing and the Israelis' current perceptions of the Arabs. Conceivably the number of observations lost by deleting these references actually detracts from validity of the final index. Additionally, it may be possible to develop indices from the Osgood dimensions less sensitive to random fluctuations, but, which still capture the essence of the concept under study. Finally, it is probably feasible to computerize the manual content analysis. Systematic exploration of these alternatives and others was beyond the scope of this project, but it should be of considerable value to machine content analysis users for international relations research.

Turning now to the hypotheses tests, we found strong support for the perceptions → behavior link. This finding has considerable policy relevance because it gives the decision maker approximately a five day warning of the outbreak of violent conflict during a crisis escalation period. We were less able to support the perceptions → perceptions link, and totally unable to find any predictive power from events to either of the threat perceptions or the subsequent events.

The latter finding is inconsistent with several empirical studies of the Middle East. Many researchers have used event variables to predict (in the statistical sense) Arab conflict from Israeli conflict and vice-versa.

Azar (1974) finds high reaction coefficients for several Middle East dyads around the 1956 and 1967 wars. Wilkenfeld (1972) finds that across an 18 year period, most of the variance in Israeli conflict would be accounted for by the conflict directed at it from the three front line Arab states. In fact, James McCormick (1975) finds Israeli conflict to be very strongly related to Egyptian conflict during the most intense phase of the 1967 crisis.

The studies noted above differ in levels of aggregation and time periods examined. One critical characteristic they have in common different from our examination of the event/interaction hypothesis is they measured the independent variable, namely; behavior received from the opponent nation, in the same time increment they measured the dependent variable. If, in fact, the relationship between Arab acts and Israeli acts is truly escalatory, we would expect to observe it when short time lags are introduced. In our analyses, the coefficients dropped precipitously when lags were introduced. If the simple stimulus-response model is to be used in a truly predictive way, there must be an observable lag between the predictor and outcome phenomena. Furthermore, there are methodological problems with evaluating the hypothesis when the variables are measured simultaneously. For one, there are problems with coding rules. Many conflict events are double-coded, virtually ensuring high correlations at t_0 . Also, the assumptions underlying ordinary correlation and regression analysis preclude unbiased parameter estimates of this type of model without using econometric estimation techniques.*

We also used Azar's scaled conflict index with a three day lag to predict the outbreak of war (Hypothesis #7). When the predicted values were

* See Choucri and North (1975, Appendix B) for a relatively understandable discussion of this problem.

plotted out, we found the model with the Azar scaled data was anticipating the mobilization period very well, but tended to seriously underpredict the outbreak of war. We suggest this is due to the fact that mobilization and verbal conflict are given more weight on the Azar scale than the Corson scale. For example, mobilizations are coded as 3 on the 1-7 Azar scale. Corson weights mobilizations in the 50-80 range on a 1-515 scale. When an exponential transformation is applied to the Azar-scaled data to reduce the emphasis on low-level conflict, the correlation at $t - 3$ reported above drops to nearly zero. These results tend to support our a priori belief in the superior validity of the Corson scale.

In addition to examining event \rightarrow event links, we attempted an explanation of perceptions as a linear function of prior events, and were notably unsuccessful. We tried a number of models in our attempts to uncover a linkage, including Zinnes' "imperfect memory" model which she used to produce some significant results from pre-World War I data (1968). On our data, this model did no better than we were able to do with simple time lags.

Despite these results, we continue to believe the sequences of events have some relationship to the originator's perceptions. Otherwise we would have to conclude the Israelis were creating threat perceptions independently of events - a highly unlikely occurrence. Assuming we have measured events and perceptions reasonably well, we must conclude the perceptual mechanism is much more complex than the linear model we used to test the relationship. Indeed, research into cognitive mapping, selective perceptions, and operational codes suggests the perceptual mechanism is quite complex.

Given that the perceptual process is obviously an intervening step between a stimulus event and the reaction to that event, it is our belief

significant effort would be devoted to unravelling it. Work in the general area of cognitive mapping (c.f., Steinbrunner (1968), Shapiro and Bonham (1973)) has showed considerable promise.

In our analysis of threat perception as a predictor of behavior, we used the Stanford studies of the outbreak of World War I as a guide.

The Arab-Israeli crisis of May-June 1967 at first glance looked like an example of a simple conflict escalation process. One side's actions appeared to lead to the other side's responses, which, in turn, created a situation from which neither side could back down. One could argue that whatever the role of perception as an intervening variable, there might be a simple linear relationship: from Arab conflict → Israeli perceptions → Israeli conflict → Arab perceptions → Arab conflict and so on until war interrupted the spiral. Such a pattern has been observed in the 1914 crisis (North, Brody, Holsti 1964). We hypothesized that the operators linking events to perceptions and perceptions to events may be sufficiently simple that perceptions could be ignored or handled by transformations of event variables to control for standard perceptual processes (c.f., Azar, 1974). Our results, however, failed to support this hypothesis. Our measures and the historical descriptions show the crisis did not keep escalating until war broke out. The mutual mobilizations had been completed by the end of May, the news accounts focused on the diplomatic efforts to end the crisis, and just before the Israeli attack on June 5, there was considerable speculation the crisis was almost over. Reflecting this general attitude, our measures of threat perception rise to a peak in late May, then drop to peak again in the first few days of the war. We originally assumed decisions to go to war have many approach-avoidance characteristics, and that periods of significant apprehension covary with major decision points.

In addition, we assumed once a critical decision is reached, a period of relative tranquility would occur until the time arrives for the decision to be acted upon.* Reading the secondary sources about the crisis (c.f., Geist, 1974), one is left with the impression Israeli leaders' opinions hardened in favor of war about May 30-June 1; the time between then and the formal decision for war on June 4 was spent tidying up domestic and diplomatic loose ends to facilitate the war. This resulted in a five day time lag between threat and conflict.

There are, of course, rival hypotheses to the argument that this reduction in threat perception (and, interestingly, time pressure) is due to an approach-avoidance phenomena. Deliberate deception, for example, would be consistent with the pattern. The most important point, however, is that the relationship between threat perception and conflict is quite strong ($R^2 = .82$) and there is a reduction in threat perception prior to the outbreak of war. This indicates that even though the crisis appears to be easing after a peak in threat perception, decision makers should not be lulled into believing the worst is over. Indeed, the reduction may be a warning that the probability of war has actually increased.

The five day lag is somewhat harder to explain for the lower-intensity conflict, but it is not difficult to imagine it would take several days to translate perceptions into concrete action. We believe that five days is only the modal time lag in the pre-war 1967 period; perceptions were probably translated into action with different lags depending on a number of things (e.g., time pressure).

* It is well known, for example, that in many suicides, the victim is quite distraught until he comes to the decision to take his life. He then proceeds to go about life as normal until he acts on his decision. Indeed, "many clinicians have noted that patients act in an organized fashion, seem to be getting better, seem to be calmed, prior to a suicidal act." Kobler (1964), page 14.

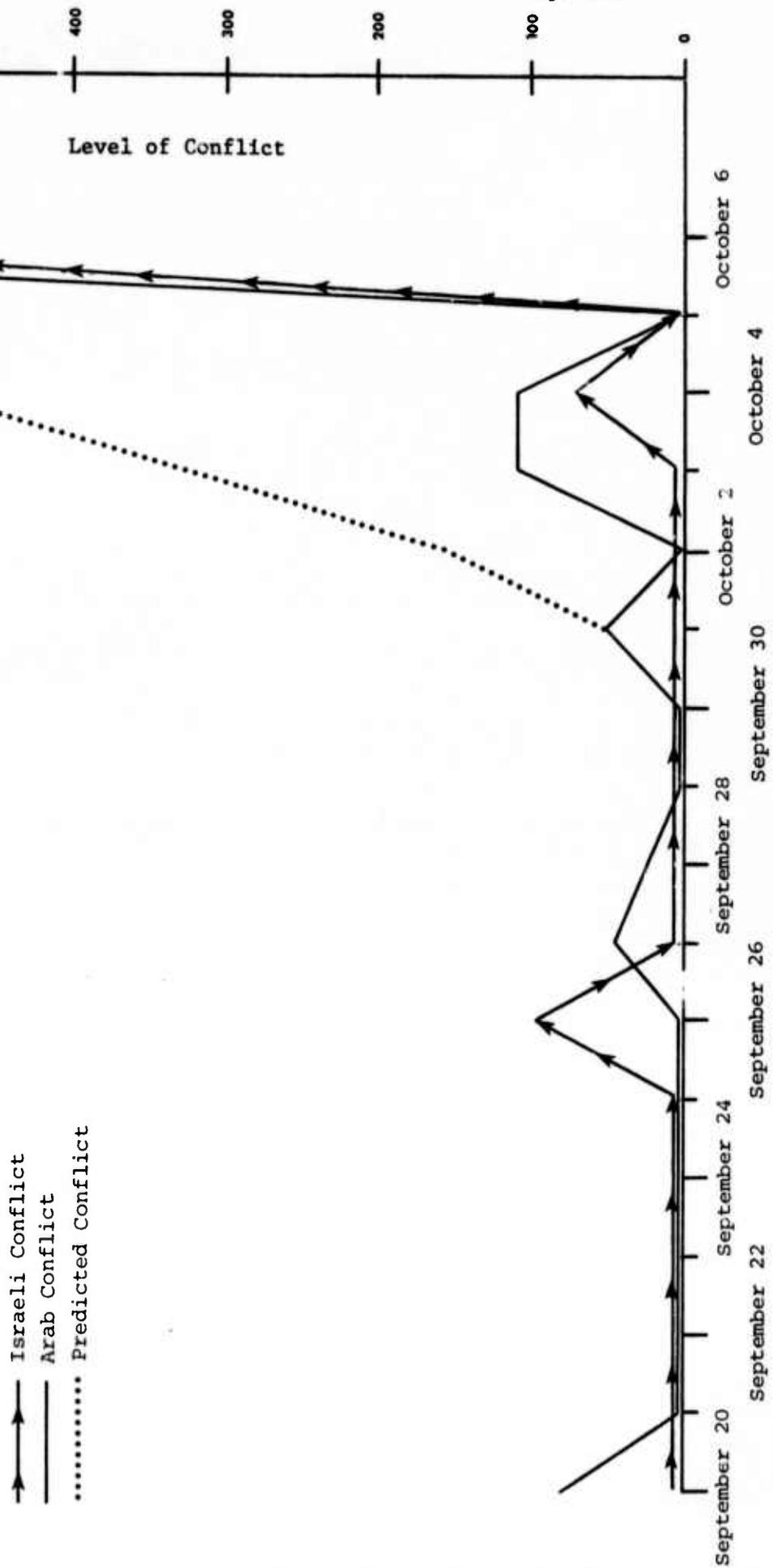
Although there are some correlations indicating threat perception is linked to conflict in the pre-war 1973 period, we are less confident that the hypothesis is supported in this case. The statistics are too ambiguous for us to place as much confidence in them.

A close examination of both history and the data reveal some important facts. While the 1967 pre-war period was a reasonable approximation of a classic escalation, the 1973 war was largely a surprise. Empirically, this means we observed conflict events in 1967 which varied throughout the pre-war period giving some admittedly complex indication that war would occur. This was far less true in 1973. Additionally, since the Israelis did not plan the attack, we could not expect the approach-avoidance phenomena to be applicable. There is reason to suspect even the Arabs may not have evinced approach-avoidance patterns prior to the 1973 war. Unlike 1967, the Arab October 6 attack followed relatively lengthy planning. They simply were not caught in a sequence of events similar to the Israelis in 1967. It is not at all certain approach-avoidance would be apparent in implementing a relatively long standing plan to attack. Given these reasons, the lack of fit between threat perception and conflict in 1973 is not surprising.

An interesting observation, however, is that despite the fact events are stable, threat perception does, in fact, increase prior to outbreak of the 1973 war as shown in Figure 6. It occurred to us this might be indicative of the outbreak of hostilities. We, therefore, used the regression coefficients from the 1967 pre-war to predict events in the 1973 war. It will be observed in Figure 6 that using these coefficients we predicted strong increases in Israeli conflict directed at the Arabs. This means that had the Israelis acted on their own threat perception in 1973 as they did in 1967, war would have broken out on or about October 5.

FIGURE 6

CONFLICT vs. PREDICTED CONFLICT - 1973



We do not want to make too much of this "prediction." It leaves a large number of questions unanswered which we hope to address. It obviously shows that the Israelis felt something was quite wrong as early as September 30. Clearly, the situation on September 30, 1973 was different from the situation on May 30, 1967 and, scientifically, we cannot transfer the results directly from one crisis to the other. Nonetheless, for the more classic escalation crisis, our findings have a very close fit with the hypotheses. We believe there is strong indication threat perception is an early warning indicator, and we have some knowledge of its dynamic characteristics during crises. For the non-escalation crisis, we do know threat increased prior to the outbreak of war but, because there was no significant event process, we do not have a good idea of how threat perception leads to war in these types of crises. The answer to this question will have to await subsequent research.

SUGGESTIONS FOR FURTHER RESEARCH

There are three research areas suggested by our findings:

1. Further examination of these two Arab-Israeli crises to clear up some ambiguities,
2. Examination of non-crises and other crises to examine generalizability of the results,
3. An examination of the process by which conflicting nations move from war to ceasefire using threat perception and operational reporting statistics data rather than event data.

Let us proceed through these points sequentially.

Point 1.

There still remain many unanswered questions concerning the Arab-Israeli wars of 1967 and 1973. The most obvious of these is that we do

not know Arab threat perceptions during the pre-war periods. It would be important to know both how they covaried with Israeli perceptions and whether and/or when they anticipated war.

Additionally, although we were unable to use Inquirer II as a predictor of conflict, it does roughly covary with the manual content analysis. However, we are concerned about its susceptibility to error on days when there are only a few words to analyze. Since, given our data set, eliminating those cases would hopelessly reduce the sample size, we were forced to include them. We believe it would be worth while to gather additional documents (probably in Hebrew) for the pre-war periods, and determine if we can produce more useful results from Inquirer II. These additional documents would, in fact, eliminate some of the degrees-of-freedom problems we experienced, particularly in the pre-war 1973 period. With sufficient documents, for example, we would not have to rely on estimates of missing data.

Finally, the analyses pointed to a number of minor revisions which should be examined before proceeding to further research. We have already alluded to the fact that, in post-war periods, our handling of the diplomatic level of effort category created problems. This should be examined and adjusted if possible.

The threat measure should be examined in the manner which an analyst would use in a real time analysis. Revising the estimates given new information using Markov processes would be a possibility.

Interval weightings of the indices were made on an ad hoc basis and should be subjected to empirical analyses.

While not involving extensive research, there are a number of questions of this type which should be answered before proceeding to further analysis.

Point 2.

The current study was performed on a sample of two crises not necessarily representative of any larger set of crises. Although we are encouraged by the results, there is more work needed to expand them to other situations. The first step should include other crises selected in a way permitting maximum comparison across categories of crises. We would want to compare our current findings with both escalating and non-escalating pre-crisis periods.

During the process of studying other crises, it would be important to include periods of tension short of crisis to avoid developing a measure having too high a false alarm rate. By checking observation of varying threat perception during both crises and non-crises, the research would attempt to uncover levels and/or patterns of threat perception characteristic of crises.

If the results of this suggested research support the current findings that manual content analysis is a good predictor of threat perception, it would be reasonable to consider computerizing as much of it as possible. Although that decision need not be made until further research is completed, it is worth keeping in mind that potential users of the measure should not be burdened by a cumbersome manual content analysis.

Point 3.

The final area of future research deals with conflict de-escalation periods. The level of threat perception may be an important factor in determining when and how peace negotiations begin. To address this question, however, the research team will require a measure superior to event data to measure the progress of ongoing war. Battle deaths, movement of fronts,

numbers of sorties, etc., should provide us with sufficient data. It seems reasonable to expect there may be combinations of threat perceptions which are optimal for negotiations. In any event, it would be useful to potential negotiators to know the threat perceptions of the various parties.

APPENDIX A
CODING RULESGeneral

Coders will fill in a full packet of coding sheets for each document. Each document will have its own set of coding sheets. The basic document information will be recorded on the "Time Code Sheet." On this sheet will be recorded the date, the speaker, the paraphraser (if any), a document number, a document reference, the coder's name and the page number.

Date: The date will be the date of the communication - not of the report of the communication.

Speaker: The name of the speaker, or if a paraphrase - list the original speaker and the paraphraser. If the speaker is not well known, identify his official position in the government.

Document Number: A pre-assigned number to identify each document. Each document has a unique number. If you come across a document you think should be coded, but for which there is no document number, give it a temporary number.

Document Reference: A reference such as "press conference" or "radio address."

Page Number: This records the page, per category (such as value threatened) per document. If only one page is needed to record

value threatened, for example, the page number is recorded 1 of 1.
If two are needed, they will be recorded 1 of 2 and 2 of 2.

Coders will always read the entire document first before going back and coding it. This will give a "sense" of the document so that phrases can be read in the proper context.

What is Codable

Decision makers frequently make reference to the historical events leading up to the time of the speech. These may be coded only according to the following rules:

If the events are behaviors of another nation(s) they are codable if:

- 1) The events are obviously continuing into the present
- and 2) These events pose some current threat to the nation of the decision maker or they provide evidence of the time pressure under which the decision maker is currently operating.

If the events are actions of the nation represented by the decision makers they are codable if:

- 1) The action refers to a decision made in the immediate past. That

is, if the speech announces or explains a recent decision to act, the actions, even if completed, are codable under level-of-effort. Actions completed in the past and associated with past decisions are not codable. Actions which were instigated in the past but are continuing in the present and are mentioned in the present are codable under the assumption that there was a conscious decision to continue the action.

Decision makers also frequently will make references to general policy. These are normally too vague to be codable. If they are specific, they are codable only if there is a clear implication for the crisis at hand. For example, "We have always sought to maintain a balance of power in South-East Asia" is codable if there is an implication that the balance of power is currently threatened.

Decision makers will often make oblique references particularly to values threatened. These are difficult judgmental decisions. Generally, we want to minimize "reading between the lines." Nonetheless, there are occasional statements containing very clear implications (usually of value threatened). Statements such as, "We love our children and are determined that they be able to attend school in a normal fashion like other children around the world," in its context implies that the situation which is described is threatened. This would be coded even though there is not an explicit statement that it is threatened. As a general rule, oblique references should be coded only if the inference is quite clear. As always, the inference must apply to the crisis at hand. Similarly, simple descrip-

tions of the opponent's military capability can be coded if, in the context of the document, it is clear that they pose a specific threat. General descriptions of the enemy as "strong" or "aggressive," etc. will not be coded here.

Finally, negative statements such as "we do not want to fight" are almost never codable unless they can be obviously translated into a statement of positive intentions.

CODING RULES - TIME

Our basic purpose in this category is to measure the perception of time pressure felt by the decision makers during a crisis. The columns to be filled in are:

Demand (D)/Perception (P)/Other (O)

Time Code

Actor-Object

Ambiguity Check

Direct Quote

Time Reference

Except for the ambiguity check, each column should be completed for each time reference.

Demand (D)/Perception (P)/Other(O) Code: This is simply a code identifying whether the time constraint is perceived by the actor to be upon himself/herself, if he/she is making a time related demand on another object nation, or if he/she is simply describing the speed with which events are moving. You should be able to translate in your mind most time references into one of the two following types of logical statements.

Demand: "A" demands that "B" perform "X" action (time reference)
e.g., Israel demands that Egypt return prisoners immediately.

Perception: "A" perceives that "A" must perform "X" action (time reference) e.g., Israel perceives that it must mobilize reserves eventually.

If the statement is of the type, "We reacted quickly" or, "They moved suddenly" or, "These rapid succession of events...", you would code it "Other" (O).

Time Code: Time pressures will be coded by three categories:

Not enough time - 0

Short time - 1

Enough time - 2

These are relatively self-explanatory but I will elaborate somewhat.

Not enough time: This requires a clear statement that there is (was) insufficient time to take some action. This should not be coded unless it is obvious that because of the time pressures, some particular action will not be or was not carried out.

Short time: This category included all references other than the above which indicate that time pressures are of concern. This will normally be apparent by words such as rapidly, immediately, urgently, soon, speedy, etc. There may be oblique references to time pressures such as actions of senior people in the middle of the night. These should be recorded with an ambiguity check (see below).

Enough time: These are time related statements in which decision makers say, or clearly imply, that they are not under time pressure. Words such as enough time, plenty of time, proceed slowly, are indicative. You should also look for statements such as, we will be prepared, react methodically, etc., as inferences of low time pressures.

Actor-Object: Use Singer-Small codes for the acting nation and the nation which is the object of the action. For time demands the actor is the nation making the demand, the recipient of the demand is the object. For time pressure there is only an actor and that is the nation which perceives that there are demands made upon itself.

Ambiguity Check: A note to identify a question on the part of the coder's mind whether the time code is correct (TC), whether it is a demand or perception (DP), who the actor-objects are (AO) or whether it is a time reference at all (TR). Any question of inference should be checked (IR).

Direct Quote: A short phrase, not necessarily a sentence, containing the time reference and enough information to convey the sense of the reference.

Task Reference: A very brief statement of the task to be completed within the time period referenced.

Date (Past Action): For actions in the past for which a time reference is given (e.g., we had to act quickly), record the date of the action - not the communication. For actions in the present or future (e.g., we do not have much time), this column will be left blank.

CODING RULES - LEVEL-OF-EFFORT

Generally, the level-of-effort category attempts to measure the effort which decision makers believe they will have to expend to overcome some threat. In your mind, a level-of-effort statement is any which can be translated into the logical statement:

"A" must expend "X" effort in order to "Y"

e.g., Israel must expend an enormous effort in order to repel the attacking Egyptians.

Actor: On the code sheet, "actor" needs to be recorded only if the nation which will have to expend the effort is different from the nation represented by the speaker.

Willingness (W)/Expectation (E): We are also trying to distinguish whether leaders are saying that they are willing to exert some given effort to allay a threat or whether they expect to have to exert some effort to allay a threat. Conceivably, they might be willing to exert more than they expect to exert (low threat), conversely they may expect to have to exert more than they are willing to (high threat). In practice, these are statements of probability or contingencies. Expectation (E) statements imply that the action is planned (or underway) and almost certain to occur barring the unexpected. With only rare exceptions "expectation" implies that the actor both expects to and is willing to exert some specific effort. "Willingness" statements are those which either explicitly state or imply that there are some factors upon which effort is conditional. These may be threats such as "We will sink your ships if you violate our blockade" or offers "We are willing to negotiate a settlement" ("If you are also willing" is implied).

Effort Type and Code: The following is the outline of the level-of-effort codes.

- I. Military
 - A. All out military action
 - B. Major military action
 - C. Moderate military action
 - D. Minor military action

- II. Social/Domestic
 - A. Total national effort
 - B. Significant national effort
 - C. Some national effort

- III. Diplomatic
 - A. Bilateral diplomacy
 - B. International organizations
 - C. Miscellaneous diplomatic

- IV. Economic Sanctions

Military Effort

- A. All out military action requires a statement of a willingness to commit the entire military to overcome some severe threat. At a minimum, this must imply war. For example:

"Declare war"
"State of war"
"Entire U.S. military"

- B. Major military action requires a statement of the deployment of reasonably sizeable military forces with either the real potential of provoking a major war and/or with the effect of causing considerable destruction of lives and/or property. For example:

"Take Soviet vessels into custody"
"Sink Soviet ships"
"Inflict grave consequences"

- C. Moderate military action includes large military alerts, non-specified military action, and the deployment of small forces intended to harass the enemy. For example:

"Armed force action"
"Double combat air patrol"
"Take up arms"

- D. Minor military action includes unspecified or smaller alerts, calling up reserves, or supplying reinforcements. For example,

"Positive but limited measures"
"Call up Air Force Reserves"

"Use force only as necessary"

"Military alert"

Social/Domestic Effort

- A. Total National Effort generally implies an unmistakably perceived national commitment to sacrifice to allay the threat. Characteristic of this effort is the breadth of its diffusion across the social structure. There should be an implication that many in the nation will be significantly affected. For example:

"To the limits of our resources"

"To the uttermost"

"Everything we've got"

"Hard work and a seven day week from all of us"

- B. Significant National Effort implies a recognition of a major national effort to allay the threat. The fact that a major effort is needed should be clear. However, there will not be any clear evidence that this sacrifice will affect a broad range of the society. For example:

"National determinism"

"Show courage and capacity to fight"

"Sacrifice"

- C. Some National Effort is characterized either by a statement that

a small segment of society will have to sacrifice or by vague statements that some unspecified effort will be required. For example:

"Practice self-discipline"

"Will stand firm"

"Our troops need our support"

Statements such as "we will do whatever is needed" will be coded A II-C unless there is a previous level of effort statement in the document. In that case, the statement will be set equal to the maximum level of effort mentioned previously.

Diplomatic Effort

- A. Bilateral efforts refers to diplomatic efforts between the nations actually involved in the fighting.
- B. International organizations refers to diplomatic efforts involving any international organization such as the U.N., OAS, or NATO.
- C. Miscellaneous Diplomatic refers to any diplomatic action not covered by the above.

Economic Sanctions

This refers to any and all international economic sanctions which can be taken against another country such as embargos and tariff restrictions.

The codes will include major type and subcategory such as,

I - C for moderate military action or

III - A for bilateral diplomatic

In addition, you will need to record the following codes.

Ambiguity check: A check if there is confusion on the coding use (ET) for confusion or the Effort Type. Use it if there is a question whether the item belongs in the level-of-effort category. This will occur if you have questions of the inferences you may have drawn on indirect statements of level-of-effort.

Direct quotation: Phrase containing the direct quotation of the level-of-effort. This need only contain enough information to permit a "feel" for the intensity being conveyed.

CODING RULES - VALUES THREATENED

In this category you are trying to identify the cognitive dimensions of value which are threatened during a crisis. You will not be measuring intensity in this portion of the analysis. Rather, you will be categorizing the substantive issue which decision makers say is being threatened. To fill out the code sheets you will need to do the following.

Actor: Needed only if the nation perceiving a threat is different from the nationality of the speaker.

Value category: There are five categories with associated codes:

Military-Security (MS)

Economic-Development (ED)

Cultural-Status (CS)

Political (Domestic) - (D)

Political (International) - (I)

Other Domestic (O)

These are further defined below.

Military-Security (MS) comprises all values which focus on questions pertaining to violence, including military alliances and weaponry, and those which are perceived by the foreign policy elite as constituting a security threat.

Economic-Development (ED) comprises all values which involve the acquisition and allocation of resources, such as trade, aid, and foreign investment.

Cultural-Status (CS) consists of those values involving cultural, educational, and scientific matters as well as international status issues as perceived by the decision-makers. (e.g., honor, dignity)

Political (Domestic) (D). Include values associated with the basic philosophy underlying the domestic political system, e.g., democracy - freedom - Communism.

Political (International) (I). Include values associated with the relationships between states or the methods by which such relationships are governed.

Other Domestic (O). Include threats to the domestic situations in countries other than that of the speaker.

Some examples of values threatened would be coded as noted in Table A-1.

Enemy Activity Code: Frequently, decision makers will imply a threat through a reference not to the value being threatened, but to the activities of the enemy which threaten some value. A reference to a blockade, for example, could mean that there is a military-security value threatened. References to Fedayeen attacks would also imply a MS threat. Where this occurs, simply check the "Enemy Activity Code" column on your code sheet and record it as normal.

Ambiguity Code: Same as above except that (VC) means the value code is ambiguous, (C) means that it is questionable whether the statement belongs in the value threatened category.

TABLE A-1

Military-Security	Economic Development	Cultural-Status	Political-Domestic	Political-International	Other Domestic
Peace *	Economic losses	Dignity Cultural values Honor	Socialism Victories of the Revolution Democracy Communism	Status quo Balance of power obligations International law Rights Order International relationships	(their) Freedom (their) Democratic government (their) Economy
Security					
Independence					
Military Assets					
Existence					
Life of Citizens					
Territory					

**Peace" can be coded as political international if it is clear that the intent is a threat to the general peaceful interactions of nations. If a threat to "peace" implies that the actor will have to go to war, a military security value is threatened.

Direct quotation: This may be a little difficult since you may occasionally have to infer the value threatened. Nonetheless, try to give the best direct quotation possible that indicates the value threatened.

APPENDIX B
COMPUTER CONTENT ANALYSIS

RATIONALE - GENERAL

Content analysis is but another way to produce systematic data. Like the use of simulation or public opinion samples or quantitative history, the use of content analysis is a way of rigorously and systematically identifying and analyzing specified characteristics--in this case the specified characteristics being a part of communications. For that is what content analysis is:

A research technique for the objective, systematic and quantitative description of the content of communication. ⁽¹⁾

Among the characteristics of content analysis noted in the above definition on which there is wide agreement are those of objectivity, systematic and quantitative. Objectivity stipulates that analysis must be carried out on the basis of explicitly formulated rules which will enable two or more persons to obtain the same results from the same documents. Systematic means that the inclusion and exclusion of content or categories is done according to consistently applied criteria of selection. This requirement eliminates analyses in which only materials supporting the investigator's hypotheses are examined. Quantitative means that the data are comparable data so as to allow statistical analysis, particularly statistical tests which permit a more accurate description of the degree of covariation of two or more attributes. ⁽²⁾

¹ John Galtung, Theory and Methods of Social Research: Data Collection, p. 1, (Oslo, Norway: Instituttet for Sosiologi, Report No. 1, 1965), p. 2.4. Bernard Berelson, "Content Analysis," in Lindzey (ed.) Handbook of Social Psychology, Vol. 1, (Cambridge: Addison-Wesley, 1954), p. 489.

² For an expanded and more detailed discussion of objectivity, systematic, and quantitative in content analysis see Ole Holsti, Content Analysis, (unpublished manuscript, Stanford University, 1965), pp. 2-8.

To return to the definition for a moment, there has developed a problem of interpretation; that is, whether content analysis must be limited to manifest content--in other words, the surface meaning of the content, or, may content analysis be used to analyze the deeper layers of meaning embedded in the content (the latent content)? The recent trend has been in the direction of a broader definition, thus including latent content as a valid means of analysis. This is primarily a result of the many recent studies which have shown that such an extension explains the content of communication more fully and more incisively than does just an analysis of manifest content.

When should one use content analysis? A fundamental reason for using content analysis in research on international conflict and integration has been succinctly stated by Dorwin Cartwright:

Social and political conflicts, although often stemming from divergent economic interests and power, cannot be fully understood without studying the words employed in the interaction of conflicting groups, and the process of mediation consists largely of talking things out.⁽³⁾

Within this context, content analysis has a number of applications. It may be used to study conflicting goals or the content of ideologies.⁽⁴⁾ It may also

³ Dorwin Cartwright, "Analysis of Qualitative Material," in Festinger & Katz (eds.), Research Methods in the Behavioral Sciences, (New York Dryden Press, 1953), pp. 421-70.

⁴ For example, content analysis has been used to measure interstate behavior, Ole Holsti, et al., Theory and Measurement of International Behavior, (Stanford: Stanford Univ., 1964); to measure international tension, Robert North et al., The Analysis of International Tension, (Stanford: Stanford Univ., 1964); and to compare Soviet and American value preferences, Robert Angell, et al., "Social Values and Foreign Policy Attitudes of Soviet and American Elites," The Journal of Conflict Resolution, 1964, 8, pp. 330-491.

be used for analysis at a somewhat different level: to gain some systematic information concerning the cognitive and evaluative, or effective states, of those persons whose decisions are binding upon the states they represent. (5)

And, finally, it may be used to describe patterns of communication in international relations. (6) In short, content analysis can be and is used in a multitude of research investigations. Whenever the problem requires precise and standardized methods for analyzing those aspects of verbal behavior which may escape casual scrutiny, content analysis will be useful.

⁵ For example, content analysis has been used to study influence and power of a party in power, G.A. Almond, *The Appeals of Communism* (Princeton: Princeton University Press, 1954); to measure influence of various groups of lobbyists, W. McPherson, "Lobbying and Communication Processes," paper read at American Political Science Association Meeting, Chicago, 1964; to infer personality traits of individuals, E.S. Shneidman, "A Psychological Analysis of Political Thinking," (Cambridge, Mass.: Harvard Univ. mimeo, 1963); to analyze the relationship between perceptions of threat and perceptions of capability during an international crisis, Dina Zinnes, et al., "Capability Threat and the Outbreak of War," in J. Rosenau (ed.) International Politics and Foreign Policy, (New York: Free Press of Glencoe, 1961), pp. 469-482; and to analyze Soviet and U.S. foreign policy attitudes, J.D. Singer, Soviet and American Foreign Policy Attitudes: A Content Analysis of Elite Articulation, (China Lake, Calif.: U.S. Naval Ordnance Test Station, 1964).

⁶ For example, the events of the 1914 crisis as analyzed in Ole Holsti, "The 1914 Case," American Political Science Review, 1965, 59, pp. 365-378; a comparison of the Cuban missile crisis to the 1914, Ole Holsti, et al., "Measuring Effects and Action in International Reaction Models: 1962 Cuban Crisis," Peace Research Society, Papers, II, 1965, pp. 170-190; trends in Sino-Soviet Relations, M.G. Zaninovich, An Empirical Theory of State Response: The Sino-Soviet Case, Unpublished Doctoral Dissertation, Stanford University, 1964, and F. Mogdis, "The Perceptual Nature of Sino-Soviet Interactions 1950-1967," paper presented at APSA, 1971.

TYPES

Content analysis can be conveniently divided into two general categories or types: (1) manual, and (2) automatic (computer). The former, in its many varying forms, has accounted for the vast majority of research done thus far in the field. It has, however, at least two major drawbacks--one is reliability and the other is cost.

An alternative to these problems seems to be offered by the use of a computer content analysis program. This approach solves the problem of reliability because every item of the same type will be automatically coded in the same way.

Although the computer algorithms to automatically content analyze verbal statements are expensive to operate, they cost considerably less than do the time consuming manual methods.*

One of the alleged disadvantages of computer content analysis is that reliability is acquired at the expense of validity. That is, while the computer will consistently provide the same answers when processing the same data, it is not certain these answers will adequately measure the underlying concepts addressed by the research.

While there is undoubtedly some truth to this position, it is also true that by increasing complexity of computer processing, validity can be improved without sacrificing reliability. This process, however, can increase costs considerably.

The General Inquirer

The General Inquirer, the forerunner of the more flexible and elaborate Inquirer II was a set of computer programs that allow the user to (a)

*Not including the initial costs of developing the computer program and analytical dictionaries.

identify systematically, within text, instances of words and phrases that belong to categories specified by the investigator: (b) count occurrences and specified co-occurrences of those categories: (c) print and graph tabulations: (d) perform statistical tests: and (e) sort and regroup sentences according to whether they contain instances of a particular category or combination of categories."⁽⁷⁾ Inquirer II is essentially a more capable program designed for the IBM S/360. The application of the Inquirer II content analysis program to any given analytical task is accomplished through three basic steps; the coding process, which is mostly a data input formatting task; the development of the dictionary; and the use of one or more of the program options.

All these steps must be guided by the underlying analytical goals of the particular research task in question. That is to say, the use of the Inquirer II, as with any analytical tool, is very much dependent upon the research needs of the investigator. Hence, efficient use of the flexibility of the program will be realized only if the investigator carefully specifies the goals of the research in considerable detail prior to undertaking any part of the process of using the Inquirer II programs. The more specific the analytical goals, the more guidance the researcher will provide for himself at each subsequent decision point.

General Description

Data and dictionary serve as input to the system. The tagging program (the program which assigns the categories) reads in the data a sentence at a time, then locates each word in the dictionary. Instructions are given by the dictionary as to what category should be assigned and/or what searches of the

⁷ Philip J. Stone, et al., The General Inquirer: A Computer Approach to Content Analysis, Cambridge and London: The MIT Press, 1966, p. 68.

context in which the word occurs should be made. The instructions are then executed. When the analysis of the sentence is completed (i.e., all the categories that are to be assigned have been assigned: all searches have been completed), the tagging program writes out the sentence and reads in the next. The process continues until such time as all the data have been tagged. The output from the tagging program is a tagged file which may be stored so that retrievals and/or tabulations of the data may be made at a later date.

Data Input Format

The input is logically equivalent to PL/1 stream I/O in that all fields are imbedded within the text, rather than utilizing a record oriented format. The text is assumed to be in upper and lower case (which allows for special handling of proper nouns). Each of the 256 different characters in the EBCDIC character set are assigned a function. For example, an alpha-numeric character (A-Z, a-z, 0-9) is considered part of a word; a blank indicates the end of the word. The period (.), plus sign (+), exclamation point (!), and question mark (?) indicate the end of a sentence. Braces ({ }), greater than and less than signs (><), and the dollar sign (\$) indicate aspects of the text which are not to be searched in the dictionary and not to be given a content category. The function of each of these input characters can be changed at tagging time if the investigator so desires. For example, the investigator may choose to have a semicolon indicate the end of a sentence in addition to or instead of the other characters.

Samples of Data Input Format

Card 1 \$# ISRAEL I PERCEPTIONS OF # EGYPT \$
Card 2 < # DAYAN SPEECH OF 9 OCTOBER >

Card 3 [CLOSED PRESS CONFERENCE]
Card 4 # THE MERE PRESENCE
Card 5 (OF TANKS ON
Card 6 THE # EGYPTIAN FRONT MEANS
Card 7 WAR [AS FAR AS # I AM CONCERNED].

The use of the dollar sign (\$) in card 1 of the example indicates that this is a title and is not to be tagged but retained and used for labeling in post processing. The pound sign (#) is a convention adopted to indicate that the following letter is to be capitalized. When no pound sign appears, lower case is assumed. The less than (<) and greater than (>) signs contain the Identification Field (ID) and may appear anywhere in the text. The appearance of a number of blanks on card 4 and card 5 illustrates that where any one blank is found any number of blanks may appear. The braces ({ }) enclose comments and indicate that the data they enclose are not to be processed by the dictionary.

Dictionary Preparation

The major task in using the Inquirer II system is the creation of a dictionary. A content analysis category (called a concept in the I/II system) consists of a number of language signs (words, idioms, phrases, and so on) that together represent a variable in the investigator's theory. The basic procedure in content analysis is to identify these signs when and if they occur in text as instances of a particular concept and score them as such. (This is the procedure known as tagging.) We were, for example, concerned with threats and threat perception in this particular research task and were interested in identifying the number of threat references within the documents chosen for analysis. Consequently, the process of tagging was used to identify, score and count each word in the next text that made reference to the concept of a threat.

The investigator seldom carries out a content analysis with a single concept. Instead, he is usually interested in examining relationships of a number of semantic categories as they appear within specific documents. Most content analysis studies, therefore, use a cluster of concepts, referred to as a content analysis dictionary. For the I/II system, the exposition of this dictionary is in a special language, Dictionary Definition Language (DDL).

Data Preparation

Prior to entering text into the computer, it is necessary to perform some data preparation. This includes two tasks: (1) inserting the correct punctuation, and, (2) eliminating irrelevant information. The correct punctuation must be inserted primarily to permit the computer to identify complete thoughts short of sentences, proper names, the end of documents, the end of logical sets of documents, etc. Irrelevant material can be treated in a number of ways.

In this project, irrelevant material included statements about actions or situations that occurred in the past but do not continue at the time of the statement, remarks about actions other than the Arabs', and introductory remarks. References to past action were deleted according to the criteria set forth in the manual coding rules (Appendix A).

Tagged Output

The initial outputs of the I/II tagging program are the original data plus the categories that have been assigned and stored for future use on some output medium specified by the investigator. If the investigator chooses, the output from category assignments may be listed so that the text can be inspected to see how well the category "fits" the data. The total text can be inspected

for category assignment and for words or combinations of words which did not receive any categories (i.e., leftover words). Those words which did not receive any categorization will be underlined in the listing so that the user knows which characteristics of the data were not handled by any of the dictionary routines. Moreover, after having inspected the listing of the output, the investigator may resubmit the original output for re-tagging by the same (usually updated) dictionary or by additional dictionaries.

Selective Listings and/or Tabulations

These listings and tabulations may be applied to the entire body of text or only certain positions. When applied to only a subset of the data, they roughly correspond to the Retrieval functions of the General Inquirer System. The conditions may specify certain content categories, certain combinations of content categories, actual words used in the text, categories and words, or may involve identification fields (ID) in combination with these content categories and/or words.

Retrieval

1. % IF WORD = 'HEALTH'
THEN (CONCEPTS;TAB)
ELSE

2. % IF TAG = HEALTH
THEN (LIST;TAB)
ELSE

The above two specifications are relatively simple. For the first, the program simple searches the document which has been tagged for the occurrence of the word "health," it will be listed with the appropriate concepts; and counts will be made for a tabulation. For the second, a similar operation occurs, but this time the instructions say look for the Concept HEALTH. The

instruction is to list (without concept names) and tabulate for those sentences which contain the Concept HEALTH.

Retrieval specifications can be quite complex. In fact, all the complexity that can be used in creating a dictionary may be used. One can check for order, position in the sentence, the presence or absence of various combinations of concepts or word or ID fields and place them in a single retrieval specification. In general, the more complex retrievals are dependent upon the investigator's hypothesis and/or familiarity with the data that were categorized.

APPENDIX C - LIST OF DOCUMENTS

1967 Crisis

<u>Date</u>	<u>Speaker</u>	<u>Reference</u>
5/13	Eshkol	Speech to Israel
5/22	Eshkol	Knesset
5/23	Cabinet	Decision to delay decision
5/23	Eshkol	Knesset speech
5/24	Rafael	U.N. S.C. statement
5/28	Eshkol	Radio address
5/29	Eshkol	Knesset
5/29	Rafael	U.N. S.C. statement
5/29	Eshkol	Knesset speech
5/30	Eban	Press conference
5/31	Rafael	U.N. S.C. statement
6/1	Eshkol	Letter to Kosygin
6/3	Rafael	U.N. S.C.
6/4	Gov't. Press Office	Press release
6/4	Cabinet	Decision to attack
6/5	Dayan	Statement to troops
6/5	Eshkol	Knesset speech
6/5	Eban	Press conference
6/5	Eshkol	Letter to Kosygin
6/5	Rafael	U.N. S.C. statement
6/5	Eshkol	Speech to Israel
6/6	Eban	U.N. S.C. statement
6/7	Eban	U.N. S.C. statement
6/7	Eshkol	Speech to Rabbis
6/8	Eban	U.N. S.C. statement
6/9	Rafael	U.N. S.C. statement
6/9	Galili	Speech to Israel
6/10	Rafael	U.N. S.C. statement
6/10	Rafael	U.N. S.C. statement
6/10	Rafael	U.N. S.C. statement
6/11	Rafael	U.N. S.C. statement
6/11	Meir	Speech to UJA in New York
6/12	Eshkol	Knesset speech
6/13	Kidron	U.N. S.C. statement
6/14	Kidron	U.N. S.C. statement
6/14	Kidron	U.N. S.C. statement
6/16	Cabinet	Reply to Soviet note
6/17	Eban	Press conference
6/19	Eban	Speech to U.N. General Assembly
6/27	Eshkol	Public statement

1973 Crisis

<u>Date</u>	<u>Speaker</u>	<u>Reference</u>
9/26	Dayan	Troops on Golan Heights
10/1	Golda Meir	Steinbock speech
10/3	Eban	United Nations
10/4	Elezar	Speech to troops
10/6	Golda Meir	Israel
10/6	Dayan	Press conference
10/7	Eban	Sec. Gen.
10/8	Eban	General Assembly
10/9	Dayan	J. Post
10/10	Golda Meir	Israel
10/10	F. Ministry	Press
10/13	Golda Meir	Press
10/16	Knesset	Resolution
10/16	Tekoah	Security C.
10/16	Golda Meir	Knesset
10/19	Tekoah	Security C.
10/20	Tekoah	Sec. Gen.
10/21	Tekoah	Security C.
10/21	Tekoah	Security C.
10/23	Eban	Africa
10/23	Golda Meir	Knesset
10/23	Cabinet	J. Post
10/24	Tekoah	Sec. Gen.
1-25	Tekoah	Security C.
10/25	Tekoah	Sec. Gen.
10/28	Golda Meir	CBS
11/1	Golda Meir	U.S. Press
11/4	Eban	Italy
11/13	Golda Meir	Knesset
11/23	Peres	J. Post
11/27	Eban	U.S. Jews
12/1	Golda Meir	Radio
12/20	Golda Meir	Knesset
12/21	Eban	Peace Conf.

BIBLIOGRAPHY

- Abelson, Robert P. and Tukey, John W., "Efficient Conversion of Non-Metric Information into Metric Information," Proceedings of Social Statistics Section of the American Statistical Association. Washington, D.C. 1959, p. 226-230.
- Azar, Edward, "Analysis of International Events," Peace Research Review, IV No. 1, 1970.
- Azar, Edward E., Bennett, James and Sloan, Thomas, "Steps Toward Forecasting International Interactions," Papers of the Peace Science Society, Vol. 23, Nov. 1974.
- Brecher, Michael, "Environmental Inputs and Decisions for War and Peace: The Israeli Experience," paper presented at the IX World Congress of the International Political Science Association, Montreal, Canada 1973.
- Brecher, Michael, Decisions in Israel's Foreign Policy, Yale University Press 1974.
- Choucrist, Nazli and North, Robert, Nations in Conflict: National Growth and International Violence, San Francisco, W.H. Freeman Co., 1975.
- Commission of Inquiry--Yom Kippur War, "Partial Report, April 1974."
- Corson, Walter, "Conflict and Cooperation in East-West Crises: Dynamics of Crisis Interaction," Ph.D. Thesis, Cambridge, Harvard University, 1970.
- Couture, Walter, Emergent Human Nature, N.Y., Alfred Knopf, 1949.
- Geist, Benjamin, The Six Day War: A Study of Decision Making Under Crisis Conditions, Ph.D. Dissertation, Hebrew University of Jerusalem, 1975.
- Herman, Charles F. (ed.), International Crises: Insights from Behavioral Research, N.Y., Free Press, 1972.
- Holsti, Ole R., Crisis Escalation and War, Montreal, McGill-Queens University Press 1972.
- Holsti, Ole R., Content Analysis for the Social Sciences and Humanities. Reading, Massachusetts: Addison Wesley Publishing Co., 1969.
- Holsti, Ole R., Brody, Richard A., and North, Robert C., "Measuring Affect and Action in International Reaction Models: Empirical Materials from 1962 Cuban Missile Crisis." Journal of Peace Research I, 1964, 170-189; and Papers of the Peace Research Society (International) II, 1965, 170-189.

- Holsti, Ole R., North, Robert C., and Brody, Richard, "Perception and Action in the 1914 Crisis," in J.D. Singer (ed.), Quantitative International Politics. New York: The Free Press, 1968, 123-158.
- Kanovsky, Eliyahu, The Economic Costs of the Six Day War, New York, Praeger, 1970.
- Kobler, Arthur and Stotland, Ezra, The End of Hope, London, Toronto, Collier Macmillan 1964.
- London Sunday Times, The Yom Kippur War, Doubleday and Company 1974.
- Mansoor, Menahem, Political and Diplomatic History of the Arab World 1900-1967; A Chronological Study, Washington, D.C., NCR/Micro Card Editions 1972.
- McClelland, Charles A., "The Acute International Crisis." World Politics, SIV (1961), 182-204.
- McClelland, Charles A. and Hoggard, Gary, "Conflict Patterns in the Interactions among Nations," in Roseman (ed.), International Politics and Foreign Policy, New York, (Revised edition) 1969 711-24.
- McCormick, James, "Evaluating Models of Crisis Behavior: Some Evidence from the Middle East," International Studies Quarterly, Vol. 17, No. 1, March 1975.
- Mogdis, Franz J. and Tidwell, Karen, Sino-Soviet Conflict: Its Implications for United States Strategic Policy in the 1970's. Bendix Corporation Final Report for Project TRIAD Contract F44620-68-C-0083 for AFXDOC/USAF August 1970.
- Namenwirth, Zvi and Brewer, Thomas, "Elite Editorial Comment on the European and Atlantic Communities in Four Continents," in P. Stone et. al. (eds.), General Inquirer, Cambridge: MIT 1966.
- Newcomb, Theodore M., Turner, Ralph and Converse, Philip. Social Psychology: The Study of Human Interaction. New York, Holt Rinehart and Winston 1965.
- North, Robert C., "Perception and Action in the 1914 Crisis," Journal of International Affairs 21 (1967), 103-22 (III-B-22 RR).
- North, Robert C., Brody, Richard and Holsti, Ole, "Some Empirical Data on the Conflict Spiral," Peace Research Society (International) Papers 1 (1964), 1-14 (III-B-22 RR).
- North, Robert C. and Choucri, Nazli, "Background Conditions to the Outbreak of the First World War," Peace Research Society (International) Papers 9 (1968), 125-37 (III-B-15).

- Osgood, Charles E., Suci, George and Tannenbaum, Percy, The Measurement of Meaning. Urbana, University of Illinois Press 1957.
- Rummel, Rudolph J., "Some Dimensions in the Foreign Behavior of Nations," Journal of Peace Research, 1966/3 201-24.
- Rummel, Rudolph J., "Indicators of Cross National and International Patterns," American Political Science Review, 63/1 (March 1969) 127-47.
- Shapiro, Michael J., and Bonham, Matthew, "Cognitive Processes and Foreign Policy Decision Making," International Studies Quarterly, June 1973 147-174.
- Singer, J. David, "Threat Perception and the Armament-Tension Dilemma," Journal of Conflict Resolution II-1 1958 90-105.
- Snider, James and Osgood, Charles (eds.), Semantic Differential Technique; A Sourcebook, Chicago, Aldin Publising Co., 1964.
- Steinbruner, John D., The Mind and Milieu of Policy Makers: A Case History of the MLF, Cambridge, Mass.: MIT, Ph.D. Dissertation, 1968.
- Stone, Philip, et. al, General Inquirer, Cambridge, MIT 1966.
- Tanter, Raymond, Modeling and Managing International Conflicts: The Berlin Crises, Beverly Hills, Sage 1974.
- Tufte, Edward, "Improving Data Analysis in Political Science," World Politics, July 1969 641-655.
- Wilkenfeld, Jonathan, "Conflict Interactions in the Middle East, 1949-1967," Journal of Conflict Resolution XVI No. 2, June 1972.
- Zadeh, L.A., "Fuzzy Sets," Information Control, Volume 8, pp. 338-353, 1965.
- Zinnes, Dina A., "Some Evidence Relevant to the Man-Milieu Hypothesis," in Roseman, James, N., Vincent Davis and Maurice East (eds.), The Analysis of International Politics, New York: Free Press 1972.

FAAC

ARPA Order Number: 2873

Contract Number: N00014-75-C-0328

Program Code Number: 5W10

Principal Investigator/
Phone Number: David M. McCormick
(313) 761-5836

Name of Contractor: First Ann Arbor Corporation
209 East Washington Street
Ann Arbor, Michigan 48108

Scientific Officer: Director, Organizational Effectiveness
Research Program
Psychological Sciences Division
Office of Naval Research
Department of the Navy
800 North Quincy Street
Arlington, Virginia 22217

Effective date of contract: 1 October 1974

Short Title of Work: DEPIC

Contract Expiration Date: 30 June 1975