ORA-TR-71-2

ORA-TR-71-2

THEORY OF DETERRENCE AND DISSUASION

Bruno J. Manz

OFFICE OF RESEARCH ANALYSIS (AFSC)

February 1971

Successor Organization:

OFFICE OF THE ASSISTANT FOR STUDY SUPPORT (AFSC)
Kirtland Air Force Base, New Mexico 87117

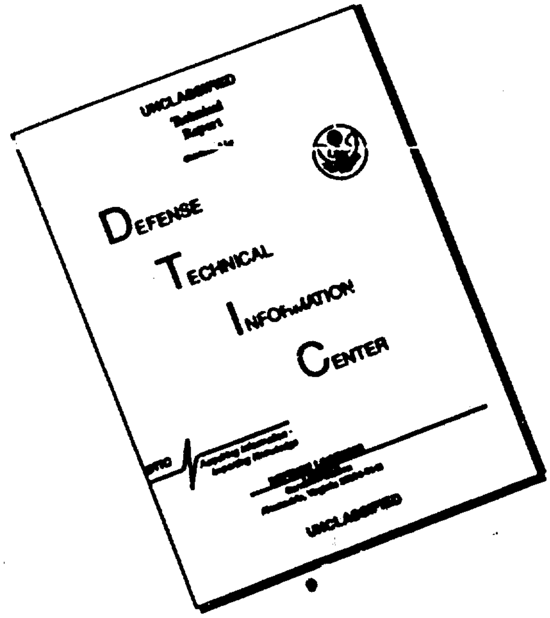
Approved for public release; distribution unlimited.

Qualified requestors may obtain additional copies from the Defense Documentation Center

R

37

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

ORA-TR-71-2

THEORY OF DETERRENCE AND DISSUASION

Bruno J. Manz

OFFICE OF RESEARCH ANALYSIS (AFSC)

February 1971

Successor Organization:

OFFICE OF THE ASSISTANT FOR STUDY SUPPORT (AFSC)
Kirtland Air Force Base, New Mexico 87117

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) OFFICE OF RESEARCH ANALYSIS (AFSC) (see 11 below)	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

THEORY OF DETERRENCE AND DISSUASION

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific. State-of-the-Art. 1 June 1961 - 30 June 1970

5. AUTHOR(S) (First name, middle initial, last name)

Bruno J. Manz

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
February 1971	22	None

8a. CONTRACT OR GRANT NO. ORA C010-00-02 (In-house Work Unit)	9a. ORIGINATOR'S REPORT NUMBER(S) ORA-TR-71-2
b. PROJECT NO. C010	
c. 63101F	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d. N/A	

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES ORA deactivated - successor organization: Ofc of the Asst for Study Support (AFSC) Kirtland AFB, NM 87117	12. SPONSORING MILITARY ACTIVITY Office of Research Analysis

13. ABSTRACT

A "theory ~~of~~ deterrence and dissuasion" is developed within the wider framework of a previously ~~developed~~ probabilistic theory of international nuclear stability. In this wider framework, it was endeavored to establish the probabilities of various types of nuclear missile wars between two opponents as functions of ~~the~~ damages and certain psycho-political parameters. This theory is essentially based on one "rationale of spontaneous belligerency" and one "rationale of preventive belligerence." The ~~present~~ report is ~~only~~ concerned with ~~the~~ spontaneous belligerency of the two opponents. A theory is presented which generates the probability that either power will attack spontaneously as a function of appropriate damages and four psycho-political parameters, viz., the aggressiveness parameters and the standards of damage of both powers. It is always endeavored to derive the probabilities logically from certain clearly exposed principles which form the postulatory basis for the theory.

FOREWORD

This report is a continuation and an improvement of work which was reported in ORA Technical Report 70-0012 entitled "Probabilistic Theory of Strategic Nuclear Stability". The latter report is one of a series of ten reports which were published under the heading "Theory and Model of Probabilities and Damages of Nuclear Missile War". To set the background for the present report, it is appropriate to repeat here the common foreword for the above mentioned series of ten reports:

This report is one in a series of ten reports that describes the results of an analysis entitled "Theory and Model of Probabilities and Damages of Nuclear Missile War". The study was conducted by the Office of Research Analyses (ORA) under Project C010, Work Unit C010-00-02 between 1 June 1969 and 30 June 1970.

The objective of the study was to evaluate the effectiveness of several strategic missile systems and to predict how those systems would affect military stability if they were deployed. The study considers all of the advanced USAF concepts for strategic missiles as well as the sensor systems necessary to acquire data. It was necessary to consider both offensive and defensive systems in order to investigate military stability since both capabilities are equally important in determining the relative strength of potential combatants. The present study differs significantly from previous studies of this type which evaluated different missile system alternatives in terms of the relative effectiveness of each system if they were used in a war. Such an evaluation overlooks an important attribute of strategic missile systems; namely, their ability to deter a war from starting. The present study makes the traditional effectiveness evaluation and also evaluates the peace-keeping ability of the different systems.

The theory presented in this report generates the probabilities of various kinds of nuclear missile wars as functions of the damages to both nuclear powers and certain psycho-political parameters.

This report has been reviewed and is approved.

Gerhard R. Eber
GERHARD R. EBER
Technical Director
Office of Research Analysis

Ernest J. Davis, Jr.
ERNEST J. DAVIS, JR.
Colonel, USAF
Commander
Office of Research Analysis

UNCLASSIFIED
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Nuclear War						
Nuclear Strategy						
Probability of War						
International Strategic Stability						
Decisions in Missile Defense						
Offense and Defense						
Belligerency						

UNCLASSIFIED
Security Classification

ABSTRACT

A "theory of deterrence and dissuasion" is developed within the wider framework of a previoulsy developed probabilistic theory of international nuclear stability. In this wider framework, it was endeavored to establish the probabilities of various types of nuclear missile wars between two cpponents as functions of the damages and certain psycho-political parameters. This theory is essentially based on one "rationale of spontaneous belligerency" and one "rationale of preventive belligerency". The present report is only concerned with the spontaneous belligerency of the two opponents. A theory is presented which generates the probability that either power will attack spontaneously as a function of appropriate damages and four psycho-political parameters, viz., the aggressiveness parameters and the standards of damage of both powers. It is always endeavored to derive the probabilities logically from certain clearly exposed principles which form the postulatory basis for the theory.

CONTENTS

SECTION I

INTRODUCTION

The report "Probabilistic Theory of Strategic Nuclear Stability" (ORA-70-0012) stands, essentially, on two legs. One is the "Rationale of Preventive Behavior" and one is the "Rationale of Spontaneous Belligerency." It has already been intimated in the above report that, from a logical point of view, the rationale of preventive behavior is satisfactory, but the rationale of spontaneous belligerency is not. The reason is simoly that the latter is substantially more difficult than the former. The reader can, probably, appreciate that it is easier to predict, in a probabilistic manner, the preventive actions of a person, than his spontaneous belligerency. In fact, one may question whether the spontaneous initiation of a nuclear war (as distinct from the preventive initiation) is at all a rational act. However, this all depends on the relation between the retaliatory damage which is the consequence of a spontaneous attack, and the damage which would materialize if the alternative course of action were taken which is to abstain from a spontaneous attack. If the latter is larger than the former, then a spontaneous attack is, indeed, a rational act. Whether or not this <u>is</u> the case or even <u>can</u> be the case, is not the subject of this report. In fact, it is not even the business of the executive branch of government, but the business of the legislature. For it is a matter of values; and the analyst's job is not to set values but to predict the consequences once they have been set. This report will strictly refrain from any value judgment and from any advice in matters of policy.

The "Theory of Deterrence and Dissuasion" presented in the present report is a considerable improvement over the "Rationale of Spontaneous Belligerency" presented in ORA-70-0012. However, it is still not the ultimate goal which would be a "Theory of Escalation". In the present report, it is still assumed that the spontaneously striking power strikes with its full force, without holding anything in reserve in anticipation of "stages of escalation". Work is presently in progress which is designed to do away with this restrictive assumption. However, it is extremely difficult, and it cannot be said at this time if and when the effort will be successful. This is the reason that an improved, but still imperfect theory of deterrence and dissuasion is presented.

For the nomenclature and the definition of symbols and concepts, the reader is referred to ORA-70-0012. Further illumination is provided by another report entitled "Theory of Preventive Behavior with an Application to International Nuclear Stability" which will be published in a journal.

The present report replaces a certain portion of ORA-70-0012. This is spelled out in more detail at the end of Section V.

The improvement of the present report over ORA-70-0012 is discussed in the conclusion of the present report.

SECTION II

THE DETERMINANTS OF SPONTANEOUS AGGRESSION

For simplicity, let us assume that the first power (sometimes designated by the symbol Po_1) contemplates a spontaneous strike against the second power (Po_2). Whether or not Po_1 will actually carry out the contemplated spontaneous strike, depends on a variety of factors which are ingredients of the strategic scenario. Not all ingredients of the scenario are relevant to the first power's decision to attack or not to attack. We call the relevant ingredients of the scenario "determinants of spontaneous aggression." It will first be necessary to list and to discuss these determinants.

For waging a spontaneous attack on the second power, the first power must have a motive and an objective. It is, again, not necessary to describe motive and objective in complete detail; only the relevant aspects need to be discussed. They will constitute the first two determinants of spontaneous aggression.

As far as the motive is concerned, all that is relevant to the theory of deterrence and dissuasion is the probability that, within a given time period τ, the first power will have a motive to wage a spontaneous attack on the second power. We denote this probability by $P_{1a}(\tau)$ and call it "probability of motivation". This probability will <u>not</u> be generated within the frame work of the present thoery. It will be seen later (see also the referenced ORA report) that we do not need this complete probability, but only the first coefficient of its Taylor development:

$$P_{1a}(\tau) = a_1 \tau + \cdots \qquad (1)$$

Here, the dots indicate terms which are of second and higher order in τ. The coefficient a_1 which has the dimension of an inverse time (which is "frequency") is called "aggressiveness parameter of the first power." More generally, a_i is the aggressiveness parameter of the ith power ($i = 1, 2$).

The relevant aspect of the objective can be described by the "damage objective" D_1^2 <u>of the first power</u>. More generally, D_i^j is the damage objective of the ith power ($i = 1, 2, i \neq j$).

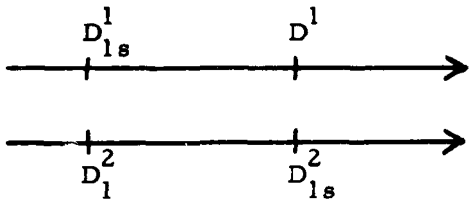
So far, we have identified the determinants $P_{1a}(\tau)$ or (or a_1) and D_1^2. We have to identify three further determinants. Obviously, the damages to both powers are determinants of spontaneous aggression. The damage to the second power which is the immediate consequence of the spontaneous strike by the first power is denoted by D_{1s}^2. The damage to the first power D_{1s}^1 is, of course, only indirectly caused by the first power herself, but is directly caused by the second power in retaliation.

Of course the damage D_{1s}^1 to the first power is closely related to the deterrence aspect. However, it describes this aspect not completely. For a certain damage may be large to a small nation and it may be small to a large nation. In other words, we need a "standard of damage" which is characteristic of the first power. We denote by D^1 the "standard of damage of the first power," i.e., D^i is the standard of damage of the ith power ($i = 1, 2$).

The following scheme lists the determinants of spontaneous aggression for the first power.

$$\begin{pmatrix} a_1 & D_1^2 & D_{1s}^2 \\ & D^1 & D_{1s}^1 \end{pmatrix} \tag{2}$$

Four of the five determinants of spontaneous aggression have the dimension "damage." These four damages can be represented on two "damage scales," one for the first and one for the second power:



Damage scale for the first power

Damage scale for the second power

The absolute values of the two damages on each scale are not important, but only their values relative to each other. We call these damage ratios "the measured damages." More precisely we define

D^1_{1s} / D^1 = "Measure of Deterrence"

D^2_1 / D^2_{1s} = "Measure of Dissuasion"

The inverse of the measure of dissuasion is called "measure of temptation":

D^2_{1s} / D^2_1 = "Measure of Temptation"

These are important concepts which need further discussion. To this end, we introduce the following concepts. First, we define three probabilities which are all relevant to the proposition that the first power will wage a spontaneous attack.

$P_{1a}(\tau)$ = Probability that, within the time period τ, Po_1 will wage a spontaneous attack, given that she is neither deterred nor dissuaded (probability of motivation)

P_{1d} = Probability that Po_1 is deterred <u>and</u> dissuaded

Q_{1d} = Probability that Po_1 is not deterred <u>and</u> not dissuaded

$P_{1s}(\tau)$ = Probability that, within the time period τ, Po_1 will wage a spontaneous attack on Po_2

It is first to be noted that $P_{1a}(\tau)$ and $P_{1s}(\tau)$ depend on the time period under consideration whereas P_{1d} and Q_{1d} do not. The latter two probabilities are, of course, related by

$$P_{1d} + Q_{1d} = 1 \tag{3}$$

It also follows from the above definitions that

$$P_{1s}(\tau) = Q_{1d} P_{1a}(\tau) \tag{4}$$

This equation expresses the fact that Po_1 will attack spontaneously if, and only if, she has a motive to do so <u>and</u> she is neither deterred nor dissuaded.

Finally, we introduce two logical symbols which stand for propositions and will facilitate otherwise lengthy expressions.

1DD = Proposition that Po_1 is deterred and dissuaded (the first character, 1, indicates the first power)

~1DD = Proposition that Po_1 is neither deterred nor dissuaded

Hence, the probability that 1DD is true is P_{1d}; and the probability that ~1DD is true is $Q_{1d} = 1 - P_{1d}$.

Let us now assume that Po_1 has a motive for a spontaneous attack. This assumption would have to be expressed by setting P_{1a} equal to one. In this situation, Po_1 has two alternative courses of action: to attack or not to attack. If she attacks she is, obviously, neither deterred nor dissuaded, i.e., the proposition ~1DD is true. Let us now inspect the consequences to Po_1 for either course of action. This is relatively easy for the alternative ~1DD. In this case, she will suffer the retaliatory damage which, however, is to be "measured" by her own standard of damage. Hence, the consequences to Po_1, if she acts according to the proposition $\sim 1DD$, is the measure of deterrence D^1_{1s} / D^1.

We turn now to the case where Po_1 decides to act according to the proposition 1DD, i.e., where she abstains from a spontaneous attack, even though she has a motive. In this case, she suffers also a damage, which, as always, is the damage one suffers if he abstains from doing something which he desires to do. What the first power desires to do is inflict on the second power a damage according to her damage capability D^2_{1s} but "measured" by her damage objective D^2_1. In other words, if the first power abstains from waging a spontaneous attack - in spite of the fact that she has a motive - then she suffers the measured damage D^2_{1s} / D_1 which we have called "measure of temptation." The following scheme summarizes the results of this discussion.

		Consequences to Po_1	
Course	$1DD(P_{1d})$:	D^2_{1s} / D^2_1	(5)
of			
Action	$\sim 1DD(1-P_{1d})$:	D^1_{1s} / D^1	

Here we listed the probabilities for either course of action in parentheses following the symbol for the respective course of action.

The above scheme is an "estimation matrix." This concept has been introduced and explained in detail in the aforementioned report, "Theory of Preventive Behavior with an Application to International Nuclear Stability."

Also in that report, a "principle of estimation" has been introduced, explained, and applied for establishing a certain probability in the mind of some "estimator" that another person, the "decision maker," will take a certain course of action. In the next section, the principle of estimation will be applied to determine the probability

P_{1d} as function of the measure of deterrence and the measure of temptation.

SECTION III

APPLICATION OF THE PRINCIPLE OF ESTIMATION

The principle of estimation serves to determine the probabilities of two alternative courses of action. Two persons (or collectives of persons) must be distinguished: The "estimator" in whose mind the probabilities exist; and the "decision maker" who decides between the two alternative courses of action. If we introduce the symbols

$$\langle D \mid 1DD \rangle^1 = D^2_{1s} \; / \; D^2_1 \qquad (6a)$$

$$\langle D \mid \sim 1DD \rangle^1 = D^1_{1s} \; / \; D^1 \qquad (6b)$$

then we may write the "estimation matrix" (5) in the more concise form

$$\begin{pmatrix} \langle D \mid 1DD \rangle^1 \\ \langle D \mid \sim 1DD \rangle \end{pmatrix} \qquad (7)$$

These are two "expected damages" to Po_1 (indicated by the superscript one) which Po_1 is expected to suffer if she chooses the courses of action which are stated in the "condition compartments" to the right of the "condition bar". If we multiply an expected damage for a certain course of action by the probability that Po_1 will take this course of action, then we arrive at the corresponding "weighted expected damage". Hence, the two weighted expected damages are

$$P_{1d} \; \langle D \mid 1DD \rangle^1 \qquad (8a)$$

$$Q_{1d} \; \langle D \mid \sim 1DD \rangle^1 = (1 - P_{1d}) \langle D \mid \sim 1DD \rangle^1 \qquad (8b)$$

We are now ready to state the principle of estimation, as it applies to the present case.

Principle of Estimation

The estimator selects the probability P_{1d} such that the two weighted expected damages to Po_1 are equal, that is

$$P_{1d} \langle D \mid 1DD \rangle^1 = (1 - P_{1d}) \langle D \mid {\sim}1DD \rangle^1 \tag{9}$$

If this equation is solved for P_{1d}, we obtain

$$P_{1d} = \frac{\langle D \mid {\sim} 1DD \rangle^1}{\langle D \mid {\sim} 1DD \rangle^1 + \langle D \mid 1DD \rangle^1} \tag{10}$$

If we resubstitute here relations (6) we obtain

$$P_{1d} = \frac{D^1_{1s} / D^1}{D^1_{1s} / D^1 + D^2_{1s} / D^2_1} \tag{11}$$

If we introduce the abbreviation

$$M_1 = \frac{D^1_{1s} \ D^2_1}{D^1 \ D^2_{1s}} \tag{12}$$

and if we observe (11) and (3) then we may write

$$Q_{1d} = \frac{1}{1 + M_1} \tag{13a}$$

$$P_{1d} = \frac{M_1}{1 + M_1} \tag{13b}$$

We call M_1 the "measure of deterrence and dissuasion" for the first power. The measure of deterrence and dissuasion for the second power is

$$M_2 = \frac{D^2_{2s} \ D^1_2}{D^2 \ D^1_{2s}} \tag{14}$$

SECTION IV

DISCUSSION OF INTERMEDIATE RESULTS

It is now no longer necessary to confine ourselves to the one case when the first power attacks spontaneously. We consider now both cases. According to equation (4), we have

$$P_{1s}(\tau) = Q_{1d} \; P_{1a}(\tau) \tag{15a}$$

$$P_{2s}(\tau) = Q_{2d} \; P_{2a}(\tau) \tag{15b}$$

If we then observe relations (12), (13a), and (14), w .y write

$$P_{1s}(\tau) = \frac{P_{1a}(\tau)}{1 + \dfrac{D^1_{1s} \; D^2_1}{D^1 \; D^2_{1s}}} \tag{16a}$$

$$P_{2s}(\tau) = \frac{P_{2a}(\tau)}{1 + \dfrac{D^2_{2s} \; D^1_2}{D^2 \; D^1_{2s}}} \tag{16b}$$

We shall first discuss these results briefly. It suffices to discuss result (16a). The following figure will facilitate this discussion.

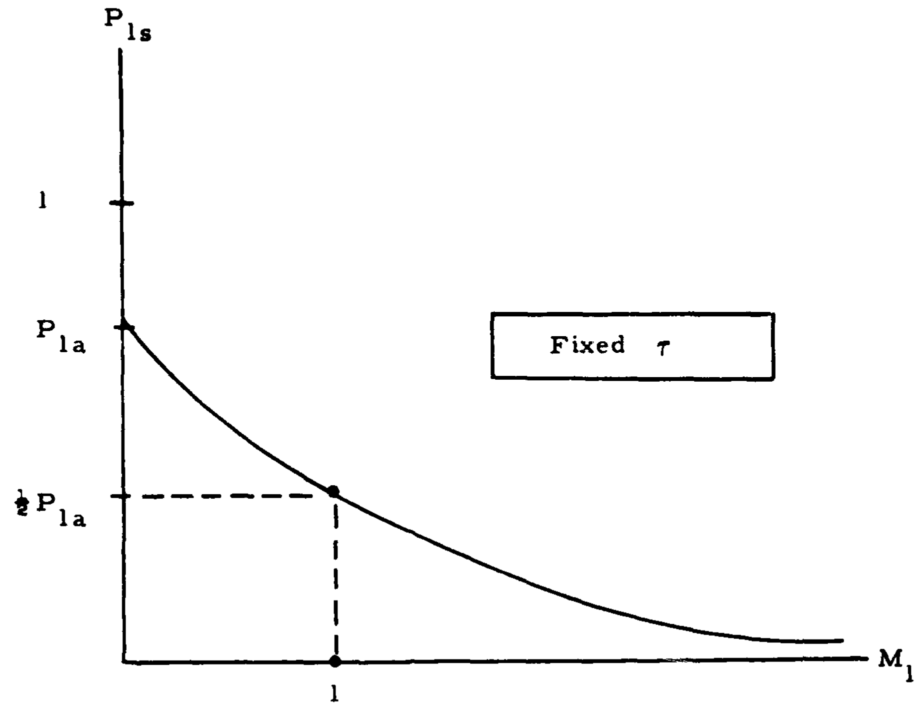


Figure: Probability of Spontaneous Attack as Function of the Measure of Deterrence and Dissuasion

It is first to be noted that the probability of spontaneous aggression never exceeds the probability of motivation:

$$P_{1s}(\tau) \leqq P_{1a}(\tau) \tag{17}$$

This must be so for, if Po_1 does not have a motive, she will not attack under any circumstances; on the other hand, if she does have a motive, she will attack only with the probability $Q_{1d} = 1 - P_{1d}$.

It is then to be noted that

$$P_{1s}(\tau) = P_{1a}(\tau) \text{ for } M_1 = 0 \tag{18}$$

In this case, Po_1 is neither deterred nor disuaded. In such a situation, she will attack if, and only if, she has a motive to do so. Hence, the two probabilities must be equal in this situation.

If now the measure of deterrence and dissuasion increases, it will be less and less likely that Po_1 will attack spontaneously. There are two causes which bring about this effect: Deterrence and Dissuasion. These two causes work in unison and have the same effect: reduction of P_{1s} from its maximum value P_{1a}.

An interesting situation arises if $M_1 = 1$. We then have

$$P_{1s}(\tau) = \tfrac{1}{2} P_{1a}(\tau) \quad \text{for} \quad M_1 = 1 \tag{19}$$

This relation may be utilized for a reasonable selection of input parameters, in this case, the "psycho-political" parameters D^1, D_1^2 and a_1.

SECTION V

THE DAMAGE OBJECTIVES

The aggressiveness parameters, the damage objectives, and the standards of damage constitute the class of "psycho-political" parameters:

$$\left.\begin{array}{c}\text{Psycho-political}\\ \\ \text{Parameters}\end{array}\right\} = \begin{pmatrix} a_1 & D_1^2 & D^1 \\ a_2 & D_2^1 & D^2 \end{pmatrix} \tag{22}$$

Principally, one may treat them as independent input parameters. That means that they are not subject to the present theory. However, one may, of course, conceive of an auxiliary or extended theory which establishes certain relations between some of the psycho-political parameters. This is the subject of this section.

The objective is to establish reasonable relations between the damage objectives and the standards of damage. It appears reasonable to suspect that relations exist.

Take, for example, the damage objection D_1^2 of the first power. Since the intent of spontaneously inflicting damage is punitive, it follows that D_1^2 should increase if the second power's insensitivity to damage increases. Since the second power's insensitivity is measured by her standard of damage D^2, it follows that

$$D_1^2 \sim D^2 \ . \tag{23}$$

With respect to the first power's insensitivity, the relation should be inverse, that is

$$D_1^2 \sim 1/D^1 \tag{24}$$

It appears therefore reasonable to postulate tne relations

$$D_1^2 = \frac{D^2}{D^1} \tag{25a}$$

$$D_2^1 = \frac{D^1}{D^2} \tag{25b}$$

That these relations make good sense can be seen best by substituting them into expressions (12) and (14):

$$M_1 = \frac{D_{1s}^1 \ D^2}{(D^1)^2 \ D_{1s}^2} \tag{26a}$$

$$M_2 = \frac{D_{2s}^2 \ D^1}{(D^2)^2 \ D_{2s}^1} \tag{26b}$$

These relations are to be used in conjunction with

$$Q_{1d} = \frac{1}{1 + M_1} \quad ; \quad P_{1d} = \frac{M_1}{1 + M_1} \tag{27a}$$

$$Q_{2d} = \frac{1}{1 + M_2} \quad ; \quad P_{2d} = \frac{M_2}{1 + M_2} \tag{27b}$$

If then the attention is focused on the first power and its probability of deterrence and dissuasion, P_{1d}, then the following trends can be seen:

1. As D_{1s}^1 increases, so does P_{1d}. That is, Po_1 becomes increasingly deterred.

2. As D^2 increases, so does P_{1d}. That is, Po_1 is increasingly dissuaded.

3. As D^1 decreases, P_{1d} increases. That is, Po_1 becomes increasingly deterred.

4. As D^2_{1s} decreases, P_{1d} increases. That is, Po_1 becomes increasingly dissuaded.

Interesting is the fact that the denominators of the expressions for M_1 and M_2 contain the squares of the standards of damage. More precisely, the expressions for M_i, P_{id} and Q_{id} contain the square of D^i but not the square of D^j $(j \neq i)$. Hence, when the i th power is faced with the decision to attack spontaneously or to abstain from such an attack, then her decision is more strongly influenced by her own standard of damage than by the other power's standard of damage. Again, this result appears to be quite resonable and desirable.

It should be mentioned in this connection that it is now mandatory that all damages and standards of damage are relative. That is, each total damage or standard of damage which carries the superscript i must be made relative by dividing it by V^i where V^i is the total destructible value of the i th power. So far, this was not important since in all formulae including ORA -70-0012 but excluding the formulae of this section, the V^i would always cancel out. In contrast, relations (26) make sense only if all damages and standards of damage are relative.

The "psycho-political relations" (25) offer the considerable advantage that the number of independent psycho-political parameters is now reduced from six to four.

SECTION VI
THE SPONTANEOUS BELLIGERENCY PARAMETERS

The definition of belligerency parameters is that they are the coefficients of the first terms in the Taylor developments of the time-dependent probabilities:

$$P_{ia}(\tau) = a_i \tau + \cdots \tag{20a}$$

$$P_{is}(\tau) = s_i \tau + \cdots \tag{20b}$$

Here a_i is the already introduced aggressiveness parameter of the i th power (i = 1, 2), and s_i is the spontaneous belligerency parameter of the i th power (i = 1, 2). For more details, see the referenced ORA report.

If relations (16) are applied to the case $\tau \to o$ and if relations (20) are heeded, then we obtain

$$s_1 = \frac{a_1}{1 + \dfrac{D^1_{1s} \; D^2_1}{D^1 \; D^2_{1s}}} \tag{21a}$$

$$s_2 = \frac{a_2}{1 + \dfrac{D^2_{2s} \; D^1_2}{D^2 \; D^1_{2s}}} \tag{21b}$$

This is the representation wanted for the spontaneous belligerency parameters. They replace relations (VI. 3a) and (VI. 3b) of the report ORA-70-0012. Formulae (VI. 4), (VI. 5), (VI. 6), and (VI. 7) of that report are obsolete, and subsequent formulae have to be modified according to formulae (21) of the present report.

SECTION VII

FURTHER REMARKS ON THE PROBABILITIES OF MOTIVATION

For an efficient discussion of this subject, it is appropriate to repeat here the following definitions (which are otherwise scattered over several reports):

$1S \equiv$ Proposition that Po_1 will attack spontaneously

$1DD \equiv$ Proposition that Po_1 is either deterred or dissuaded or both

$\sim 1DD \equiv$ Proposition that Po_1 is neither deterred nor dissuaded

$$P_{1s} = \text{Probability}\ \{ 1S \} \tag{22}$$

$$P_{1d} = \text{Probability}\ \{ 1DD \} \tag{23}$$

$$Q_{1d} = \text{Probability}\ \{ \sim 1DD \} \tag{24}$$

$$P_{1a} = \text{Probability}\ \{ 1S \mid \sim 1DD \} \tag{25}$$

Applying the product rule in the form

$$P\{A \wedge B\} = P\{A\}\ P\{B \mid A\} \tag{26}$$

we may write

$$P\{1S \wedge \sim 1DD\} = P\{\sim 1DD\}\ P\{1S \mid \sim 1DD\} \tag{27}$$

Since

$$1S \supset \sim 1DD \tag{28}$$

it follows that

$$P\{1S \wedge \sim 1DD\} = P\{1S\} \tag{29}$$

and therefore

$$P_{1s} = Q_{1d}\ P_{1a} \tag{30}$$

We may then introduce the following definitions:

$$a_1 = s_1^* = s_1 \, (Q_{1d} = 1) \tag{31}$$

$$b_1^* = b_1 \, (Q_{1d} = 1) \tag{32}$$

$$b_2^* = b_2 \, (Q_{1d} = 1) \tag{33}$$

In other words, the symbols carrying asterisks are the respective belligerency parameters subject to the condition $Q_{1d} = 1$ which is "Po_1 is neither deterred nor dissuaded." We may then write

$$P_{1a} = P_{1s}^* \tag{34}$$

If we then heed (see ORA 70-0012)

$$P_{1s} = \frac{s_1}{b_1 + b_2} \left\{ 1 - e^{-(b_1 + b_2)\tau} \right\} \tag{35}$$

we have

$$P_{1a} = \frac{a_1}{b_1^* + b_2^*} \left\{ 1 - e^{-(b_1^* + b_2^*)\tau} \right\} \tag{36}$$

For $\tau \to 0$, we obtain from (30), (35), and (36)

$$s_1 = Q_{1d} \, a_1 \tag{37}$$

The $P_{1a}(\tau)$ and $P_{2a}(\tau)$ are genuine probabilities, however, they play here mainly an auxilliary role and are, for applications, not of particular interest. In contrast, the relations

$$s_i = Q_{id} \, a_i \qquad (i = 1, 2) \tag{38}$$

are essential for the entire probabilistic theory of nuclear stability.

SECTION VIII

CONCLUSION

The improvement of the "Theory of Deterrence and Dissuasion" presented in this report over the "Rationale of Spontaneous Belligerency" presented in ORA-70-0012 is as follows. According to the formulae of the latter report, it is, in certain cases, possible that $P_{is}(\tau)$ is exactly zero even though $P_{ia}(\tau)$ is not. Clearly, this is an "overstatement" If $P_{ia}(\tau)$ is not zero, then nobody can say with certainty that Po_i will not attack spontaneously. Such a proposition would be "illegitimate" in the sense of inductive logic. But $P_{is}(\tau) = o$ precisely amounts to such an illegitimate proposition. This is what is meant by "overstatement".

The present theory avoids such overstatements. If $P_{ia}(\tau)$ is not zero, then $P_{is}(\tau)$ is also different from zero, except in the limiting case $M_i \to \infty$, which is quite reasonable and, in fact, necessary.

Hence, the present theory is much more compatible with the principles of inductive logic. It is certain that it is not completely compatible with those principles. But that is simply asking for too much.

Nevertheless, there is still much room for improvement. Probably the most important and most urgently needed improvement is to develop the present theory into a "theory of escalation", as has already been intimated in the introduction. In this case, the spontaneously striking power would strike with only part of her force, holding the other part in reserve as "deterrence against excessive retaliation" by the other power. The other power, in turn, would generally retaliate also with only part of her force, holding the other part in reserve as "deterrence against excessive counter-retaliation". This process may continue through several "stages of escalation" until it is either "stopped" or it ends in a full fledged nuclear catastrophe. In this context, the analysis, the quantification and the understanding of the mechanism of the "escalation stopping power" seems to be one of the most rewarding and most urgently needed endeavors. No doubt that it is extremely difficult, but no doubt, either, that there are not many opportunities for analysts and theoreticians where so little effort could accomplish so much.