

A laboratory experiment in retaliation¹

MORRIS F. FRIEDEL

Department of Sociology, University of Michigan

Introduction

This paper summarizes an investigation of the effects of psychological and situation variables in an experimental game study of retaliation which was designed to simulate aspects of accidental nuclear war. Additional detail may be found in Friedell (1965).

The study of retaliation is of prime importance in the context of accidental war. This is because an accidental war (like any other) can be thought of as the result of *decisions*—probably the result of an escalating sequence of decisions to retaliate. It is only the initial stimulus, such as an unauthorized nuclear detonation, that is an “accident.” Of course, findings on individual decision-making with regard to retaliation in the laboratory, reported here, are not directly applicable to nations; nonetheless they may supply background for future work.

Retaliation—aggression instigated by aggression—has an obvious prominence in

interpersonal conflict. But frustration (a stimulus different from attack) and non-cooperation (a response different from counter-attack) have attracted most of the relevant research. Hence the present study is somewhat exploratory.

An experimental game with options to attack and to retaliate was presented to 196 college students. A $2^3 \times 3$ factorial design explored the effects of accidental attack, expected iteration, social influence, and timing conditions upon the incidence of retaliation. Measures of authoritarianism, orientation to military-political strategy, and attitudes relating to nuclear war were obtained. Statistically significant findings ($p \leq .05$) were that expected iteration and social influence inhibited retaliation while authoritarianism fostered it. The other four variables did not have significant effects. The most important finding was that the possibility of accidental attack did not appreciably inhibit retaliation.

Hypotheses

The principal independent variables were selected partly on the basis of plausible relevance to nuclear decision-making. For-

¹I wish to thank Jack Sawyer and Joe Harding for their extensive help, and William Gamson and Michael Inbar for their thoughtful comments.

mulation of hypotheses stemmed from the underlying perspectives that retaliation was more likely when (a) it could be considered as deserved, (b) the decision-maker was tense and lacking an alternative channel of discharge of tension (cf. Thibaut and Coules, 1952), and (c) the decision-maker had a "rightist" political orientation.

Situational hypotheses were:

(1) The "accident" hypothesis—the probability of retaliation decreases if the attack is thought possibly unintentional.

(2) The timing hypothesis—the probability of retaliation increases if it is known that there is only a short period after the attack during which retaliation is possible; conversely, the probability decreases if there is an enforced delay before the subject can retaliate.

(3) The expected iteration hypothesis—expectation of continuing interaction after retaliation affects the probability of that retaliation. Here the expected direction of effect is unclear. Retaliation may be fostered by the thought that it might deter future attack; it may be inhibited by the thought that it might provoke future attack.

(4) The social hypothesis—a person will be less likely to retaliate if another person shares his situation, in particular sharing all gains or losses, but can independently decide to retaliate. Experimentation on this hypothesis is related obliquely to classic social-psychological research on the individual and the group. It differs both from together-and-apart experiments and from group decision-making experiments in that the outcome for all individuals is determined by the decision of some single member (who yet cannot be considered an autocratic leader). Such a social situation, rare in everyday life, may be of importance for the development of social-psychological theory. It is illustrated by a military situa-

tion in which any private's firing can involve his platoon in battle.

The principal personality hypothesis was that authoritarianism is positively correlated with retaliation. "Absolutist" rather than "pragmatic" politico-military strategic perspective (the SP scale; cf. Janowitz, 1960) and assent to clichés such as "Better dead than Red" (the NPNS scale) were also predicted to encourage retaliation.

Method

THE BASIC EXPERIMENTAL SITUATION

In order to study in the laboratory factors affecting retaliation, it is necessary to create a situation in which subjects have the options of retaliating or declining to retaliate (declining, for short). Since retaliation may be defined as "return of evil for evil," the situation must include culpable attack. It may be assumed that the victim will hold the attacker blameworthy if he believes the attacker knows that the victim will lose more than the attacker gains. There must be a plausible motive for the attack—we do not consider lunatics clearly evil nor do we retaliate vigorously against them. On the other hand, it is desirable that there be no motive in the absence of attack for the behavior representing retaliation. Otherwise, it would possess an instrumental value as well as that of requital. Nuclear retaliation against cities may exemplify such non-instrumental retaliation.

Most previous laboratory experimentation in retaliation has involved verbal attack and response (Buss, 1961, ch. 3). There are several methodological difficulties in the use of verbal aggression. Responses are hard to scale and attacks are difficult to replicate exactly. If the experimenter provides the attack, his role as an authority figure may inhibit retaliation. The use of an accomplice obviates this difficulty but

TABLE 1
PAYOFF MATRIX

	B (column)	
	Not retaliate	Retaliate if attacked
Not attack A (row)	\$1.00 = a_{11} , \$1.00 = b_{11}	\$1.00 = a_{12} , \$1.00 = b_{12}
Attack	\$1.50 = a_{21} , \$0.15 = b_{21}	\$0.15 = a_{22} , \$0.10 = b_{22}

is expensive. And it might be hard to make verbal aggression appear accidental. The experimental game method, in which subjects deprive each other of money rather than insult each other, does not pose those problems and appears to be a superior approach.

The game employed is perfectly symmetric, but let us assume for simplicity of exposition that player A, the row-chooser, can only attack, and player B, the column-chooser, can only retaliate. The payoff matrix is presented in Table 1. The following considerations partially determined the payoff values:

(1) It was required that $a_{21} > a_{11}$ (A is motivated to attack).

(2) $b_{11} - b_{21} > a_{21} - a_{11}$ (B's loss from the attack is greater than A's gain).

(3) $b_{22} < b_{21}$ (retaliation costs B something).

Corresponding to several scenarios for accidental nuclear war:

(4) $a_{22} < a_{11}$ (B's retaliation causes A to experience a net loss).

Finally, following a commonly used model for the international situation and most previous experimental game research, the players should be in symmetrical roles. Hence:

(5) $a_{11} = b_{11}$, and either player can attack or retaliate.

The payoffs were to be fixed, if possible, so that attack would occur with probability 1.0—to maximize the opportunity for re-

taliation—and retaliation would occur with probability .5—to maximize variance. The payoff values were of some concern; it was feared that either there would be few retaliators among those who had the opportunity or that potential attackers would think retaliation highly probable and be deterred. But the above matrix performed fairly well by these criteria.

Note the following characteristics of this matrix:

(1) Attack alone reduces the total wealth of the subjects by 17.5 percent.

(2) After attack, the attacker has ten times as much as the victim.

(3) After retaliation, the attacker still has five cents more than, or one and one-half times as much as, the victim. To strike first continues to provide a *relative* advantage even after the victim has retaliated.

(4) After retaliation, the total wealth of the subjects is only 12.5 percent of its original value.

It is seen that from the utilitarian perspective of the total wealth of the players the retaliatory attack is far more severe than the initial attack. Furthermore, while retaliation reduces the difference between the players it does not entirely even the score. This contrasts with competition in the Prisoners' Dilemma, which can profit the competitor.

The conditions under which subjects were to play this game had to be prescribed. Subjects were told that they were participat-

ing in an experiment on decision-making. Because of its frivolous connotations the word "game" was never used. Neither was any reference made to war, since it was feared that this might introduce symbolic behavior as an uncontrolled and unmeasured variable.²

In order to simulate accidental war, the subjects were allowed to communicate, as heads of state do. Only communication by the laborious procedure of writing messages was allowed, and subjects were isolated from face-to-face contact during the experiment. The self-recording character of the written messages was an additional advantage of this technique. The international analogue is the "hot line," which is a teletype system.

Simulating an ongoing situation by one which must begin at a fixed time is always a problem. Here, the exact time at which the game began was unknown to the subjects because there was no starting signal. They were told that the experiment would begin within a minute after a switch was closed and that earlier their buttons (see below) would be inoperative. This indeterminacy was introduced in order to minimize incentive for a race to preempt—the subjects might have structured the situation as a contest in speed if there had been a starting signal. Except in the time-pressure treatments a trial was ended by a bell-signal from a timer (which the subjects could see) after ten minutes, but a subject was always given at least five minutes to retaliate, lest an attack be delayed until just before the bell in order not to give the victim a chance to respond.

The subjects acted by pushing buttons:

² A few subjects had had a course which enabled them to recognize the game character of the experiment, but I saw very little awareness of any reference to war.

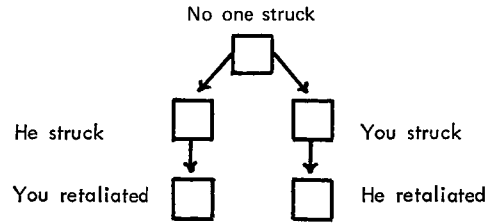


FIG. 1. Schematic view of apparatus box.

if a subject's button was pushed first it meant attack (a "strike"), if second it meant retaliation. What had happened was indicated to a subject by which one of five lights was lit on an apparatus box which separated the subjects. The arrangement of the lights is indicated schematically in Figure 1.³ The arrows show possible transitions. A trial could terminate in any of the five positions. Only one strike and one retaliation were allowed. The electrical apparatus also provided a convenient way to produce accidents, and perhaps a curiosity motive for attack.

In short, retaliation was operationalized as pushing the button second, where the retaliator knew that this would cost him 5 cents and cost the other \$1.35. The average subject indicated that he would work over an hour to earn \$1.35; this appeared to be long enough to make the loss of this amount an "evil." Conceivably the victim is rich and does not care, but the retaliator has no way of knowing this.

Support for the proposition that pushing the button second is a valid instance of retaliation also comes from retaliators' responses to the question, "In your decision to push your button second, how important were these factors?" Eighty-two percent gave responses definitely indicating requital.

³ Actually, the labels for the lights indicated who had pushed and repeated the payoff scheme; the words "struck" and "retaliated" were not used.

SITUATIONAL VARIATIONS AND
PERSONALITY MEASURES

The hypotheses were tested by varying the basic experimental situation and by use of questionnaires.

In "accident" (A) treatments subjects were read the following⁴ in their instructions:

There is a small chance that the lights will show that one of you pushed your button *first* when it didn't really happen. In that case one person will gain and the other lose just as if a person had actually pushed his button. And the person who lost can push second.

The only lie subjects were told was about the "small chance"—actually, an accident would usually occur if no subject attacked, about three minutes after the starting switch was operated. (Whether an accident occurred or not, "accident" treatments were distinguished from others by introducing the *possibility* of accident.) There was frequently considerable doubt as to whether or not an accident had occurred.⁵

In the time-pressure (P) treatments subjects were told:

The experiment ends when the bell rings, about 10 minutes after it starts, *except* that you will have always and only about 25 *seconds* to decide whether to push your button second. Then the lights on the box will go out.

This time limit was designed to put considerable pressure on the subject while yet allowing him ample opportunity to act.

In the enforced-delay or "wait" (W) treatments the instructions were:

⁴ The instructions were modified slightly during the experiment. This and other minor irregularities are discussed in Friedell (1965).

⁵ After the trial, subjects in A treatments who were victims of strikes or accidents were asked to indicate on a 5-point scale their responses to: "Do you think he really pushed his button?" Answers were: No = 1, 12 Ss; 2, 2 Ss; 3, 5 Ss; 4, 5 Ss; 5 = Yes, 4 Ss.

The experiment ends when the bell rings, about 10 minutes after it starts, except that you will never have less than 5 minutes to decide whether to push your button second. However, you must wait about 2 *minutes* after the other person has pushed his button *in order to press your button second*.

In the "free" (F) treatments, where neither time-pressure nor enforced-delay obtained, subjects were told only the first sentence in the foregoing instructions.

In the expected iteration (I) treatments, subjects were told: "The experiment will be done over again *once*, with another dollar [as stake]." (The trial was actually iterated to ascertain what effect their experience would have.) Otherwise, the instruction was: "*The experiment will not be repeated.*"

The instructions for the "social" (D) treatments, with a dyad on each side of the apparatus box, had to be somewhat more complicated. Before receiving the general instructions, the subjects were told:

First, would each of you pretend that you are the only person on your side and that there is only one person on the other side, while I read the instructions.

This permitted the tetrads in social treatments to be given the same instructions as others, insofar as possible. After receiving all the instructions for the corresponding nonsocial treatment, they were told:

Now, there are actually two of you on each side. Either person on a side can press his side's button, in which case he gains or loses just as if he were the *only* person on his side, and each person on the *other* side loses just as if he were the only person on that side. It *does not matter* which person on a side actually presses the button—*both persons on a side always gain or lose equally*.

If you wish to press the button, the other person on your side does *not* have to agree. You can't talk with him, but you can exchange messages. If you agree with a message he sends to the other side, sign your initials to it.

Each member of a tetrad began with his own \$1.00 stake. "Partners" were told to communicate in writing to provide a record of their interaction; this also functioned to prevent the other side from hearing them. (They seemed generally to conceive of themselves as a team.)

The attitude items, and some others to be discussed below, were administered in a mail questionnaire sent to the subjects from a few days to about a month after they had participated in the experiment. The delay was intended to obviate a possible halo effect from the trial. The return rate was very high (92 percent).

EXPERIMENTAL DESIGN AND PROCEDURE

The twenty-four possible treatments obtainable from the situational variations were to be presented in a $2^3 \times 3$ factorial design with four replications of nonsocial treatments and two replications of social treatments in order to maximize information obtained. For various reasons this design was not perfectly implemented, but the number of replications of each treatment approximates that given.

The subjects met the experimenter (the author) in a large room in which the experiment was to be conducted or at the door to the building at a time prearranged by telephone. They were seated on opposite sides of a table holding the apparatus box. (Two were on each side in social treatments.) In social treatments where a pair of subjects knew each other, this pair was assigned to the same side. Subjects were then asked their names, and (if this could not be ascertained from observation) whether they knew each other. Trials were excluded where Ss on opposite sides were well acquainted.

In front of each subject were placed \$1.00 in change and a colored pen and

colored paper for writing messages. Each side had a spindle on which to place messages, a reminder card succinctly stating the payoff scheme, and a pushbutton.

The subjects were then read the following general instructions:

Thank you for coming to take part in this decision experiment. Each of you has a pushbutton; we want to know whether or not you push the button. I will tell you what the buttons do. It is complicated, but it is important for you to completely understand. So feel free to ask questions at any time.

Each of you has \$1.00 to start with. It is on the table in front of you. How much you take home depends on what you do with the buttons. You can push them whenever you want, or not at all.

You will not know *just* when the experiment starts, but it will start a little less than a minute after I operate the switch. *Before that the buttons will not do anything.* So, if you press your button right after I operate the switch, nothing will happen.

After the experiment has started, the person who pushes his button first *gets an extra 50 cents*, and the other person *loses 85 cents*. The person who pushes his button first then has \$1.50, and the other is down to 15 cents.

You can push your button second, if the other person has pushed his button first. Then, *you lose another 5 cents*, but he loses \$1.35. *So you have 10 cents left*, and he has 15 cents left.

Remember that pushing your button first gives you 50 cents, and takes 85 cents away from the other person. Pushing your button second costs you 5 cents and takes \$1.35 away from the other person.

The lights on the box show what has happened. The white light will be on to start with because neither of you have pressed your buttons. The arrows show what *might* happen next.

Now I am going to ask each of you some questions to make sure you understand what the buttons do.

Suppose you wait for a minute and push your button before the other person pushes his. How much do you get?

If you push your button right away when I operate the switch, what will happen?

Suppose the other person pushed his button

and made you lose 85 cents. Then if you push your button, how much will it cost you?

Suppose the other person pushed his button and made you lose 85 cents. Then, if you push your button, how much will it cost him?

Here are some details to remember:

You are not allowed to talk, except to me, but you can send messages. Use the pen and paper in front of you. Give each message to me and I will give it to the other person. Put all messages you receive on your spindle.

[Then the timing, iteration, and accident instructions were given, according to the treatment. Finally, subjects were told:]

One more detail—if the other person has just pushed his button, you have to wait a few seconds before yours will do anything.⁶

Remember:

At no time are you required to push the button.

All amounts gained or lost come from or go to the experimenter, not the other person.

There is a small chance that the lights will show that one of you pushed first when he didn't. It still counts as if he had.⁷

[Final instructions to tetrads were inserted here.]

Are there any questions?

Now I will turn the switch and we will begin.

After questions were answered the trial began. During the experiment, button-pushing and its timing were noted, as well as other behaviors of particular interest. After the completion of a trial, the subjects were given a postinterview asking the reasons for their behaviors, whether they had heard anything about the experiment, their perception of the experimenter's bias, their image of the situation, their surprise at events in the game, whether they thought an attack was accidental (in A treatments) and what their behavior would have been in certain contingencies. In I treatments each subject was given another dollar before the repetition, and there was another post-

interview. Finally, the subjects were told not to talk about the experiment, asked to fill out and return the mail questionnaire (measuring the personality variables, as well as demographic characteristics and some miscellaneous items), and dismissed.

SUBJECTS

The design called for 192 subjects. (Each subject was used only once.) There was no feasible way to make participation compulsory, so it was necessary to obtain a large, fairly homogenous population from which to recruit volunteers. The population of male undergraduates at Northwestern University satisfied this requirement, and subjects were recruited and the experiment conducted there. Approximately 10 to 15 percent of the subjects were obtained from announcements made in political science classes; the rest were obtained from notices in dormitories. Both groups were told that it would be a laboratory decision-making experiment, and the importance of these experiments was sketched. They were told the experiment would last from 30 to 45 minutes, and that they could gain up to \$3.00 for participating. The political science students were also told that this would be an opportunity to gain first-hand experience with this research method. It is estimated that about 20 percent of the population reached by the announcements volunteered. Subjects were recruited and run in the spring and summer of 1964.

What were the 196 participants like? The mail questionnaire revealed a modal age of 19, approximately equal proportions of liberal Democrats, liberal Republicans, and conservative Republicans, and a slight average tendency to disagree with the 14 F-scale items used (mean: $-.45$; s.d. : $.88$). Fifty-three percent listed money as a primary reason for volunteering; 36 percent

⁶ In F and P treatments, in order to avoid counting an intended attack as retaliation.

⁷ In A treatments.

TABLE 2
MARGINALS: PROPORTIONS OF RETALIATING TRIALS*

		A'	A	I'	I	F (Free)	P (Pressure)	W (Delay)	Combined
Nontetrad	D'	.50 (12)	.43 (23)	.56 (18)	.35 (17)	.58 (12)	.36 (11)	.42 (12)	.46 (35)
Tetrad	D	.20 (10)	.50 (8)	.44 (9)	.22 (9)	.50 (6)	.29 (7)	.20 (5)	.33 (18)
Nonaccident	A'			.41 (12)	.30 (10)	.71 (7)	.13 (8)	.29 (7)	.36 (22)
Accident	A			.60 (15)	.31 (16)	.45 (11)	.50 (10)	.40 (10)	.45 (31)
Noniterated	I'					.56 (9)	.50 (8)	.50 (10)	.52 (27)
Iterated	I					.56 (9)	.20 (10)	.14 (7)	.31 (26)
	Combined					.56 (18)	.33 (18)	.35 (17)	.415 (53)

* The number of trials with opportunities to retaliate is in parentheses.

listed curiosity or the interest of the experiment as primary, with money perhaps as a secondary reason. Surprisingly, 62 percent of the respondents said that they would work an hour or longer to earn a dollar.

Findings

TESTING THE HYPOTHESES

The effect of the principal independent variables upon retaliation will be examined by comparing two disjoint subgroups of the subject population—retaliators and decliners.⁸ Numbers of retaliating and declining trials will also be compared. The distinction between subjects and trials is made because, in social treatments, a declining trial produced two decliners, while in non-social treatments a declining trial produced one decliner; only persons who, themselves, pushed second are counted as retaliators. There were 22 retaliators, 31 declining trials and 43 decliners.

The Data. The most important data are shown in Tables 2 and 3.

Table 2 shows that the accident treatment had an effect opposite to that expected; expected iteration inhibited retaliation; along with enforced-delay, time-pressure inhibited retaliation; social treatments inhibited individual retaliation. Table 3 shows that the F-score had a relatively strong positive

correlation with retaliation; the effect of strategic perspective was weaker; NPNS seemed actually to inhibit retaliation.

Statistical Analysis. Before interpreting these results it is necessary to ascertain whether the effects of one variable were suppressed or enhanced by those of another and to test for statistical significance.

A powerful statistical procedure which is unequivocally applicable to the present hypotheses and data has not yet been developed. Therefore, a variety of techniques were used.

First, the personality variables will be examined. If it is assumed that accident can never increase retaliation and that NPNS and SP can never decrease it, regardless of the values of the other variables, one can infer that NPNS had no independent effect on retaliation. When F-score is partialled out, retaliation correlates only .042 with accident and .050 with strategic perspective.

As would be expected from the experimental design, little correlation appeared between F-score and assignment of subjects to different timing, iteration, and social conditions. The correlation $r(38,71) = .339$ may now be tested for significance by the z-test ($z = r(N-1)^{1/2}$), which yields $z = 2.65$; a t-test on the difference between F-scale means (.587) yields $t_{60df} = 2.75$. Either test indicates significance at better than the .005 level, one-tailed.

⁸ Behavior during the repetition in I treatments will not be discussed except in "Miscellaneous Findings" below.

TABLE 3
CORRELATIONS AMONG PRINCIPAL VARIABLES FOR RETALIATORS AND DECLINERS

	-76*	-42	-43	44	71	72	74	38
-76* (Social)	—	-.257	.047	.078	-.084	-.136	.040	-.271
-42 (Accident)	-.257	—	.076	-.040	.303	.163	.243	.140
-43 (Iteration)	.047	.076	—	.112	.103	.176	.035	-.206
44 (Time Pr.)	.078	-.040	.112	—	-.147	-.068	-.121	-.027
71 (F-scale)	-.084	.303	.103	-.147	—	.310	.367	.339
72 (SP)	-.136	.163	.176	-.068	.310	—	.274	.153
74 (NPNS)	.040	.243	.035	-.121	.367	.274	—	-.132
38 (Retaliation)	-.271	.140	-.206	-.027	.339	.153	-.132	—

Sample size = 57-65.

* Var. -76 = 1 in D treatments, = 0 in D' treatments

-42 = 1 in A treatments, = 0 in A' treatments

-43 = 1 in I treatments, = 0 in I' treatments

44 = 2 in P treatments, = 1 in F treatments, = 0 in W treatments

38 = 1 for retaliators, = 0 for decliners.

The variables are numbered as in Friedell (1965) to facilitate reference.

Now the treatment variables will be considered. When the 12 signs of differences in proportions retaliating in treatments differing only on the I variable are compared, expected iteration is found significantly to inhibit retaliation ($p = .04$, two-tailed).

With regard to the social variable, the hypothesis of particular interest is not that of absence of social effect on probability of retaliation per trial, but rather that of no effect on probability of retaliation per subject. Since two subjects share the opportunity to retaliate in each D trial considered here, this null hypothesis states that x , the probability of retaliation in a D' trial, corresponds to the probability $1-(1-x)^2$ of retaliation in a comparable D trial.⁹ It appears reasonable to pool all the I and all the I' observations in testing the social hypothesis. Then sample sizes are large enough so that χ^2_y (chi-square with Yates' correction) may be computed, with expected observations determined by the maximum likelihood probabilities under the null hypothesis; the χ^2_y values from the I and the I' observations can be added to give an overall χ^2_y with two degrees of free-

dom. The value obtained is $\chi^2_y = 6.00$ ($p = .05$).

Discussion. It was found that authoritarianism increased the propensity to retaliate, that partnership inhibited individual retaliation,¹⁰ and that expected iteration decreased retaliation. None of these results is surprising. Nor is it surprising, in retrospect, that the P effect was opposite to that predicted (although it must be noted that, when retaliation occurred in F treatments, it usually occurred before a P trial would have terminated). The experimenter's informal observation supported the alternative hypothesis that the subjects were somewhat immobilized, though probably more because they were trying to digest what was going on than because of felt pressure to make a decision. It should be noted that they were distracted by the experimenter passing money back and forth, although this occurred in all timing conditions. As for NPNS, and the weak and nonindependent effect of SP: as will be discussed below, it does not seem that retaliators are "rightist." It will be argued that this is nevertheless consistent with their high F-scores. Compare observations by McClintock *et al.*,

⁹ In two of the 18 D trials in Table 2, those able to retaliate knew each other; retaliation occurred in one of them. These trials are pooled with the others in the analysis.

¹⁰ Compare a somewhat similar finding in Pylyshyn *et al.* (1966).

using an asymmetric game, which suggested "that Internationalists may be more responsive to variations in their opponents' strategy. *They tended, more than Isolationists, to punish a previously competitive opponent and to reward a previously cooperative one*" (Gallo and McClintock, 1965; italics mine).

The lack of effect of the A treatment is the most unexpected finding.

OTHER FINDINGS

Communication and Retaliation. Messages and their absence play a vital part in the process leading to retaliation or declining (cf. Deutsch, 1958). In A treatments there can be communications about whether or not the attack was intentional. Messages can convey a very reasonable alternative to retaliation, an offer to split 50-50. (This would involve the striker giving 67.5 cents to the victim, and indeed most offers to split suggested amounts close to this. There seemed to be little concern over odd pennies or nickels.) A less adequate alternative to retaliation that was occasionally suggested was "alternation"—that the victim push first next time. This occurred when subjects did not realize it would be best not to push at all.

In A treatments an accident occurred fourteen times and a strike occurred seventeen times. No subject lied about the occurrence of strike. In ten of the fourteen trials in which an accident occurred, the profitters sent spontaneous messages to the effect that they did not push the button. In two cases, the victims asked whether the other side pushed the button. In one of the above cases both occurred simultaneously. In these eleven cases the assertion of accident tended to be believed. In the three cases where an accident occurred and there was no discussion of this sort, there was more

doubt. When a strike occurred there was only one case in which the possibility of accident was discussed—it was admitted that a strike had happened, but the victim declined. Nobody asked if it was an accident, received the answer "No," and then retaliated. There was no meaningful correlation between retaliation and the belief that a strike had occurred.

It was thought that the passage of messages between sides might reduce retaliation. This effect existed but was not statistically significant. The occurrence of an accident did stimulate messages.

There were nineteen cases in which a suggestion about splitting or alternating was received before the end of a trial. Figure 2 shows who initiated discussion of an arrangement and what the outcome of the trial was. Once discussion was initiated, retaliation rarely ensued. The three exceptions occurred in tetrad trials. In one there were no messages to give a clue to the reasons for rejection; in fact, it is possible that retaliation occurred before the message was received. In another case the striking side rejected a demand for alternation; in the third the victims rejected an unusually low offer of 15 cents each by the striking side.

To conclude: the occurrence, rather than the mere possibility, of accident did tend to depress retaliation. (Ratios on a per trial basis were .36 when accident occurred and .50 otherwise, excluding the three trials covered in note b to Figure 2.) And when retaliation occurred after accident, the process tended to involve the victim retaliating before he got a message that it was an accident, let alone a suggestion for a deal. Further, the victim of an accident suggested a deal only once, although in cases where strike occurred the victim often did so. This is probably due to the greater

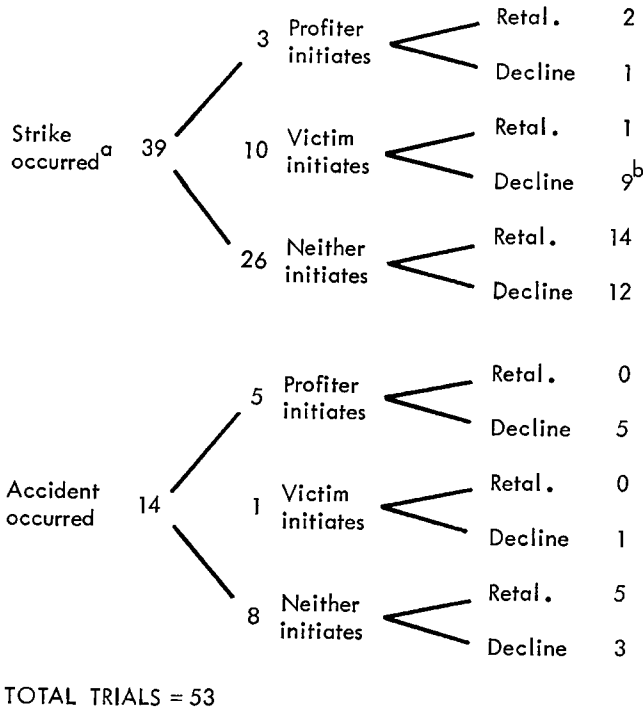


FIG. 2. Initiators and outcomes of arrangements. Notes: (a) Includes both A and A' treatments. (b) Three trials are included where the strike was suggested by the victim as part of an arrangement to alternate. In no trial did a subject suggest that he himself strike first as part of such an arrangement.

readiness of profilers from accident to initiate negotiations.

Miscellaneous Findings. Surprise by the attack was correlated with retaliation ($r = .284, N = 65$).

Retaliators ranked about equally the following explanations: "It wouldn't be right for him to go away with \$1.50;" "showing him that he couldn't get away with that kind of behavior;" and "not wanting to sit and do nothing." Next came anger. Curiosity and fulfillment of commitment ranked lowest as reasons for retaliation. In I trials, deterrence was the most prominent reason. Actually, retaliation tended to be somewhat provocative.

Subjects, such as the strikers, who were

not confronted with the decision whether to retaliate were asked what they would have done. There was little similarity ($r = .153$) between correlations with retaliation and with "would-be" retaliation.

The absence of similarity between would-be retaliation and real retaliation casts doubt on the applicability of findings such as those in Deutsch (1960) to makers of real decisions. This dissimilarity may be relevant to the absence of correlation between retaliation and SP and NPNS, which also do not deal with overt aggressive behavior.

Portrait of the Retaliator. At the beginning of this investigation it was thought that retaliators tended to have impulsive, aggressive personalities. Thus the prediction that

their hostility against outgroups (in which they would presumably place the strikers) would result in a high F-score. It was thought that, since retaliation on the personal and the national level probably shared some common roots, high NPNS and SP scores should result. Since most of these propositions were disconfirmed, it is clear that the personality of the retaliator must be reexamined. The data obtained will suggest a more accurate and refined ideal type for the retaliator, although one that is, of course, highly tentative.

The revised picture of the retaliator, in contrast to the decliner, is that of a person with a belief in force, a tendency toward moralism, and a tendency toward interactive isolation from others. He tends to believe that:

An insult to our honor should always be punished.

America is getting so far away from the true American way of life that force may be necessary to restore it.

People can be divided into two distinct classes: the weak and the strong.

Most pointedly, he does not consider non-violent demonstrations to promote racial equality desirable, but he *does* feel that racial desegregation should be accomplished throughout the US by all means in the power of the federal government, including force if necessary.¹¹ (The common factor is obvious in retrospect, although entirely unanticipated.) This belief in force must not be equated with destructiveness and cynicism. The retaliator would deny, more than the decliner, the proposition that, human nature being what it is, there will always be war and conflict. Nor should it be equated with a rightist political atti-

¹¹ Fifty-nine percent of the retaliators, in contrast to 30 percent of the decliners and 31 percent of all subjects who responded to the question on segregation, circled this response.

tude, as is shown by the results on that question.

The retaliator's moralism is shown by his opposition to people who do not love their parents and, most strikingly, by his conception of the situation as a problem in trust and even as a problem in cooperation, in contrast to a business or bargaining problem. He was surprised when the other side pushed first. While anger was important in his decision to retaliate, even more important was the feeling that it wouldn't be *right* for the other person to go away with \$1.50.

While the retaliator takes a somewhat personal approach to the situation, in the sense that "revenge" is more personal than "business," he would tend to use rather abstract terms in characterizing the problem it poses. He is more likely to view the situation as either "cooperation" or "competition" than in the more concrete interactive fashion suggested by the concept of "bargaining." He does not see the situation as a problem in "human relations" and he does not initiate suggestions about splitting.

The moralism, emphasis on force, and deemphasis of interaction, indeed, are aspects of the authoritarian personality, but the retaliator in this study does not appear to be cynical or projective. (If projective, he presumably would not be surprised at the strike.)

Recall that retaliators were a self-selected pacific group in that they usually chose not to attempt a first strike. Compared to strikers, they viewed the situation as one of trust and cooperation. (Reaction to perceived betrayal of trust, indeed, may be a major factor in retaliation.) This kind of self-selection is of social importance, however; perhaps when a potential retaliator is found in a peaceful social situation he usually is not also the equivalent of a striker

—otherwise the situation would not often be peaceful. This apparent contradiction describes, of course, the national posture of the US: pledged to retaliate, but not to strike first.

The Striker. Other variables generally did not differentiate between those who struck and those who kept their original dollar ("cooperators"). The 39 strikers, however, tended to view the situation more competitively than others and not to be would-be retaliators. On the other hand, those who cited gaining 50 cents as an important reason for striking were more authoritarian ($r = .430$), more "rightist" in political attitudes, and higher on NPNS than were those who gave this reason less importance. In contrast to the correlations of retaliation with F-scale items related to force, striking to gain 50 cents correlated most strongly with F-scale items related to discipline. These data suggest that strikers are more typical authoritarians than retaliators.

Military Subjects. Nine trials of the experiment were performed with managerial Air Force officers. Only one retaliation occurred, and that one stemmed from restlessness and curiosity. The officers were thus less likely to retaliate than the North-western subjects.

Concluding Discussion

In this final section, findings from the study will be compared with previous research and theory and certain of its implications will be examined.

ACCIDENTAL ATTACK AND PSYCHOLOGICAL THEORY

The finding that the possibility of accident did not appreciably inhibit retaliation may be used to support the frustration-aggression hypothesis—for it is perhaps just as frustrating to be suddenly deprived of

85 cents (for which, recall, the subjects were willing to work a considerable length of time) when the deprivation is accidental as when it is deliberate. However, research has strongly challenged the applicability of that hypothesis to situations like this. There are two related criticisms. First, it has been proposed that attack is an antecedent of aggression, and, indeed, a stronger instigator than frustration. Karl Menninger has stated, "Anyone who has had his toe stepped on, which is certainly not a frustration, knows how inadequate such a formula [the frustration-aggression hypothesis] is" (quoted by Berkowitz, 1962, p. 29). There has been little empirical research comparing attack to frustration as an instigator of aggression, but Buss (1961, p. 32) cites a dissertation by J. Gillespie as providing results that "are clear in indicating that verbal attack is a more potent antecedent of aggression than frustration." However, attack can be conceived as a form of frustration, e.g., "the interruption of an internal response sequence" (Berkowitz, 1962, p. 30). Regardless of whether attack is subsumed under frustration, it would appear from this that more retaliation should be expected in nonaccident than in accident treatments.

The Dollard theory has also been accused of neglecting arbitrariness of frustration as a variable affecting the instigation of aggression. Pastore and others have found that "nonarbitrary" frustrating situations elicit significantly less aggression than when the frustration is arbitrary or unreasonable. Pastore (1952) and Rothaus and Worchel (1960) included unintentional frustrations in their nonarbitrary lists. Thus their findings would suggest that accident depresses retaliation. Three objections to this conclusion can be raised. First, exposure to arbitrary frustration may be more frustrating, blocking more response sequences, than

exposure to nonarbitrary frustration—because, if frustration is less expected, more internal response sequences may be anticipatorily set in motion. This objection clearly applies to one of Pastore's ten arbitrary frustrating situations, and perhaps to more. Berkowitz (1962, pp. 68–72) interprets the findings of other studies to substantiate the hypothesis that expected frustrations instigate less aggression. In the present experiment, surprise was just as great in the accident as in the nonaccident treatments. Thus, insofar as unexpectedness rather than arbitrariness is the variable to which subjects were responding in the Pastore and in the Rothaus and Worchel experiments, their results were not inconsistent with mine.

A second objection is that the Pastore and the Rothaus and Worchel experiments used only questionnaire descriptions of hypothetical situations. The question then is raised whether a social desirability factor biased their responses toward ones that seemed acceptable. Note that in the present study the correlates of would-be retaliation were quite different from those of actual retaliation. A study by Allison and Hunt (1959) is quite interestingly related to the accident finding. Subjects were given hypothetical frustrating situations that were "justified," "unjustified," and "unspecified," and were divided into two groups on the basis of low or high propensity to give socially desirable responses. The aggression scores were in the expected directions. Both groups rated the unspecified situations closer to the unjustified than to the justified, this effect being most pronounced in the low social desirability group. If the responses of the low group are taken as more suggestive of behavioral responses than those of the high group, and accident treatment is taken as "unspecified" while the

nonaccident treatment is "unjustified" frustration, the results approach consistency with mine.

A final objection to Pastore's conclusions is that some of Pastore's and Rothaus and Worchel's nonarbitrary frustrators were, so to speak, unfortunate. Four of Pastore's ten nonarbitrary frustrating situations fell into this category. (E.g., "Your date phones at the last minute and breaks the appointment because she [or he] had suddenly become ill.") Aggression should be inhibited here as an injury to a "loved" object. It is quite likely that aggression toward unfortunate frustrators is often inhibited, because of conscience and social norms, and because aggressive needs are satisfied covertly by contemplating the misfortune. (Aggression toward minority groups appears to be an exception, but consider the myths of the "happy Negro" and the "rich Jew.") Pastore's published data do not permit isolation of the effects of unfortunate frustrators from others. While differences on all his items were in the expected direction, most of the effect may have resulted from those involving unfortunate frustrators. In the present experiment, in both accident and nonaccident treatments, the frustrator profited. Thus, to the extent that misfortune accounted for the variance in Pastore's responses, his study and mine can be reconciled.

On the whole, the foregoing discussion suggests that the A treatment should decrease retaliation, but perhaps not very much; sampling variation may complete the explanation of the data.

IMPLICATIONS FOR EXPERIMENTAL-GAME METHODOLOGY

The present research bears on the following five questions:

(1) How does the subject's utility for money affect his behavior? This is related

to the unanswered question of how well the results obtained with games played for pennies could be replicated in games played for dollars or in situations with "real-life" stakes. "There is at least some evidence that changing the size of, say, monetary rewards (within the limits of what experimenters can afford) does not alter the behavior of subjects significantly" (Snyder, 1963, p. 13). In this study it was found that a scale measuring marginal utility for money as contrasted with utility for time showed slight correlation with retaliation, with strike, and with striking for gain. A moderate though not significant negative correlation was found with striking from fear of loss, which can probably be disregarded. These results support the hypothesis that, within limits, utility for money does not affect game behavior.

(2) How much of game behavior can be explained by differential mathematical ability, a factor that may be irrelevant in many contexts? The number of mathematics courses the present subjects had taken was no more than slightly correlated with any behavior. This is consistent with an impressionistic finding of Scodel's reported by Lutzker (1960).

(3) How much does the fact that subjects frequently are volunteers affect experimental game results? The distinguishing personality characteristics of volunteers vary with what they think they are volunteering for.

Martin and Marcuse (1958) found no significant ($p < .05$) difference on the traits of intelligence, anxiety, ethnocentrism, self-sufficiency, introversion-extroversion, dominance-submission, or sociability between volunteers and nonvolunteers for a learning experiment. However, differences between volunteers and nonvolunteers for hypnosis and personality experiments emerged. The

present decision-making experiment is probably perceived to be more like a learning experiment than like the others if the relevant dimension is that of being emotionally loaded. Thus the cited study suggests that little relevant selective effect occurred from the use of volunteers.

A more direct investigation of the meaning of volunteering is possible. In the mail questionnaire, subjects were asked why they volunteered. Responses were grouped into those giving money as a primary reason and those who mentioned interest or curiosity (some of whom also gave money as a reason). If the fact of volunteering is important, one would expect that the reason for volunteering would also be. Conversely, if the reason for volunteering does not covary with behavior, this tends to suggest that volunteers are not a special group.

There was no correlation between reason for volunteering and retaliation. If strike and striking for gain are considered, there is still little correlation with reason for volunteering. Curiously, the correlation between reason for volunteering and utility for money was .000. Volunteering for money increases the correlation between authoritarianism and retaliation to .422 ($N = 34$) from .147 ($N = 26$), but the data generally support the impression from the Martin and Marcuse study.

(4) The previous topics have dealt with characteristics of the subjects. The remaining two deal with the experimental situation. Is it important whether subjects play with real or imaginary money? The latter situation is generally, although not inevitably, the least expensive for the experimenter. Gallo and McClintock's (1965) results in a Deutsch-type simulated trucking game suggest that, where cooperation is profitable but competition is interesting, subjects will tend to be competitive with low or imagi-

TABLE 4

Would-be retaliator?	Strikers	Cooperators
Yes	26%	41%
Maybe	21%	15%
No	54%	44%
	(N = 39)	(N = 54)

nary payoff scales. This is pertinent to topic (1) above, and also to the distinction between the present would-be retaliators and retaliators. In the two-person trials, retaliation occurred during 46 percent of the opportunities; 46 percent to 57 percent of the subjects are estimated to be potential retaliators.¹² Among strikers and cooperators we have the results shown in Table 4. The figures are comparable with those for actual retaliation but, as indicated above, the causes may be different.

(5) Finally, there is the question of how "demand characteristics" (cf. Orne, 1961)—the subjects' perceptions of the experimenter's desires—affect outcomes. There was a trend toward correspondence between button-pushing and the belief that I, the experimenter, wanted the button pushed. If negative, neutral, and positive responses to the question "Do you think I wanted you to push the button?" are scored -1, 0, and 1, respectively, the means for retaliators, strikers, and others are .27, .13 and .02, respectively. Comparison of within-trial and between-trial responses suggests that differential perceptions of my desires were internally motivated. Would-be retaliation was uncorrelated with the question about what I wanted. Assuming that a trend toward congruence between behaviors and the experimenter's perceived desires is validly reflected here, whether the latter

¹² The discrepancy is explained by the selective process leading to the opportunity to retaliate. See Friedell (1965).

should be controlled for depends on the object of inquiry. In most cases, as long as the stimulus from the experimenter is actually constant, the subjects' perceptions should be allowed to vary freely as an intrinsic correlate of behavior. Of interest is the moralistic response from several of the subjects that I shouldn't have cared whether they pushed the button. Others asked whether the "I" in the question referred to me, although there was no one else to whom it could have referred. It seems that the raising of this question conflicted with a narrow cognitive set about psychological experiments possessed by these students.

Implications for the Problem of Accidental War

After all this speculation, theory, and data on interpersonal retaliation and related topics, the reader may wonder if the results have any relevance for problems of accidental war. The difficulty in generalizing is obvious, but one topic of especial relevance to the relationship between the experimental findings and nuclear war will be discussed.

How are individuals who vary on the measures that have been discussed here differentially selected to major decision-making roles, particularly that of the Presidency? Lasswell (1954, pp. 221-23) suggests that authoritarian personalities are likely to be relegated to minor political roles in both democratic and established, industrialized totalitarian societies. He argues mainly from the rigid compulsiveness of the authoritarian, and his argument is thus more appropriate to the discipline-oriented striker than to the force-oriented retaliator who may be more flexible in responding to others. One could speculate on the moralism characterizing America's political stance,

its refusal self-consciously to initiate aggression, and its relatively blind pugnacity when highly aroused. Such speculation would be entirely inconclusive as evidence that retaliators are selected as leaders by American political processes. The point, rather, is that such a possibility cannot be deprecated by pointing to mechanisms that reject typical authoritarians.

The most statistically significant finding from this research was that retaliators were higher F-scorers, on the average, than decliners. The most important finding, however, if importance is taken to be the product of unexpectedness, degree of verification, and meaningful implications, is probably that introducing the possibility of accident did not inhibit retaliation.

REFERENCES

- ALLISON, J., and D. HUNT. "Social Desirability and the Expression of Aggression under Varying Conditions of Frustration," *Journal of Consulting Psychology*, 23 (1959), 528-32.
- BERKOWITZ, L. *Aggression: A Social Psychological Analysis*. New York: McGraw-Hill, 1962.
- BUSS, A. H. *The Psychology of Aggression*. New York: Wiley, 1961.
- DEUTSCH, M. "Trust and Suspicion," *Journal of Conflict Resolution*, 2, 4 (Dec. 1958), 267-79.
- . "Trust, Trustworthiness and the F Scale," *Journal of Abnormal and Social Psychology*, 61 (1960), 138-40.
- FRIEDEL, M. "A Laboratory Experiment in Retaliation." Unpublished Ph.D. dissertation, University of Chicago, 1965.
- GALLO, P. S., JR., and C. G. McCLINTOCK. "Cooperative and Competitive Behavior in Mixed-Motive Games," *Journal of Conflict Resolution*, 9, 1 (March 1965), 68-78.
- JANOWITZ, M. *The Professional Soldier: A Social and Political Portrait*. Glencoe: Free Press, 1960.
- LASSWELL, H. D. "The Selective Effect of Personality on Political Participation." In R. Christie and M. Jahoda (eds.), *Studies in the Scope and Method of "The Authoritarian Personality"*. Glencoe: Free Press, 1954.
- LUTZKER, D. R. "Internationalism as a Predictor of Cooperative Behavior," *Journal of Conflict Resolution*, 4, 4 (Dec. 1960), 426-30.
- MARTIN, R. M., and F. L. MARCUSE. "Characteristics of Volunteers and Nonvolunteers in Psychological Experimentation," *Journal of Consulting Psychology*, 22 (1958), 475-79.
- ORNE, M. "On the Social Psychology of the Psychological Experiment." Paper presented at the American Psychological Association Convention, New York, 1961.
- PASTORE, N. "The Role of Arbitrariness in the Frustration-Aggression Hypothesis," *Journal of Abnormal and Social Psychology*, 47 (1952), 728-31.
- PYLYSHYN, Z., et al. "Comparison of Individuals and Pairs as Participants in a Mixed-motive Game," *Journal of Conflict Resolution*, 10, 2 (June 1966), 211-20.
- ROTHAUS, P., and P. WORCHEL. "The Inhibition of Aggression under Non-arbitrary Frustration," *Journal of Personality*, 28 (1960), 108-17.
- SNYDER, R. C. "Some Perspectives on the Use of Experimental Techniques in the Study of International Relations." In H. Guetzkow et al., *Simulation in International Relations: Developments for Research and Teaching*. Englewood Cliffs, N. J.: Prentice-Hall, 1963.
- THIBAUT, J. W., and J. COULES. "The Role of Communication in the Reduction of Interpersonal Hostility," *Journal of Abnormal and Social Psychology*, 47 (1952), 770-77.