

A. EFFECTS OF OBSERVING ATHLETIC CONTESTS ON HOSTILITY

In May, 1964, a riot, precipitated by a referee's decision, erupted at a soccer match in Lima, Peru, killing a number of spectators; the war between El Salvador and Honduras has been traced to a soccer match between those two countries (Lever, 1969); and additional outbreaks of violence have occurred at soccer matches in Great Britain and at boxing matches in New York's Madison Square Garden.

Of course, where large numbers of people gather for public events many of the preconditions for collective behavior exist (cf. Milgram & Toch, 1969; Turner & Killian, 1957; Zimbardo, 1969). Nevertheless, the nature of such competitive and aggressive sports may, in itself, increase spectators' predispositions to engage in violent behavior. The present study examines the arousal of hostility among spectators at athletic contests.

Although a number of laboratory studies have examined hostility and aggression as a function of observing violence (e.g., Bandura, Ross & Ross, 1963; Berkowitz, Corwin & Hieronimus, 1963; Feshbach, 1961; Geen & O'Neal, 1969), the natural setting contains many characteristics not present in such situations. Spectators at a sports event are likely to be more involved than laboratory subjects since school ties and ego identity increase their stake in the outcome of the event. Observers are apt to be more committed in the field setting since they must pay a price for admission. The observed event in the natural setting differs from that typically used in laboratory studies; a sports event, such as a football or soccer game, is a sanctioned and carefully regulated form of interaction in which penalties are imposed for the violation of rules and in which some kinds of violence are condoned. In laboratory experiments the observed aggression is usually a film of an overt aggressive act

Source: Jeffrey H. Goldstein and Robert L. Arms, "Effects of Observing Athletic Contests on Hostility," *Sociometry*, 1971, 34, 83-90.

This study was supported by a grant to the first author from the Bolton Research Fund, Temple University. We thank the Department of Recreation, Philadelphia, for their cooperation. The assistance of Jerry Suls and the comments of Robert Lana, Thomas Ostrom, and Ralph Rosnow on an earlier draft of this paper are greatly appreciated.

which cannot usually be interpreted as justifiable. Given these several differences between the laboratory and field settings, a priori predictions concerning changes in hostility among observers at athletic contests are not possible. However, an examination of theoretical interpretations of aggression leads to several mutually exclusive hypotheses.

Recently a number of books have appeared which suggest that the observation of competitive aggressive sports will serve to reduce hostility among spectators (Ardrey, 1966; Lorenz, 1966; Storr, 1968). This type of vicarious hostility catharsis has been reported in only a few laboratory experiments (cf. Bramel, 1969; Feshbach, 1955, 1956, 1961). If observation of aggressive sports does result in a hostility catharsis, then spectators' post-game hostility should be less than pre-game.

Most laboratory experiments on the observation of aggression report that observers are likely to be more aggressive after viewing violence than before. Frustration-aggression theories (Dollard, Doob, Miller, Mowrer & Sears, 1939; Berkowitz, 1969) suggest that aggression is most likely to occur in observers who are angry, frustrated, or whose goal-directed behavior has been thwarted. If watching one's preferred team lose a game can be regarded as frustrating, then observers whose preferred team loses should be more aggressive than those whose preferred team wins and than those who have no team preference.

Other approaches to aggression are not dependent upon a prior state of frustration and suggest that witnessing violence may reduce the strength of inhibitions against aggressive behavior (Bandura & Walters, 1963). Thus, all spectators at an aggressive athletic contest, regardless of their team preferences, should show an increase in hostility.

To determine the relative merits of these theoretical positions, a field study was conducted in which spectators at a football game were interviewed before or after the game. As a control condition, spectators were also interviewed at a competitive, though nonaggressive sport. The interview was designed to assess hostility, team preference, and additional demographic data.

METHOD

Interviews were conducted at the 1969 Army-Navy football game and also at an Army-Temple gymnastics meet held during the same month. The Army-Navy football game was chosen for study because it is played on "neutral" territory (Kennedy Stadium, Philadelphia), thus assuring a relatively even split among observers' team preferences. The Army-Navy game is more than just a "game" to most spectators; it is a traditional rivalry and emotional involvement in its outcome is quite high.

Interviewers (*Is*) were 13 paid undergraduate students who, several weeks prior to the game, received detailed instructions concerning interview

procedures. Each *I* memorized a prepared introductory speech to be presented to each subject. The speech explained the study as a survey of spectators' attitudes at various intercollegiate athletic contests conducted by Temple University. Pairs of *Is* were randomly assigned to entrances of the Kennedy Stadium and were to conduct interviews only with subjects about to enter their assigned gates. The subjects consisted exclusively of adult males.

Immediately after arrival at the assigned entrance, *I* was to interview the first adult male approaching that gate. After completing an interview, *I* was to approach the very next eligible subject. This procedure was employed to eliminate any systematic factors which may have biased subject selection. The *Is* also recorded the number of people who refused to be interviewed.

The interview began with the introductory speech. All subjects were assured that the interviews would be anonymous. A number of demographic questions were then asked. These items served two purposes: to check on the equivalence of the various groups in the study and to engage the subject's involvement in the interview. The questions concerned distance traveled to the game, frequency of attendance at football games, cost of tickets, and number of other people accompanying the subject. Subjects were also asked to indicate their preferred team, if any, and in the former case, how upset they would be if their preferred team lost the game.

Following these items were three scales taken from the Buss-Durkee inventory, designed to measure hostility (Buss & Durkee, 1957). Hostility was used as the dependent variable in the study because it was felt to be sensitive to influence by situational factors. Hostility is used here as one index of overt aggression. Each hostility scale consists of a number of statements to be answered "true" or "false" by the subject. The scales employed were the indirect hostility (9 items), resentment (8 items), and irritability (11 items) scales, each of which was found to have satisfactory reliability in a number of independent investigations reported by Buss (1961). Included among these 28 hostility items were eight filler questions concerning football, placed at random intervals throughout the hostility portion of the interview schedule. These items were designed to minimize suspicion about the true nature of the study.

Following completion of the hostility items, subjects were asked to state their reactions to the study and to indicate what they felt the study was about. Subjects were then thanked for their cooperation and dismissed. The total time required for each interview was approximately 10 minutes. The post-game interview was identical to the pre-game interview, except that the tense was changed in a few items where that was appropriate.

In the control condition, five *Is* were employed at the Army-Temple gym meet. They were to interview only male spectators and the interview schedule was nearly identical to that employed at the Army-Navy football game. The only difference between the two interviews was that gymnastics items were inserted for football items on the appropriate questions.

RESULTS

Comparability of Groups

Football Game Data

Before the game, 59 eligible subjects refused to be interviewed from a total of 156 subjects approached (37.8%). There were 44 post-game refusals of the 97 approached (45.5%). A chi-square with 1 *df* was not significant. A total of 150 subjects completed the interview, 97 pre- and 53 post-game.¹

A 2 x 3 analysis of variance was computed for each dependent measure, the factors being Time of Interview (pre- or post-game) and Preferred Team (Army, Navy, or no preference). Since unequal *ns* resulted, an unweighted means solution was used. Analysis of demographic data revealed no significant differences among any of the six groups, for distance traveled, frequency of attendance at football games, cost of tickets, or number of companions. Thus, the six experimental samples were considered to come from the same population. The subjects' comments about the interview indicated no suspicion of the true purpose of the study.

Gymnastics Meet Data

Because the gymnastics meet was not held on "neutral" territory as the football game was, over 90% of the subjects favored the home team (Temple). Because of the small number of subjects who had no team preference or who preferred Army, these groups were combined and only before-after comparisons were made.

A total of 49 pre- and 32 post-game interviews were completed. Only 4% of the subjects approached refused to be interviewed, and there were no differences in refusal rate for the before and after groups. The higher refusal rate at the football game may have been due to the difficulty encountered in getting into and out of the stadium, which placed a greater premium on time at that event.

There were no differences between the before and after gym meet groups on distance traveled, age of subjects, or number of companions with subjects. However, pre-meet subjects attended gym meets significantly more often than post-meet subjects ($t = 2.84$, $df = 79$, $p < .01$). This may reflect the fact that the less involved (those who attend meets less frequently) left the gym meet earlier and were, therefore, overrepresented in the post-meet sample. On the whole, the pre- and post-meet groups are considered to be equivalent.

1. This difference reflects the greater variance in arrival time than in departure time. Subjects began arriving for the game as much as two hours prior to its scheduled starting time. Nearly all subjects left the stadium within 30 minutes of the termination of the game.

Comparability of Football and Gym Meet Groups

The before and after subjects at the football game and the gymnastics meet were compared on distance traveled to attend the event, frequency of attendance at such events, age, and number of companions with the subject, in 2 x 2 unweighted means analyses of variance. These results indicate no differences among any of the four groups on distance traveled or number of companions. Subjects attending the football game, however, were significantly older than those at the gymnastics meet ($p < .01$) and they attended football games significantly more often ($p < .05$). The only interaction obtained was for frequency of attendance, in which post-gym meet subjects attended meets less than pre-meet subjects, and this finding has been discussed above. Taken as a whole, the pre-post football and pre-post gym meet subjects are considered equivalent.

Hostility Data

Football Game

Group means for each of the three hostility subscales from the Buss-Durkee inventory were highly intercorrelated, and separate analyses for each subscale lead to similar findings. Therefore, a single score, the sum of the three subscales, was computed for each subject. The possible range of hostility scores was from 0 to 28, with the higher figure representing maximum hostility.

The mean hostility scores by group are presented in Table 1. Analysis of these data indicate that, regardless of subjects' preferred team, post-game hostility was greater than pre-game ($F = 5.29$, $df = 1/144$, $p < .025$). Neither the main effect for Preferred team ($F < 1.0$) nor the interaction effect ($F < 1.0$) was significant.

Gymnastics Meet

As with the football game hostility data, the three Buss-Durkee subscales were combined into a single hostility score. The pre-meet hostility

TABLE 1
Mean Football Game Hostility Scores by Condition

	Preferred			Total
	Army (winning team)	Navy (losing team)	No preference	
Pre-game	n = 38 10.42 ^a	n = 47 11.72	n = 12 11.67	11.20
Post-game	n = 18 13.33	n = 30 13.17	n = 5 15.00	13.40

^aThe higher the score, the greater the hostility.

mean (12.00) was not significantly different from the post-meet hostility mean (12.71, $t = .66$, $df = 79$, $p < .20$). Thus, hostility did not significantly increase as a result of observing the gymnastics meet.

DISCUSSION

Hostility data collected at the football game indicates that, regardless of team preference and the outcome of the game, subjects were significantly more hostile after observing the game than before. A number of alternative explanations for this finding may be eliminated on the basis of data obtained at the gymnastics meet. The gym meet includes many similarities to the football game: subjects are seated for two to three hours in a large and compact crowd where outbursts of cheering and applause occur. Therefore, the relative increase in hostility found at the football game, but not at the gym meet, cannot be attributed to any of these factors, although it should be borne in mind that the spectators at the football game were slightly older and attended football games more often than their counterparts at the gym meet. In addition, there was a slight, though nonsignificant, increase in hostility among the gymnastics spectators. There are a number of differences between the two events, however, which cannot be eliminated as possible explanations for the present findings: the absolute number as well as the density of others was greater at the football game; norms for expressive behavior at a football game differ from those at a gym meet, where the crowd is usually less vociferous, and where the general activity level is lower owing to the absence of vendors, bands, and cheerleaders. It seems, however, that one major difference between the nature of the two events is that a football game involves multiple players in direct physical contact with one another, while a gym meet involves individual performances in which no contact can occur. It seems likely, therefore, that the increase in hostility is due to the nature of the observed event; watching an aggressive sport leads to an increase in hostility among spectators.

One methodological difficulty in the present study is the possibility of a subject selection bias: it may be that more hostile observers are attracted to a football game than to a gymnastics meet. This may be indicated by the higher refusal rate at the football game. However, it might be reasonably assumed that subjects who refuse to be interviewed are more hostile than those who are cooperative, and thus the differential refusal rate would lead one to expect fewer hostile responses at the football game among the cooperative subjects.

No support for a catharsis effect is obtained in the present study, contrary to the many popular notions (Ardrey, 1966; Lorenz, 1966; Storr, 1968) that such an effect would occur. Elicitation of hostility catharsis, if it does occur, may require more intense or more direct aggression than that present in a football game.

The failure to find an interaction of preferred team and game outcome seems to support a general disinhibition notion. That is, the act of observing an aggressive sport may reduce subjects' inhibitions against aggression and result in increased hostility. Whether disinhibition is a result of an increase in aggressive drive or is due to the heightened salience of hostility cannot be determined from the present data.

REFERENCES

- Ardrey, R. *The territorial imperative*. New York: Dell, 1966.
- Bandura, A., Ross, Dorothea, & Ross, Sheila. Imitation of film-mediated aggressive models. *Journal of Abnormal and Social Psychology*, 1963, *66*, 3-11.
- Bandura, A., & Walters, R. H. *Social learning and personality development*. New York: Holt, Rinehart & Winston, 1963.
- Berkowitz, L. The frustration-aggression hypothesis revisited. In L. Berkowitz (ed.), *Roots of aggression*. New York: Atherton Press, 1969, pp. 1-29.
- Berkowitz L., Corwin, R., & Hieronimus, R. Film violence and subsequent aggressive tendencies. *Public Opinion Quarterly*, 1963, *27*, 217-229.
- Bramel, D. The attraction and reduction of hostility. In J. Mills (ed.), *Experimental social psychology*. New York: Macmillan, 1969, pp. 33-63.
- Buss, A. H. *The psychology of aggression*. New York: Wiley, 1961.
- Buss, A. H., & Durkee, Ann. An inventory for assessing different kinds of hostility. *Journal of Consulting Psychology*, 1957, *21*, 343-348.
- Dollard, J., Doob, L., Miller, N., Mowrer, O. & Sears, R. *Frustration and aggression*. New Haven: Yale University Press, 1939.
- Feshbach, S. The drive reducing function of fantasy behavior. *Journal of Abnormal and Social Psychology*. 1955, *50*, 3-11.
- Feshbach, S. The catharsis hypothesis and some consequences of interaction with aggressive and neutral play objects. *Journal of Personality*, 1956, *24*, 449-462.
- Feshbach, S. The stimulating versus cathartic effects of a vicarious aggressive activity. *Journal of Abnormal and Social Psychology*, 1961, *63*, 381-385.
- Geen, R. G., & O'Neal, E. C. Activation of cue-elicited aggression by general arousal. *Journal of Personality and Social Psychology*, 1969, *11*, 289-292.
- Lever, Janet. Soccer: Opium of the Brazilian people. *Transaction*, 1969, *7*(2), 36-43.
- Lorenz, K. *On aggression*. New York: Harcourt, Brace & World, 1966.
- Milgram, S., & Toch, H. Collective behavior: Crowds and social movements. In G. Lindzey and E. Aronson (eds.), *Handbook of social psychology*. 2nd. ed. Reading, Mass.: Addison-Wesley, 1969.
- Storr, A. *Human aggression*. New York: Atheneum, 1968.
- Turner, R. H., & Killian, L. M. *Collective behavior*. Englewood Cliffs, N.J.: Prentice-Hall, 1957.
- Zimbardo, P. G. The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. In W. J. Arnold and D. Levine (eds.), *Nebraska Symposium on Motivation*. Lincoln: University of Nebraska Press, 1969.

B. CONDUCTING FIELD RESEARCH ON AGGRESSION: NOTES ON "EFFECTS OF OBSERVING ATHLETIC CONTESTS ON HOSTILITY"

Jeffrey H. Goldstein

ORIGINS OF THE STUDY

Human aggression is both a fascinating and an important topic, and as a psychologist I have been interested in it since my days as a graduate student. There were no courses taught on human aggression when I was a student at Ohio State, and there are few offered in it now around the country, so the student who wants to learn something about the psychology of aggression must devote some of his spare time to reading the literature on his own. I was particularly interested in the effects of violence in television and movies and in sports and humor. Much of what I had read in the professional journals failed to provide satisfactory answers to my questions about aggression, and I began to suffer from the feeling that many freshmen experience upon taking their first course in psychology. I was curious about problems that seemed pertinent to my experiences and interests only to discover that most psychologists spent their time running rats and studying what appeared to me then, at least at first glance, to be trivial phenomena. Although there are philosophical, historical and practical reasons for this tendency—and people should, after all, be free to study what they wish—still it seemed that my questions were amenable to study by scientific means. Yet the psychological and sociological literature was quite meager on this subject. Either the research was too artificial, making it difficult to generalize beyond the laboratory to the real world, or it was theoretically empty, with no hint of the underlying dynamics of aggression.

I have always felt that data from an experiment must serve two functions: first, they must test, build, or revise theory; and second, they must relate, as closely as possible, to real and important human behaviors. In other words, research should serve to provide explanations, and not merely descriptions, of behavior. I also believe, though many of my colleagues do

Source: Prepared especially for this volume.

not share this view, that the behavior under study must be one of some consequence. A good deal of human behavior, though it may follow discoverable and empirical principles, is simply too trivial to study. For example, it may be that each person puts on the same shoe first every day; if one put on his left shoe first this morning, he is likely to put the left shoe on first every morning. Now this could quite easily be studied, and I suppose that some sort of abstract explanation could be built and tested around this observation. But the behavior has no consequences for other actions and is in this sense "unconnected" to other behaviors and, therefore, trivial. A behavior worth spending time and energy studying must be one with both theoretical implication for and some "connectedness" or interrelation with one or more other behaviors. The study of human aggression is connected to such issues as war, crime, prejudice, hatred, and less directly to political and economic behavior, and it is this connectedness which makes it a nontrivial and worthwhile phenomenon to study.

The available research in the journals suffered primarily from the unconnectedness of the behavior studied in the laboratory to real-life behaviors and was also, for the most part, theoretically unimportant as well. In the course of discussing this state of affairs with Robert Arms, then a graduate student at Temple University and now a member of the psychology faculty at the University of Lethbridge, we decided to conduct an aggression experiment outside the laboratory. We sought to test aggression theory using real episodes of violence rather than contrived situations. Our primary interest was in determining, first, whether the results obtained in the laboratory could also be obtained outside the laboratory, using a real aggressive episode. Second, we wanted to study a fairly subtle but pervasive kind of violence rather than the type used in laboratory research. In most laboratory research on aggression, subjects are exposed to an excerpt from a movie, usually of only a few minutes' duration, which is either extremely violent or (for subjects in control groups) totally benign. What would happen if researchers used an entire aggressive film, such as *Clockwork Orange* or *Straw Dogs*, rather than a brief excerpt?¹ Would they still discover that watching violence leads to an increase in aggression among observers? What would happen if the violence were not quite so intense?

PARING DOWN THE ALTERNATIVES

Our interest was essentially to examine the effects on observers of witnessing a fairly lengthy and realistic kind of violence in a natural setting.

1. A study recently completed and submitted for publication was designed to answer this question (Jeffrey H. Goldstein, Ralph L. Rosnow, Tamas Raday, Irwin Silverman, & George D. Gaskell). Punitiveness in response to films varying in content: A cross-sectional field study of aggression, and an unpublished study just called to my attention has also sought to answer it (Martin, Gray, Smoke, & Wilson).

We expected, in the absence of substantial laboratory evidence to the contrary, that observers would show an increase in aggressiveness after watching an aggressive event. This hypothesis is consistent with the bulk of laboratory research on the effects of witnessing violence. However, there were theories of aggression which made quite different predictions, and so we were able to test these theories at the same time, most notably those of Ardrey, Lorenz, and Storr (see the article reprinted above).

After considerable discussion with several of our colleagues (once you obtain a Ph.D. your friends mysteriously turn into "colleagues"), we decided to use a football game as the aggressive event for the study and to use spectators at the game as our research subjects. We next had to decide which football game to use. We thought it would be ideal to find a game where half the fans were partisan toward one team and the remaining half toward the other; in this way we could get some subjects whose preferred team won and others whose preferred team lost the game. We felt this would be desirable because the effects of the game might well depend upon whether subjects were pleased or disappointed over the outcome of the game. A football game played on "neutral" territory would certainly satisfy this requirement. As it turns out, one of the rare advantages of living in Philadelphia is that it is the site of the annual Army-Navy football game. We expected that, since Philadelphia is about halfway between West Point and Annapolis, nearly half the fans would be pro-Army and half pro-Navy.

At this point we knew only that we wanted to measure aggression as a result of watching the Army-Navy game. With over 100,000 people expected at the game, we knew we would have to use an easily administered measure of aggression, since our subjects would probably be in a hurry to get to their seats on the way into the game and back to their cars on the way out. We also knew that it would be impractical to measure aggression in the same spectators both before and after the game, so we decided to have two groups of subjects, one in which we measured aggression prior to the game, and one in which we measured aggression afterwards. These two groups could then be divided into subgroups according to whether they wanted Army or Navy to win, or whether they didn't care about the outcome. Deciding on the measure of aggression to use was more difficult. Of course, we could have asked each subject how aggressive he felt, but that is a vague and not too meaningful question. We eventually decided to use three subscales from the Buss-Durkee Hostility Inventory. This measure has several advantages: it can be administered verbally, so subjects don't have to fill out any forms; its reliability and correlation with other measures of aggression are known; and, by using three subscales rather than the whole inventory, the measure could be administered in only a few minutes.

By now, the study was beginning to take shape. We knew what we wanted to do and how we wanted to do it; aggression would be measured in spectators using the Buss-Durkee inventory before and after the Army-Navy

game. What if the subjects interviewed after the game were different in aggressiveness from those interviewed before the game? Well, of course, we would want to attribute the difference to the violent nature of the game itself. But it was possible that those who would agree to be interviewed after the game might differ in some major way from those who agreed to be interviewed before the game, and this uncontrolled factor might have caused our difference in aggressiveness, and not the football game. That was a realistic possibility, and to help eliminate it we decided to ask all subjects a number of questions, such as their age, place of residence, and so on, in order to determine whether the before and after groups were similar on these demographic variables. If the mean age, occupation, residence, and other variables were the same for all groups of subjects, we could say with some confidence that the groups came from the same population. Then, if a before-to-after change in aggression were found, it could more safely be attributed to the football game.

Or could it? Although it didn't seem likely, it was possible that people who went to a football game were different from other people, and that they were more likely to become aggressive as a result of watching a game than others. It was also plausible that not just football, but *any* game—even a chess match—might cause people to become more aggressive. Clearly a control group was needed which might help eliminate these possible alternative explanations for an increase in aggression at the football game. The same week as the Army-Navy game, an Army-Temple gymnastics meet was scheduled at Temple University, and we decided to measure aggression before and after the gym meet among spectators as a control group. A gym meet is clearly a nonaggressive sport, and, if an increase in aggression were found at the football game but not at the gym meet, it was likely to be caused by the violence of the football game.

While the questionnaires for the study were being typed and mimeographed, we asked undergraduate psychology students if they would like to serve as interviewers for our study. About a dozen students agreed to participate, and a meeting was arranged with them. Because the study was scheduled for Thanksgiving vacation, students would have to interview during their free time, and so we wanted to compensate each interviewer for his time and expenses. I applied for, and graciously received, a small grant from Temple University to meet these expenses.

At the meeting, we explained in considerable detail how the interviews were to be conducted, and went over the entire interview with the students, discussing possible difficulties that might arise: what they were to do if someone decided not to complete the interview once he had begun, for example, or what should be recorded if more than one person answered a question directed to a particular respondent. Of primary importance was who should be selected as subjects for the study. We feared that our interviewers might tend to approach spectators in such a way as to systematically

eliminate (or systematically include) only those who appeared least aggressive; in other words, they might selectively choose whom to interview, and this nonrandom selection would bias our results. In order to minimize this possibility, interviewers were instructed to approach the very first male spectator entering the stadium (or leaving it, for postgame subjects) and then to interview the first male they saw as soon as the previous interview had been completed. Finally, interviewers were told to keep an accurate count of the number of people approached who refused to be interviewed.

It should be clear from the above that we anticipated several possible problems before they ever arose and had decided (though sometimes rather arbitrarily) what to do about each of them. I think that any piece of research can be improved considerably by trying to anticipate every conceivable problem that can arise. In the case of our study, such considerations led us to add the Temple-Army gym meet as a control group, to determine the comparability of our groups by use of demographic questions, to keep a tally of the refusals for each condition of the study, and to specify rules for the selection of subjects. We also decided to restrict our subjects to male spectators, since females at a football game are generally accompanied by a male, and it is unclear whether they are there by choice or coercion. One additional bit of trouble that could arise was that the police might object to interviewers at the football game. So we wrote to the Philadelphia Department of Recreation, caretakers of the J. F. Kennedy Stadium, site of the Army-Navy game, and obtained permission to conduct the study.

With seating diagrams in hand, we randomly assigned our interviewers to entrances of the JFK Stadium and to entrances of the Temple gymnasium, where the Army-Temple gym meet would be held. After arming them with an introductory speech about the study, we waited for the day of the football game.

THE DATA AND DATA ANALYSIS

Interviewers arrived at JFK Stadium about two hours before game time and began interviewing as soon as they arrived and until the game began. This resulted in 97 completed pregame interviews. As soon as the final gun sounded, they began the postgame interviews, and a total of 53 completed interviews were obtained. A few days later, five interviewers collected 49 pre- and 32 post-gym-meet interviews.

We now had 150 interviews from the football game and 81 from the gym meet, each of which contained demographic information about the respondent, his preferred team, and answers to 28 true-false hostility questions. We coded the questionnaires by indicating on each interview whether the respondent was at the football game or gym meet, whether he was interviewed before or after the event, and which team, if any, he wanted to win. We then recorded all this information, plus the respondent's age, residence,

cost of tickets, and number of companions, onto IBM punch cards, along with the total hostility score from the Buss-Durkee inventory. The analysis of the data was fairly simple and straightforward; we wanted to know whether subjects showed any change in hostility as a result of watching the football game or the gym meet, and we wanted to know whether any of the groups in the study differed from the others on any of the demographic variables. These analyses were done by computer but could easily have been done by hand.

The data analysis revealed that those at the football game were more aggressive after the game, whereas there was no statistically significant change in level of aggression for those at the gym meet. Since the results were not unexpected, we didn't have to spend a great deal of time in reanalysis of the data trying to find out what did happen to spectators' levels of aggression. The data made sense and were consistent with previous research on the effects of witnessing violence. The final task was to write up the study in the form of a journal article, have several colleagues comment on the manuscript, and, after making some suggested changes,² submit the paper for publication. Since my training in social psychology was both in psychology and in sociology, my choice of journal was *Sociometry*, the social-psychology journal of the American Sociological Association.

THE IDEAL STUDY

It should be obvious that the study has several weaknesses, some of which are mentioned in the paper itself and some of which are implied above. Its major shortcoming is that it is open to at least two alternative explanations: first, that the subjects interviewed before and after the football game and gymnastics meet may not have been equivalent on their pregame levels of hostility, and second, that the excitement inherent in a football game, and not the violence of the game, caused the obtained increase in hostility. The ideal study would involve random assignment of subjects to the various groups in the study; that is, people would be randomly assigned to watch either a football game or a nonviolent sport which was equivalent to the football game in excitement. In this way it could be assumed that, prior to these two events, all groups of subjects were equivalent in aggressiveness. This would permit assessment of aggression after the two events, with no need for a before measure. Then if the post-football-game hostility were greater than the post-control-game hostility, the difference between the two could be confidently attributed to the violence of the football game.

One difficulty with all aggression research is the measurement of aggression. Laboratory research tends to rely heavily on the administration of shock

2. Helene Feinberg, Louise Kidder and Ralph Rosnow were kind enough to comment on a draft of this paper as well, and I am grateful to them for their suggestions.

by subjects to another person as a measure of aggression, while field research relies on fairly lengthy interviews or questionnaires. What is needed is a brief interview or easily observable behavior which is reliable and is correlated highly with aggression in naturalistic situations.

A NOTE ON INDIVIDUAL DIFFERENCES

It is obvious to any psychologist that, while observing violence may lead to an increase in subjects' levels of aggression, it leads to greater increases in some subjects than in others. There may even be some people who show a decline in aggressiveness after watching violence. The problem I am raising here is generally referred to under the heading of "individual differences." As a *social* psychologist, my interest is in studying general effects of social variables; that is, dealing with factors which have an influence on people in general, or on most people. It is a perfectly legitimate enterprise to study characteristics of people which make them respond differently to a situation than others do. However, there are two disadvantages to this type of study, from my point of view. First, personality traits can be *measured*, but they often cannot be experimentally *manipulated*; thus, individual-difference studies tend to be correlational rather than experimental. Second, even if a particular trait were found which enabled the researcher to determine who responded aggressively to observed violence and who did not, such a finding would not in itself provide an explanation for this effect. In other words, such correlational studies are often descriptive, but only rarely explanatory.

EXPLANATION AND DESCRIPTION

The data from our study tell us what happened, but they do not tell us, to any great extent, why it happened. The study indicates that aggressiveness increased as a result of watching a football game but not as a result of a gymnastics meet. It is also clear that the catharsis effect predicted by Ardrey, Lorenz and Storr failed to occur; watching a violent sport did *not* lower observers' aggressiveness. But why does watching a football game increase aggressiveness? The article above contains a brief discussion of several possible explanations: (1) there could be a general reduction of inhibitions against aggressing as a result of watching aggression on the playing field; (2) watching aggression might increase the strength of observers' aggressive drives; (3) watching aggression might lead to a general increase in spectators' levels of arousal, and the arousal in turn may lead to increased aggression; (4) the general excitement of the football game might cause the increased aggression, and not necessarily the violence of the game; and (5) watching violence might make aggression more salient to observers, and the salience could cause the increase in aggressiveness.

Since I have come to believe that there is no such thing as an aggressive "drive," the most plausible explanations to me are the arousal and salience

interpretations. An increase in observers' arousal might well account for the increased aggression, and there is some experimental evidence which supports this explanation (Zillmann, 1971). Support also has been generated for the notion that increasing the salience of a topic leads to a preference for that and related topics (Goldstein, 1972; Goldstein, Suls, & Anthony, 1972). Since these explanations seem the most likely, I have been engaged in further research on them for the past three years to find out more precisely why we obtained the results we did (Goldstein, Davis & Herman, in press; and the Goldstein, Rosnow, Raday, Silverman & Gaskell study cited in footnote 1 above).

There is one other interpretation of the data which occurred to me only after the study was published, based on the well-known work of Bandura and his associates (1971, 1972). If you will examine the means in Table 1 of the study, you will notice that the smallest increase in hostility at the football game occurred among those whose preferred team lost. For the Navy fans, there was an increase of only about 1.5 points, though for the Army fans the increase was nearly twice that. Surely this is an unexpected finding; one might expect that those whose preferred team lost the game would be more upset and, therefore, more frustrated and aggressive. If it can be argued that both Army and Navy players were aggressive on the field, and that winning the game was a reward for Army's aggression while losing was a punishment for Navy's aggression, then this finding can be explained. The vicarious punishment served to inhibit the aggressiveness of pro-Navy spectators, while the vicarious reward served to increase the aggressive level of pro-Army spectators. This interpretation is based on that developed by Bandura, Ross and Ross in their 1963 study in which aggressive models were either rewarded or not rewarded for aggression. Observers were more aggressive when they had seen the model rewarded than when they had seen the model punished.

THE AFTERMATH

The study reprinted above appeared in *Sociometry* in the Spring of 1971, and after its appearance there were the usual inquiries and requests for reprints from people in the social sciences. Over a year later, the public relations office of Temple University prepared a brief press release which was a summary in nontechnical terms of the study and its findings. Two days later I was contacted by the Associated Press and other news sources which were preparing stories about the study.

Since that time I have been invited to appear on talk shows and have received letters from scores of people around the country inquiring about the study. I have been besieged by letters from wives who are angry that their husbands spend too much time watching football on TV, by jocks of every description wondering whether they should give up their sport, by people writing term papers or theses for physical education, sociology, psychology,

or health. I am asked questions about topics from child rearing to the massacre at the 1972 Munich Olympics.

I have neither the physical nor intellectual resources to answer all these questions, and so I have generally responded by sending a reprint of the paper along with a note of apology at my inability to offer much practical advice. But there are some lessons to be learned from this experience. People are genuinely interested in psychological research when it proves pertinent to their own lives; they seem willing to consider empirical findings as one factor in their decision making. It seems possible to conduct psychological research which is meaningful both to other researchers and to laymen. I must confess that I was surprised at the level of sophistication of questions asked by both journalists and laymen concerning aggression. Obviously one doesn't have to be a trained social scientist to be interested in, or to know a good deal, about human behavior. There were, of course, some exceptions, such as the sports magazine article which reported the study as comparing reactions at a football game with those of people at the movie *Bambi* and which implied that hostility increased at the football game because spectators were interviewed *during* the game. But generally the journalists reported the study accurately, and peoples' questions about aggression were psychologically meaningful.

No one study is able to provide very many answers to practical questions, but I think research can be done which is theoretically meaningful and of value to other professionals and at the same time "connected" enough to everyday experience to be of interest, and perhaps of some use, to all of us. Certainly this should come as no surprise to anyone, least of all to social psychologists, who are generally familiar with Kurt Lewin's concept of "action research." But if Lewin could survey the bulk of contemporary social-psychological research, he would surely ask where the action was. By conducting theoretically meaningful studies in field settings, the action can be put back into action research, and we can then pay more than lip service in our debt to Lewin. Steps in this direction are being taken, if slowly, by social psychologists, and there are now enough such studies to have warranted an anthology (Bickman & Henchy, 1972; see also Webb, Campbell, Schwartz, & Sechrest, 1966).

SOME CONCLUDING NOTES

Although the discussion above indicates in a fairly specific fashion why the Army-Navy game study was done in the manner in which it was, I have not made much mention of research as a whole. I would not spend the better part of my days engaged in research if it weren't personally satisfying. First, systematic inquiry is the *only* way I know that enables one to find the answer to a previously unanswered question, and to retain some confidence in the answer. Short of divine revelation, the supply of which seems to have run out in the 18th Century, empirical research can provide answers to questions

available by no other means. Most important, if others doubt your answer, they are free to repeat the research for themselves.

Second, I find that doing research is fun and challenging. A study may begin with an observation, with a problem for which there is no known solution, or with a question derived from a theory. Regardless of its origin, I begin with some abstract idea which can usually be stated in the form of a hypothesis without much difficulty. From then on, I am free to devise what seems the best test of that hypothesis, given my limitations of time, money, research assistants, subjects, and patience. Although the practical tasks of questionnaire construction, data collection, and analysis are less interesting than the creation of the experimental design, they are, of course, indispensable to the overall project. Once the data have been collected and analyzed, I have already seen an idea develop from a mere thought to a tangible set of findings. If the data fail to support the hypothesis, there is the frustrating but challenging problem of making maximum sense of that data.

However, regardless of the results, conducting a study is a demanding and creative enterprise, not unlike painting a picture or taking a photograph, hobbies in which I engage with considerable enthusiasm if not talent. In each case, one starts with an idea and the raw materials of the craft. The challenge is to translate the idea, given the usually severe limitations of the tools, into a concrete form. The results cannot always be anticipated in advance, as witness some of my paintings, photographs, and research, but the act in itself provides the real enjoyment.

REFERENCES

- Albert Bandura. *Social learning theory*. Morristown, N.J.: General Learning Press, 1971.
- Albert Bandura. *Psychological modeling*. New York: Atherton, 1972.
- Albert Bandura, Dorothea Ross, & Sheila A. Ross. Vicarious reinforcement and imitative learning. *Journal of Abnormal and Social Psychology*, 1963, 66, 601-607.
- Leonard Bickman & Thomas Henchy. *Beyond the laboratory: Field research in social psychology*. New York: McGraw-Hill, 1972.
- Jeffrey H. Goldstein. Preference for aggressive movie content: The effects of cognitive salience. Unpublished manuscript, Temple University, 1972.
- Jeffrey H. Goldstein, Roger Davis, & Dennis Herman. The escalation of aggression: Experimental studies. *Journal of Personality & Social Psychology*, in press.
- Jeffrey H. Goldstein, Jerry M. Suls, & Susan Anthony. Enjoyment of specific types of humor content: Motivation or salience? In J. H. Goldstein & P. E. McGhee (Eds.) *The psychology of humor*. New York: Academic Press, 1972. Pp. 159-171.
- J. D. Martin, L. N. Gray, G. L. Smoke, & F. D. Wilson. Mass media violence and overt behavior: A natural experiment. Unpublished manuscript, Washington State University.
- Eugene J. Webb, Donald T. Campbell, Richard D. Schwartz, & Lee Sechrest. *Unobtrusive measures: Nonreactive research in the social sciences*. Chicago: Rand McNally, 1966.
- Dolf Zillmann. Excitation transfer in communication-mediated aggressive behavior. *Journal of Experimental Social Psychology*, 1971, 7, 419-434.