Review of General Psychology
1997, Vol. 1, No. 1, 19–41

1089-2680/97/$3.00

External Validity of "Trivial" Experiments: The Case of Laboratory Aggression

Craig A. Anderson
University of Missouri—Columbia

Brad J. Bushman
Iowa State University

The external validity of artificial "trivial" laboratory settings is examined. Past views emphasizing generalizability of relations among conceptual variables are reviewed and affirmed. One major implication of typical challenges to the external validity of laboratory research is tested with aggression research: If laboratory research is low in external validity, then laboratory studies should fail to detect relations among variables that are correlated with aggression in "real-world" studies. Meta-analysis was used to examine 5 situational variables (provocation, violent media, alcohol, anonymity, hot temperature) and 3 individual difference variables (sex, Type A personality, trait aggressiveness) in real-world and laboratory aggression studies. Results strongly supported the external validity of trivial laboratory studies. Advice is given on how scholars might handle occasional descrepancies between laboratory and real-world findings.

Consider the following two scenarios.

Scenario 1: It is a hot July afternoon. Frank has had a difficult day at the auto plant. While working on the line, he accidentally smashed his left elbow—nothing serious, but it still hurts. Instead of sympathy, or even a little good natured ribbing, Frank's immediate supervisor yells at him for this mishap, remarking that he didn't want any clumsy oaf fouling up the safety record of his area. On the way home, Frank stops at a local bar for a brew and a quick look at the Cubs baseball game. The game is a close one, so Frank stays for several innings and several beers. Unfortunately, the Cubs blow a two-run lead in the ninth and lose the game 5–4. Disgusted, Frank goes home. Because of the game and the beers, he's a bit late, and he knows it. Upon arrival at home, Frank's wife lets him have it. She's mad, she says, because she works full time too, and still has to pick up the kids and make dinner, and then he shows up late and has been out drinking. A fight ensues, like many others they have had in the past, only, this time it becomes physical. Who slapped whom first is unclear, but in the end, she has a broken jaw, a black eye, and several minor injuries.

Scenario 2: Emilio arrives promptly at 3 p.m. for the psychology experiment "Effects of Alcohol and Stress on Reaction Times." The experimenter introduces herself, goes over the consent form with Emilio, and then describes the study. "We are looking at the effects of alcohol and stressors on people's reaction times," she says. "In this experiment, you will compete with another participant, in the next room, on a reaction-time task for 25 trials. Both of you will first consume an alcoholic beverage and then wait 30 minutes while the alcohol is absorbed into your bloodstream. Electric shocks will be used as stressors. You will both be hooked up to this shock generating machine. The loser on each trial will receive an electric shock. Before each trial, you get to set the level of shock that your opponent will receive if he loses. We want to see how alcohol and stressors affects peoples' reaction times. Any questions?" Emilio agrees to participate. Although all participants are told that they will receive an alcoholic beverage, half receive an alcoholic beverage and half receive a placebo beverage. Emilio, who is randomly assigned to receive an alcoholic beverage, drinks the beverage and then waits for 30 minutes while the alcohol is absorbed into his bloodstream. Emilio then begins the reaction-time procedure. Emilio notices early on that his opponent seems to be increasing the shock levels he sets. Emilio decides that "This guy is asking for it," and subsequently raises the shock levels that he selects, eventually using the highest level possible for the last five trials.

What is the relation between these two scenarios? Specifically, what is the relation between wife battering and laboratory aggres-

Craig A. Anderson, Department of Psychology, University of Missouri—Columbia; Brad J. Bushman, Department of Psychology, Iowa State University.

We thank Ann Bettencourt for sharing her recent meta-analytic review and for aiding us in our specific use of her review. We also thank Leonard Berkowitz, Douglas Mook, and several anonymous reviewers for their comments on earlier versions of this article.

Correspondence concerning this article should be addressed to Craig A. Anderson, Department of Psychology, University of Missouri, Columbia, Missouri 65211, or to Brad J. Bushman, Department of Psychology, Iowa State University, Ames, Iowa 50011-3180. Electronic mail may be sent via the Internet to psycaa@showme.missouri.edu or to bushman@iastate.edu. Specific questions concerning the meta-analyses reported in this article should be addressed to Brad J. Bushman.

sion? Is there any reason to believe that the "trivial" laboratory aggression experiment can inform us about domestic violence, or any other kind of "real-world" aggression? At a more general level, the question becomes one of external validity of theories derived and tested largely in the artificial confines of the experimental laboratory. External validity questions are not restricted to aggression research, of course. Almost all areas of psychology that involve the experimental laboratory are at least occasionally challenged on external validity grounds (e.g., Harré & Secord, 1972; Neisser, 1978).

In this article, the external validity issue is examined with a special emphasis on aggression. We first present the standard defenses of laboratory research in psychology, which largely revolve around the goals of theory construction and testing, and the generalizability of theory-based understandings about psychological phenomena. We then propose an empirical justification, based on the similarity of findings between real-world and laboratory studies of aggression. Finally, we present two new assumptions about validity issues, assumptions that we believe will promote a more useful way of handling laboratory/real-world discrepancies when they arise, and illustrate their utility in furthering theory in the temperature/aggression literature.

We focus on aggression for three reasons; one is practical whereas the other two are strategic. The practical reason is that we spend a lot of our time doing aggression research, publishing aggression articles, reviewing aggression manuscripts, and teaching aggression classes. In other words, among the hundreds of possible research domains for an examination of the external validity question, this is the one with which we are most familiar. One strategic reason for choosing the aggression domain is that it is frequently targeted as an example of how "those laboratory folks waste time and resources on trivial studies that have no bearing on the real world." For instance, Gottfredson and Hirschi (1993) argue that aggression research produced in psychological laboratories confuses various types of aggression to such an extent that researchers "working on violence and aggression should not attempt to incorporate the findings of this laboratory research into their theories" (p. 64; see also Kane, Joseph, & Tedeschi, 1976). The second strategic reason for examining the external validity issue in the aggression domain is that virtually everyone (including both major parties in the U.S. Congress) agrees that aggression in modern society is a serious problem. In other words, contributions toward understanding aggression are desperately needed.

External and Internal Validity

Though few readers of this journal are unaware of external and internal validity issues, a brief reminder will facilitate later discussions. Our focus is primarily on external validity, but external validity cannot be properly discussed or understood without reference to internal validity, so that is where we begin. (For detailed discussions of internal and external validity, see Aronson & Carlsmith, 1968; Carlsmith, Ellsworth, & Aronson, 1976; Campbell, 1957; Campbell & Stanley, 1966; Cook & Campbell, 1979.)

Internal Validity

If the design and structure of a study are such that one can confidently conclude that the independent variable caused systematic changes in the dependent variable, then the study is said to have *high internal validity.* On the other hand, if a study leaves us with plausible alternative interpretations of the observed relation between the independent and dependent variables, then it is said to have *lower internal validity.* The level of internal validity depends on the number of plausible alternative explanations for the results, which is largely determined by the design and structure of the study.

Correlational studies generally have lower internal validity than do experimental studies, because of the former's relatively higher probability that there exists some additional unmeasured (and confounded) variable that actually caused the observed relation between the independent and dependent variables. Experimental studies, in which the independent variable is manipulated and participants are randomly assigned to levels of the independent variable, generally have higher levels of internal validity. If the independent variable is systematically related to the dependent variable in an experimental study, then one can be fairly confident that the independent variable (rather than some unmeasured third variable) caused the differences in the dependent variable. Obviously, some correlational studies are better than others, just as some experimental studies are better than others.

In recent years, examination of mediational processes has become a popular way to increase the internal validity of studies (see Baron & Kenny, 1986, for a discussion of how to test mediational models). A study showing that theoretically derived mediational variables relate to the independent and dependent variables in an expected way is more convincing than is a similar study that shows the same independent–dependent variable relation but has not assessed the mediational process. The reason for the superiority of mediational approaches to internal validity assessment has not changed, however: The number of plausible alternative explanations for the relation between the independent and dependent variables is reduced by the successful demonstration of the proposed mediational process.

External Validity

External validity typically refers to the generalizability of the results of a study to other (usually real world) settings or populations. All else being equal, external validity is assumed to be a function of the similarity of the sample, setting, and empirical realizations of the variables in the target study to the population, setting, and empirical realizations of the variables in the target setting or population to which one wishes to generalize.

The level of external validity also is usually assumed to be higher in field studies (correlational or experimental) than in laboratory experiments.[1] This assumption is generally defended by noting that the laboratory setting has some artificial features lacking in the real world (e.g., the research participant knows he or she is in an experiment), and that it doesn't have some of the extraneous features present in the real world (e.g., multiple levels of extraneous causal factors).

These assumptions are somewhat controversial for at least two very different reasons. First, if a study has low internal validity—if it doesn't clearly demonstrate a causal relation between the independent and dependent variables—then there is nothing to generalize (e.g., Banaji & Crowder, 1989; Carlsmith et al., 1976). Many field studies are low in external validity precisely because the complex mix of uncontrolled factors that make them seem so attractively "real" also ruin their internal validity.[2]

Second, the assumptions are called into question when one considers what it is that is supposed to generalize. If one is asking whether the specific name-calling anger manipulations that work on 5-year-old children will have the same effect on 35-year-old adults, then it is unlikely that high external validity (in this case, generalization across participant populations) will exist regardless of whether the target study was done in an artificial laboratory or in a naturalistic field setting. "Pooh-pooh head" just doesn't pack the same "punch" for 5- and 35-year-old people. However, if we are asking whether a general psychological process—such as "insults increase aggressive behavior"—holds across participant populations, then lexically different but conceptually equivalent insult manipulations allow an empirical test of the external validity. In this psychological process sense, the external validities of studies done in laboratory versus field settings are entirely an empirical matter, and there appears to be no reason to assume that the latter type of study generalizes across settings or populations with greater relative frequency than does the former (cf, Dipboye & Flannigan, 1979). Indeed, if the field studies are (on average) lower in internal validity than are laboratory studies, then they almost certainly will be lower (on average) in external validity as well (Banaji & Crowder, 1989; Berkowitz & Donnerstein, 1982; Mook, 1983). We comment further on implications of the "what is supposed to generalize" question in a subsequent section. Our primary concern is with the external validity of laboratory experiments, rather than with the external invalidity of field studies, so we turn to that issue next.

From our perspective, there are two key points in understanding the fallacy of most attacks on the external validity of artificial

[1] Throughout this article, we will use the terms *field, naturalistic, real world,* and *ecologically valid* interchangeably.

[2] Of course, field experiments that clearly demonstrate the causal effects of the independent variables on the dependent variables are particularly impressive because their high internal validity occurred in spite of the noise created by other uncontrolled factors.

laboratory experiments. One concerns a set of supposedly rampant experimenter and participant biases that traumatized social psychology for more than a decade. This set of biases is usually discussed in terms of experimenter demands, participant evaluation apprehension, unrepresentativeness of participant populations, and participant guessing of hypotheses. Although such biases can occur in any type of research and therefore require attention in the planning and implementation of any research project, there now appears to be no justification for the extreme view that they are so prevalent and powerful as to make laboratory research automatically suspect (Berkowitz & Donnerstein, 1982; Carlsmith et al., 1976; Kruglanski, 1975). Indeed, the evidence suggests that demand characteristics and evaluation apprehension work against the experimenter's hypothesis, at least when undesirable behaviors such as aggression are the target of laboratory investigations (Berkowitz & Donnerstein, 1982).

The second key point concerns a misunderstanding about what is supposed to generalize, as mentioned previously. Defenders of laboratory experiments have clearly outlined the view that the primary goal of most laboratory research is the development of theories designed to explain underlying processes and mechanisms. Furthermore, it is these theoretical principles that one wishes to generalize, not the specific characteristics of the sample, setting, manipulation, or measure (e.g., Banaji & Crowder, 1989; Berkowitz & Donnerstein, 1982; Henshel, 1980; Kruglanski, 1975; Mook, 1983). Berkowitz and Donnerstein (1982) also noted that ecological validity facilitates population estimates, but that laboratory research is seldom concerned with establishing such estimates. Instead, most laboratory studies are designed to test causal theoretical propositions, a task that is frequently impossible for ecologically valid designs.

Mook (1983) described four cases in which the artificial lab setting is not only acceptable but is actually preferred to the real world setting:

> First, we may be asking whether something can happen, rather than whether it typically does happen. Second, our prediction may . . . specify something that ought to happen in the lab. . . . Third, we may demonstrate the power of a phenomenon by showing that it happens even under unnatural conditions that ought to preclude it. Finally, we may use the lab to produce conditions that have no counterpart in real life at all. . . . (p. 382)

In other words, the goal of most laboratory research is to discover theoretical relations among conceptual variables that are never sufficiently isolated in the real world to allow precise examination. In those rare cases where the researcher can conduct highly controlled field experiments on appropriate participant populations and in appropriate settings, the field experiment might well be preferred over its laboratory counterpart. However, in many cases lab-based research may be the only way, the only affordable way, or the only ethical way to test key theoretically derived propositions. This last feature, concerning ethics, is particularly relevant to aggression research. How could one ethically induce real-world participants to actually aggress against others in natural settings? For certain fairly simple hypotheses, ingenious field studies of aggression can be conducted (e.g., Kenrick & MacFarlane, 1984; Reifman, Larrick & Fein, 1991), but in general both the ethical and financial costs are prohibitive.

Thus, the key feature of most laboratory experimentation—high internal validity—makes it the research setting of choice for much of the important work done in psychological science. In laboratory experiments, one can create and test theories of how conceptual variables interrelate. We believe that one can also usually expect the relations among conceptual variables to generalize. This expectation, in turn, provides the basis for modifying or (in some cases) constructing real-world environments to produce desired results. Whether one is designing brief psychotherapies for alleviation of depression, group interactions for reduction of racial prejudice, or mathematics instruction modules for teaching of algebra, an understanding of the causal relations among the conceptual variables is the crucial first step. That step is often best taken in the artificial laboratory environment. Indeed, to the extent that increasing the similarity of the laboratory to the real world interferes with the internal validity of the study, external dissimilarity (along with internal validity) is strongly preferred (Banaji & Crowder, 1989; Berkowitz & Donnerstein, 1982; Henshel, 1980; Kruglanski, 1975; Mook, 1983; Roediger, 1991).

External Validity of Trivial Experiments Revisited

One critical implication of the "theories generalize" defense of laboratory research is that theoretical relations among variables in laboratory settings do in fact generalize. Framed differently, many external validity attacks on laboratory experimentation are based on the assumption that laboratory findings do not generalize. Our basic question in this article, illustrated by the wife battering and the electric shock scenarios, concerns whether relations among conceptual variables examined in the artificial confines of the laboratory really do generalize to the real world. If the answer is "no," then the value of the trivial laboratory experiment must rely solely on the other defenses sketched in the preceding section and detailed by the authors cited in that section. However, if laboratory experiments do, in general, produce findings about relations among conceptual variables that are similar to those found in comparable real-world settings, then researchers might be encouraged to increasingly emphasize laboratory studies, in large part because of their ethical acceptability, economy, precision, and internal validity. To our knowledge, no authors have systematically examined laboratory-based and real-world-based studies in a broad research domain to see if the emergent conceptual relations tend to be the same in the two types of studies, though such comparisons have appeared for specific sets of studies (e.g., Bem & Lord, 1979; Berkowitz & Donnerstein, 1982; Dipboye & Flanagan, 1979). For instance, Berkowitz & Donnerstein (1982) briefly summarized a few studies of the relation between laboratory aggression and real-world measures of aggression. They also summarized a few real-world and laboratory studies of the weapons effect. In both cases, the artificial laboratory aggression phenomena also occurred in their real-world counterparts.

Our empirical approach to the external validity question in the aggression domain is similar, and is analogous to the null hypothesis strategy common throughout psychology and other sciences. Specifically, we examine the null hypothesis that trivial laboratory studies of aggression do not yield the same independent–dependent variable relations that naturalistic studies have convincingly shown to exist. To test this null hypothesis, one first must find aggression domains in which the naturalistic studies have yielded convincing (i.e., apparently internally valid) results. This proved to be a difficult task, precisely because naturalistic studies tend to be low on internal validity. Indeed, various reviewers of this work sometimes offered alternative explanations for the naturalistic findings that we had deemed sufficiently convincing. Ignoring this problem for the moment, we come to the second task: One must find laboratory studies of the same aggression phenomenon and assess the comparability of the findings between laboratory and naturalistic setting studies.

Classification of Aggression

There are many classification schemes used in the study of aggression, especially if aggression by nonhuman species (e.g., predatory, defensive) is considered. This article focuses exclusively on human aggression. Geen (1990) noted that human aggression can be defined by the presence of three features: (a) one person (the perpetrator) delivers noxious stimuli to another person (the victim), (b) the perpetrator delivers the noxious stimuli with the intent to harm the victim, and (c) the perpetrator expects that the noxious stimuli will have their intended effect.

Even this definition includes several types of aggression that can be usefully distinguished. One useful distinction is between physical and verbal aggression. *Physical aggression* typically involves direct physical attacks on the victim using body parts (e.g., limbs, teeth) or weapons (e.g., clubs, knives, guns), but may also involve more indirect attacks such as stealing or destroying the victim's property. *Verbal aggression* typically involves verbal attacks on the victim's self-image via face-to-face confrontation, but may also involve more indirect attacks designed to harm the victim either by causing the occurrence of some negative event or by preventing the occurrence of some positive event. Office politics and playground "tattling" are examples of indirect verbal aggression. To the extent that physical and verbal aggression mean the same thing to people, the distinction becomes less crucial. However, if different types of people differentially use these two types of aggression, then the distinction becomes important. As we shall see, men and women seem to differentially use these two types of aggression in both real-world and laboratory settings.

Prototypical Methods of Assessing Laboratory Aggression

In this section we briefly describe the major laboratory paradigms used to assess physical and verbal aggression. Special attention is paid to laboratory measures of physical aggression because they have received the most criticism.

Physical Aggression

The aggression machine paradigm (Buss, 1961) has been the primary laboratory procedure used to measure direct physical aggression, although alternative procedures have been developed and been found useful (e.g., Berkowitz, 1962; Taylor, 1967). In the aggression machine paradigm, a participant and a confederate are told that the study is concerned with the effects of punishment on learning. Using a rigged lottery, the real participant is selected to be the teacher and the confederate is selected to be the learner. The participant presents stimulus materials to the confederate who attempts to learn them. Before the learning task begins, the participant is sometimes angered by the confederate. When the confederate makes an incorrect response on a trial, the participant is told to punish him or her by means of electric shock. By using different buttons, the participant can control the intensity of shock given to the confederate. The shocks, for example, may range in intensity from *just perceptible* (e.g., button 1) to *excruciatingly painful* (e.g., button 10). The dependent variable is the intensity of shock the participant gives the confederate.

In some experiments, shock duration is also controlled by holding down the shock button for the desired duration. Some researchers have used noxious stimuli other than electric shocks (e.g., noise blasts, heat pulses). In other modifications of the aggression machine paradigm, the participant subtracts money from the confederate's pay whenever he or she makes an incorrect response on a trial.

Verbal Aggression

In the laboratory, verbal aggression is often assessed by recording a participant's vocal comments to a confederate, and counting the frequency of attacks or other negative verbal statements. In a study by Wheeler and Caggiula (1966), for example, male naval recruits evaluated opinions expressed by a confederate on various topics (e.g., religion, war, sex, liquor). On most of the topics, the confederate expressed socially undesirable opinions. On the topic of religion, for instance, the confederate said: "I think my religion is best, and I don't think the others are worth a damn. . . . If I had my way, all other religions would be illegal." The participant then was given an opportunity to comment on the confederate's opinions. Because the confederate could presumably overhear the participant's evaluations of him, it was possible for the participant to make direct verbal attacks against the confederate. The dependent variable was whether or not the participant made extremely aggressive evaluations of the confederate (e.g., that the confederate was an "ass," "idiot;" that he was "crazy," "nuts," "insane," etc.; that he should be "locked up," "shot," "deported," "beaten up," "tortured," etc.).

Indirect measures of verbal aggression are more common in laboratory experiments than are direct measures. Generally, a confederate or experimenter first provokes the participant. Rather than confronting the confederate or experimenter face to face, the participant uses a pencil-and-paper measure to evaluate him or her. The participant is led to believe that negative ratings will harm the confederate or experimenter in some way. In one study (Rohsenow & Bachorowski, 1984), for example, the participant was told to trace a circle as slowly as possible. After this task was completed, the experimenter burst in the room, introduced himself as the supervisor who had been observing through a one-way mirror, and contemptuously stated, "Obviously, you don't follow instructions. You were *supposed* to trace the circle as slowly as possible without stopping but you clearly didn't do this. Now I don't know if we can use your data." The experimenter paused, then continued (interrupting the participant if he or she tried to respond), "Do it over again." After the experiment, the participant completed an evaluation form for each member of the lab staff, including the obnoxious experimenter. The form asked the participant to rate each staff member on 7-point scales as to whether he or she was effective in performing duties, was a capable employee, was likeable, made the participant feel comfortable, showed respect for the participant, and should be rehired. The evaluations were placed in a sealed envelope and were allegedly sent to the principal investigator to be used in future hiring decisions. The participant could therefore harm the experimenter's chances of being rehired by evaluating him in a negative manner.

Convergence of Laboratory Aggression Measures

In a quantitative review of laboratory aggression measures, Carlson, Marcus-Newhall, and Miller (1989) found that correlations between effect-size estimates of various measures of physical aggression (i.e., number, duration, and intensity of punishments) ranged from .70 to .88. Written and physical response measures of aggression were positively correlated (overall $r = .71$). The correlation between written and oral aggression effect-size estimates also was positive ($r = .42$). Thus, these different laboratory measures of aggression appear to be tapping the same conceptual variable. So whatever these various laboratory aggression paradigms are assessing, they appear to be doing so consistently. The larger question concerns whether these paradigms assess true aggression; the next section examines this question empirically.

Measures of Real-World Aggression

In the real world, most extreme acts of aggression are violent crimes, which the Federal Bureau of Investigation (FBI) classifies as murder, aggravated assault, forcible rape, and robbery (U.S. Department of Justice, 1994). Laboratory measures of aggression have few surface features in common with violent crimes. It is this lack of surface similarity that, in our view, leads to the overly pessimistic critiques of the value of laboratory aggression paradigms. However, if laboratory aggression does reflect the participant's intention to harm the confederate, then laboratory studies can help us understand the causes and correlates of violent crime.

Method

General Strategy

If the trivial laboratory paradigms described above do in fact measure aggression in an externally valid way, if laboratory aggression means the same thing to participants as real-world aggression, then there should be considerable correspondence between the effects of the same conceptual independent variables on laboratory and real-world aggression measures. Moreover, individual differences in aggressiveness observed in the real world should also be observed in the laboratory. We expected to find considerable correspondence between the results from studies using laboratory and real-world measures of aggression.

We began this study by creating a list of situational and individual difference variables that we believed a priori to relate to real world aggression. We then examined the published literature to compare the effects of these variables on real-world and laboratory aggression. Table 1 lists the situational and individual difference variables we examined. Table 1 also indicates for each variable whether there was sufficient empirical support to warrant a comparison.

Meta-analytic procedures were used to integrate the results from studies conducted inside and outside the laboratory.[3] Although meta-analytic procedures can be used to combine the results from two studies,

[3] To obtain a weighted average of the sample correlations, r_+, we first obtained a weighted average of Fisher's r-to-z transformation values in which each z value was weighted by the inverse of its variance (i.e., $N - 3$). Thus, correlations based on larger sample sizes received more weight than did correlations based on smaller sample sizes. Once a 95% confidence interval was obtained for the population z value, it was transformed to a 95% confidence interval for the population correlation, ρ (see Hedges & Olkin, 1985, pp. 235–236).

The standardized mean difference was defined as $d = (M_E - M_C)/SD$, where M_E and M_C are the sample means for the experimental and control groups, respectively, and SD is the pooled estimate of the population standard deviation. To obtain a weighted average of the sample standardized mean differences, d_+, each d value was weighted by the inverse of its variance, $[2(n_E + n_C)n_En_C]/[2(n_E + n_C)^2 + n_En_Cd^2]$. A 95% confidence interval also was calculated for the population standardized mean difference, δ (see Hedges & Olkin, 1985, pp. 110–113).

Cohen (1988) has offered conventional values for "small," "medium," and "large" effect-size estimates. For the standardized mean difference, the conventional values for small, medium, and large effects are $d = 0.20$, $d = 0.50$, and $d = 0.80$, respectively. For the correlation coefficient, the conventional values for small, medium and large effects are $r = .10$, $r = .30$, and $r = .50$, respectively. According to Cohen, most of the effects in the social sciences are small to medium.

Table 1
Review Status of the Individual Difference and Situational Variables Originally Selected for Review

Individual difference variables			Situational variables		
Variable	Keywords	Review status	Variable	Keywords	Review status
Sex	NA	Kept	Provocation	NA	Kept
Trait aggressiveness	Buss, BDHI, Aggression Questionnaire, AQ, trait aggress*	Kept	Alcohol	NA	Kept
			Media violence	NA	Kept
			Anonymity	Anonym,* deindividuat*	Kept
Type A personality	Type A, Type B	Kept	Temperature	NA	Kept
Sex role orientation	Sex, gender, sex role,* gender role*	Dropped[a]	Frustration	Frustrat*	Dropped[b]
			Self-awareness	Self-awarene,* mirror	Dropped[a]
Attitudes toward violence	Attitud*	Dropped[c]	Weapons effect	Weapon,* gun*	Dropped[a]
Rape myth beliefs	Rape myth,* attitud,* belief*	Dropped[d]			
Biological differences	XYY, hormone,* insulin, serum cholesterol, glucose genetic,* biolog*	Dropped[d]			

Note. NA = not applicable because an extant meta-analysis was used.
[a]These variables were dropped because there were too few field studies with similar measures to laboratory experiments. [b]This variable was dropped because it is frequently defined in different ways, and because a more precise variable, provocation, was kept. [c]This variable was dropped because only two laboratory experiments were located, and meta-analytic procedures were deemed inappropriate for such a small number of studies. [d]These variables were dropped because only one laboratory experiment was located.

in this article meta-analytic procedures were used only if there were at least three independent studies. When possible, we also tested whether the mean effect-size estimates differed for the different types of studies. A .05 significance level was used for all tests.

Literature Search Procedure

Whenever possible, extant meta-analytic literature reviews were used. In those cases in which such reviews did not already exist, we attempted to conduct an exhaustive search of the literature. Table 1 also indicates whether a new literature search was conducted for each variable. The PsycINFO computer database (1974–1996) was searched using the key words *aggress** and *violen**. The asterisk at the end of the key word indicates that all forms of the key words (e.g., *aggress, aggressive, aggressiveness, aggression, aggressed, aggressor*) are included in the search. The *aggress** and *violen** key words were paired with the key words for the individual difference and situational variables requiring a new search, as shown in Table 1. The search was restricted to studies published in English and to studies that used human participants.

Results

Individual Differences and Aggression

In this section we assess the correspondence between findings from field studies and laboratory studies, focusing on individual difference variables. The following individual difference variables were examined: sex differences, trait aggressiveness, and Type A coronary-prone behavior pattern.

Sex Differences

Studies conducted outside the laboratory. Archival data on violent crime rates clearly show that men commit more murders and assaults than do women. This sex effect occurs in virtually every murder and assault rate data set that can be found. For instance, a century ago, Dexter (1899) showed that the male rate of assault in New York city in the years from 1891–1897 was more than 11 times larger than the female rate. More recently, the 1993 FBI *Uniform Crime Report* showed that the male murder rate was almost 10 times larger than the female rate (U.S. Department of Justice, 1994).

To investigate sex effects in field studies, we recombined studies from Bettencourt and Miller's (1996) meta-analysis.[4] The results showed

[4] The main focus of the Bettencourt and Miller (1996) meta-analysis is sex differences in aggression as a function of provocation. For instance, they showed that males are more aggressive than females in both neutral and provoked conditions, but that the difference is smaller in provoked

that males aggressed more than females when the aggression was a physical act such as horn honking, but not when the aggression was a verbal act such as making negative remarks (see Table 2).

Studies conducted inside the laboratory. To investigate sex effects in laboratory studies, we also recombined studies from Bettencourt and Miller's (1996) meta-analysis. The results showed that male participants were more physically and verbally aggressive than were female participants. Sex differences in verbal aggression, however, were quite small.

In summary, our analysis suggests that men are more physically aggressive than are women both inside and outside the laboratory. Sex differences in verbal aggression are small to trivial in both contexts. The setting (i.e., inside versus outside the laboratory) did not significantly influence the magnitude of sex differences found for either type of aggression.

Trait Aggressiveness

Informal observation suggests that some people are especially likely to become involved in aggressive interactions. The personality trait of aggression is referred to as *trait aggressiveness.* Trait aggressiveness can be operationally defined using (a) self-report personality scales; (b) aggression nominations by others (e.g., peers, teachers, counselors); or (c) violent histories. The most widely used self-report measure of trait aggressiveness is the Buss-Durkee Hostility Inventory (BDHI; Buss & Durkee, 1957). Sample items from the BDHI include "Once in a while I cannot control my urge to harm others" and "I often find myself disagreeing with people."

In most field studies, participants are individuals with violent histories. In most laboratory studies, participants are college students. Because violent individuals were expected to have more aggressive personalities than college students, stronger relations were expected for field studies than for laboratory studies.

Studies conducted outside the laboratory. The BDHI has been used to successfully discriminate between violent and nonviolent criminals (Gunn & Gristwood, 1975; Selby, 1984; Syverson & Romney, 1985); between domestically violent and nonviolent men (Maiuro, Cahn, Vitaliano, Wagner, & Zegree, 1988); between violent and nonviolent patients (Lange, Dehghani, & De Beurs, 1995; Maiuro et al., 1988); between violent and nonviolent alcoholics (Renson, Adams, & Tinklenberg, 1978); and between violent and nonviolent adolescent offenders (Boone & Flint, 1988; Lothstein & Jones, 1978). Scores on the revised BDHI also are positively correlated with peer-nominated aggression in college students (Buss & Perry, 1992) and with self-reported involvement in physical fights (Archer, Holloway, & McLoughlin, 1995; Stanford, Greve, & Dickens, 1995). Meta-analysis revealed a substantial positive correlation between trait aggressiveness, as measured by the BDHI, and real-world aggression (see Table 2).

Studies conducted inside the laboratory. Scores on self-report trait aggression questionnaires have been found to correlate positively with laboratory measures of physical aggression (Bushman, 1995; Giancola & Zeichner, 1995; Hammock & Richardson, 1992; Knott, 1970; Larsen, Coleman, Forbes, & Johnson, 1972; Leibowitz, 1968; Scheier, Buss, & Buss, 1978; Shemberg, Leventhal, & Allman, 1968), although null results have been reported (Muntaner et al., 1990). The level of laboratory measures of physical aggression has been found to be higher for male adolescent delinquents with a history of violence than for male adolescent delinquents with no history of violence (Hartman, 1969), and for young male offenders in a maximum security penal facility than for male college students (Wolfe & Baron, 1971). The level of laboratory measures of physical aggression also has been found to be higher for high school students nominated by their counselors to be aggressive than for high school students nominated by their counselors to be nonaggressive (Shemberg et al., 1968), and for third-graders nominated by their peers to be aggressive than for third-graders nominated by their peers to be nonaggressive (Williams, Meyerson, Eron, & Semler, 1967). Meta-analytic procedures found a medium-sized correlation between trait aggressiveness and laboratory aggression (see Table 2).

conditions. They also showed that the sex difference is larger when the aggressive behavior involves delivery of an aversive physical stimulus to the victim than when it involves some type of verbal aggression. Their analyses did not focus on laboratory versus field studies, so we used their tabled results to conduct our own meta-analyses.

Table 2
Mean Effect Size Estimates and Confidence Intervals (CI) for Individual Difference and Situational Variables

Variable	N	Effect-size estimate	95% CI
Individual difference variables			
Sex[a]			
Studies conducted outside the laboratory			
Physical aggression	6	$d_+ = 0.40_b$	[0.25, 0.55]
Verbal aggression	3	$d_+ = 0.03_a$	[−0.15, 0.22]
Studies conducted inside the laboratory			
Physical aggression	37	$d_+ = 0.31_b$	[0.23, 0.38]
Verbal aggression	18	$d_+ = 0.13_a$	[0.03, 0.24]
Trait aggressiveness			
Studies conducted outside the laboratory	16	$r_+ = .42_b$	[.37, .47]
Studies conducted inside the laboratory	11	$r_+ = .25_a$	[.19, .30]
Type A personality			
Studies conducted outside the laboratory	3	$d_+ = 0.97_b$	[0.71, 1.23]
Studies conducted inside the laboratory	9	$d_+ = 0.34_a$	[0.18, 0.49]
Situational variables			
Provocation[a]			
Studies conducted inside the laboratory	66	$d_+ = 0.76$	[0.66, 0.85]
Alcohol[b]			
Studies conducted inside the laboratory	86	$d_+ = 0.37$	[0.29, 0.45]
Media violence[c]			
Studies conducted outside the laboratory	556	$d_+ = 0.42_b$	
Studies conducted inside the laboratory	586	$d_+ = 0.87_a$	
Anonymity			
Studies conducted outside the laboratory	5	$d_+ = 0.47_a$	[0.29, 0.65]
Studies conducted inside the laboratory	19	$d_+ = 0.57_a$	[0.45, 0.69]
Temperature[d]			
Studies conducted inside the laboratory			
Overall	28	$d_+ = 0.06$	[−0.11, 0.23]
Neutral context	12	$d_+ = 0.25$	[−0.03, 0.53]
Extra-negative context	14	$d_+ = -0.09$	[−0.33, 0.15]

Note. N = number of effect size estimates. Statistical test was the unit of analysis for the media violence variable; study was the unit of analysis for all variables. Mean effect-size estimates (inside the lab vs. outside the lab comparisons) having the same subscript are not significantly different for that variable at the .05 level. CI = confidence interval; r_+ = weighted average of the sample correlations; d_+ = weighted average of the sample standardized mean differences.
[a]Data from Bettencourt & Miller (1996). [b]Data from Bushman (in press). [c]Data from Paik and Comstock (1994). [d]Data from K. Anderson & Anderson (1996).

In summary, trait aggressiveness was positively correlated with aggression in both field studies and laboratory studies. As expected, stronger correlations were found for field studies than for laboratory studies.

Type A Coronary-Prone Behavior Pattern

Type A pattern is characterized by three major behavioral components: excessive competitive achievement striving, exaggerated time urgency, and aggression or hostility (Glass, 1977). The latter component has the most relevance to the present discussion (and to heart disease also). Type A personality can be assessed using either a self-report personality test or a structured interview. The most popular self-report personality test is the Jenkins Activity Survey (JAS; Jenkins, Zyzanski, & Rosenman, 1979). A sample item from the college student form of the JAS is "When you are studying and somebody interrupts you, how do you usually feel inside?" Response options include: (a) "I feel OK because I work better after an occasional break," (b) "I feel only mildly annoyed," or (c) "I really feel irritated because most such interruptions are unnecessary." Type B's tend to choose response (a), whereas Type A's tend to choose response (c). For reasons similar to those given for trait aggressiveness, we expected larger aggression differences between Type A's and Type B's in field studies than in laboratory studies.

Studies conducted outside the laboratory. Strube and his colleagues (Strube, Turner, Cerro, Stevens, & Hinchey, 1984) compared JAS scores for violent and nonviolent women. The sample of violent women was selected from a population of women under treatment for child abuse; the sample of nonviolent women was selected from a population of mothers of preschoolers in the same city. The nonviolent and violent women were matched according to the age of their child. The violent women were classified as Type A more often than were nonviolent women. In another study (Schell, Cachon, Ganjavi, & Porporino, 1986), inmates with a violent criminal background were classified as either assaulters or nonassaulters depending on whether they had been charged with some act of physical aggression (murder, attempted murder, or a physical assault of a guard or fellow inmate) over the past year. Inmates completed the Behavior Activity Profile (Matteson & Ivancevich, 1982)—a self-report measure of Type A pattern—as part of a battery of questionnaires. The results showed that most assaulters were classified as Type A's, whereas most nonassaulters were classified as Type B's. In another study (Hurlbert, Whittaker, & Munoz, 1991), abusive husbands were classified as Type A's, as measured by the JAS, significantly more often than were nonabusive husbands. Meta-analysis of the results from studies conducted outside the laboratory found a strong relation between Type A personality and aggression (see Table 2).

Studies conducted inside the laboratory. Most laboratory studies have found that Type A's, in comparison with Type B's, are more physically aggressive (Baron, Russell, & Arms, 1985; Carver & Glass, 1978; Check & Dyck, 1986; Holmes & Will, 1985; Llorente, Bernardo, de Flores, & Valdes, 1985; Strube et al., 1984), although a few studies have found null results (Berman, Gladue, & Taylor, 1993; Muntaner, Llorente, & Nagoshi, 1989). Meta-analytic procedures found that, on average, Type A's behaved significantly more aggressively in the laboratory than did Type B's (see Table 2).

In summary, Type A's behaved more aggressively than did Type B's both inside and outside the laboratory. As expected, stronger effects were obtained for field studies than for laboratory studies.

Situation Variables and Aggression

In this section we assess the correspondence between findings from field studies and laboratory studies, focusing on situational variables. The effects of the following situational variables on human aggression were examined: provocation, alcohol, media violence, anonymity, and temperature.

Provocation

By *provocation,* we mean acts of harm committed by the target against the person whose aggressive behavior is eventually assessed. In the real world, provocations are quite common. They may involve cutting off another driver on the freeway, stealing someone's property, verbally insulting someone, or physically attacking someone. In most laboratory studies, provocations consist of physical attacks (e.g., painful shocks or noise blasts) or verbal insults.

Studies conducted outside the laboratory. Crime statistics clearly demonstrate that provocation is the major source of real-world aggression. In the breakdown of murders, for instance, the vast majority are the result of some intense, personal provocation. In the United States in 1993, only 27% of all murders were the result of some other felony activity, such as robbery. Of the remaining 1993 murders for which the circumstances are known, 73% were classified by the FBI as being due to arguments. Another 5% were due to romantic triangle disputes, and another 7% resulted from alcohol- and drug-related brawls (U.S. Department of Justice, 1994). Thus, the common circumstances surrounding murder involve attacks on one's self-esteem, public image, or family structure, all of which are types of provocation.

Several more formal studies have examined effects of provocation on violence in natural settings. For instance, Curtis (1974) examined a U.S. national sample of police reports, and found that provocation was common in homicide and aggravated assault, less common in robbery, and least common in forcible rape. Similarly, Davis (1991) examined psychiatric inpatient violence, and found provocation to be an important situational predictor.

Searching for tests of provocation in field experiments is a difficult task because field experiments do not typically include a "no provocation" control condition. Indeed, Bettencourt and Miller (1996) could find no field experiment with a control condition to include in their meta-analysis of sex differences in the effects of provocation on aggression.

Studies conducted inside the laboratoray. The meta-analysis by Bettencourt and Miller (1996) also examined provocation effects in the context of sex effects. They found that high provocation increased aggression for all possible combinations of sex of confederate and sex of participant. Overall, the provocation effect was quite large (see Table 2). In summary, provocation has a large effect on aggression both inside and outside the laboratory.

Alcohol

Studies conducted outside the laboratory. Several correlational studies have found a strong relation between alcohol intoxication and violent crime. These studies generally find that approximately 50% of the assailants were intoxicated at the time the violent crimes were committed (e.g., Beck, 1991; Beck, Kline, & Greenfield, 1988; Greenberg, 1981; Innes, 1988; MacDonald, 1961; Murdoch, Pihl, & Ross, 1990; Pernanen, 1991). Using regression analysis of U.S. violent-crime data from the period 1979 to 1988, Cook and Moore (1993) predicted that if alcohol consumption per capita decreased by just 10%, there would be a corresponding 1% decrease in homicides, 6% decrease in rapes, 6% decrease in assaults, and 9% decrease in robberies.

Studies conducted inside the laboratory. Numerous laboratory studies have investigated the relation between alcohol and aggression. Meta-analytic reviews of these studies have found that intoxicated participants are significantly more aggressive than are sober participants (Bushman, 1993, in press; Bushman & Cooper, 1990; see Table 2). The type of aggression measure used does not appear to influence the results (Bushman, in press; Bushman & Cooper, 1990). Larger effects might be obtained if ethical considerations did not prevent researchers from using higher alcohol doses in their laboratory studies (i.e., the target blood alcohol level is, at most, .10). In any case, alcohol appears to increase aggression inside and outside the laboratory.

Media Violence

The effect of violent media on aggression was expected to be larger for laboratory studies than for field studies. There are at least three reasons for making this prediction. First, laboratory studies are more effective at controlling extraneous variables than are field studies. Second, the violence is generally more concentrated in laboratory studies than in field studies. Third, the time between exposure to violent media and measurement of aggression is generally shorter in laboratory studies than in field studies.

Studies conducted outside the laboratory. It is not hard to find anecdotal examples that suggest a relation between exposure to violent media and real-world aggression. Consider the following news stories:

> On April 22, 1974, three people were murdered in a store in Ogden, Utah, by two armed men who forced them to drink liquid Drano, a caustic drain cleaner. In the court proceedings, the Assistant State Attorney General said that the accused murderers "had seen the movie *Magnum Force,* in which liquid Drano was used to kill a woman, the same month of the killings and took Drano to the [store] as a premeditated lethal weapon" ("Selby Makes One Last Plea," 1987).
>
> Another witness testified that the two men saw *Magnum Force* "three times in one day" the same month of the killings. ("Still at a Loss for 'Why'," 1987)

In a recent meta-analysis, Paik and Comstock (1994) reported that violent media has a small to medium effect on aggression in field studies (see Table 2).

Studies conducted inside the laboratory. Paik and Comstock (1994) reported that violent media has a large effect on aggression in laboratory studies (see Table 2). In one study (Bushman, 1995), for example, undergraduate psychology students were randomly assigned to view a 15-minute videotaped film segment that was either violent or nonviolent. The two videotapes were selected from a large pool of tapes because they were judged to be equally exciting but differentially violent. In addition, there were no significant differences between the two tapes on cardiovascular measures of arousal (i.e., systolic blood pressure, diastolic blood pressure, heart rate). After viewing the videotape, participants competed with an ostensible opponent on a reaction time task in which the slower responding person received a blast of noise. The results showed that participants who had seen the violent videotape set significantly higher noise levels for their "opponent" than did participants who had seen the nonviolent videotape. Viewing violence increases laboratory aggression as well as real-world aggression.

In summary, violent media increased aggression both inside and outside the laboratory. As expected, the effect of violent media on aggression was larger for laboratory studies than for field studies.

Anonymity

Consider the following news story:

> A gang of young men wielding knives and bats went on a Halloween rampage Wednesday night, assaulting several homeless people on the foot-bridge to Wards Island and leaving one of them dead among the garbage-strewn weeds, his throat slashed. The group of about 10 young men, some wearing Halloween masks, apparently attacked the homeless men for thrills. (McKinley, 1990)

What factors influence people to engage in such uninhibited antisocial behaviors as those described above? Festinger, Pepitone, and Newcomb (1952) proposed that when group members are not seen as individuals, a state of deindividuation may result, with a consequent lowering of social restraints. The terms *deindividuation* and *anonymity* often are used interchangeably (Lightdale & Prentice, 1994). An individual can achieve anonymity by being part of a group, by wearing a mask, or by performing behaviors in the dark rather than in the light. Deindividuation leads to a reduced sense of accountability. This may partially account for why bank robbers and members of the Ku Klux Klan wear masks when committing violent crimes, why crowds attending sporting events sometimes become violent (e.g., Dunand, 1986), and why violent crimes are more frequently committed during nighttime hours than during daytime hours (e.g., Meyer, 1982; Tamura, 1983).

Studies conducted outside the laboratory. We found four studies that examined the role of deindividuation on real-world aggression. Mullen (1986) conducted an archival analysis to determine whether the atrocities committed by lynch mobs could be accounted for in terms of self-attention processes. Sixty newspaper reports of lynching events were coded for information regarding group composition (i.e., number of victims, numbers of lynchers) and atrocity (i.e., occurrence or nonoccurrence of hanging, shooting, burning, lacerating, or dismembering the victim, as well as the duration of the lynching). The results showed that as lynchers became more numerous relative to the victims, atrocities increased. Mann (1981) analyzed 21 cases in which crowds were present when a disturbed person threatened to jump off a building, bridge, or tower. Analysis of newspaper accounts of the episodes showed that relatively large crowds were more likely to taunt and urge the victim to jump than were relatively small crowds. In addition, more baiting episodes occurred in the nighttime hours than in the daytime hours. Wilson and Brewer (1993) reported that the amount of conflict police encountered while on patrol was higher when a large number of bystanders (six or more) were present than when a small number of bystanders (five or less) were present. Ellison, Govern, Petri, and Figler (1995) found that drivers in convertibles or 4×4s with their tops up honked more at a stalled confederate than did drivers with their tops down. Meta-analysis of these results yielded a medium-sized effect (see Table 2).

Studies conducted inside the laboratory. In an early laboratory study on deindividuation (Zimbardo, 1969), female participants were randomly assigned to one of two groups. In the deindividuated group, participants wore large lab coats, wore hoods over their heads, and were not referred to by name. In the individuated group, participants did not wear lab coats, wore large name tags, and were were referred to by name. The results showed that participants in the deindividuated group gave a confederate longer shocks than did participants in the individuated group. A number of subsequent laboratory studies have shown that aggression is increased when participants are placed in a deindividuated state (Diener, 1976; Diener, Dineen, Endresen, Beaman, & Fraser, 1975; Lightdale & Prentice, 1994; Mann, Newton, & Innes, 1982; Paloutzian, 1975; Prentice-Dunn & Rogers, 1980, 1982; Prentice-Dunn & Spivey, 1986; Rogers, 1980; Rogers & Ketchen, 1979; Rogers &

Prentice-Dunn, 1981; Spivey & Prentice-Dunn, 1990; Taylor, O'Neal, Langley, & Butcher, 1990; Worchel, Arnold, & Harrison, 1978), although a few null results have been reported (Propst, 1979; Worchel & Andreoli, 1978). Meta-analysis of these results showed that anonymous participants behaved significantly more aggressively in the laboratory than did nonanonymous participants. The average effect-size estimate for laboratory experiments was medium in size (see Table 2).

Thus, anonymity increased aggression both inside and outside the laboratory. The type of setting (i.e., inside versus outside the laboratory) did not significantly influence the magnitude of the effect of anonymity on aggression.

Temperature

We intentionally selected temperature as an independent variable for this article because we knew from the outset that the laboratory and real-world aggression studies sometimes produced different outcomes (C. A. Anderson, 1989). We specifically wanted at least one set of contradictory findings in order to facilitate our discussions of when researchers should expect different results from laboratory and real-world measures of aggression. Such discrepancies can result in improved theory and understanding of both types of aggression. We first present the temperature results, and demonstrate that there is an inconsistency between laboratory and field studies. We follow this with a more detailed analysis of the possible sources of this discrepancy and demonstrate how this analysis led to new research designed to reconcile the laboratory and field study discrepancies.

Studies conducted outside the laboratory. Two comprehensive reviews of temperature effects on real-world aggression found striking consistency across type of study (C. A. Anderson, 1989; C. A. Anderson & K. B. Anderson, in press). Although formal meta-analytic statistical procedures were not used in those reviews, the meta-analytic strategies of exhaustively sampling the literature (with some methodological restrictions, of course), partitioning studies by important features, and combining results within these partitions were used.

Time-period studies, in which aggression rates are compared across time periods that differ in temperature, showed that numerous types of aggressive behaviors such as murders, rapes, assaults, and wife battering were relatively more frequent during the hotter periods of time. More recent time-period studies have shown that in the United States (from 1950–1995) violent crime rates are higher during hotter years than during cooler years, and that the usual summer increase in violent crime is magnified in hotter years (Anderson, Bushman, & Groom, 1996).

Similarly, geographic region studies from several different countries found that hotter regions tend to have higher aggression rates than do cooler regions. More recent analyses of U.S. violent crime rates (C. A. Anderson & K. B. Anderson, in press) show that this region effect occurs even when steps are taken to control for possible regional differences in "culture of honor" (Nisbett, 1993).

Finally, the two concomitant studies of real-world aggression reviewed by C. A. Anderson (1989) in which temperature and aggression were simultaneously assessed also yielded significant temperature effects on aggression. A more recent study of this type also found the same results. Specifically, Reifman et al. (1991) showed that major league baseball pitchers were more likely to hit batters during hotter games than during cooler games, even when amount of pitcher control (i.e., number of walks) was held constant.

In summary, the field studies show consistent temperature effects. Hot temperatures produce increases in aggression.

Studies conducted inside the laboratory. Laboratory studies of aggression have yielded inconsistent effects. Hot temperatures sometimes increase and sometimes decrease aggression (C. A. Anderson, 1989). To examine these inconsistent effects, K. B. Anderson and C. A. Anderson (1996) conducted a meta-analysis on the laboratory studies of aggression. When all of the laboratory effects of hot versus comfortable temperatures were combined, the average effect size estimate was not significantly different from zero (see Table 2).

One possible factor that might account for this null effect concerns whether there were other manipulated variables designed to raise or lower the participant's feelings of anger, annoyance, or friendliness. Specifically, Baron's (1979) negative affect escape model (NAE) predicts that hot temperatures will increase aggression when there are no other negative factors present, but will decrease aggression when there are other factors present that, taken together, would tend to heighten negative feelings. NAE postulates that aggressive motivation outweighs the desire to escape from the situation when the total amount of negative affect produced by the situation is in the low to moderate range. It further postulates that at some intermediate level of negative affect, the escape motive begins taking on relatively more importance than the aggression motive. Thus, hot temperatures should decrease aggressive behavior when other negative factors are at work, because the total amount of negative affect is so high that escape motivation overrides aggressive inclinations (see C. A. Anderson & DeNeve, 1992; C. A. Anderson, Dorr, K. B. Anderson, & DeNeve, 1996 for fuller accounts of this theory). In laboratory studies, the most common factor used to heighten negative affect has been an anger manipulation.

To test the NAE model, K. B. Anderson & C. A. Anderson (1996) categorized 26 separate effects of hot temperatures (i.e., in the 90s °F versus low to mid 70s °F) on the basis of whether other experimental factors could be expected to produce a net increase in negative affect. The results were in the predicted direction, but were not significantly different from zero. The hot temperatures almost produced a significant increase in aggression in the neutral context studies, but the predicted decrease in aggression due to hot temperatures in the negative context studies did not approach significance (see Table 2). In short, these results confirmed the earlier conclusions (Anderson, 1989) that the laboratory studies of the hot temperature effect are inconsistent among themselves, and are inconsistent with the findings from field studies.[5]

Dealing With Real-World/Laboratory Discrepancies

The main point of this article was to see whether "trivial" laboratory paradigms have good external validity when the conceptual relations (rather than specific operationalizations) are to be generalized. Our review found that convincing real-world independent–dependent variable relations in the aggression domain are usually replicated in artificial laboratory paradigms. A second important point, however, was to suggest a useful way of dealing with cases where the findings from laboratory and real-world studies are discrepant, such as in the temperature–aggression literature.

Two imprudent approaches involve simple rejection. One can reject the real-world findings as being the result of the confounds or the lack of control that typifies such studies. Alternatively, one can reject the laboratory findings as being the result of suspicion problems, trivial manipulations, or trivial measures of aggression. The former approach appears to characterize the theoretical (basic, experimental) perspective, whereas the latter appears to characterize the applied (nonexperimental) perspective.

Our view (shared by many, we hope) is that such discrepancies should serve as signals that additional conceptual work is needed, and should be followed by additional empirical work. In other words, rather than take the perspective that one "side" or the other is wrong, it may be more prudent to try to locate the source of the discrepancies in psychological processes that may differ in the two settings. One then could try to discover the conditions that lead to one versus the other type of finding (cf, Greenwald, Pratkanis, Leippe, & Baumgardner, 1986).

In general, discrepancies between field and laboratory studies should arise when key conceptual variables or processes (a) are prevalent and operate freely in the real world but are controlled in the laboratory, or (b) are prevalent and operate freely in the laboratory, but are infrequent or less prevalent in the real world.[6]

[5] Two additional points are worth noting here. First, the Neutral context and Extra-Negative context effect sizes reported in Table 2 do not include two effects from Palamarek and Rule (1979) because additional attribution and motive variables in that study showed that the behavioral results could not have resulted from the mediating factors proposed by the NAE model (see C. A. Anderson, 1989). If these two effect sizes are included, one each in the Neutral and Extra-Negative contexts, the results change slightly but the conclusions do not. Specifically, the temperature effect in the neutral context increases enough to reach significance, $d_+ = 0.26$, 95% confidence interval (0.001, 0.53). The temperature effect in the Extra-Negative context changes little, $d_+ = -0.10$, 95% confidence interval (−0.33, 0.13). Second, we believe that the NAE model is probably correct under certain circumstances.

[6] Of course, there also are many uninteresting ways the discrepancies can arise, such as poor measurement of key variables, the introductions of suspicion and experimenter demand problems, and the use of inappropriate control

One can reduce the process differences either by finding field settings that possess the same hypothesized process characteristics as the target laboratory setting, or by modifying laboratory procedures so that the extant processes better match the target field setting. In most cases—in the temperature domain as well as in virtually all theoretical domains of psychology—it is easier, cheaper, and more ethically sound to modify the laboratory setting than the real-world setting.

Discrepancies in the Temperature Domain

Step 1: Conceptual Analysis

Many studies of real-world temperature-related aggression suffer from two problems. First, most real-world studies do not control people's attempts to regulate their temperature-based discomfort. This control failure tends to reduce the effects of temperature on aggression, but only to the extent that most people can and do engage in temperature regulation. In fact, most people are able to avoid experiencing lengthy exposure to uncomfortably cold temperatures using combinations of clothing and heating equipment. However, many people are considerably less successful at avoiding uncomfortably hot temperatures. Thus, one should expect to see little evidence of cold discomfort increasing aggression in real-world settings, but considerably stronger evidence of heat discomfort increasing aggression. This is precisely what happens in real-world studies.

Second, many field studies do not assess aggression and temperature simultaneously. Thus, detailing the functional relation between temperature and aggression becomes impossible in field studies. For instance, the systematic increase in violent crimes during hotter days could be the result of anger that was exacerbated by the maximum temperature of the day (e.g., 95 °F), but it could also result from anger produced during the moderately hot part of the day (e.g., 85 °F). In other words, moderately uncomfortable and extremely uncomfortable temperatures are confounded in the real world, so in studies in which it is not known when (and at what temperature) the primary instigating incident took place, it is difficult to rule out some alternative explanations of the heat–aggression relation (C. A. Anderson, 1989).

Laboratory studies may also contribute to the discrepancy between laboratory and real-world temperature-related aggression. First, laboratory studies are seldom designed to assess accurately the kind of spontaneous outburst that uncomfortable temperatures are most likely to instigate. Second, even though the temperature effect presumedly "works" primarily when the situational context is somewhat ambiguous, laboratory studies of temperature effects do not typically include an ambiguous provocation condition. Third, in some laboratory studies the aggressive behavior has both harm and control features built in to the aggression measure. For example, in the reaction-time paradigm (Taylor, 1967), the most commonly used one in laboratory studies of the heat effect, participants compete on a series of trials and deliver aversive stimulation to an ostensible opponent after each trial. Participants' attempts to control their "opponent's" punishment levels (e.g., a tit-for-tat strategy) may overwhelm effects of relatively weak instigators of aggression, such as temperature.[7]

One or more of these conceptual process differences between studies from laboratory and real-world settings might well underlie the observed differences in heat effects in field versus laboratory studies. This theoretical analysis, instigated by observing lab–field discrepancies, suggests that the two settings might produce similar effects when these conceptual process differences are reduced or eliminated. Specifically, one should expect an increase in aggression at uncomfortable temperatures in laboratory settings when participants cannot avoid the uncomfortable temperatures, when spontaneous outbursts are assessed, and when alternative behavioral goals do not impede aggression. Furthermore, this analysis suggests that under these conditions one should find an increase in outburst aggression at uncomfortably cold as well as uncomfortably hot temperatures.

groups. Although important in some of these studies, these factors will not be systematically discussed in this article.

[7] Other more powerful instigators of aggression, such as provocation, may override control concerns. In some cases, use of a tit-for-tat strategy can actually exaggerate the effects of provocation. Studies in which provocation is manipulated by having the opponent deliver high (versus low) levels of shock to the participant have exactly this characteristic.

Step 2: Empirical Follow-Up

Recent empirical work on these issues has confirmed at least some aspects. C. A. Anderson, Dorr, K. B. Anderson, and DeNeve (1996; see also C. A. Anderson & K. B. Anderson, in press) report a series of laboratory experiments designed to reduce the conceptual differences between real-world and laboratory studies of temperature-related aggression. In brief, they modified the Taylor (1967) competitive reaction-time paradigm and obtained the predicted increases in aggression at both hot and cold temperatures. Interestingly, this effect occurred only when the participants were provoked in an ambiguous way, were not threatened by possible retaliation, and could not escape the aversive environment. Furthermore, this temperature effect occurred only on an aggression measure that was predicted to be the best indicator of an anger-based outburst—the first trial on which the participant was allowed to aggress against his or her "opponent."

This one set of studies does not resolve all the issues in the temperature aggression literature, of course. We discuss it in this article merely to demonstrate the benefit of using the occasional laboratory–real world discrepancies to further the understanding of the phenomenon in general—across time, settings, populations, and specific empirical realizations of conceptual variables.

General Discussion

At the outset, we presented two aggression scenarios—one real-world example of wife battering and one laboratory example of trivial aggression—and asked if there was any reason to believe that the findings from the latter could inform us about the former. In our view, the answer must be a resounding "yes." When careful conceptual analyses of both types of situations are conducted, and when solid empirical research methods are used, findings about the relations between conceptual variables will generalize from the laboratory to the real world, and vice versa.

Summary of the Real–Trivial Comparisons

The various aggression literatures sampled for this article provide strong empirical support for the laboratory researchers' faith in their trivial laboratory aggression paradigms. All of the individual difference variables (sex, trait aggressiveness, Type A personality) and most of the situational variables (provocation, alcohol, media violence, anonymity) consistently influenced aggressive behavior both inside and outside the laboratory. Such a convergence of findings in such disparate settings confirms the validity of both types of studies.

Even in the one case where real-world and trivial aggression differed, the temperature domain, the differences appear to be a function of different psychological processes at work. It is not the case that laboratory studies are less externally valid than are field studies, nor is it the case that field studies are necessarily flawed with fatal confounds. A more accurate view of this particular discrepancy between field and laboratory studies is that the theoretical analysis of the temperature aggression phenomenon was (and may still be) incomplete. There are processes at work in each setting that are relatively unimportant in the other. Once such processes were identified and at least partially equated, comparable findings emerged.

What Generalizes?

It is important to note that real-world aggression measures (e.g., violent crime) share few surface features with laboratory aggression measures (e.g., delivery of electric shock). However, these aggression measures do share the conceptual features of delivering a noxious stimulus to a victim with the intent and expectation of harming the victim. As noted by numerous defenders of laboratory research, what one should expect to generalize are theories. In other words, the conceptual relations among variables are expected to be similar in quite dissimilar situations. The aggression literature, often the most volatile domain in this external validity debate, clearly shows considerable consistency between the real-world and the trivial aggression measures. In summary, we believe the studies that we have reviewed conclusively demonstrate that the trivial laboratory paradigms of aggression are not at all trivial; they are quite high in external validity at this conceptual level of generalizability. Furthermore, we believe that other so-called trivial laboratory paradigms in other domains can accurately be seen as high in external validity. It would be interesting to sample several other domains where the schism between the laboratory-based researchers and their real-world-research cousins is large, but such a task is beyond our expertise and our page limits.

When Laboratory and Real-World Findings Collide

When real-world and laboratory measures (of aggression or of other conceptual dependent variables) do not yield comparable results, the research and application communities would be best served not by dismissing one or the other set of findings. Rather, such discontinuities should stimulate additional conceptual and empirical work designed to discover what additional forces may be at work. Certainly, it is useful to look for standard methodological "problems" in each domain, such as demand characteristics or suspicion problems in laboratory settings and third variable confounds in real-world settings. But we suspect that the methodological perspectives and preferences of scholars committed to one or the other of these two types of settings (which roughly corresponds to those with an applied vs. theoretical orientation) frequently leads to biased interpretations of relevant studies. The applied researchers may too quickly posit and accept demand characteristics explanations of laboratory findings, whereas the theoretical researchers may too quickly posit and accept confounded variable explanations of real-world findings.

These same preferences, and the hypothesis confirming processes instigated by them, may frequently steer researchers away from potentially valuable insights about additional processes that are active in one setting and dormant in the other. In a sense, the external validity debate itself forces many scholars to take a position in favor of one type of study over the other, and thereby makes it harder to engage in productive conceptual reanalysis of domains in which the two types of data conflict.

What we are arguing for, instead, is a new set of default assumptions. First, instead of assuming that laboratory-based research is low in external validity, we should assume that it is high in external validity at the level of conceptual relations among variables. Second, instead of assuming that discrepancies between real-world and laboratory studies are due to the internal invalidity of the field studies, we should assume that the discrepancy itself is informative of additional factors at work in at least one of the research contexts. These assumptions naturally lead to a reexamination of the phenomena under study and a search for additional processes that differentiate the research contexts. Finding and testing such additional processes is then likely to improve the theoretical underpinnings and, eventually, the applied aspects of the phenomenon.

Caveats

Although we argue for this new pair of assumptions about external and internal validity, we recognize that researchers also need to be aware of potential invalidating features. In the laboratory domain, one must be sure that participants understand the dependent variable in the way intended by the experimenter. If delivery of electric shock is supposed to measure aggression and only aggression, then the conditions must be set up so that participants believe that the shocks they deliver will harm the victim. This match between the conceptual variable and the specific realization (or operationalization) of it is crucial for both conceptual independent and dependent variables (Carlsmith et al., 1976). It also is as crucial for real-world variables as it is for laboratory variables. One must similarly be aware of the internal validity problems characteristic of field studies. What additional processes might be operating that could be confounded with the independent variable of interest?

Roles of Trivial Laboratory and Real-World Studies

This shift in perspective on the external validity of laboratory-based findings implies that laboratory studies are more valuable than usually is granted by many people in psychology, in other sciences, in funding agencies, in Congress, and in the general population. This bias against laboratory studies is not unique to psychology, of course; the occasional public ridicule heaped on medical researchers for using animal studies to make judgments about potential carcinogens demonstrates the same bias against generalization of general theoretical relations. Some people still argue that cigarette smoking has not been proven to cause cancer in humans, but only in white rats.

But if the laboratory-based scholars succeed in refurbishing the image of laboratory research in psychology, at least among active scholars, doesn't this also tarnish the image of field research? We think not. Each type of study still has major roles to play in the development of theory and practical applications. The relative strengths and weaknesses, as described in various methodology textbooks and by the various defenders of laboratory research cited in our introduction, remain almost unchanged. The laboratory is still the best place to test causal hypotheses derived from theories. The laboratory is still the best place to investigate boundary conditions on the phenomena under consideration. The laboratory is still the best place to investigate relations among variables that do not co-occur naturally in the real world, or do so too infrequently to have a major impact. The laboratory is sometimes the only place to conduct certain kinds of research for ethical reasons. Finally, the laboratory is usually the least expensive way, in terms of both person-hours and actual dollars, to test our major theories.

On the other hand, the real world is often the best place to observe and define new phenomena, and to create new theoretical propositions about human psychology. The real world is still the best place to devise, test, refine, and put into practice specific applications. For both practical and ethical reasons, such applications must rest on firm theoretical foundations. The interplay between research in laboratory and real-world settings, and between theory and practice, is essential to the effective development of our understanding of and interventions in the aggression domain, and to the field of psychology in general. We hope that our view of the external validity of the clearly *nontrivial* laboratory approaches to aggression will enhance this interplay, and that a similar view of laboratory work in general will aid the development of other areas of psychology.

References

References marked with an asterisk indicate studies included in the meta-analyses.

Anderson, C. A. (1989). Temperature and aggression: Ubiquitous effects of heat on occurrence of human violence. *Psychological Bulletin, 106,* 74–96.

Anderson, C. A., & Anderson, K. B. (1996). Violent crime rate studies in philosophical context: A destructive testing approach to heat and southern culture of violence effects. *Journal of Personality and Social Psychology, 70,* 740–756.

Anderson, C. A., & Anderson, K. B. (in press). Temperature and aggression: Paradox, controversy, and a (fairly) clear picture. Chapter to appear in R. Geen & E. Donnerstein (Eds.), *Human aggression: Theories, research and implications for policy.*

Anderson, C. A., Bushman, B. J., & Groom, R. (1996). Hot years and violent crime: Empirical tests of the heat hypothesis. *Manuscript under review.*

Anderson, C. A., & DeNeve, K. M. (1992). Temperature, aggression, and the negative affect escape model. *Psychological Bulletin, 111,* 347–351.

Anderson, C. A., Dorr, N., Anderson, K. B., & DeNeve, K. (1996). Temperature-aggression inconsistencies resolved: Laboratory evidence of multiple functional relations between uncomfortable temperatures and aggressive behavior. *Manuscript under review.*

Anderson, K. B., & Anderson, C. A. (1996). Laboratory effects of hot temperatures on aggressive behavior: A meta-analysis. *Manuscript under review.*

*Archer, J., Holloway, R., & McLoughlin, K. (1995). Self-reported physical aggression among young men. *Aggressive Behavior, 21,* 325–342.

Aronson, E., & Carlsmith, J. M. (1968). Experimentation in social Psychology. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (2nd ed., pp. 1–79). Reading, MA: Addison-Wesley.

Banaji, M. R., & Crowder, R. G. (1989). The bankruptcy of everyday memory. *American Psychologist, 44,* 1185–1193.

Baron, R. A. (1979). Aggression and heat: The "long hot summer" revisited. In A. Baum, J. E. Singer, & S. Valins (Eds.), *Advances in environmental psychology* (pp. 57–84). Hillsdale, NJ: Erlbaum.

*Baron, R. A., Russell, G. W., & Arms, R. L. (1985). Negative ions and behavior: Impact on mood, memory, and aggression among Type A and Type B persons. *Journal of Personality and Social Psychology, 48,* 746–754.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182.

Beck, A. J. (1991). *Profile of jail inmates, 1989.* Washington, DC: Bureau of Justice Statistics.

Beck, A. J., Kline, S. A., & Greenfield, L. A. (1988). *Survey of youth in custody, 1987.* Washington, DC: Bureau of Justice Statistics.

Bem, D. J., & Lord, C. G. (1979). Template matching:

A proposal for probing the ecological validity of experimental settings in social psychology. *Journal of Personality and Social Psychology, 37,* 833–846.

Berkowitz, L. (1962). *Aggression: A social psychological analysis.* New York: McGraw-Hill.

Berkowitz, L., & Donnerstein, E. (1982). External validity is more than skin deep: Some answers to criticism of laboratory experiments. *American Psychologist, 37,* 245–257.

*Berman, M., Gladue, B., & Taylor, S. (1993). The effects of hormones, Type A behavior pattern, and provocation on aggression in men. *Motivation and Emotion, 17,* 125–138.

Bettencourt, B. A., & Miller, N. (1996). Gender differences in aggression as a function of provocation: A meta-analysis. *Psychological Bulletin, 119,* 422–447.

Boone, S. L., & Flint, C. (1988). A psychometric analysis of aggression and conflict-resolution behavior in Black adolescent males. *Social Behavior and Personality, 16,* 215–226.

Bushman, B. J. (1993). Human aggression while under the influence of alcohol and other drugs: An integrative research review. *Current Directions in Psychological Science, 2,* 148–152.

*Bushman, B. J. (1995). Moderating role of trait aggressiveness in the effects of violent media on aggression. *Journal of Personality and Social Psychology, 69,* 950–960.

Bushman, B. J. (in press). Effects of alcohol on human aggression: Validity of proposed explanations. In D. Fuller, R. Dietrich, & E. Gottheil (Eds.), *Recent developments in alcoholism: Alcohol and violence* (Vol. XIII). New York: Plenum.

Bushman, B. J., & Cooper, H. M. (1990). Effects of alcohol on human aggression: An integrative research review. *Psychological Bulletin, 107,* 341–354.

Buss, A. H. (1961). *The psychology of aggression.* New York: Wiley.

Buss, A. H., & Durkee, A. (1957). An inventory for assessing different kinds of hostility. *Journal of Consulting Psychology, 21,* 343–349.

*Buss, A. H., & Perry, M. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology, 63,* 452–459.

Campbell, D. T. (1957). Factor relevant to validity of experiments in social settings. *Psychological Bulletin, 54,* 297–312.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

Carlsmith, J. M., Ellsworth, P. C., & Aronson, E. (1976). *Methods of research in social psychology.* Reading, MA: Addison-Wesley.

Carlson, M., Marcus-Newhall, A., & Miller, N. (1989). Evidence for a general construct of aggression. *Personality and Social Psychology Bulletin, 15,* 377–389.

*Carver, C. S., & Glass, D. C. (1978). Coronary-prone behavior pattern and interpersonal aggression. *Journal of Personality and Social Psychology, 36,* 361–366.

*Check, J. V., & Dyck, D. G. (1986). Hostile aggression and Type A behavior. *Personality and Individual Differences, 7,* 819–827.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cook, P. J., & Moore, M. J. (1993). Violence reduction through restrictions on alcohol availability. *Alcohol, Health, and Research World, 17,* 151–156.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings.* Boston: Houghton Mifflin.

Curtis, L. A. (1974). Victim precipitation and violent crime. *Social Problems, 21,* 594–605.

Davis, S. (1991). Violence by psychiatric inpatients: A review. *Hospital and Community Psychiatry, 42,* 585–590.

Dexter, E. G. (1899). Conduct and the weather. *Psychological Monographs, 11(10),* 1–103.

*Diener, E. (1976). Effects of prior destructive behavior, anonymity, and group presence on deindividuation and aggression. *Journal of Personality and Social Psychology, 33,* 497–507.

*Diener, E., Dineen, J., Endresen, K., Beaman, A. L., & Fraser, S. C. (1975). Effects of altered responsibility, cognitive set, and modeling on physical aggression and deindividuation. *Journal of Personality and Social Psychology, 31,* 328–337.

Dipboye, R. L., & Flanagan, M. F. (1979). Research settings in industrial and organizational Psychology: Are findings in the field more generalizable than in the laboratory? *American Psychologist, 34,* 141–150.

Dunand, M. A. (1986). Violence and panic at the Brussels football stadium in 1985: Social psychological approach to the event. *Cahiers de Psychologie Cognitive, 6,* 235–266.

*Ellison, P. A., Govern, J. M., Petri, H. L., & Figler, M. H. (1995). Anonymity and aggressive driving behavior: A field study. *Journal of Social Behavior and Personality, 10,* 265–272.

Festinger, L., Pepitone, A., & Newcomb, T. (1952). Some consequences of de-individuation in a group. *Journal of Abnormal and Social Psychology, 47,* 382–389.

Geen, R. G. (1990). *Human aggression.* Pacific Grove, CA: Brooks/Cole.

*Giancola, P. R., & Zeichner, A. (1995). Construct validity of a competitive reaction-time aggression paradigm. *Aggressive Behavior, 21,* 199–204.

Glass, D. C. (1977). *Behavior patterns, stress, and coronary disease.* Beverly Hills, CA: Sage.

Gottfredson, M. R., & Hirschi, T. (1993). A control theory interpretation of psychological research on aggression. In R. B. Felson & J. T. Tedeschi (Eds.), *Aggression and violence: Social interactionist perspectives* (pp. 47–68). Washington, DC: American Psychological Association.

Greenberg, S. W. (1981). Alcohol and crime: A methodological critique of the literature. In J. J. Collins (Ed.), *Drinking and crime: Perspectives on the relationships between alcohol consumption and criminal behavior.* New York: Guilford.

Greenwald, A. G., Pratkanis, A. R., Leippe, M. R., & Baumgardner, M. H. (1986). Under what conditions does theory obstruct research progress? *Psychological Review, 93,* 216–229.

*Gunn, J., & Gristwood, J. (1975). Use of the Buss-Durkee Hostility Inventory among British prisoners. *Journal of Consulting and Clinical Psychology, 43,* 590.

*Hammock, G. S., & Richardson, D. R. (1992). Predictors of aggressive behavior. *Aggressive Behavior, 18,* 219–229.

Harré, R., & Secord, P. R. (1972). *The explanation of social behaviour.* Oxford, England: Blackwell.

*Hartman, D. P. (1969). Influence of symbolically modeled instrumental aggression and pain cues on aggressive behavior. *Journal of Personality and Social Psychology, 11,* 280–288.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* New York: Academic Press.

Henshel, R. L. (1980). The purpose of laboratory experimentation and the virtues of deliberate artificiality. *Journal of Experimental Social Psychology, 16,* 466–478.

*Holmes, D. S., & Will, M. J. (1985). Expression of interpersonal aggression by angered and nonangered persons with Type A and Type B behavior patterns. *Journal of Personality and Social Psychology, 48,* 723–727.

*Hurlbert, D. F., Whittaker, K. E., & Munoz, C. J. (1991). Etiological characteristics of abusive husbands. *Military Medicine, 156,* 670–675.

Innes, C. A. (1988). *Drug use and crime.* Washington, DC: U.S. Department of Justice.

Jenkins, C. D., Zyzanski, S. J., & Rosenman, R. H. (1979). *The Jenkins Activity Survey, Form C.* New York: The Psychological Corporation.

Kane, T. R., Joseph, J. M., & Tedeschi, J. T. (1976). Person perception and the Berkowitz paradigm for the study of aggression. *Journal of Personality and Social Psychology, 33,* 663–673.

Kenrick, D. T., & MacFarlane, S. W. (1984). Ambient temperature and horn-honking: A field study of the heat/aggression relationship. *Environment and Behavior, 18,* 179–191.

*Knott, P. (1970). A further methodological study of the measurement of interpersonal aggression. *Psychological Reports, 26,* 807–809.

Kruglanski, A. W. (1975). The human subject in the psychology experiment: Fact and artifact. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 8, pp. 101–147). New York: Academic Press.

*Lange, A., Dehghani, B., & De Beurs, E. (1995). Validation of the Dutch adaptation of the Buss-Durkee Hostility Inventory. *Behavioral Research Therapy, 33,* 229–233.

*Larsen, K. S., Coleman, D., Forbes, J., & Johnson, R. (1972). Is the subject's personality or the experimental situation a better predictor of a subject's willingness to administer shock to a victim? *Journal of Personality and Social Psychology, 22,* 287–295.

*Leibowitz, G. (1968). Comparison of self-report and behavioral techniques of assessing aggression. *Journal of Consulting and Clinical Psychology, 32,* 21–25.

*Lightdale, J. R., & Prentice, D. A. (1994). Rethinking sex differences in aggression: Aggressive behavior in the absence of social roles. *Personality and Social Psychology Bulletin, 20,* 34–44.

*Llorente, M., Bernardo, M., de Flores, T., & Valdes, M. (1985). Type A behavior and Buss's instrumental aggression paradigm (BIAP). *Activitas Nervosa Superior, 27,* 106–109.

*Lothstein, L. M., & Jones, P. (1978). Discriminating violent individuals by means of various psychological tests. *Journal of Personality Assessment, 42,* 237–243.

MacDonald, J. M. (1961). *The murderer and his victim.* Springfield, IL: Charles C. Thomas.

*Maiuro, R. D., Cahn, T. S., Vitaliano, P. P., Wagner, B. C., & Zegree, J. B. (1988). Anger, hostility, and depression in domestically violent versus generally assaultive men and nonviolent control subjects. *Journal of Consulting and Clinical Psychology, 56,* 17–23.

*Mann, L. (1981). The baiting crowd in episodes of threatened suicide. *Journal of Personality and Social Psychology, 41,* 703–709.

*Mann, L., Newton, J. W., & Innes, J. M. (1982). A test between deindividuation and emergent norm theories of crowd aggression. *Journal of Personality and Social Psychology, 42,* 260–272.

*Matteson, M. T., & Ivancevich, J. M. (1979). Behavior activity profile. In J. L. Gibson, J. M. Ivancevich, & J. H. Donnelly, Jr. (Eds.), *Organizations: Behavior, structure, processes* (pp. 169–173). Plano, TX: Business.

Meyer, C. K. (1982). An analysis of factors related to robbery-associated assaults on police officers: I. *Journal of Police Science and Administration, 10,* 1–27.

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist, 38,* 379–387.

*Mullen, B. (1986). Atrocity as a function of lynch mob composition: A self-attention perspective. *Personality and Social Psychology Bulletin, 12,* 187–197.

*Muntaner, C., Llorente, M., & Nagoshi, C. (1989). Evaluation instructions and interpersonal aggression in the Type A behavior pattern. *Aggressive Behavior, 15,* 161–170.

*Muntaner, C., Walter, D., Nagoshi, C., Fishbein, D., Haertzen, C. A., & Jaffe, J. H. (1990). Self-report vs. laboratory measures of aggression as predictors of substance abuse. *Drug and Alcohol Dependence, 25,* 1–11.

Murdoch, D., Pihl, R. O., & Ross, D. (1990). Alcohol and crimes of violence: Present issues. *International Journal of the Addictions, 25,* 1065–1081.

Neisser, U. (1978). Memory: What are the important questions? In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 3–24). London: Academic Press.

Nisbett, R. E. (1993). Violence and U.S. regional culture. *American Psychologist, 48,* 441–449.

Paik, H., & Comstock, G. (1994). The effects of television violence on antisocial behavior: A meta-analysis. *Communication Research, 21,* 516–546.

Palamarek, D. L., & Rule, B. G. (1979). Effects of ambient temperature and insult on the motivation to retaliate or escape. *Motivation and Emotion, 3,* 83–92.

McKinley, J. C., Jr. (1990, November 2). Gang kills homeless man in Halloween rampage. *New York Times,* A1, B3.

*Paloutzian, R. F. (1975). Effects of deindividuation, removal of responsibility, and coaction on impulsive and cyclical aggression. *Journal of Psychology, 90,* 163–169.

Pernanen, K. (1991). *Alcohol in human violence.* New York: Guilford Press.

*Prentice-Dunn, S., & Rogers, R. W. (1980). Effects of deindividuating situational cues and aggressive models on subjective deindividuation and aggression. *Journal of Personality and Social Psychology, 39,* 104–113.

*Prentice-Dunn, S., & Rogers, R. W. (1982). Effects of public and private self-awareness deindividuating situational cues and aggressive models on subjective deindividuation and aggression. *Journal of Personality and Social Psychology, 39,* 104–113.

*Prentice-Dunn, S., & Spivey, C. B. (1986). Extreme deindividuation in the laboratory: Its magnitude and subjective components. *Personality and Social Psychology Bulletin, 12,* 206–215.

*Propst, L. R. (1979). Effects of personality and loss of anonymity on aggression: A reevaluation of deindividuation. *Journal of Personality, 47,* 531–545.

Reifman, A. S., Larrick, R. P., & Fein, S. (1991). Temper and temperature on the diamond: The heat-aggression relationship in major league baseball. *Personality and Social Psychology Bulletin, 17,* 580–585.

*Renson, G. J., Adams, J. E., & Tinklenberg, J. R. (1978). Buss-Durkee assessment and validation with violent versus nonviolent chronic alcohol abusers. *Journal of Consulting and Clinical Psychology, 46,* 360–361.

Roediger, H. L., III. (1991). They read an article? *American Psychologist, 46,* 37–40.

*Rogers, R. W. (1980). Expressions of aggression: Aggression-inhibiting effects of anonymity to authority and threatened retaliation. *Personality and Social Psychology Bulletin, 6,* 315–320.

*Rogers, R. W., & Ketchen, C. M. (1979). Effects of anonymity and arousal on aggression. *Journal of Psychology, 102,* 13–19.

*Rogers, R. W., & Prentice-Dunn, S. (1981). Deindividuation and anger-mediated interracial aggression: Unmasking regressive racism. *Journal of Personality and Social Psychology, 41,* 63–73.

Rohsenow, D. J., & Bachorowski, J. (1984). Effect of alcohol and expectancies on verbal aggression in men and women. *Journal of Abnormal Psychology, 93,* 418–432.

*Scheier, M. F., Buss, A. H., & Buss, D. M. (1978). Self-consciousness, self-report of aggressiveness, and aggression. *Journal of Research in Personality, 12,* 133–140.

*Schell, B. H., Cachon, J., Ganjavi, O., & Porporino, F. (1986). A pilot study assessing Type A behavior in violence-prone inmates. *Psychological Reports, 59,* 371–382.

*Selby, M. J. (1984). Assessment of violence potential using measures of anger, hostility, and social desirability. *Journal of Personality Assessment, 48,* 531–544.

Selby makes one last plea. (1987, August 15). Ogden Standard Examiner, p. 1A.

*Shemberg, K. M., Leventhal, D. B., & Allman, L. (1968). Aggression machine performance and rated aggression. *Journal of Experimental Research in Personality, 3,* 117–119.

*Spivey, C. B., & Prentice-Dunn, S. (1990). Assessing the directionality of deindividuated behavior: Effects of deindividuation, modeling, and private self-consciousness on aggressive and prosocial responses. *Basic and Applied Social Psychology, 11,* 387–403.

*Stanford, M. S., Greve, K. W., & Dickens, T. J., Jr. (1995). Irritability and impulsiveness: Relationship to self-reported impulsive aggression. *Personality and Individual Differences, 19,* 757–760.

Still at a loss for "why". (1987, August 26). *Ogden Standard Examiner*, p. 1A.

*Strube, M. J., Turner, C. W., Cerro, D., Stevens, J., & Hinchey, F. (1984). Interpersonal aggression and the Type A Coronary-Prone Behavior Pattern: Theoretical distinction and practical implications. *Journal of Personality and Social Psychology, 47,* 839–847.

*Syverson, K. L., & Romney, D. M. (1985). A further attempt to differentiate violent from nonviolent offenders by means of a battery of psychological tests. *Canadian Journal of Behavioural Science, 17,* 87–92.

Tamura, M. (1983). Changes in the patterns of criminal homicide for recent three decades. *Reports of National Research Institute of Police Science, 24,* 149–161.

*Taylor, S. L., O'Neal, E. C., Langley, T., & Butcher, A. H. (1990). Anger arousal, deindividuation, and aggression. *Aggressive Behavior, 17,* 193–206.

Taylor, S. P. (1967). Aggressive behavior and physiological arousal as a function of provocation and the tendency to inhibit aggression. *Journal of Personality, 35,* 297–310.

U.S. Department of Justice. (1994). *Uniform crime reports for the United States: 1993.* Washington, DC.

Wheeler, L., & Caggiula, A. R. (1966). The contagion of aggression. *Journal of Experimental Social Psychology, 2,* 1–10.

*Williams, J. F., Meyerson, L. J., Eron, L., & Semler, I. J. (1967). Peer-rated aggression and aggressive responses elicited in an experimental situation. *Child Development, 38,* 181–190.

*Wilson, C., & Brewer, N. (1993). Individuals and groups dealing with conflict: Findings from police on patrol. *Basic and Applied Social Psychology, 14,* 55–67.

*Wolfe, B. M., & Baron, R. A. (1971). Laboratory aggression related to aggression in naturalistic social situations: Effects of an aggressive model on the behavior of college student and prisoner observers. *Psychonomic Science, 24,* 193–194.

*Worchel, S., & Anderoli, V. (1978). Facilitation of social interaction through deindividuation of the target. *Journal of Personality and Social Psychology, 36,* 549–556.

*Worchel, S., Arnold, S. E., & Harrison, W. (1978). Aggression and power restoration: The effects of identifiability and timing on aggressive behavior. *Journal of Experimental Social Psychology, 14,* 43–52.

Zimbardo, P. G. (1969). The human choice: Individuation, reason, and order, versus deindividuation, impulse, and chaos. In W. Arnold & D. Levine (Eds.), *Nebraska Symposium on Motivation, 1969.* Lincoln, NE: University of Nebraska Press.

Received October 4, 1995
Revision received March 12, 1996
Accepted May 14, 1996 ■