Consider the following news story (McKinley, 1990).

A gang of young men wielding knives and bats went on a Halloween rampage Wednesday night, assaulting several homeless people on the footbridge to Ward's 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.

Now consider the following description of a laboratory aggression experiment. In this experiment (Zimbardo, 1969), college women were told that the researchers were studying "empathic responses to strangers." Individuals participated in groups of four. By the flip of a coin, participants were assigned to anonymous or identifiable conditions. In the anonymous condition, participants wore large laboratory coats, wore hoods over their heads, and were not referred to by name. In the identifiable condition, participants did not wear laboratory coats, wore large name tags, and were referred to by name. The room lights were turned off in the anonymous condition or only dimmed in the identifiable condition. Participants then listened to one of two tape-recorded interviews between the experimenter and another "participant" (who was actually an accomplice to the experimenter). In one of the tapes, the accomplice behaved in a nice, altruistic, and accepting manner. In the other tape, she behaved in an obnoxious, self-centered, and critical manner.
Participants then shocked the accomplice by holding down a button as long as they wanted her to be shocked. The results showed that participants in the anonymous condition gave the accomplice longer shocks than did participants in the identifiable condition, especially when the accomplice was obnoxious.

What is the relation between the news report and the laboratory aggression experiment? Is there any reason to believe that "artificial" laboratory studies of aggression can inform us about gang violence or any other kind of "real world" aggression? That is, do the measures of aggression that are typically used in laboratory studies truly measure "aggression"? There are two related issues embedded in this question. First, do different laboratory measures of aggression measure the same underlying conceptual variable? Second, do laboratory measures of aggression measure "real" aggression, as found in the natural environment? The major purpose of this chapter is to answer these important questions. First, however, we must discuss what is meant by "aggression" and how aggression is measured in the laboratory.

MEASUREMENT OF AGGRESSION

Buss (1961, p. 1) defined aggression as "a response that delivers noxious stimuli to another organism." Geen (1990) clarified this definition by adding two elements: (a) the aggressor delivers the noxious stimuli with the intent to harm the victim, and (b) the aggressor expects that the noxious stimuli will have their intended effect. Buss further proposed that acts of human aggression can be classified using combinations of three dichotomous variables: physical versus verbal, direct versus indirect, and active versus passive. Although there are eight possible combinations of the three dichotomous variables proposed by Buss, none of the four "passive" types of aggression are common in experimental studies of aggression. This review therefore focuses on the four "active" types of human aggression. In physical aggression, noxious stimuli delivered to the victim are pain and injury, whereas in verbal aggression, noxious stimuli delivered to the victim are rejection and threat. In direct aggression the aggressor is easily identified by the victim, whereas in indirect aggression the aggressor is not easily identified by the victim. There are two ways in which an aggressive act can be indirect. First, the victim is not present and noxious stimuli are delivered via the negative reactions of others. Second, the victim is not injured or threatened, but his or her belongings are stolen or damaged. The next section describes prototypical procedures for measuring each type of active aggression.

DIRECT PHYSICAL AGGRESSION

The aggression machine paradigm (Buss, 1961) has been the primary laboratory procedure used to measure direct physical aggression. In this procedure, 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" whereas the confederate is selected to be the "learner." The participant presents stimulus materials to the confederate who (supposedly) attempts to learn them. In some experiments, the participant is angered by the confederate prior to the beginning of the learning task.

When the confederate makes an incorrect response on a trial, the participant is supposed to punish him or her by means of electric shock. By using different buttons, the participant can control the intensity and duration 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). In some experiments, shock duration is controlled by holding down the shock button for the desired amount of time. The two measures of aggression then are intensity and duration of shock "given" to the confederate by the participant. Some researchers have used noxious stimuli other than electric shocks, such as noise blasts and heat pulses.

Another commonly employed method to study direct physical aggression is to place the participant and the confederate in a situation that requires the confederate to evaluate the participant and later require the participant to evaluate the confederate. In Berkowitz's (1962) paradigm, for example, participants are led to believe that they will be evaluating another student's performance on an assigned task. Solutions are evaluated using anywhere from 1 to 10 electric shocks, where 1 shock indicates a very favorable evaluation and 10 shocks indicates a very unfavorable evaluation. First, the confederate evaluates the participant's solution. Generally, half of the participants receive a positive evaluation from the confederate (e.g., 1 shock), whereas the other half receive a negative evaluation (e.g., 7 shocks). After exposure to some treatment (e.g., a violent or nonviolent film), the participant then evaluates the confederate's solution. The measure of aggression is the number of shocks the participant gives the confederate.

The competitive reaction time paradigm (Taylor, 1967) is a third common method employed in the laboratory to study direct physical aggression. In this procedure, the participant competes with an ostensible opponent on a reaction time task in which the slower responding person receives electric shock. At the beginning of each trial the participant sets the intensity of shock he or she wants the opponent (confederate) to receive if the opponent's response is slower. At the end of each trial, the participant is informed of the level of shock the opponent set for him or her to receive on the trial. The slower responding person then supposedly receives the shock. In actuality, the experimenter determines both who wins and loses and the feedback/shocks delivered. Sometimes provocation is manipulated by increasing the intensity of shock set by the "opponent" across trials on the reaction time task. The measure of aggression is the intensity of shock the participant sets for the opponent. Some researchers have included duration as a second measure of aggression. Other researchers have used noise rather than shock as the noxious stimuli.
INDIRECT PHYSICAL AGGRESSION

The laboratory paradigms used to measure direct physical aggression also have been modified to measure indirect physical aggression. In one study (Barnett, 1979), for example, male college students were given $2.00 and course credit for their participation. Participants were told to subtract between 0 cents (button 0) and 9 cents (button 9) from a confederate whenever he made a mistake on a trial. This paradigm measures indirect physical aggression because the participant takes the confederate's belongings (i.e., his money).

The free-operant paradigm (e.g., Cherek, 1981) is another method commonly employed to measure indirect physical aggression. In this procedure, the participant can press one of two buttons on an apparatus. Pressing button A results in the accumulation of points exchangeable for money. Pressing button B results in the subtraction of points from a fictitious second participant. Sometimes provocation is manipulated by subtracting points from the participant; the point loss is attributed to the fictitious second participant. The fixed ratios associated with each button also can be manipulated (e.g., a fixed ratio of 100 responses might be required for button A, whereas a fixed ratio of 10 responses might be required for button B).

DIRECT 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. For example, in a study by Wheeler and Caggia (1966), 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 example, 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 (that he was an “ass,” “idiot,” “crazy,” “nuts,” “insane,” etc.; that he should be “locked up,” “shot,” “deported,” “beaten up,” “tortured,” etc.).

INDIRECT VERBAL AGGRESSION

Indirect measures of verbal aggression are more common in laboratory experiments than are direct measures of verbal aggression. Generally, a confederate or experimenter first provokes the participant. Rather than confronting the confeder-
Physical and written measures of aggression were significantly positively related to each other across a wide range of laboratory studies. For instance, correlations of number, duration, and intensity of punishments ranged from .70 to .88. Written and physical response measures of aggression were also positively correlated (overall $r = .71$).

Overall, the Carlson et al. (1989) results clearly demonstrate that the most commonly used (and "artificial") laboratory measures of aggression tap the same underlying conceptual variable, which leads us to the second question posed at the outset: Do laboratory measures of aggression measure "real" aggression, as found in the natural environment?

### Similarity of Real-World and Laboratory Aggression

Table 2.1 gives "real world" and laboratory examples of each of the four types of active aggression proposed by Buss (1961). In the real world, most extreme acts of aggression are violent crimes, which the FBI classifies as murder, aggravated assault, forcible rape, and robbery. According to Buss’s framework, murder, forcible rape, and aggravated assault are examples of direct physical aggression, whereas robbery is an example of indirect physical aggression. As can be seen in Table 2.1, laboratory measures of aggression have few surface features in common with real world measures of aggression. It is this lack of surface similarity that leads to the overly pessimistic critiques of the value of laboratory aggression paradigms.

<table>
<thead>
<tr>
<th>Type of aggression</th>
<th>&quot;Real world&quot; examples</th>
<th>Laboratory examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct physical</td>
<td>Assaulting someone with body parts (e.g., limbs, teeth) or weapons (e.g., clubs, knives, guns)</td>
<td>Shocking a confederate on a task</td>
</tr>
<tr>
<td>Indirect physical</td>
<td>Stealing or damaging someone’s property</td>
<td>Subtracting money from a confederate on a task</td>
</tr>
<tr>
<td></td>
<td>Setting a booby trap for someone</td>
<td>Hire an assassin to kill someone</td>
</tr>
<tr>
<td>Direct verbal</td>
<td>Criticizing, derogating, or cursing someone</td>
<td>Making negative verbal statements to a confederate</td>
</tr>
<tr>
<td>Indirect verbal</td>
<td>Threatening someone</td>
<td>Evaluating a confederate negatively on a questionnaire</td>
</tr>
<tr>
<td></td>
<td>Spreading vicious rumors about someone</td>
<td></td>
</tr>
</tbody>
</table>
Now back to the question at hand: Is there any reason to believe that the results from laboratory aggression studies can inform us about "real world" aggression? We believe that the answer to this question depends on whether aggression in the laboratory paradigm has the same meaning as it does in the real world (also see Berkowitz & Donnerstein, 1982). If laboratory aggression "means" the same thing to people 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. Likewise, individual differences in aggressiveness observed in the real world should also be observed in the laboratory.

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 (see Anderson & Bushman, 1997). We then examined the published literature to compare the effects of these variables on real world and laboratory aggression. Table 2.2 lists the situational and individual difference variables we examined. Also indicated for each variable in Table 2.2 is 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.1 Although meta-analytic procedures can be used to combine the results from two studies, in this chapter 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 studies conducted inside and outside the laboratory. A .05 significance level was used for all tests.

Because laboratory researchers study more homogeneous populations (e.g., college students), individual differences should vary more outside the laboratory. In other words, individual difference variables have a restricted range in laboratory studies. Thus, stronger effects for individual difference variables are expected outside the laboratory than within it. The opposite pattern of results is expected

---

1To obtain a weighted average of the sample correlations, $r_2$, we first obtained a weighted average of Fisher’s $z$-to-$r$ 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, $r$ (see Hedges & Olkin, 1985, pp. 235 and 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, $d_2$, each $d$ was weighted by the inverse of its variance, $[2(n_e + n_c)h^2]/[2(n_e + n_c)^2 + n_e n_c d^2]$. A 95% confidence interval was also calculated for the population standardized mean difference, $d_0$ (see Hedges & Olkin, 1985, pp. 110–113). Cohen's (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 2.3 Mean Effect Size Estimates and Confidence Intervals for Individual Difference and Situational Variables Studied Inside and Outside the Laboratory

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Effect-size estimate</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual difference variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Studies conducted outside the laboratory</td>
<td>6</td>
<td>$d_s = 0.40^b$</td>
<td>[0.25, 0.55]</td>
</tr>
<tr>
<td>Physical aggression</td>
<td>3</td>
<td>$d_s = 0.03^a$</td>
<td>[-0.15, 0.22]</td>
</tr>
<tr>
<td>Verbal aggression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Studies conducted inside the laboratory</td>
<td>37</td>
<td>$d_s = 0.31^b$</td>
<td>[0.23, 0.38]</td>
</tr>
<tr>
<td>Physical aggression</td>
<td>18</td>
<td>$d_s = 0.13^a$</td>
<td>[0.03, 0.24]</td>
</tr>
<tr>
<td>Verbal aggression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trait aggressiveness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Studies conducted outside the laboratory</td>
<td>16</td>
<td>$r_s = .42^b$</td>
<td>[0.38, 0.47]</td>
</tr>
<tr>
<td>Studies conducted inside the laboratory</td>
<td>13</td>
<td>$r_s = .24^a$</td>
<td>[0.18, 0.29]</td>
</tr>
<tr>
<td>Type A personality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Studies conducted outside the laboratory</td>
<td>3</td>
<td>$d_s = 0.97^b$</td>
<td>[0.71, 1.23]</td>
</tr>
<tr>
<td>Studies conducted inside the laboratory</td>
<td>9</td>
<td>$d_s = 0.34^a$</td>
<td>[0.18, 0.49]</td>
</tr>
<tr>
<td>Situational variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provocation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Studies conducted inside the laboratory</td>
<td>66</td>
<td>$d_s = 0.76$</td>
<td>[0.66, 0.85]</td>
</tr>
<tr>
<td>Alcohol</td>
<td>130</td>
<td>$r_s = 0.16^a$</td>
<td>[0.14, 0.17]</td>
</tr>
<tr>
<td>Weapons effect</td>
<td>42</td>
<td>$r_s = 0.26^a$</td>
<td>[0.21, 0.30]</td>
</tr>
<tr>
<td>Studied conducted inside the laboratory</td>
<td>5</td>
<td>$d_s = 0.17^a$</td>
<td>[-0.05, 0.39]</td>
</tr>
<tr>
<td>Media violence</td>
<td>16</td>
<td>$d_s = 0.21^a$</td>
<td>[0.01, 0.41]</td>
</tr>
<tr>
<td>Anonymity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Studies conducted outside the laboratory</td>
<td>556</td>
<td>$d_s = 0.42^b$</td>
<td>[0.25, 0.63]</td>
</tr>
<tr>
<td>Studies conducted inside the laboratory</td>
<td>586</td>
<td>$d_s = 0.87^a$</td>
<td>[0.45, 0.69]</td>
</tr>
<tr>
<td>Temperature</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Studies conducted inside the laboratory</td>
<td>28</td>
<td>$d_s = 0.06$</td>
<td>[-0.11, 0.23]</td>
</tr>
<tr>
<td>Overall</td>
<td>12</td>
<td>$d_s = 0.25$</td>
<td>[-0.03, 0.53]</td>
</tr>
<tr>
<td>Neutral context</td>
<td>14</td>
<td>$d_s = -0.09$</td>
<td>[-0.33, 0.15]</td>
</tr>
</tbody>
</table>

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. CI, confidence interval. $r_s$, weighted average of the sample correlations. $d_s$, weighted average of the sample standardized mean differences. 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.

*Data from Bettencourt & Miller (1996).

$^b$Data from Lipsey et al. (1997).

$^c$Data from Carlson, Marcus-Newhall, & Miller (1990).

$^d$Data from Paik and Comstock (1994).

For situational variables because the greater variability in individual differences outside the laboratory makes situational variables less predictive.

LITERATURE SEARCH PROCEDURE

Whenever possible, extant meta-analytic literature reviews were used. In those cases where such reviews did not already exist, we attempted to conduct an exhaustive search of the literature. Table 2.2 also indicates whether a new literature search was conducted for each variable. The PsycINFO computer database was searched (1974–1996) using the key words aggress* and violen*. The asterisk at the end of the key word gives all forms of the key words (e.g., aggress, aggressive, aggressiveness, aggression, aggressed, aggressor). 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 2.2. The search was restricted to studies published in English and to studies that used human participants.

RESULTS

INDIVIDUAL DIFFERENCES IN AGGRESSION

This section assesses the correspondence between findings from studies using real world and laboratory measures of aggression, focusing on the following individual difference variables: sex, 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 males commit more murders and assaults than do females. This sex effect occurs in virtually every murder and assault rate data set that can be found. For example, Dexter (1899) showed that the male rate of assault in New York city in the years from 1891 to 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. The results showed that males aggressed more than females in both neutral and provoked conditions, but that the difference is smaller in provoked 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. As their analyses did not focus on laboratory versus field studies, we used their tabled results to conduct our own meta-analyses.

2The 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 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. As their analyses did not focus on laboratory versus field studies, we used their tabled results to conduct our own meta-analyses.
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 2.3).

**Studies Conducted Inside the Laboratory**

To investigate sex effects in laboratory studies, we also recombined studies from Bettencourt and Miller's (1996) meta-analysis (see Table 2.3). The results showed that males were more physically and verbally aggressive than females. Sex differences in verbal aggression, however, were quite small.

In summary, our analysis suggests that males are more physically aggressive than females both inside and outside the laboratory. Sex differences in verbal aggression are small to trivial in both settings. The setting 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 defined operationally 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 this scale 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 histories of violence. In most laboratory studies, participants are college students. Because of greater variability in trait aggressiveness in field studies, 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, & Zeegree, 1988), between violent and nonviolent patients (Lange, Dehghani, & Debeurs, 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 2.3).

**Type A Coronary-Prone Behavior Pattern**

The 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 as well). A 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, Zyzinski, & 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 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 preschool population 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 assailants 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, 1979), a self-report measure of the Type A pattern, as part of a battery of questionnaires. The results showed that most assailants 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 2.3).

**Studies Conducted Inside the Laboratory**

Most laboratory studies have found that Type A's, in comparison to 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 2.3).

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**

This section assesses 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, the presence of weapons, 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 example, 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.

**Alcohol**

**Studies Conducted Outside the Laboratory**

Co-occurrence statistics are often used to establish a relation between alcohol use and violent crime. For example, numerous studies have reported that at least 50% of the perpetrators of violent crimes were intoxicated at the time of the offense (e.g., Beck, 1991; Beck, Kline, & Greenfield, 1988; Greenberg, 1981; Innes, 1988; MacDonald, 1961; Murdoch, Pihl, & Ross, 1990; Permanen, 1991). The problem with co-occurrence statistics is that they provide no base rate information about the level of alcohol use among comparable people who were not violent. Lipsey, Wilson, Cohen, and Derzon (1997) correctly point out that without the “other half” of the data, one cannot determine the strength of the relation between alcohol use and violent crime. Thus, Lipsey and his colleagues combined the results from studies that examined whether individuals with higher alcohol use also exhibit higher levels of violent behavior. The results of their meta-analysis found a small to medium-sized correlation between alcohol and violent behavior (see Table 2.3).  

Lipsey et al. (1997) provided separate correlations for criminals with chronic alcohol consumption, criminals with acute alcohol consumption, and domestic offenders with chronic alcohol consumption. We combined these correlations.
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 sober participants (Bushman, 1993, 1997; Bushman & Cooper, 1990; Ito, Miller, & Pollock, 1996; Lipsey, Wilson, Cohen, & Derzon, 1997; see Table 2.3). The type of aggression measure used does not appear to influence the magnitude of effects (Bushman, 1997; Bushman & Cooper, 1990). The correlation between alcohol and aggression is medium-sized in laboratory studies. Larger effects might be obtained if ethical considerations did not prevent researchers from using higher doses of alcohol in their laboratory studies (i.e., the target blood alcohol level is at most .10). In summary, alcohol appears to increase aggression inside and outside the laboratory.

Weapons Effect

According to Berkowitz (1968), “Guns not only permit violence, they can stimulate it as well. The finger pulls the trigger, but the trigger may also be pulling the finger” (p. 22). Carlson, Marcus-Newhall, and Miller (1990) conducted a meta-analytic review of the effects of aggression-related cues, including weapons, on aggression. We reanalyzed data from studies testing the effects of weapons on aggression inside and outside the laboratory. Carlson et al. (1990) found that the weapons effect was reversed when participants were suspicious or experiencing evaluation apprehension, so studies with such artifactual features were not included in our analysis.

Studies Conducted Outside the Laboratory

A few studies have investigated the presence of weapons in naturalistic settings (Boyanowski & Griffiths, 1982; Turner, Layton, & Simons, 1975, Studies 2 and 3). For example, in one study (Turner et al., 1975, Study 2), a confederate in a pickup truck stalled at a traffic signal light for 12 seconds. For some of the motorists, a .303-caliber military rifle was placed in a gun rack mounted on the back window of the confederate’s truck. Among those motorists who saw the gun, half saw a bumper sticker attached to the tailgate that said “VENGEANCE” and half saw a bumper sticker that said “FRIEND.” Motorists in the control group saw no gun or bumper sticker. The results showed the highest level of horn honking among motorists who saw a gun and the VENGEANCE sticker, followed respectively by motorists who saw a gun and the FRIEND sticker and by motorists who saw no gun or sticker. Four of the five comparisons yielded positive effects. Meta-analysis of the five results from these studies showed a small effect, but it was not significantly different from zero (due to the small number of studies). These results are depicted in Table 2.3.

 Media Violence

There are more television sets in the United States than there are toilets. Over 98% of American homes have at least one television set (APA, 1993). In the United States, adults spend more time watching television than they spend on any other activity except sleeping and working, and children spend more time watching television than they spend at school (Huston et al., 1992). About 60% of television programs contain violence (National Television Violence Study, 1996, 1997). By the time the average child graduates from elementary school, he or she will have seen at least 8,000 murders and more than 100,000 other assorted acts of violence on television (Huston et al., 1992). 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 field studies. Second, the violence shown 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 story.
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 have a small to medium effect on aggression in field studies (see Table 2.3).

**Studies Conducted Inside the Laboratory**

Paik and Comstock (1994) reported that violent media have a large effect on aggression in laboratory studies (see Table 2.3). 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.

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**

What factors influence people to engage in such uninhibited antisocial behaviors as those described in the news story at the beginning of this chapter? 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 (Lighthdale & Prentice, 1994). An individual can achieve anonymity by being part of a group, by wearing a mask, or by performing behaviors in the dark. 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., Tamura, 1983; Meyer, 1982). Deindividuation leads to a reduced sense of accountability.

**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, number 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 large crowds were more likely to taunt and urge the victim to jump than were small crowds. In addition, more baiting episodes occurred in nighttime hours than in 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 x 4s with their tops up honked more at a stalled motorist than did drivers with their tops down. Meta-analysis of these results yielded a medium-sized effect (see Table 2.3).
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 aggressive behaviors such as murders, rapes, assaults, and wife battering were relatively more frequent during hotter periods of time than during cooler periods of time. More recent time period studies have shown that violent crime rates in the United States (from 1950 to 1992) 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, 1997). 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 (Anderson & Anderson, 1995) show that this region effect occurs even when numerous 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 Anderson (1989), in which temperature and aggression were simultaneously assessed, also yielded significant temperature effects on aggression. In sum, real world aggression studies show that hot temperatures produce increases in aggression.

**Laboratory Aggression**

Laboratory studies of aggression have yielded quite inconsistent effects. Hot temperatures sometimes increase and sometimes decrease aggression (Anderson, 1989). To interpret these discrepancies, Anderson and Anderson (1997) conducted a meta-analysis on the laboratory studies of aggression. When all of the laboratory effects of hot versus comfortable temperatures are combined, the average effect size is about zero (see Table 2.3). One possible factor that might account for this null effect concerns whether there were other manipulated variables that raised or lowered the participant's feelings of anger, annoyance, or friendliness. Specifically, Baron’s (1979) negative affect escape (NAE) model predict that hot temperatures will increase aggression when there are no other negative factors present, but will decrease aggression when there are other factors present which, taken together, would tend to heighten negative feelings. In laboratory studies, the most common factor used to heighten the negative affect has been an anger manipulation.

To test the NAE model, Anderson and Anderson (1997) categorized the 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 the negative affect. The results were in the direction predicted by the NAE model, but they were not statistically significant (due to the small number of studies). These results are depicted in Table 2.3.

**DEALING WITH REAL-WORLD/LABORATORY DISCREPANCIES**

Discrepancies between real world 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. When the conceptual variables or processes are the same, parallel results should be obtained in the laboratory and real world. Two imprudent approaches to dealing with discrepancies between real world and laboratory findings involve simple rejection. One can reject real world findings as being the result of the confounds or the lack of control that typifies such studies. Alternatively, one can reject laboratory findings as being the result of suspicion problems, trivial manipulations, or trivial measures of aggression. The former approach appears to characterize the theoretical, experimental perspective, whereas the latter appears to characterize the applied, nonexperimental perspective.

Our view is that such discrepancies should serve as signals that additional conceptual work is needed, to be followed by additional empirical work on the new understandings that result from the additional conceptual 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). The analysis of the temperature/aggression discrepancies was the result of such a conceptual reanalysis.

**GENERAL DISCUSSION**

At the outset, we presented two aggression scenarios—one real world example of gang violence 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 employed, findings about the relations between conceptual variables will generalize from the laboratory to the real world, and vice versa.

The various aggression literatures sampled for this chapter 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 pattern) and most of the situational variables (provocation, alcohol, the presence of weapons, media violence, anonymity) consistently influenced aggressive behavior in real world and laboratory paradigms in the same way. 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. Once such processes were identified and at least partially equated, comparable findings emerged.

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 causing harm. As noted by Mook (1983), Berkowitz and Donnerstein (1982), and others, what we 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 real world and 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.

REFERENCES

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


METHODOLOGY IN THE STUDY OF AGGRESSION


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


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


