Oxford Handbooks Online

Research Methods, Design, and Statistics in Media Psychology

Sara Prot and Craig A. Anderson

The Oxford Handbook of Media Psychology
Edited by Karen E. Dill

Print Publication Date: Dec 2012
Online Publication Date: Jan 2013

Subject: Psychology, Psychological Methods and Measurement
DOI: 10.1093/oxfordhb/9780195398809.013.0007

[–] Abstract and Keywords

This chapter provides an overview of contemporary research methods used in the field of media psychology. Basic scientific principles are discussed. Commonly used research designs are described. Some methodological pitfalls in media psychology research are explained and suggestions are given on how to avoid them. Finally, guidelines are given on how to convey scientific methodology and findings to the general public (see Chapter 26). We hope that this chapter will aid readers from other fields in becoming informed consumers of media psychology research and aid media psychology researchers in continuing the trend toward better methodological quality in the field.

Keywords: media psychology, research designs, research methods

The 20th century witnessed a mass media explosion after the invention of television, digital computers, and the Internet. This rapid technological development was followed by rapid growth in the field of media psychology.

Researchers have gone from asking relatively simple, basic questions such as, “Does observing filmed aggression increase aggressive behavior?” (Bandura, Ross, & Ross, 1963a) to asking complex and highly specific questions, such as, “Through which cognitive and affective mechanisms do violent media exert their influence on aggression?” (Anderson et al., 2003), “How robust and consistent are media violence effects on different outcomes?” (Anderson & Bushman, 2001; Anderson et al., 2010) and “What are the long-term consequences of habitual media violence exposure?” (Huesmann et al., 2003; Bartholow, Bushman, & Sestir, 2005).

This chapter offers a broad review of contemporary methodology used in the field of media psychology in studying *the effects of exposure to media violence on the consumer of such media*. Although the basic principles and ideas described here apply more broadly to other domains of media-related research, such as motivations underlying media choices and preferences (e.g., Ryan, Rigby, & Przybylski, 2006), we focus on the *effects of exposure* domain—and within this domain, we focus on *media violence* effects.

We discuss basic scientific principles that are at the foundation of all psychological research. An overview of widely used research designs is given. Common methodological pitfalls in media psychology are described as well as some suggestions on how to avoid them. Finally, guidelines are given on how to convey scientific methodology and findings to the general public. We hope that this chapter will aid media psychology researchers in continuing the trend toward better methodological quality in the field, aid journal editors and reviewers in doing a better job of screening out weak and promoting (p. 110) strong research, and aid readers from other fields in becoming informed consumers of media psychology research.

Science, Causality, and Media Psychology

Empirical Research and Theory Development

Test/Revise/Test/Revise Cycle

Research in the field of media psychology can generally be divided into two approaches: quantitative and qualitative. Qualitative methods (e.g., content analyses, ethnographic studies and phenomenological studies) generate descriptive findings and are usually conducted without forming a priori hypotheses (for a discussion of qualitative methods see Chapters 8 and 23). The majority of media effects research, however, is quantitative and follows a different pattern, progressing through a cyclic interaction between theory and empirical research. Researchers identify a question of interest (e.g., What effects does media violence have on viewers?). One or more hypotheses are generated (e.g., Observing media violence will increase the likelihood of later aggression. Exposure to media violence will decrease helping.) and tested using multiple research methods. Empirical results lead to revisions and refinement of the original hypotheses. Over time, a set of related hypotheses and empirical findings is developed, a set that can be integrated into a larger conceptual model or theory. The theory can then be used to develop novel hypotheses that can be tested further through empirical research. The cycle is repeated, leading to further refinement of the theory. This extensive test/revise/test/revise process leads to the development of theoretical models based on sound principles which are unlikely to be invalidated by future research. For example, the general aggression model (GAM) (Anderson & Bushman, 2002b; Anderson & Huesmann, 2003; DeWall & Anderson, 2011) and the general learning model (GLM) (Buckley & Anderson, 2006; Barlett, Anderson, & Swing, 2009; Gentile et al., 2009) integrate a number of earlier models and are based on more than 100 years of psychological research on learning, emotion, cognition, and behavior. Well-tested models such as these provide a solid foundation for interpreting findings, making new predictions, and developing interventions. Nonetheless, specific interpretations can always be changed as a result of new discoveries. It is for this reason that scientists are reluctant to use the words *fact*, or *proven*, or *truth*, even when speaking with audiences and individuals who do not understand this perpetual cycle of theory and data. Thus, the general public may view the “theory” of evolution as a mere guess or hypothesis, whereas the scientific community knows that the basic tenants of the theory are as well established and as factual and basic as the law of gravity. This differential understanding of the meaning of “theory” and other common words leads to much unnecessary miscommunication among scientists and nonscientists, a topic that is addressed in a later section of this chapter.

Translations from Conceptual to Empirical

One frequently overlooked (or underevaluated) aspect of scientific theory development and testing concerns the multiple translations that take place between the conceptual/theoretical level and the specific procedures used to conduct empirical tests. That is, one must translate the conceptual hypothesis into specific empirical realizations of the independent and dependent variables (Carlsmith, Ellsworth, & Aronson, 1976; Anderson & Anderson, 1996). Figure 7.1 illustrates some of the multiple levels and translations that underlie an experimental manipulation of violent versus nonviolent violent game exposure. As can be seen, there are lots of levels between the most basic (and therefore the conceptually broadest) theoretical level and the specific manipulation that a researcher creates in an empirical study. Keep in mind that a similar set of translations are needed to get from the conceptual dependent variable (e.g., aggression) to its empirical realization. Thus, there are lots of ways one can test the same conceptual hypothesis. Furthermore, although theory provides many constraints on what should be considered reasonable tests of any given conceptual hypothesis, there is no such thing as a perfect empirical realization of that hypothesis. For this reason, multiple studies using multiple methods give a better overall picture of the validity of any conceptual hypothesis than any single method or study can give. Further discussion of this appears in the next section.

Causality

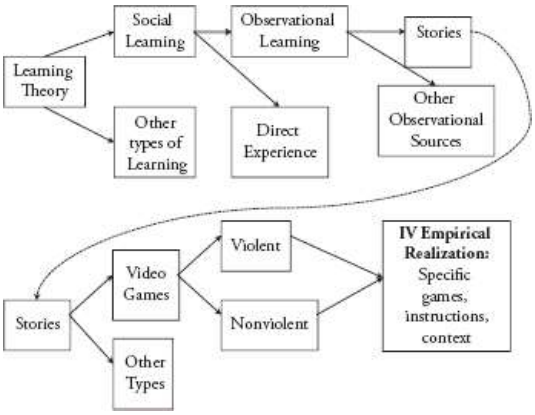


Figure 7.1 Illustration of Multiple Translation Levels from Learning Theory to Empirical Realization of the Independent Variable: Experimental Manipulation of Video Game Violence.

The majority of scientific theories and models in nomothetic scientific disciplines (those that seek to uncover general laws that underlie phenomena, such as natural sciences and psychology; Münsterberg, 1899) are causal. Widely used theories in media effects psychology such as social learning theory and social cognitive theory (Bandura, 1973, 1983), general aggression model (Anderson & Bushman, 2002b; Anderson & Huesmann, 2003), cultivation (p. 111) theory (Comstock & Scharrer, 2007), and social information processing theory (Crick & Dodge, 1994) all imply causal relationships among variables. The central characteristics of any good theory, the ability to predict and control outcomes, require causal models. Of course, establishing causality is often a difficult task, one that is seldom understood by public policy makers or the general public, and too often is misunderstood even by members of the scientific community. What follows is a partial listing of the most common difficulties.

Scientific Causality Is Probabilistic

The old Logic 101 principles of establishing causality do not apply to most modern science (Anderson & Bushman, 2002a). Scientific causality is probabilistic, instead of "necessary and sufficient." Stating that X causes Y means that variable X causes an increase in the likelihood of outcome Y (Anderson, 2004). For example, saying that smoking causes lung cancer means that smoking increases the likelihood of developing lung cancer. This does not mean that all smokers get lung cancer; many do not (a violation of the principle of "sufficient" causality). Furthermore, some nonsmokers do get lung cancer (a violation of the principle of "necessary" causality). Correspondingly, saying that violent video games cause aggression does not mean that every person who plays violent video games will necessarily become aggressive, or that all aggressive behavior is a result of violent video game play. It means that exposure to violent video games increases the likelihood of future aggression.

Probabilistic causality is a result of the fact that most (if not all) biological outcomes, disease processes, and human behaviors are multicausal (Gentile & Sesma, 2003). Complex behaviors of interest, such as prosocial behavior and aggression, are influenced by a large number of factors (e.g., genetic predispositions, parental practices, cultural influences) (Anderson & Huesmann, 2003; DeWall, Anderson, & Bushman, 2012). Media use is just one of many relevant factors that influence the likelihood of these behaviors. In most cases, it is neither a necessary nor a sufficient cause. Nonetheless, media effects are not negligible and have important practical consequences in many domains, including aggression (Anderson & Dill, 2000; Gentile et al., 2004; Anderson et al., 2010), helping (Greitemeyer, 2009; Greitemeyer & Osswald, 2010), risk taking (Fischer et al., 2011), and school performance (Sharif & Sargent, 2006; Anderson, Gentile, & Buckley, 2007; Rideout, Foehr, & Roberts, 2010), among others.

A methodological difficulty in the field of media psychology stems from the fact that many media effects are subtle, cumulative, and unintentional. For example, advertisements can have a subtle influence on viewers without their awareness (Gentile & Sesma, 2003). Although such short-term influences may be small, over time they can produce large cumulative effects. To use the cigarette smoking analogy, although short-term effects of smoking are relatively harmless and transient, long-term cumulative effects of this risk factor are lasting and severe. Likewise, although effects of watching (p. 112) a single violent TV show dissipate fairly quickly, habitual exposure to violent media has long-lasting effects on desensitization to violence (Bartholow, Bushman, & Sestir, 2005), hostile attribution biases (Anderson, Gentile, & Buckley, 2007), development of an aggressive personality (Bartholow, Sestir, & Davis, 2005), and aggressive behavior (Huesmann, Moise-Titus, Podolski, & Eron, 2003;

Anderson, Sakamoto, Gentile, Ihori, Shibuya, Yukawa, Naito, & Kobayashi, 2008; Möller & Krahé, 2009).

An interesting solution to the methodological difficulty of studying cumulative effects of media exposure is proposed by Potter (see Chapter 23). The author suggests that, instead of measuring group differences in effects of media exposure, as is done in the majority of media effects studies, attention should be shifted to patterns of effect score changes for individuals over time. This approach would allow researchers to directly examine the course of long-term changes produced by media influences, identifying how mass media influences gradually change a person's baseline. Indeed, this is essentially what longitudinal studies of media violence do (e.g., Huesmann, Moise-Titus, Podolski, & Eron, 2003; Anderson, Sakamoto, et al., 2008).

Role of Plausible Alternative Explanations

Testing scientific theories largely involves creating alternative explanations for a given phenomenon and then empirically testing whether the originally hypothesized relations among variables fits the data better than the alternative explanation. In essence, establishing causality involves testing and ruling out *plausible* alternative explanations. We emphasize "plausible" because the total number of alternative explanations—plausible + implausible—approaches infinity. Furthermore, only alternative explanations that are empirically testable (at least in principle) qualify as plausible. Alternative explanations that cannot be empirically tested (e.g., god did it) usually fall outside the realm of science. Of course, technological advances often create opportunities to test alternative hypotheses that previously had been untestable, which is why the "in principle" aspect of plausible alternative explanations is important. For example, recent advances in genetics and in neuroimaging have allowed tests of numerous new hypotheses about aggression and violence (DeWall, Anderson, & Bushman, 2011, 2012).

Relevant empirical results can cast doubt on alternative explanations and lend support to the target theory. Or, such tests can support an alternative explanation, thereby pointing to parts of the theory that need further revision. As the number of plausible alternative explanations is reduced, the strength of the remaining theoretical explanation increases.

Triangulation and Alternative Explanations

No single test of a theory-based hypothesis is definitive, irrespective of whether it confirms or disconfirms the prediction (Anderson & Anderson, 1996). One reason for this is because theoretical models involve abstract conceptual variables, whereas empirical tests involve concrete operationalizations of those variables (Carlsmith, Ellsworth, & Aronson, 1976). In other words (and as noted earlier), there are multiple levels of interpretation and specification between the theoretical model and empirical tests of the implications of that model (see Anderson & Anderson, 1996, for an example concerning the heat/ aggression hypothesis). Operationalization of a conceptual hypothesis involves making several assumptions concerning the empirical methods being used (e.g., reliability and validity of the measures, adequacy of the sample of variables and participants). Because of this, null results are often less informative than confirming results, especially when new measures or procedures are used. Findings that are in line with theory-based predictions give support not only to the target conceptual hypothesis, but also to various implicit assumptions made in the study. If, on the other hand, the study fails to support the hypothesis, a common reaction of researchers is to acknowledge that there are many possible reasons for those findings. The unexpected results possibly reflect the fact that the original hypothesis is wrong, but also might be a product of methodological weaknesses of the specific study. Typically, for null results to be informative and result in a major change in theory they have to be replicated many times, shown to be not the result of mere poor methods or small samples, and have to lead to a more comprehensive theory that accounts not only for the null results but also accounts for the many results explained by the original theory (Kuhn, 1962). Unfortunately, occasional unreplicated null results based on small samples or poor methods, are frequently misinterpreted in the media violence domain as evidence of a lack of effects overall. Nonsignificant findings from specific studies have been regularly used by media industry apologists to question the validity of studies showing significant harmful media effects, a criticism media violence scholars have faced many times (Huesmann & Taylor, 2003; Bushman, Rothstein, & Anderson, 2010). For example, one of the methodologically poorest media violence studies ever published (Williams & Skoric, 2005) is frequently promoted by the video game industry and gamers as proof that violent games don't increase aggression. They conveniently ignore that fact that that study didn't measure aggression, had severe dropout rate problems, had differential dropout rates in the two conditions, and failed to show that the "violent" and "nonviolent" game conditions actually differed on the amount of exposure to violent

games.[1]

Far greater support for a conceptual hypothesis is given if conceptual relations are repeatedly tested and confirmed using different methodologies in different contexts. This is the logic of multiple operationalism or triangulation (Campbell & Fiske, 1959; Anderson, 1987, 1989; Anderson, Gentile, & Buckley, 2007). Different types of research designs make different methodological assumptions, so if a conceptual relationship is confirmed time after time in studies using different designs, it is extremely unlikely that the results are just a byproduct of methodological flaws. Similarly, conceptual relationships that yield similar results using different (but theoretically compatible) measures or manipulations greatly strengthen one's confidence in the basic conceptual model. When weaknesses of a particular type of study do not apply to other types, this enables researchers to triangulate or home in on a true causal factor (Anderson, 1989). When a hypothesis survives many potential falsifications using varied methods, a robust effect is established. For example, Bandura's initial findings concerning the effect of televised violence on modeling of aggressive behavior may have been falsified by several possible alternative explanations (Bandura, Ross, & Ross, 1961, 1963a). However, today researchers have no doubt that televised violence increases aggression because this effect has been repeatedly shown using correlational studies (Eron, Huesmann, Lefkowitz, & Walder, 1972; McLeod, Atkin, & Chaffee, 1972), experimental studies (Bjorkqvist, 1985; Josephson, 1987), and longitudinal studies (Huesmann, Moise-Titus, Podolski, & Eron, 2003). The interpretation of Bandura's early studies has changed slightly as a result of changes in definitions of aggression. But our main point in this example is that when studies using various research designs and measures, done in a number of different contexts and with samples from diverse populations all converge on the same answer, we can be much more confident that this answer is indeed true. In the words of Richard Cardinal Cushing when asked about the propriety of calling Fidel Castro a communist, “When I see a bird that walks like a duck and swims like a duck and quacks like a duck, I call that bird a duck” (*The New York Times*, 1964).

Although this chapter is focused almost exclusively on quantitative research methods, it is important to emphasize that qualitative methods also play a significant role in the field of media effects, providing rich knowledge on the content of media messages and people's individual experiences that cannot be obtained through experimentation or other forms of quantitative research (see Chapter 8). Also note that the line between qualitative and quantitative research is sometimes quite blurry.

Outcome Measures, Research Designs, and Review Types

Outcome Measures

Choice of outcome measure is crucial to any study, both because it influences the likelihood that the study will yield useful results and because of theoretical relevance. A measure of aggressive behavior that is appropriate in one research context may well be inappropriate in another. For example, a count of how often each child trips, pushes, or bites another child in a daycare setting can be a useful measure of physical aggression in that research context (i.e., young children at daycare), but would not be a valid measure of physical aggression for college students in a laboratory setting. A less obvious but equally important example frequently arises in the study of violent video game effects. Because violent video games involve a lot of physical aggression and almost no indirect or relational aggression, the dominant theoretical models of social learning and development all predict that playing such games is most likely to influence physical aggression. Measures of verbal and indirect aggression are unlikely to provide sensitive tests of the main hypothesis that exposure to violent video games increases the likelihood of aggressive behavior. Similarly, the measure of physical aggression has to match the age of the participants and the research context. For example, a measure of trait physical aggression in which the participant reports the frequency of aggressive acts over the past year is inappropriate as the main outcome measure in a short-term experimental study in which participants have just played a randomly assigned violent or nonviolent video game. The recent game play cannot change the frequency of aggressive acts that the person committed in the year before the game play, unless of course time travel is involved. Of course, such a trait physical aggression measure might be influenced by the content (violent (p. 114) versus nonviolent) of a recently played game, but in such a case it would represent some type of memory or reporting bias, not a true measure of the video game effect on physical aggression (Anderson & Bushman, 2001). Yet, several short-term experimental studies have used traitlike measures of aggression as the main dependent variable.

It is impossible to succinctly describe all of the measures that have been used or could be used in the study of the effects that media have on consumers. We focus here on a few of the measures related to antisocial effects (e.g., aggressive behavior, cognition, and affect) and to prosocial behaviors, cognitions, and affect.

Aggressive Behavior Measures

Because definitions of conceptual variables such as "aggression" and "violence" differ somewhat between disciplines and even over time, clarity of definition is critical in theory development and in translating conceptual variables into empirical realizations (e.g., Carlsmith et al., 1976). Social psychologists have come to rely on a specific definition that is much narrower than what is used by the general public and in some other areas of psychology. Specifically, human aggression is "… any behavior directed toward another individual that is carried out with the *proximate* (immediate) intent to cause harm. In addition, the perpetrator must believe that the behavior will harm the target, and that the target is motivated to avoid the behavior" (Anderson & Bushman, 2002b, p. 28; see also Berkowitz, 1993; Baron & Richardson, 1994; Geen, 2001). *Aggression* and *aggressive behavior* are used interchangeably throughout this chapter. It is important to note that aggression is always a behavior; it is not an emotion, thought, or desire. Also note that it is not the outcome of a behavior that defines it as aggressive or not, but the intent of the behavior, that is, the intent to harm. Thus, shooting an arrow at another person with the intent to harm them is an act of aggression, regardless of whether the arrow strikes or missed the target person. A shortcoming of many media effects studies arises from failure to use this definition.

Physical Aggression in a Lab Setting

Numerous methods have been developed that allow direct observation and measurement of aggressive behavior in laboratory settings. A common procedure used to measure physical aggression is the *teacher/learner paradigm*, sometimes known as the *Buss aggression machine* paradigm (Buss, 1961; Geen & O'Neal, 1969; Milgram, 1974; Donnerstein & Berkowitz, 1981). In this procedure, participants are told that purpose of the study is to explore effects of punishment on learning. They are paired with a supposed second participant (actually a confederate). The real participant is selected to be the "teacher" and the confederate is selected to be the "learner." The participant presents stimuli to the confederate who seemingly tries to learn them. When the "learner" gives an incorrect response, the participant is supposed to punish him or her with an electric shock. Aggression is measured by the intensity and/or the duration of the shock the participant chooses to give the confederate. For example, Donnerstein and Berkowitz (1981) used this procedure to measure effects of combining violent and sexual content on aggression of males toward a female target. Participants who had viewed a violent, sexual film delivered shocks of a higher intensity to a female "learner" than did those who viewed films containing only violent or sexual content. There have been many variations of this task, including use of different types of punishments (e.g., hand in ice water instead of electric shock) (Ballard & Lineberger, 1999) and different rationales for why the participant is delivering punishments (Baron & Richardson, 1994, pp. 69–75).

Another common method of measuring physical aggression in the laboratory is the *competitive reaction time task* (Taylor, 1967; Bushman, 1995; Giancola & Parrott, 2008). Participants in this task compete against a supposed opponent on a reaction time task in which the winner delivers aversive stimulation (an electric shock or a noise blast) to the loser. In actuality, the pattern of wins and losses is predetermined by the experimenter. Provocation can be manipulated by increasing the intensity of shocks set by the "opponent." Aggression can be measured as the intensity, duration, or number of high-intensity blasts given. For example, Anderson and Carnagey (2009) used this paradigm to test the effects of violent and nonviolent sports video games on aggression. They found that playing violent sports games increased aggressive behavior, even after controlling for competitiveness. In other words, competitive reaction time task measures aggression, not competitiveness (Gaebelein & Taylor, 1971; Bernstein, Richardson, & Hammock, 1987). Like the teacher/learner paradigm described earlier, the competitive reaction time task has been used in various modified forms in hundreds of studies, and is one of the most extensively validated measures of physical aggression.

(p. 115) 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* requires the participant to evaluate the confederate. In the *evaluation paradigm* (Berkowitz, 1962), 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 one to ten electric shocks, in which one shock indicates a very favorable

evaluation and ten shocks indicates a very unfavorable evaluation. In some studies, the confederate evaluates the participant's solution. Generally, half of the participants receive a positive evaluation from the confederate (e.g., one shock), whereas the other half receive a negative evaluation (e.g., seven 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.

Barlett, Branch, Rodeheffer, and Harris (2009) used a more recently developed laboratory measure of physical aggression, the *hot sauce paradigm* (developed by Lieberman, Solomon, Greenberg, & McGregor, 1999) to measure how long the effects of brief exposure to violent video games persist. In this procedure, participants decide how much hot sauce another person (known to dislike spicy food) must consume. Alternatively, one can have the participant determine the degree of hotness of the sauce that the other person must consume. Aggression is measured as the amount of hot sauce given to the target and/or the degree of hotness of the sauce selected.

Indirect, Verbal, and Other Laboratory Aggression Measures

Some laboratory based studies use *verbal aggression* measures. For example, in some studies the participant is given the opportunity to provide a potentially harmful written or verbal evaluation of another person (e.g., another participant, a confederate, or the experimenter), and does so knowing that the evaluation could hurt the other person. Sometimes the verbal aggression is *direct*, meaning that the participants believe that the target of their harmful evaluations will see or hear it. For example, Wheeler and Caggiula (1966) had participants listen and later evaluate another person's (actually, a confederate's) extreme and socially undesirable statements. The participants believed that the other person would hear their evaluations, so anything negative in the evaluations would presumably cause some harm. These evaluations were recorded and later coded for the degree of hostility.

Sometimes the evaluation is in the form of ratings that the target will not see, but that the participant believes will harm the target *indirectly*. For example, Berkowitz (1970) randomly assigned some female undergraduates to an anger induction condition (in which they listened to a job applicant's insulting statements about university women) or a control condition. Half in each condition then listened to either a hostile or a nonhostile comedian. All participants then rated the job applicant on several measures, with the knowledge that their ratings could affect the applicant's chances of getting the job. Interestingly, the women who had heard the hostile humor gave the applicant worse ratings than those who had heard the neutral humor. Other similar indirect verbal aggression measures have ranged from ratings of competence, to liking, job performance, and grades (Obuchi, Kameda, & Agarie, 1989; Dill & Anderson, 1995).

Perhaps the most recent addition to the list of laboratory aggression tasks is the *tangram task* (Gentile et al., 2009). In one study Gentile et al. randomly assigned participants to play a violent video game, a prosocial video game, or a game that was neither violent nor prosocial. Later, participants assigned an anonymous partner a set of 11 easy, moderately complex, or difficult tangram puzzles to attempt to solve within 10 minutes. Participants were led to believe that the partner would win a prize if they completed a sufficient number of puzzles in 10 minutes. The number of hard puzzles chosen constituted a measure of aggression, whereas the number of easy puzzles measured helping behavior. As expected, the violent video games increased aggressive choices, whereas the prosocial games increased helpful choices. Because this measurement task is the newest, it has received less empirical attention that the older measures described earlier, and thus does not yet have an extensive network of validation studies.

Aggression Measures Outside the Lab

The variety of ways that one can measure aggressive behavior outside of controlled laboratory settings is huge, limited only by the combination of the conceptual definition and the creativity of researchers. Generally, they can be categorized as self-reports, other reports, and archival.

Self-reports may be very specific, such as reporting how many physical fights one has been in during the past school year. Or, they may be broad traitlike measures of habitual aggressiveness. They may (p. 116) include any type of aggression (e.g., verbal, physical) at any severity level (e.g., said mean things about a classmate, attacked a peer with a knife or gun). Common self-report measures of trait aggression in the media effects domain include the physical and verbal aggression subscales of the Buss and Perry (1992) Aggression Questionnaire and the

physical violence subscale from the National Youth Survey (Elliot, Huizinga, & Ageton, 1985; Anderson & Dill, 2000). Other commonly used self-report trait aggression scales that include relatively more items that are not strictly aggressive behaviors are the Caprara Irritability scale (Caprara et al., 1985) and the Cook-Medley Hostility Inventory (Cook & Medley, 1954). Of course, there are many self-report measures, and researchers create new ones as the empirical and theoretical need arises.

Others' reports of aggression include a wide range of measures, usually subcategorized into peer reports, teacher/supervisor reports, parent reports, and direct observation. Peer reports are frequently used in pre–high school settings. Often these involve asking each student in a classroom to rate each of their classmates on specific behaviors, or to nominate classmates who do certain aggressive behaviors. For example, it is common to ask, "Who pushes, shoves, or hits other kids to get what they want?" Teacher and supervisor reports ask similar questions about those under their care or supervision. Parent reports often ask about the frequency with which their child has done specific aggressive behaviors; other parent reports are vaguer, asking for ratings of "how aggressive" is your child. Direct observation studies often involve the recording of behavior in some naturalistic setting, followed by standardized coding of the recorded behavior. Sometimes, however, trained observers watch and code behaviors directly in the setting, such as while watching children on a playground.

Archival measures are derived from written records, such as crime reports and school incident records. Frequently archival measures are combined with other types of aggression measures.

Aggressive Cognition Measures

Exposure to violent media has a host of cognitive consequences, which in turn can lead to aggressive behavior. A number of outcome measures have been used to assess influences of violent media on cognition in both short- and long-term contexts.

Aggressive Cognition in Lab Settings

Laboratory experiments measure short-term influences of exposure to violent media on cognitive processing. Such short-term effects mainly occur through priming of aggressive knowledge structures, making them more accessible (Anderson & Huesmann, 2003). Various methods have been successfully used in laboratory settings to measure these increases in aggressive thinking.

A number of studies have shown an increased frequency of aggressive thought content during or immediately after media violence exposure. For example, Calvert and Tan (1994) used a *thought-listing* questionnaire to measure differences in aggressive thoughts while observing or playing a violent game in virtual reality. In a study by Bushman (1998), participants made *free associations to nonaggressive words and to homonyms* with one meaning more aggressive than the other (e.g., cuff, mug). More aggressive associations were made to both types of words by participants who had just watched a violent video.

Several studies have used a *word completion task* to measure accessibility of aggressive thoughts (Anderson, Carnagey, & Eubanks, 2003; Anderson, Carnagey, Flanagan, Benjamin, Eubanks, & Valentine, 2004; Barlett et al., 2008). In this kind of task, participants are given a list of word fragments and are asked to fill in the missing letters to form the word. Some of the fragments can be completed to form either an aggressive word or a nonaggressive word (e.g., "h_t" can become *hit* or *hat*). Aggressive thought accessibility can be calculated as the proportion of word completions that were aggressive. Similar tasks have been commonly used to measure implicit memory (e.g., Roediger, Weldon, Stadler, & Riegler, 1992), and have been used to assess accessibility of prosocial thoughts as well (e.g., Greitemeyer, 2011).

A number of studies have used *reading reaction times* to aggressive and nonaggressive words as a measure of accessibility of aggressive cognitions (also called the *word pronunciation task*) (e.g., Bushman, 1998; Anderson & Dill, 2000; Anderson, Carnagey, & Eubanks, 2003). In the reading reaction time task (e.g. (Anderson et al., 1996; Anderson, 1997; Anderson, Benjamin, & Bartholow, 1998), participants are timed as they read aggressive and nonaggressive words. Average reaction times to aggressive and nonaggressive words can be compared to assess relative accessibility of aggressive thoughts. An advantage of this task is that suspicion or hypothesis-related demand characteristics are unlikely to influence responses because participants are taxed with trying to read all words as quickly as possible (Anderson, 1997). Furthermore, attempts by suspicious participants to bias the results

in either direction can be (p. 117) detected by examining the distribution of reaction times, because such biasing attempts typically yield unusually long reaction times.

In a somewhat different approach, Uhlmann and Swanson (2004) measured the effects of violent video game play on automatic aggressive thoughts using the implicit association test. This study showed that media violence exposure can teach a person to automatically associate the self with aggressive traits and actions. More recently, Saleem and Anderson (in press) have used another version of this task to assess anti-Arab bias.

Interesting methods have been used to assess cognitive biases that result from media violence exposure. For example, to assess hostile expectation bias Bushman and Anderson (2002) had participants read ambiguous story stems about potential interpersonal conflicts. Participants were then asked to list what the main character will think, feel, say, and do next and their responses were coded for aggressive content. Several media effect studies (Kirsh, 1998; Anderson, Gentile, & Buckley, 2007; Möller & Krahé, 2009) have also used ambiguous provocation stories to assess hostile attribution bias. In each story, an actor causes a negative event to happen, but the intent of the actor is unclear. After each story, participants are asked a series of questions concerning the provocateur's intent. It has been shown that exposure to media violence leads to the development of a hostile attribution bias, a tendency to interpret ambiguous behaviors of others as malevolent (Kirsh, 1998).

Yet another method of assessing accessibility of aggressive cognitions is the word pair similarity rating task. This task was originally developed by Bushman (1996) to assess individual differences in aggressive cognitive-associative networks. But a minor revision to the task has been used to examine the effects of short-term experimental manipulations of variables, including pain (K. Anderson, Anderson, Dill, & Deuser, 1998), cooperative versus competitive video game instructions (Anderson & Morrow, 1995), and violent versus nonviolent music lyrics (Anderson, Carnagey, & Eubanks, 2003). This task consists of rating the degree of meaning similarity of each paired combinations of 20 words. Ten of these words have both aggressive and nonaggressive connotations (e.g., bottle, night, stick). These words are referred to as ambiguous words. The remaining ten words are more obviously related to aggression (e.g., butcher, choke, hatchet). Ratings of each word pair are made on a 1 to 7 scale of how “similar, associated, or related” they are. Each participant gets three scores, the average similarity rating of all ambiguous/aggressive word pairs, ambiguous/ambiguous pairs, and aggressive/aggressive word pairs. Anderson, Carnagey, and Eubanks (2003) found that participants who had just listened to songs with violent lyrics gave higher similarity ratings to ambiguous/aggressive word pairs than did participants who had just listened to nonviolent songs. In other words, the violent lyrics increased the accessibility of the aggressive meaning of the ambiguous words.

Aggressive Cognition Outside the Lab

Correlational studies and longitudinal studies provide an opportunity to explore long-term influences of violent media on cognition. Repeated exposure to media violence strengthens aggression-related knowledge structures and can make them chronically accessible. Additionally, long-term exposure reinforces normative beliefs that violence is common and appropriate (Carnagey & Anderson, 2003; Bushman & Huesmann, 2006). Dependent variables in correlational and longitudinal studies of aggressive cognition often include normative beliefs about violence (Gerbner, Gross, Jackson-Beeck, Jeffries-Fox, & Signorelli, 1978; Gerbner, Gross, Morgan, & Signorelli, 1980; Bryant, Carveth, & Brown, 1981), positive attitudes toward violence (Funk et al., 2004; Anderson, Gentile, & Buckley, 2007) and hostile attribution bias (Anderson, Gentile, & Buckley, 2007). These long-term consequences can be assessed using self-report measures, such as the Normative Aggressive Beliefs Scale (Anderson, 2004; Anderson, Gentile, & Buckley, 2007), Huesmann's NOBAGS scales (Huesmann et al., 1992), Funk's Attitudes toward Violence Scales (Funk, Elliott, Urman, Flores, & Mock, 1999), and the Revised Attitudes toward Violence Scale (Anderson, Benjamin, Wood, & Bonacci, 2006). Some studies also use trait measures of aggressive cognition, such as the hostility subscale of the Buss-Perry Aggression Questionnaire (Anderson, Carnagey, Flanagan, Benjamin, Eubanks, & Valentine, 2004; Shibuya, Sakamoto, Ihori, & Yukawa, 2004; Bartholow, Sestir, & Davis, 2005).

Aggressive Affect Measures

Another route through which violent media can increase aggression is by producing feelings of anger and hostility (Anderson et al., 2003; Swing & Anderson, 2010). Brief exposure to media violence has been shown to lead to temporary increases in aggressive affect (Barlett et al., 2009), whereas chronic exposure leads to the

development of a (p. 118) hostile personality (Bartholow, Sestir, & Davis, 2005; Bushman & Huesmann, 2006).

Aggressive Affect in Lab Settings

Experimental studies in laboratory settings measure effects of media violence exposure on short-term affective states. Short-term affective consequences are most often assessed using self-report scales such as the State Hostility Scale (SHS) (Anderson, Deuser, & DeNeve, 1995), the Multiple Affect Adjective Checklist (Zuckerman, 1960; Zuckerman, Lubin, Vogel, & Valerius, 1964), or the State Anger subscale of the State-Trait Anger Expression Inventory (STAXI) (Spielberger, 1988). Many other studies have used the Positive and Negative Affect Schedule (PANAS) (Watson, Clark, & Tellegen, 1988). This widely used self-report scale has the advantage of assessing both positive and negative affect, as well as several more specific subtypes of affect. However, it is a less sensitive measure of hostility/anger, most likely because of fewer items for this more specific affective state (Anderson, Deuser, & DeNeve, 1995; Anderson, Anderson, & Deuser, 1996; Anderson, Anderson, Dorr, DeNeve, & Flanagan, 2000).

Such self-report measures can be complemented with physiological indicators such as heart rate, blood pressure, or skin conductance (Ballard, Hamby, Panee, & Nivens, 2006; Carnagey, Anderson, & Bushman, 2007). Additionally, researchers have started examining neural bases of short- and long-term media effects on emotional processing by using techniques such as magnetic resonance imaging (Weber, Ritterfeld, & Mathiak, 2006) and event-related brain potentials (Bartholow, Bushman, & Sestir, 2005; Bailey, West, & Anderson, 2011a).

Aggressive Affect Outside the Lab

Long-term changes in affective processing as a result of habitual media violence exposure can be assessed outside the laboratory using trait measures such as the Caprara Irritability Scale (CIS) (Caprara et al., 1985), the Cook-Medley Hostility Inventory (Cook & Medley, 1954), and the anger subscale of the Buss-Perry Aggression Questionnaire (Buss & Perry, 1992). Once again, more general trait-affect scales may be appropriate in some research contexts, but researchers need to be aware that general measures of positive and negative affect usually are less sensitive measures of any give specific affect type.

Physiological Arousal Measures

For most people, exposure to media violence tends to produce physiological arousal (Anderson et al., 2003; Swing, Gentile, & Anderson, 2008). Arousal can be measured in experimental studies using indicators such as heart rate, blood pressure, or skin conductance (Ballard & Wiest, 1996; Fleming & Rickwood, 2001; Anderson et al., 2004; Barlett et al., 2008).

How lasting are these effects? Barlett et al. (2009) showed that heightened arousal immediately after playing a violent video game lasts between 4 and 9 minutes. However, these short-term changes can start aggression promoting processes that last much longer than 4 to 9 minutes, such as long-term desensitization to violence. Even within a short-term context, exposure to violent media may increase arousal (and hostile affect) for longer periods of time if the violent media episode increases processes that typically last long and that increase anger arousal, such as rumination on a perceived unjust harm.

A popular belief in our culture is that playing violent video games or watching violent television and films allows people to “vent” their aggression, decreasing arousal and reducing subsequent aggressive behavior (Anderson, Gentile, & Buckley, 2007). According to the catharsis hypothesis, engaging in real or imagined aggression helps relieve angry feelings, leaving us emotionally calmed (Dollard et al., 1939; Campbell, 1993). However, the bulk of research evidence opposes the catharsis hypothesis (Mallick & McCandless, 1966; Geen, Stonner & Shope, 1975; Geen & Quanty, 1977; Bushman, Baumeister, & Stack, 1999; Geen & Quanty, 1977 Bushman, 2002). Although physiological arousal can decrease after the initial aggressive act, later aggressive behavior does not (Geen & Quanty, 1977). Instead, studies show that viewing, thinking about or performing aggressive acts increases the likelihood of aggressive behavior (Dill & Dill, 1998; Geen, 2001).

Desensitization/Empathy Measures

Repeated exposure to violence can lead to desensitization, best defined as a reduction in emotional and

physiological reactivity to violence (Carnagey, Anderson, & Bushman, 2007). Empathy can be defined as the degree to which a person identifies and commiserates with a victim and feels emotional distress (Anderson et al., 2010). A small number of high-quality studies exist in this domain (Anderson et al., 2010). However, media violence has been clearly linked to both short-term desensitization as a result of brief exposure (Carnagey, Anderson & Bushman, 2007), and chronic desensitization and decreased empathy as a result of habitual, long-term (p. 119) exposure (Mullin & Linz, 1995; Funk et al., 2003; Bartholow, Bushman, & Sestir, 2006).

Short-Term Effects on Desensitization/Empathy

Desensitization to violence after brief periods of exposure is typically explored in experimental studies using physiological indicators such as heart rate, blood pressure, and galvanic skin response (Thomas, Horton, Lippincott, & Drabman, 1977; Linz, Donnerstein, & Adams, 1989; Carnagey, Anderson & Bushman, 2007). For example, participants in the Carnagey et al. (2007) experiment played a violent or nonviolent video game, and then watched film clips of real violent behavior, including shootings, stabbings, and fights. Heart rate and skin conductance were recorded before and during video game play, and during observation of the violent film clips. Both physiological indicators of emotional arousal increased in both game conditions while playing the assigned video games, but only those who had played a violent game showed decreases in arousal while watching the violent film clips.

Long-Term Effects on Desensitization/Empathy

Neurological evidence of chronic desensitization to violence through playing video games also exists. Bartholow, Bushman, and Sestir (2005) found that habitual violent game players have reduced amplitudes of the P300 component of the event-related brain potential while viewing violent images. Other laboratory studies have found similar effects (Kronenberger et al., 2005; Bailey et al., 2011a). Outside the laboratory, long-term effects on desensitization and empathy can be measured using self-report scales such as the Basic Empathy Scale (Jolliffe & Farrington, 2006), the Interpersonal Reactivity Index (Davis, 1980), the Index of Empathy for Children and Adolescents (Bryant, 1982) or Children's Empathic Attitudes Questionnaire (Funk et al., 2008). Indeed, longitudinal studies have yielded evidence of long-term changes in desensitization/empathy as a result of media violence exposure (see Anderson et al., 2010, for the video game case).

Prosocial Behavior/Helping Measures

Prosocial behavior involves helping or rewarding others, especially when this behavior brings no benefit to the helper (Barlett, Anderson, & Swing, 2009). Effects of violent media on prosocial behavior have been less frequently explored than effects on aggression. In spite of this, several measures have been developed that make it possible to perform reliable and valid measurement of prosocial behavior and helping.

Prosocial Behavior in Lab Settings

Several procedures have been used in media violence research that allow direct observation and measurement of prosocial behavior in the laboratory. For example, Chambers and Ascione (1987) showed that children who had played a violent game donated less to charity. Ballard and Lineberger (1999) employed a variation of the teacher/learner paradigm in which participants could award jelly beans to their partner. The number of jelly beans awarded served as a measure of helping and it was shown that participants who had just played a violent game tended to award a smaller number of jelly beans.

Bushman and Anderson (2009) simulated a fight in a laboratory experiment and found that participants who had played a violent video game were less likely to help and took more time to help the “victim.” These participants perceived the fight as less serious and were less likely to notice the fight than the participants who played a nonviolent game.

Sheese and Graziano (2005) used a prisoner's dilemma game in which participants were given a choice to cooperate with their partner for mutual gain, exploit their partner for their own benefit or withdraw. Participants in the violent condition were significantly more likely to choose to exploit their partner.

Gentile et al. (2009) developed a new task to measure helping—the previously mentioned *tangram task*. In this

task, participants are asked to assign easy, moderately complex or difficult tangram puzzles to an anonymous partner. Participants are told that the partner will win a prize if they complete a sufficient number of puzzles in 10 minutes. The number of easy puzzles represents a measure of helping behavior. Participants who had just played a prosocial video game assigned the most easy tangrams, where as those who had just played a violent game assigned the fewest.

Prosocial Behavior Outside the Lab

The most common type of measure chosen outside laboratory settings are self-report questionnaires such as the Prosocial Orientation Questionnaire (Cheung, Ma, & Shek, 1998). For example, a correlational study by Gentile et al. (2009) assessed video game habits of a large sample of children, along with several prosocial measures. Playing violent video games was negatively related to helping behavior, whereas prosocial gaming was positively associated with helping.

(p. 120) A longitudinal study by Anderson, Gentile, and Buckley (2007) measured children's media violence exposure and prosocial behaviors twice during a school year and showed that video game violence at time 1 significantly predicted a relative decrease in prosocial behavior over time. In this study, prosocial behavior was measured using teacher ratings and peer ratings.

Another less common measurement procedure is naturalistic observation. In an unusual field experiment by Bushman and Anderson (2009), violent and nonviolent movie attendees saw a young woman with an injured leg struggle to pick up her crutches. Participants who had just watched a violent movie took longer to help than those who had just watched a nonviolent movie. Violent and nonviolent moviegoers did not differ in their helpfulness before seeing the movie.

It is important to emphasize that prosocial and antisocial behaviors are not simply opposite sides of the same coin. Measures of aggressive and prosocial behavior tend to be negatively correlated, but not strongly so. One can be high both in helpful and in hurtful behaviors—for example, hostile toward enemies and kind toward friends (Gentile et al., 2009).

Research Designs

Researchers generally use three broad types of research designs: experimental studies, cross-sectional correlational studies, and longitudinal studies (Anderson & Bushman, 2001; Swing & Anderson, 2010). Each design has its own advantages and disadvantages and is appropriate for certain kinds of research problems. Findings from different kinds of studies complement each other and help researchers form a complete picture of media effects.

Experimental Studies

Advantages

In experimental studies, researchers manipulate exposure to media content and measure brief, short-term effects. Participants are randomly assigned to treatment and control groups; for example, playing a violent or nonviolent video game (Anderson & Dill, 2000). With all other factors controlled, a difference between two groups, for example, in aggression, establishes a causal link between violent media and subsequent aggression. Random assignment ensures that there are no preexisting differences between the two comparison groups (within certain statistical limits) and allows researchers to rule out a host of alternative explanations. If a difference in aggressive behavior of the two groups is found, it is very likely that this difference was caused by experimental manipulation (exposure to video game violence). It is very improbable (although not impossible) that highly aggressive individuals just happened to be randomly assigned to the experimental group and nonaggressive individuals were assigned to the control group. The larger the sample size, the less likely it is that a disproportionate percentage of highly aggressive people were randomly assigned to any one condition, just as tossing a coin 100 times is less likely to yield 80% “heads” than tossing it only ten times.

If the researcher has additional information about the research participants before they are assigned to condition, information that may be relevant to the dependent variables of interest such as gender or trait aggressiveness,

they may decide to "block" on these other variables in the random assignment procedure. For example, they may separately randomize males and females to the different experimental conditions to ensure that each gender is represented equally across the conditions; but the logic and power of true experiments does not require this.

Methodologically sound experimental studies in the field of media psychology share several characteristics—they are designed so that they control for many possible alternative explanations (i.e., high internal validity), have adequate sample sizes, employ effective experimental manipulations, and use a reliable and valid measure of the dependent variable.

High-quality laboratory experiments use well-validated paradigms to test relevant hypotheses. For example, Anderson and Dill (2000) conducted a laboratory experiment to test short-term effects of playing a violent video game on aggressive thoughts and behavior. In this study, a large sample of 227 college students participated. Participants were randomly assigned to play a violent or a nonviolent game. Games used in the study were carefully pretested and matched on several relevant dimensions (e.g., difficulty, frustration, and the physiological arousal levels they produce). Aggressive behavior was measured using a modified version of the Competitive Reaction Time Task (Taylor, 1967), a widely used measure of aggressive behavior that has well-established internal and external validity (Carlson, Marcus-Newhall, & Miller, 1989; Anderson & Bushman, 1997; Giancola & Chermack, 1998; Anderson, Lindsay, & Bushman, 1999). Aggressive cognition was measured with a reading reaction time task that had been successfully used in previous aggression studies (Anderson, 1997; Anderson, Benjamin, & Bartholow, 1998) as well as in many studies in cognitive psychology. Violent (p. 121) video game play led to significant increases in aggressive cognition and aggressive behavior. This study made an important contribution to the violent video game effects literature because previous experimental studies in this area had methodological weaknesses that put their results into question. A number of early experiments testing for violent video game effects (Cooper & Mackie, 1986; Silvern & Williamson, 1987; Schutte, Malouff, Post-Gorden, & Rodasta, 1988; Irwin & Gross, 1995) did not match violent and nonviolent games on important dimensions and thus could not rule out the possibility that other variables such as arousal, difficulty, or frustration caused the observed difference in aggressive behavior.

High-quality field experiments use measures of real-life behavior in natural settings. For example, as mentioned earlier, Bushman and Anderson (2009) tested effects of media violence on helping behavior by staging a minor emergency outside movie theaters that were showing either a violent or a nonviolent movie. Moviegoers saw a young woman with a wrapped ankle "accidentally" drop her crutches outside the theater and struggle to pick them up. The emergency was staged either before the movie (to control for helpfulness of people who choose to view violent versus nonviolent movies) or after the movie (to test for the effect of viewing media violence on helping). In this case, the randomization was whether the measure of helpfulness occurred before or after viewing the movie. Before watching a movie, no differences in helping were found between those going to a violent versus nonviolent movie. However, after the movie, participants who had just viewed a violent movie took significantly longer to help the confederate than those who had viewed a nonviolent movie.

Disadvantages

The main advantage of experimental studies is that they enable strong causal inferences. A potential disadvantage concerns the ability to generalize results to real-life conditions. Field experiments don't suffer this concern. But, because most experiments are conducted in the laboratory, the generalizability of findings from such studies to real-world settings is sometimes questioned. However, such doubts have been challenged and refuted both by rational arguments (e.g., Mook, 1983) and empirical studies of external validity of laboratory experiments (e.g., Anderson & Bushman, 1997).

The main purpose of most laboratory studies is to explore conceptual relationships between variables and thus test and develop theories. The goal is to be able to generalize these underlying theoretical principles, not specific features of the sample, manipulation, or measure (Berkowitz & Donnerstein, 1982; Henshel, 1980; Mook, 1983; Banaji & Crowder, 1989; Anderson & Bushman, 1997). Conceptual relationships between variables generalize, even if specific operationalizations do not.

The external validity of laboratory experiments is also supported by empirical findings from several studies. For example, in the aggression domain it has been shown that laboratory measures of aggression are positively associated with each other, and that variables that influence aggression and violence in the real world have the

same kind of effects on laboratory measures of aggression (Carlson, Marcus-Newhall, & Miller, 1989; Anderson & Bushman, 1997; Bushman & Anderson, 1998). Similarly, Anderson, Lindsay, and Bushman (1999) explored the consistency between findings obtained in laboratory and field settings across several domains in psychology (e.g., aggression, helping, leadership style, social loafing, self-efficacy, depression, and memory). This study found considerable correspondence between lab- and field-based effect sizes, suggesting that laboratory experiments have high external validity.

Laboratory settings also enable researchers to explore relationships between variables that may never be sufficiently isolated in real life to enable precise testing (Mook, 1983). If increasing the similarity of the laboratory situation to real-world conditions interferes with the internal validity of the study, external dissimilarity (to achieve high internal validity) is strongly favored (Anderson & Bushman, 1997).

There are two additional potential disadvantages of experimental designs in media effects studies. Both involve ethical considerations. First, one cannot ethically conduct an experiment in which one of the experimental treatments is expected to increase a seriously harmful behavior, such as aggravated assault or homicide. One can't randomly assign a group of 10 year olds to play either a violent or nonviolent video game, then give each a handgun, and turn them loose on the playground to see which group does the most killing during recess. For this reason, alternative measures of aggressive behavior have been developed and used. Field experiments typically measure milder forms of physical aggression, such as hitting, pushing, shoving, and biting. Laboratory experiments use a variety of measures of aggression, including measures of physical and verbal aggression. And as noted earlier, these measures have been well validated, showing high levels of external validity.

(p. 122) The second potential disadvantage of experimental designs in this domain concerns duration of the manipulation. It is not ethical to intentionally expose a group of participants to a long-term high media violence diet to see whether this randomly assigned group becomes more aggressive than a randomly assigned control group. One can't randomly assign a group of 4 year olds to grow up in either a high or low media violence household, and then measure their level of aggressiveness in school or criminal records at age 18. One can, however, use a long-term experimental design to see if an intervention designed to reduce exposure to media violence has any effect on aggression. A few such experimental intervention studies have been done (e.g., Huesmann, Eron, Klein, Brice, & Fischer, 1983; Robinson, Wilde, Navracruz, Haydel, & Varady, 2001), and have found that such interventions can reduce aggression.

Cross-sectional Correlational Studies

Advantages

Cross-sectional correlational studies explore the direction and magnitude of associations among relevant variables. The independent variable is measured instead of manipulated, and both the independent and dependent variable are measured once, usually at the same point in time. Strengths of correlational studies include the ability to measure real-world outcomes, test different alternative explanations, and suggest new hypotheses about causal relationship.

Disadvantages

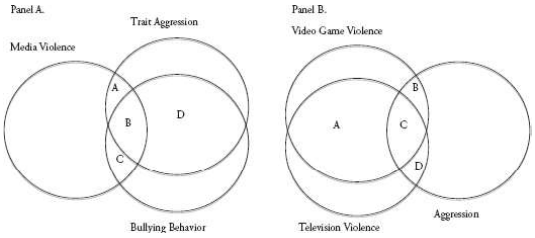
The main weakness of correlational research is difficulty in establishing causality. Results of a single correlational study in which variables are measured at the same single point in time cannot ascertain cause-and-effect relationships. In other words, correlational studies generally have lower internal validity than experimental studies (Anderson & Bushman, 1997). Of course, some correlational studies are more informative about causality than others. For example, some of the early violent video game effect studies had serious methodological difficulties (Dominick, 1984; Lin & Lepper, 1987; Fling et al., 1992). These studies showed significant associations between playing video games and aggression, but did not distinguish between playing violent versus nonviolent games. In contrast, Anderson, Gentile, and Buckley (2007) tested the strength of the association between aggression and violent video game play, while controlling for several key competitor variables (total screen time, normative aggression beliefs, positive orientation toward violence and sex). This example leads us to the important concept of destructive testing.

Destructive Testing

Because of the critical role played by testing plausible alternative explanations in theory development, even cross-sectional correlational studies can play an important part in testing causal hypotheses. They can provide an opportunity for falsification of the causal hypothesis as well as for testing and ruling out alternative hypotheses. Well-designed correlational studies can measure many theoretically relevant variables along with the target independent variable and the target dependent variable, and then statistically control for effects of those other variables. For example, Anderson and Dill (2000, Study 1) used the *destructive testing approach* (Anderson & Anderson, 1996) to assess the strength of the relationship between violent video game exposure and aggression. In this approach, a predicted relationship between variables is first established. Then one attempts to break the relationship by adding competitor variables. The key question is not whether the relationship can be broken—even strong truly causal links can eventually be rendered nonsignificant in a correlational study by adding more correlated predictors into the model. Instead, the focus of destructive testing is on how difficult it is to break the relation, considering the theoretical and empirical strength of the competitor variables used to test it. If the inclusion of several relevant competitor variables fails to break the relationship, this gives strong support to the validity of the target link. For example, in the study by Anderson and Dill (2000), the effect of violent video game play on aggression remained significant even with the inclusion of variables such as time spent playing any kind of video game and sex. Statistically controlling for these covariates invalidated several possible alternative explanations of the video game violence effect, thereby strongly supporting the authors' prediction that playing violent games will increase aggression. When using destructive testing, relevant covariates may include confounds (e.g., sex), potential competitors (e.g., total time spent playing), and potential mediators (e.g., aggressive personality). Occasionally, researchers also have mistakenly included as covariates variables that are better conceived as additional outcome (dependent) variables.

If the target link is broken by a single competitor variable or a single confounded variable, this puts the validity of the original causal hypothesis into (p. 123) question. However, mediating variables and secondary outcome variables have a very different theoretical status in correlational studies. Mediating variables are those that theoretically link the predictor variable to the outcome variable. In essence, they are another outcome variable of the same independent variable. For example, repeated exposure to violent video games (the predictor or independent variable) may increase aggressive behavior (outcome or dependent variable) because such exposure increases trait aggressiveness (the mediator variable). Thus, a proposed mediator variable should significantly weaken or even break the link between the predictor and outcome variables, even when that link is causal. When this happens, it lends support for the predicted theoretical model. Unfortunately, some gamers/media researchers (e.g., Ferguson et al. 2010) either don't understand this principle or they choose to ignore it when promoting their position. They have incorrectly concluded that when mediator variables such as trait aggressiveness weaken the correlation between habitual exposure to violent video games and aggressive behavior, this weakening of the key link contradicts the main theoretical hypothesis; in reality, such a result supports the causal model.

Figure 7.2A displays this issue with a Venn diagram. The three circles represent the variance of three variables, media violence, trait aggression, and bullying behavior. The area represented by sections A + B represents the correlation (or overlap) between media violence and trait aggression. C + B represents the correlation between media violence and bullying. B + D is the correlation between trait aggression and bullying behavior. Significance tests of the various relations can be thought of tests of whether overlapping areas are significantly greater than zero. If media violence truly causes an increase in the likelihood of bullying behavior, and it does so at least in part because it increases trait aggression as a mediating variable, then the theoretically most appropriate test of whether media violence is significantly related to bullying is the B + C area. But when trait aggression is treated as a nuisance variable that is statistically controlled, then the test of the hypothesis includes only area C, an unrealistically conservative test. By adding more restrictions on what gets counted as media violence/bullying variance, such as by adding additional covariates that themselves are theoretical outcomes of high media violence exposure, one can further inappropriately reduce the "unique" overlap between media violence and bullying.



Click to view larger

Figure 7.2 Inappropriate Uses of Covariates in Regression Analyses. A. When a mediator variable is added as a covariate, it can significantly weaken or even break the link between the predictor and outcome variables, even when that link is causal. B. When two conceptually related predictors are used in the same regression model, considerable predictive variance is removed which results in an overly conservative significance test.

A related problem occurs when two conceptually related predictors are used in the same regression model. For example, one study included both television violence and video game violence as separate predictors of aggression (Ferguson, San Miguel, & Hartley, 2009). This also removes considerable predictive variance inappropriately because television violence and video game violence are highly correlated (in that sample: r (602) = .47, p ⟨ .001) yet both contribute to the same theoretical explanation (i.e., media violence increases aggression). Figure 7.2B displays this problem. Testing the video game effect on aggression after controlling for the television violence effect, that is, testing area B, is overly (p. 124) conservative. It may be appropriate to include two conceptually and empirically overlapping predictors in a model if one wants specifically to compare the unique contributions of these two predictors (as in Anderson et al., 2007 tests of the relative strength of old versus new media), but otherwise the hypothesis that media violence increases aggression is better tested by including these two variables in separate regression models or by statistically combining them into a single "media violence" predictor. Thus, models that test the hypothesis of media violence effects on aggression with more than one distinct media violence predictor are unnecessarily biased against this hypothesis, losing a substantial portion of the effect(s) of interest to a covariate that is not an alternative explanation at all.

A third version of this problem concerns cases in which a control variable (e.g., sex of participant) is correlated with both the main independent variable (e.g., video game violence) and the main dependent variable (e.g., physical aggression). Males spend more time playing violent video games than females, and also are more likely to use physical aggression in many contexts. If playing violent video games is truly a causal risk factor for later physical aggression, then at least part of the confounded variance predicting physical aggression truly belongs to the violent video game effect. Controlling for the sex effect in essence overcorrects for the confound between sex and exposure to violent video games. Thus, correlational studies that control for sex likely underestimate the true effect size of violent video games on physical aggression (Anderson et al., 2010).

This problem is not unique to the media violence domain; indeed, this is pretty basic to research design, statistics, and methodology. This diagram also illustrates another point. Although the media violence critics are quick to note that "correlation is not causation," they seem to miss the necessary counterpoint that "lack of correlation is not lack of causation." That is, the same third variable problems that make it risky to conclude on the basis of one or several cross-sectional studies showing significant overlap between X and Y that "X causes Y," also make it risky to conclude that "X is not a cause of Y" based on studies showing that X and Y do not significantly overlap, especially if theoretically inappropriate covariates are first controlled for the overlap tests.

Longitudinal Studies

Advantages

In longitudinal studies, independent and dependent variables are measured at two or more points in time. Such studies provide an opportunity to assess real-life consequences of long-term media exposure. Causality is easier to establish in longitudinal studies than in cross-sectional correlational studies because temporal relations among variables make it possible to rule out a host of alternative explanations. For example, media habits and school performance can be assessed both early and late in a school year (as was done in a study by Anderson, Gentile, & Buckley, 2007). Results can be analyzed to see if the amount of habitual entertainment screen time (television,

film, video games …) at measurement Time 1 predicts school performance at Time 2 after statistically controlling for Time 1 school performance. The finding that total habitual screen time measured at Time 1 is a significant negative predictor of grades at Time 2 provides much stronger support for the hypothesis that time spent on television and video games has a negative effect on school performance than results from cross-sectional correlational studies showing a significant association at a single point in time (Anderson & Dill, 2000; Gentile et al., 2004; Sharif & Sargent, 2006).

In cases in which experimentally manipulating a particular independent variable would be difficult or unethical, longitudinal studies represent an excellent way for making sound causal inferences. For example, in a study by Hopf, Huber, and Weiß (2008), cumulative long-term influences of media violence exposure on adolescents' violence and delinquency were investigated—two behaviors that cannot be ethically investigated in an experimental study. The frequency of adolescents' exposure to media violence was measured over a 2-year period as well as exposure to eight other risk factors. Exposure to media violence at age 12 was a significant predictor of violence ($b = .28$) and delinquency ($b = .39$) at age 14, even after controlling for earlier levels of violence and delinquency and several other relevant variables.

Disadvantages

The main disadvantages of longitudinal designs are that they are time consuming and expensive. Repeated measurement requires researchers to keep track of participants and pay them to stay in the study. Large samples need to be taken to compensate for dropout rates. Another potential concern is nonrandom attrition. For example, in a 3-year study of television violence effects commissioned by the NBC television company (Milavsky, Kessler, Stipp, & Rubens, 1982), data from a large portion of the most aggressive participants in the sample were deleted because they allegedly had not (p. 125) given accurate reports of their television viewing. Although the original authors concluded that there was little evidence of a television violence effect, closer examination of this study reveals different conclusions (Kenny, 1984; Anderson et al., 2003; Huesmann & Taylor, 2003).

Mixed Designs

Many high-quality media effects studies combine multiple design features. Adding a correlational component in experimental designs can have several advantages. Including measures of relevant covariates makes it possible to perform additional tests of key hypotheses and explore effects of possible mediating and moderating variables. Including individual difference variables can also decrease error variance and increase the precision or power of statistical analyses, maximizing the likelihood that true effects will be detected. For example, an experimental study by Konijn, Bijvank, and Bushman (2007) elucidated the role of wishful identification as a possible moderator/mediator of violent video game effects. In this study, exposure to video game violence was experimentally manipulated. In addition to the dependent variable (aggressive behavior), several relevant covariates were measured (e.g., trait aggressiveness, general exposure to video games, immersion, and wishful identification). It was shown that playing a violent video game had the strongest effect on aggression for participants who wished they were like a violent character in the game. Furthermore, identification was associated with realism of the game and with immersion.

Some experimental studies also include a longitudinal component. For example, Huesmann, Eron, Klein, Brice, and Fischer (1983) conducted a 2-year intervention study that aimed to mitigate effects of television violence on aggressive behavior of school-aged children. Children selected because of their high exposure to violent television were randomly assigned either to a control group or an experimental group that received treatments designed to decrease effects of television violence (lessons about the unreality of television violence and an attitude change treatment). After the intervention, children in the experimental group were rated as significantly less aggressive by peers and showed a lower association between viewing violence and aggression.

A potential methodological difficulty in long-term experiments concerns the effective manipulation of the independent variable and control and measurement of possible confounds over a period of time. For example, an experimental study by Williams and Skoric (2005) attempted to measure effects of violence in a massively multiplayer online role-playing game (MMORPG, a type of online game in which a large number of players interact and play the roles of different characters) on aggression after 1 month of game play. However, exposure to other violent games was not controlled or measured during this 1-month period so no evidence existed that participants

in the violent game condition actually spent more time playing violent video games than participants in the control condition. Furthermore, the MMORPG used in this study was not very popular, which apparently resulted in players being unable to do much fighting in the game because of a lack of opponents. The participants in this study were recruited from online gaming sites. Furthermore, the overall dropout rate was huge, especially in the control condition, thus ruining the main advantage of experimental studies. Therefore, it is possible that during the study period participants in the control condition were exposed to as much (or even more) violent video game play than those in the violent game condition.

Scientific Literature Reviews

Each research design plays an important role in the study of media effects. Sound causal conclusions are based on consistent results across each of these designs (Abelson, 1995; Swing & Anderson, 2010). When a sufficient number of studies have been done on a specific topic, the results can be combined in a literature review. Such a review can answer additional questions, support or refute theoretical models, and point toward areas that are in need of further research. Reviews enable researchers to draw more advanced conclusions than would be possible on the basis of results from any single study. Two types of reviews can be performed—narrative and meta – analytic reviews.

Narrative Reviews

In traditional narrative literature reviews findings from relevant published empirical studies are described, categorized, and summarized. Possible goals of narrative reviews include providing an overview and integration of an area, theory evaluation and development, identification of weaknesses or contradictions in a specific field of investigation, and generating new problems and hypotheses (Baumeister & Leary, 1997). By searching for connections among a large number of empirical findings, narrative reviews can address much wider questions than any single empirical study. The major (p. 126) strength of narrative literature reviews is their focus on conceptual relationships between key variables that can lead to rich theoretical and methodological insights (Anderson, Gentile, & Buckley, 2007).

However, different studies necessarily yield somewhat different findings. Even if a study was replicated perfectly using the exact same methods, the results would be different because of effects of random factors. How should these differing results be interpreted and what conclusions should be drawn? A weakness of narrative literature reviews is that many critical decisions that are made while selecting and interpreting studies are subjective. This opens the door for reviewer biases that can result in drastically different interpretations of the empirical findings by different reviewers. People generally have a tendency to disregard evidence that contradicts their beliefs (Lord, Ross, & Lepper, 1979; Kunda, 1990; Koehler 1993), and reviewers are not exempt from such biases. Differing theoretical and empirical orientations can lead reviewers to form different inclusion criteria and organizational schemes, leading to different conclusions (Dill & Dill, 1998; Griffiths, 1999). Because of this, it's important that reviewers pay attention to counterexamples and allow themselves to be led by evidence rather than rigidly imposing a priori beliefs and expectations (Baumeister & Leary, 1997).

Meta-analytic Reviews

Meta-analytic reviews use statistical techniques to combine the results of a number of empirical studies that tested the same hypothesis. Meta-analyses describe the typical strength of an effect, its variability, its statistical significance, and variables that moderate it (Rosenthal, 1995). When a sufficiently large number of studies are available that tested the same hypothesis and a meta-analysis is usable, it is generally the preferred review method (Baumeister & Leary, 1997). By combining results from multiple studies, meta-analytic reviews can resolve inconsistencies caused by small sample sizes. The main strength of meta-analytic reviews is objectivity (Anderson, Gentile & Buckley, 2007). Unlike narrative reviews, meta-analyses done to answer a particular research question tend to give similar answers irrespective of different perspectives held by different reviewers. However, the meta-analytic reviewer still has to make important decisions concerning what studies to include and what studies to exclude from the sample. Thus, poorly conducted meta-analyses, those that do not include all relevant studies (Ferguson et al., 2010), can be just as misleading as a biased narrative review. The major weakness of meta-analyses is that the focus on statistical aspects sometimes leads the researchers to ignore important conceptual aspects.

In well-conducted meta-analyses, researchers attempt to find all available published and unpublished studies that might be eligible for inclusion in the sample, construct a clear and explicit set of inclusion criteria, and conduct publication bias analyses. For example, the most recent and comprehensive meta-analysis in the violent video game effects domain was conducted by Anderson et al. (2010). This meta-analysis combined a total of 136 research papers with 381 effect size estimates involving more than 130,000 participants from Eastern and Western countries. Six outcome variables were included in the meta-analysis: aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, desensitization/low empathy, and prosocial (helping) behavior. Newer studies of higher methodological quality made it possible to use more stringent inclusion criteria in this meta-analytic review than in prior reviews, and allowed tests of the effects of a number of relevant moderators (e.g., sex, culture, player's point of view). Both the best practices sample and the full sample yielded the same results: Violent video games had significant effects on all six outcome variables, showing that video game violence is indeed a causal risk factor for increased aggression and decreased prosocial behavior.

Methodological Pitfalls in the Field of Media Psychology

Conducting Studies in a "Theoretical Vacuum"

When attempting to understand underlying processes of media effects, it's important to keep in mind general knowledge in the field of psychology concerning mechanisms of memory, learning, social cognition, and development. Media effects research is informed by extensively replicated findings and well-validated theoretical models from several disciplines, including, among others, cognitive psychology, developmental psychology, personality psychology, social psychology, and neuroscience. Well-tested and generally accepted theories such as schema theory (Alba & Hasher, 1983; Schmidt & Sherman, 1984), social learning theory and social cognitive theory (Bandura, 1973, 1983), script theory (Huesmann, 1986, 1988, 1998), and risk and resilience models (Glantz & Johnson, 1999; Gentile & Sesma, 2003) provide a solid foundation for predicting and explaining findings in the field (p. 127) of media psychology. As Kurt Lewin (1951) put it, "There is nothing so practical as a good theory." However, even though there are well-developed and well-validated theoretical models behind media violence effects (Anderson et al., 2003), some other research domains involving media effects suffer from a lack of theoretical focus (see Chapter 23).

A dangerous error sometimes made by media effects researchers is planning studies and interpreting results as if they are completely disconnected from the field's general knowledge of psychological functioning. For example, researchers who deny the existence of media violence effects on aggression are ignoring reliable and extensively replicated findings regarding priming (Bargh, 1982; Bargh & Pietromonaco, 1982), observational learning and imitation (Bandura, Ross, & Ross, 1961, 1963a,b; Meltzoff & Moore, 1977), excitation transfer (Zillmann, 1971, 1983) and desensitization (Wolpe, 1982). Given this extensive literature concerning ways that aggressive and nonaggressive social behaviors are learned and induced, it would indeed be very surprising if media violence did not affect us.

Using Inadequate Sample Sizes

Because most effects found in media psychology are small or medium in size, adequately large samples are needed to detect media effects. If the average effect size is about $r = .20$ (Anderson & Bushman, 2001), a sample of at least 200 participants should be taken to have .80 power. If sample sizes being used are too small, this will lead to results that are unstable and seemingly inconsistent. More reliable estimates can be obtained through combining such studies using meta-analytic techniques, of course, but researchers need to use adequate sample sizes in every study.

Using Inappropriate Experimental Manipulations

Any experimental manipulation represents an attempt by the researcher to construct a valid empirical realization of the conceptual independent variable (Carlsmith et al., 1976). Ideally, the various experimental manipulations: (1) differ from each other on the conceptual independent variables that they are supposed to represent, and (2) do not differ on other aspects that might (theoretically) influence responses on the dependent variable. For example,

an experimental study to test the theoretical hypothesis that violent video game content increases the likelihood or amount of physical aggression minimally requires two conditions that differ in the amount of violent content (one should have a lot, the other should have none). Some early experiments (which shall remain nameless) did not successfully do this, in part because the researcher used an inappropriate definition of "violent content." That is some experiments used violent video games in the nonviolent control condition, because the researcher defined violent content as content that contained blood and gore, rather than the now-accepted definition of violent content as content in which player–characters try to harm other game characters. Also, recall our earlier comments on the failure of the Williams and Skoric (2005) "experiment" to appropriately manipulate exposure to violent video games.

The second requirement of the ideal case, that the relevant comparison conditions do not differ in aspects that might influence the dependent variable, also requires careful attention. We know, for example, that variables such as excitement, arousal, and frustration can sometimes increase aggressive behavior in some circumstances. Therefore, such extraneous factors (extraneous to the violent content/physical aggression hypothesis) need to be controlled.

There are two basic strategies for controlling such extraneous factors. One is to pretest several possible empirical realizations of the independent variable on the extraneous factors (using the same participant population as will be used in the main experiment), and then choose those that meet the theoretical and empirical requirements for use in the main experiment. For example, in a pilot study one could use several violent and several nonviolent video games, measure excitement, arousal, and frustration, and then select games that differ in violent content but that do not differ in induced excitement, arousal, and frustration for use in the main experiment (e.g., Anderson et al., 2004).

The second strategy is to measure the extraneous factors in the main experiment on the main participants, and then statistically control for those factors in analyses of the violent content manipulation on aggressive behavior. If it turns out that an extraneous factor (e.g., excitement) doesn't contribute significantly to aggressive behavior, then one doesn't need to control for it. However, if it does relate to aggressive behavior, then one can use the measure of excitement as a covariate in the statistical analysis. And of course, both of these strategies can be used in the same program of research as has been done in many of the methodologically strongest studies (e.g., Anderson & Dill, 2000, Study 2; Anderson et al., 2004).

It is important to keep in mind that the comparison conditions still will likely differ in other (p. 128) ways. This is especially likely in media psychology studies and any domain in which the stimuli tend to be selected rather than created. Although beyond the scope of this article, one solution to this issue is to use multiple stimuli (e.g., games) of each type (violent and nonviolent) such as in Anderson and Carnagey's (2009) study of violent and nonviolent competitive games. In addition, if one uses many examples of each game type, one could use random effects statistical models rather than the more common fixed-effect models.

Yet another approach to equating experimental conditions on extraneous factors is to use the same stimuli with changes only to the violent content. For example, Carnagey and Anderson (2005) reprogrammed the violent video game *Carmageddon* (a driving game in which one gets points for running over pedestrians) so that in the nonviolent condition there were no pedestrians to kill. Similarly, Anderson et al. (2004, Experiment 3) modified the violent video game *Marathon 2*. However, even this approach does not guarantee that the comparison games will be "equal" on relevant extraneous factors. So, it is useful to control such factors by pretests, measuring and statistically controlling for them in the main study, or both.

Finally, combining the effects of well-designed experiments in a meta-analysis also helps eliminate alternative explanations based on potential extraneous factors coinciding with the experimental variable of interest. Because different researchers have used a wide range of violent and nonviolent video games in their experiments, the likelihood that some extraneous factor existing in all or most of them is quite remote. Indeed, if one comes up with a plausible alternative explanation that might account for some of the results, one can test that alternative in a meta-analysis. For example, some gamer/scholars have proposed that the violent video game effect in experimental studies only works with the competitive reaction time task. Anderson et al. (2010) tested this alternative hypothesis, and found that the average effect size of such CRT studies is actually slightly smaller than the average effect found in the other experimental studies of violent video games, thus disproving that alternative explanation.

It is important to keep in mind that this type of reasoning, development, and assessment of experimental manipulations, and theory testing can and should be done in other media psychology domains, once sufficient numbers of studies are available. We use the media violence domain as an example because it is large, has had many excellently designed and executed studies published, has had a number of poorly executed studies published, and also because we are most familiar with this domain.

Using Poor or Inappropriate Measures

Differences in the direction of findings and in effect sizes can sometimes be a result of different measures of the independent, the dependent, or the control variables in particular studies. To detect effects and accurately assess their magnitude, reliable and valid measures need to be used. For example, the meta-analysis by Anderson et al. (2010) showed that the way one measures violent video game exposure in nonexperimental studies significantly influences the magnitude of effects found. Using specific measures of the length of exposure and violence levels in particular games (Anderson & Dill, 2000) yielded larger effect sizes than did other methods of assessing exposure to violent games.

Another potential pitfall involves using dependent measures that don't fit the research context. This can happen in multiple ways. For example, some short-term experimental studies of violent media effects have used traitlike measures as the dependent variable. Such traitlike measures essentially assess how frequently one has behaved aggressively in recent years. How can a 15-minute experimental manipulation today (violent versus nonviolent video game) influence how often one has behaved aggressively before today? Another version of this problem concerns what is an appropriate measure of aggression. Is having an argument with a friend or spouse a measure of aggression, as claimed by Williams and Skoric? Is the proximal intent of such an argument to harm the friend or spouse? In most cases, the answer is probably “no,” so this is a very poor measure of the conceptual variable “aggression.” It is even more inappropriate in a study designed to test the effects of violent video games on the kinds of aggression most frequently modeled in violent games, physical aggression. And it is even more inappropriate when most of the participants don't have a spouse with which to argue (Williams & Skoric, 2005). Certainly, there is evidence that school children arguing with teachers and other authority figures is one valid aspect of antisocial tendencies, but that is very different from using arguments with friends/spouses as a measure of video game–induced aggression in adult participants.

Often, the most important findings are acquired by using multiple measures. For example, in a longitudinal study of media violence effects, Anderson, Gentile, and Buckley (2007) obtained multiple (p. 129) measures of children's aggressive behavior through self-report, peer nominations, and teacher nominations. Sometimes such measures can be usefully combined into an overall index of aggression (Study 3).

An interesting recent direction in the media psychology field concerns examining neurocognitive bases of media effects through techniques such as event-related brain potentials (ERPs) and functional magnetic resonance imaging (fMRI). For example, Bartholow, Bushman, and Sestir (2005) showed that habitual violent video game players have reduced amplitudes of the P300 component of the event-related brain potential while viewing violent images and that this reduced response predicts more aggressive behavior. Research by Kronenberger et al. (2005) has shown similar fMRI and Stroop attention differences in conduct disordered and high violent gaming adolescents (Mathews et al., 2005; Weber et al., 2006). Similarly, Bailey, West, and Anderson (2010, 2011a,b) have used ERPs, Stroop tasks, and photo rating tasks to compare high and low action gamers on their attention control and emotional reactions to violence. Relative to low gamers, high gamers show deficits in proactive control, other more general attention deficits, and brain activation patterns suggesting desensitization to violent images. Overall, these various findings, each using multiple ways to measure theoretically related processes, provide converging support on desensitizion and decreased empathy as results of media violence exposure (Mullin & Linz, 1995; Dexter, Penrod, Linz, & Saunders 1997).

However, obtaining multiple measures sometimes comes at a cost. A potential pitfall stems from the fact that measurement of one variable of interest may influence the values of other related variables. Similar to the Heisenberg uncertainty principle in physics, the *psychological uncertainty principle* states that that measurement of one variable may change the psychological processes at work and thus change the values of downstream variables (Lindsay & Anderson, 2000). For example, measuring attitudes toward aggression after watching a violent movie may reveal the purpose of the study to participants and influence their later behavior. The possibility

of such an influence can be controlled by experimentally varying the order in which variables are assessed and then testing for order effects. If significant order effects are found, this shows that the psychological uncertainty principle is at work. To test for mediation effects in such cases, multiple experiments need to be conducted, each of which assesses one of the key variables (Lindsay & Anderson, 2000).

Significance Testing

A problematic statistical practice employed in many media violence studies consists of using null-hypothesis significance testing without reporting effect sizes and confidence intervals. This widely used approach (in psychology as well as other social sciences) has been the subject of much criticism (Rozeboom, 1960; Cohen, 1994; Kirk, 1996; Thompson, 1998; Bonett & Wright, 2007). Unfortunately, null hypothesis tests are often misinterpreted (Nickerson, 2000). Failing to reject the null hypothesis is frequently viewed as proof that the null hypothesis is true, whereas rejection of a null hypothesis is taken as evidence of a practically and theoretically relevant finding (Bonett & Wright, 2007).

In the media violence domain, in which effect sizes are in a small to medium range (Anderson & Bushman, 2001; Anderson et al., 2010), interesting findings may be overlooked because of Type II errors (failure to reject the null hypothesis when it is true) and may go unpublished. The absence of significant differences found in particular studies are sometimes misinterpreted as evidence that there indeed are no effects, without taking into account other possible reasons for the nonsignificant result (e.g., inadequate control of extraneous variables, inappropriate overcontrol of mediating outcome variables, unreliable measurement techniques, and small sample sizes). A wide confidence interval immediately indicates to the reader that the sample estimate may not be reliable and may be quite different from the true effect in the population. Meta-analytic technique can then be used to combine such studies and enable researchers to draw firmer conclusions.

The American Psychological Association (APA) Task Force on Statistical Inference advocated for an improvement of statistical practices by including effect size estimates along with confidence intervals more than 10 years ago (Wilkinson & the Task Force on Statistical Inference, 1999). However, these changes have not yet been widely implemented in psychology journals (Finch et al., 2004; Cumming & Finch, 2005; Cumming et al., 2007). As the APA *Publication Manual* now strongly encourages authors to include confidence intervals (APA, 2011), it is our hope that this change in reporting styles will reduce miscommunication and misunderstanding in the media violence literature.

Effect Size Interpretation

Media effects research has sometimes been criticized on the grounds that effect sizes found in most studies are small and are thus unimportant (p. 130) (Ferguson & Kilburn, 2010). However, it is dangerous to assume that just because most studies find small effect sizes, media do not have important practical consequences.

The effect sizes found in most media effects studies conform to the range of effect sizes usually found in social psychology studies in general (Richard, Bond, & Stokes-Zoota, 2003). Because complex behaviors are determined by a multitude of personal and situational factors, no one causal factor by itself can explain more than a small proportion of the variance in a particular behavior. Because of this, some authors suggest that effect size conventions should be revised so that $r = .1$ is small, $r = .2$ is medium, and $r = .3$ is large (Hemphill, 2003).

Some of the effects found in media psychology are, in fact, considerably larger than effect sizes found in medical research that are seen as extremely important (Bushman & Huesmann, 2001). For example, the effect of violent video games on aggression outweighs effects of substance use, abusive parents, and poverty (U.S. Department of Health and Human Services, 2001) and is larger than the effects of passive smoking on lung cancer and the effect of calcium intake on bone mass (Anderson, 2004). Furthermore, because a large proportion of the population is exposed to violent mass media, even small statistical effects can have important societal consequences (Abelson, 1985; Rosenthal, 1986; Prentice & Miller, 1992; Anderson et al., 2003).

Communicating Research Findings and Methodology to the General Public

An important role for many scientists involves disseminating knowledge gained from their research not only among

the scientific community, but also among the general public. Indeed, several American Psychological Association presidents have urged its members to "give psychology away" to the public. One of the goals of media psychology as an applied field is to benefit society with its insights, a goal that requires effective communication between media researchers and the media, public policy makers, parents, teachers, and so on. Unfortunately, the scientific community has not always been successful in communicating research findings to the general public. For example, a content analysis of research papers and newspaper articles conducted by Bushman and Anderson (2001) revealed a large disparity between news reports and the actual state of scientific knowledge concerning media violence effects.

Researchers often do not see themselves as public educators. Differences in terminology and basic assumptions between scientists and nonscientists can impede effective communication and contribute to misinterpretation of scientific findings in the general public. Additionally, public involvement comes with costs (e.g., time, effort, money, and personal costs)—a price that researchers frequently are unwilling to pay. The costs are especially large when the research suggests that certain products are harmful (e.g., lead, tobacco, violent media), and when there is a large and committed group of product users and industry leaders who are highly threatened (e.g., by threats to self-image, profits) by the research findings. There is a long history of industries in the United States spending large sums of money attacking research findings that they don't like, attacking the integrity or scientific reputations of researchers whose work discovered the harmful effects. There is such a history in the television and film violence domain. For example, both Albert Bandura and Leonard Berkowitz were excluded from key governmental review panels on media violence because of pressure brought by the entertainment media industry. Similar attacks are widespread in the video game violence domain, and with the rise of the Internet, the personal attacks on and outright fabrications about key researchers has taken on a new dimension. One need only Google the names of the leading video game violence researchers to find such fabrications about them and their research.

However, it is our belief that the benefit of effective communication between scientists and the general public outweighs such costs. Therefore, a final task of successful researchers in the field of media psychology is to be able to clearly and effectively inform general audiences concerning their findings and methods used to obtain them.

References

Abelson, R. P. (1985). A variance explanation paradox: When a little is a lot. *Psychological Bulletin*, *97*, 129–133.

Abelson, R. P. (1995). *Statistics as Principled Argument*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Alba, J., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, *93*, 203–231.

American Psychological Association (2011). *Publication Manual of the American Psychological Association*, 6th ed. Washington, DC: Author.

Anderson, C. A. (1987). Temperature and aggression: Effects on quarterly, yearly, and city rates of violent and nonviolent crime. *Journal of Personality and Social Psychology*, *52*, 1161–1173.

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

Anderson, C. A. (1997). Effects of violent movies and trait hostility on hostile feelings and aggressive thoughts. *Aggressive Behavior*, *23*, 161–178.

Anderson, C. A. (2004). Violence in the media: Its effects on children. An edited transcript of a seminar presented on September 11, 2003, in Melbourne, Australia. Published by Young Media Australia (Glenelg, South Australia) and the Victorian Parenting Centre (Melbourne, Victoria).

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., & Deuser, W. E. (1996). Examining an affective aggression framework: Weapon and temperature effects on aggressive thoughts, affect, and attitudes. *Personality and Social Psychology Bulletin*, *22*, 366–376.

Anderson, C. A., Anderson, K. B., Dorr, N., DeNeve, K. M., & Flanagan, M. (2000). Temperature and aggression. *Chapter in Advances in Experimental Social Psychology*, *32*, 63–133.

Anderson, C. A., Benjamin, A. J., & Bartholow, B. D. (1998). Does the gun pull the trigger? Automatic priming effects of weapon pictures and weapon names. *Psychological Science*, *9*, 308–314.

Anderson, C. A., Benjamin, A. J., Wood, P. K., & Bonacci, A. M. (2006). Development and testing of the Velicer Attitudes Toward Violence Scale: Evidence for a four-factor model. *Aggressive Behavior*, *32*, 122–136.

Anderson, C. A., Berkowitz, L., Donnerstein, E., Huesmann, R. L., Johnson, J., Linz, D., Malamuth, N., & Wartella, E. (2003). The influence of media violence on youth. *Psychological Science in the Public Interest*, *4*, 81–110.

Anderson, C. A., & Bushman, B. J. (1997). External validity of "trivial" experiments: The case of laboratory aggression. *Review of General Psychology*, *1*, 19–41.

Anderson, C. A., & Bushman, B. J. (2001). Effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, and prosocial behavior: A meta-analytic review of the scientific literature. *Psychological Science*, *12*, 353–359.

Anderson, C. A., & Bushman, B. J. (2002a). Media violence and the American public revisited. *American Psychologist*, *57*, 448–450.

Anderson, C. A., & Bushman, B. J. (2002b). Human aggression. *Annual Review of Psychology*, *53*, 27–51.

Anderson, C. A., & Carnagey, N. L. (2009). Causal effects of violent sports video games on aggression: Is it competitiveness or violent content? *Journal of Experimental Social Psychology*, *45*, 731–739.

Anderson, C. A., Carnagey, N. L., & Eubanks, J. (2003). Exposure to violent media: The effects of songs with violent lyrics on aggressive thoughts and feelings. *Journal of Personality and Social Psychology*, *84*, 960–971.

Anderson, C. A., Carnagey, N. L., Flanagan, M., Benjamin, A. J., Eubanks, J., & Valentine, J. C. (2004). Violent video games: Specific effects of violent content on aggressive thoughts and behavior. *Advances in Experimental Social Psychology*, *36*, 199–249.

Anderson, C. A, Deuser, W. E., DeNeve, K. (1995). Hot temperatures, hostile affect, hostile cognition, and arousal: Tests of a general model of affective aggression. *Personality and Social Psychology Bulletin*, *21*, 434–448.

Anderson, C. A., & Dill, K. E. (2000). Video games and aggressive thoughts, feelings, and behavior in the laboratory and in life. *Journal of Personality and Social Psychology*, *78*, 772–790.

Anderson, C. A., Gentile, D. A., & Buckley, K. E. (2007). *Violent Video Game Effects on Children and Adolescents: Theory, Research, and Public Policy*. New York: Oxford University Press.

Anderson, C. A., & Huesmann, L. R. (2003). Human aggression: A social-cognitive view. In Hogg, M. A., & Cooper, J. (Eds.), *Handbook of Social Psychology*. London: Sage, pp. 296–323.

Anderson, C. A., Lindsay, J. J., & Bushman, B. J. (1999). Research in the psychological laboratory: Truth or triviality? *Current Directions in Psychological Science*, *8*, 3–9.

Anderson, C. A., & Morrow, M. (1995). Competitive aggression without interaction: Effects of competitive versus cooperative instructions on aggressive behavior in video games. *Personality and Social Psychology Bulletin*, *21*, 1020–1030.

Anderson, C. A., Sakamoto, A., Gentile, D. A., Ihori, N., & Shibuya, A., Yukawa, S., Naito, M., & Kobayashi, K. (2008). Longitudinal effects of violent video games aggression in Japan and the United States. *Pediatrics*, *122*, e1067–e1072.

Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., Rothstein, H. R., & Saleem, M. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in Eastern and Western countries. *Psychological Bulletin*, *136*, 151–173.

Anderson, K. B., Anderson, C. A., Dill, K. E., & Deuser, W. E. (1998). The interactive relations between trait hostility, pain, and aggressive thoughts. *Aggressive Behavior*, *24*, 161–171.

Bailey, K., West, R., & Anderson, C. A. (2010). A negative association between video game experience and proactive cognitive control. *Psychophysiology*, *47*, 34–42.

Bailey, K., West, R., & Anderson, C. A. (2011a). The association between chronic exposure to video game violence and affective picture processing: An ERP study. *Cognitive, Affective, and Behavioral Neuroscience*, *11*, 259–276.

(p. 132) Bailey, K., West. R., & Anderson, C. A. (2011b). The influence of video games on social, cognitive, and affective information processing. In Decety, J., & Cacioppo, J. (Eds.), *Handbook of Social Neuroscience*. New York: Oxford University Press, pp. 1001–1011.

Ballard, M. E., Hamby, R. H., Panee, C. D., & Nivens, E. E. (2006). Repeated exposure to video game play results in decreased blood pressure responding. *Media Psychology*, *8*, 323–341.

Ballard, M. E., & Lineberger, R. (1999). Video game violence and confederate gender: Effects on reward and punishment given by college males. *Sex Roles*, *41*, 541–558.

Ballard, M. E., & Wiest, J. R. (1996). Mortal Kombat (tm): The effects of violent videogame play on males' hostility and cardiovascular responding. *Journal of Applied Social Psychology*, *26*, 717–730.

Banaji, M. R., & Crowder, R. G. (1989). The bankruptcy of everyday memory. *American Psychologist*, *44*(9), 1185–1193.

Bandura, A. (1973). *Aggression: A Social Learning Analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1983). Psychological mechanisms of aggression. In Geen, R. G., & Donnerstein, E. (Eds.), *Aggression: Theoretical and Empirical Reviews*. New York: Academic Press, pp. 1–40.

Bandura, A., Ross, D., & Ross, S. A. (1961). Transmission of aggressions through imitation of aggressive models. *Journal of Abnormal and Social Psychology*, *63*, 575–582.

Bandura, A., Ross, D., & Ross, S. A. (1963a). Imitation of film-mediated aggressive models. *Journal of Abnormal Social Psychology*, *66*, 3–11.

Bandura, A., Ross, D., & Ross, S. A. (1963b). Vicarious reinforcement and imitative learning. *Journal of Abnormal Social Psychology*, *67*, 601–607.

Bargh, J. A. (1982). Attention and automaticity in the processing of self-relevant information. *Journal of Personality and Social Psychology*, *43*, 425–436.

Bargh, J. A., & Pietromonaco, P. (1982). Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation. *Journal of Personality and Social Psychology*, *43*, 437–449.

Barlett, C. P., Anderson, C. A., & Swing, E. L. (2009). Video game effects confirmed, suspected and speculative: A review of the evidence. *Simulation & Gaming*, *40*, 377–403.

Barlett, C. P., Branch, O. L., Rodeheffer, C. D., & Harris, R. H. (2009). How long do the short-term violent video game effects last? *Aggressive Behavior*, *35*, 225–236.

Barlett, C. P., Rodeheffer, C. D., Baldassaro, R. M., Hinkin, M., & Harris, R. J. (2008). The effect of advances in video game technology and content on aggressive cognitions, hostility, and heart rate. *Media Psychology*, *11*, 540–565.

Baron, R. A., & Richardson, D. R. (1994). *Human Aggression*, 2nd ed. New York: Plenum Press.

Bartholow, B. D., Bushman, B. J., & Sestir, M. A. (2006). Chronic violent video game exposure and desensitization to violence: Behavioral and event-related brain potential data. *Journal of Experimental Social Psychology*, *42*(2), 283–290.

Bartholow, B. D., Sestir, M. A., & Davis, E. (2005). Correlates and consequences of exposure to video game violence: Hostile personality, empathy, and aggressive behavior. *Personality and Social Psychology Bulletin*, *31*, 1573–1586.

Baumeister, R. F., & Leary, M. R. (1997). Writing narrative literature reviews. *Review of General Psychology*, *1*(3), 311–320.

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

Berkowitz, L. (1970). Aggressive humor as a stimulus to aggressive responses. *Journal of Personality & Social Psychology*, *4*, 710–717.

Berkowitz, L. (1993). *Aggression: Its Causes, Consequences, and Control*. 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.

Bernstein, L. S., Richardson, D., & Hammock, G. (1987). Convergent and discriminant validity of the Taylor and Buss measures of physical aggression. *Aggressive Behavior*, *13*, 15–24.

Bjorkqvist, K. (1985). *Violent Films, Anxiety, and Aggression*. Helsinki: Finnish Society of Sciences and Letters.

Bonett, D. G., & Wright, T. A. (2007), Comments and recommendations regarding the hypothesis testing controversy. *Journal of Organizational Behavior*, *28*, 647–659.

Bryant, B. K. (1982). An index of empathy for children and adolescents. *Child Development*, *53*, 413–425.

Bryant, J., Carveth, R. A., & Brown, D. (1981). Television viewing and anxiety: An experimental examination. *Journal of Communication*, *31*, 106–119.

Buckley, K. E., & Anderson, C. A. (2006). A theoretical model of the effects and consequences of playing video games. In Vorderer, P., & Bryant, J. (Eds.), *Playing Video Games—Motives, Responses, and Consequences*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 363–378.

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

Bushman, B. J. (1996). Individual differences in the extent and development of aggressive cognitive-associative networks. *Personality and Social Psychology Bulletin*, *22*, 811–819.

Bushman, B. J. (1998). Priming effects of violent media on the accessibility of aggressive constructs in memory. *Personality and Social Psychology Bulletin*, *24*, 537–545.

Bushman, B. J. (2002). Does venting anguish feed or extinguish the flame? Catharsis, rumination, distraction, anger, and aggressive responding. *Personality and Social Psychology Bulletin*, *28*, 724–731.

Bushman, B. J., & Anderson, C. A. (1998). Methodology in the study of aggression: Integrating experimental and nonexperimental findings. in Geen, R., & Donnerstein, E. (Eds.), *Human Aggression: Theories, Research and Implications for Policy*. San Diego: Academic Press, pp. 23–48.

Bushman, B. J., & Anderson, C. A. (2001). Media violence and the American public: Scientific facts versus media misinformation. *American Psychologist*, *56*, 477–489.

Bushman, B. J., & Anderson, C. A. (2002). Violent video games and hostile expectations: A test of the general aggression model. *Personality and Social Psychology Bulletin*, *28*, 1679–1686.

Bushman, B. J., & Anderson, C. A. (2009). Comfortably numb: Desensitizing effects of violent media on helping

others. *Psychological Science*, *20*, 273–277.

Bushman, B. J., Baumeister, R. F., & Stack, A. D. (1999). Catharsis, aggression, and persuasive influence: Self-fulfilling or self-defeating prophecies? *Journal of Personality and Social Psychology*, *76*, 367–376.

Bushman, B. J., & Huesmann, L. R. (2001). Effects of televised violence on aggression. In D. Singer & J. Singer (Eds.), (p. 133) *Handbook of children and the media* (pp. 223–254). Thousand Oaks, CA: Sage Publications.

Bushman, B. J., & Huesmann, L. R. (2006). Short-term and long-term effects of violent media on aggression in children and adults. *Archives of Pediatrics & Adolescent Medicine*, *160*, 348–352.

Bushman, B. J., Rothstein, H. R., & Anderson, C. A. (2010). Much ado about something: Violent video game effects and a school of red herring: Reply to Ferguson and Kilburn. *Psychological Bulletin*, *136*, 182–187.

Buss, A. H. (1961). *The Psychology of Aggression*. New York: Wiley.

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

Calvert, S. L., & Tan, S. (1994). Impact of virtual reality on young adults' physiological arousal and aggressive thoughts: Interaction versus observation. *Journal of Applied Developmental Psychology*, *15*, 125–139.

Campbell, A. (1993). *Men, Women and Aggression*. New York: Basic Books.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105.

Caprara, G. V., Cinanni, V., D' Imperio, G., Passerini, S., Renzi, P., & Travablia, G. (1985). Indicators of impulsive aggression: Present status of research on irritability and emotional susceptibility scales. *Personality & Individual Differences*, *6*, 665–674.

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.

Carnagey, N. L., & Anderson, C. A. (2003). Theory in the study of media violence: The General Aggression Model. In Gentile, D. (Ed.), *Media Violence and Children*. Westport, CT: Praeger, pp. 87–106.

Carnagey, N. L., & Anderson, C. A. (2005). The effects of reward and punishment in violent video games on aggressive affect, cognition, and behavior. *Psychological Science*, *16*, 882–889.

Carnagey, N. L., & Anderson, C. A., & Bushman, B. J. (2007). The effect of video game violence on physiological desensitization to real-life violence. *Journal of Experimental Social Psychology*, *43*, 489–496.

Chambers, J. H., & Ascione, F. R. (1987). The effects of prosocial and aggressive video games on children's donating and helping. *Journal of Genetic Psychology*, *148*, 499–505.

Cheung, P. C., Ma, H. K., & Shek, T. L. D. (1998). Conceptions of success: their correlates with prosocial orientation and behavior in Chinese adolescents. *Journal of Adolescence*, *21*, 31–42.

Cohen, J. (1994). The earth is round (p ⟨ .05). *American Psychologist*, *49*(12), 997–1003.

Comstock, G., & Scharrer, E. (2007). *Media and the American Child*. San Diego: Academic Press.

Cook, W. W., & Medley, D. M. (1954). Proposed hostility and pharisaic virtue scales for the MMPI. *Journal of Applied Psychology*, *38*, 414–418.

Cooper, J., & Mackie, D. (1986). Video games and aggression in children. *Journal of Applied Social Psychology*, *16*, 726–744.

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information processing mechanisms in children's adjustment. *Psychological Bulletin*, *115*, 74–101.

Cumming, G., Fidler, F., Leonard, M., Kalinowski, P., Christiansen, A., Kleinig, A., et al. (2007). Statistical reform in psychology: Is anything changing? *Psychological Science*, *18*, 230–232.

Cumming, G., & Finch, S. (2005). Inference by eye: Confidence intervals and how to read pictures of data. *American Psychologist*, *60*(2), 170–180.

Davis, M. H. (1980). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*, 113–126.

DeWall, C. N., & Anderson, C. A. (2011). The General Aggression Model. In Shaver, P., & Mikulincer, M. (Eds.), *Human Aggression and Violence: Causes, Manifestations, and Consequences*. Washington, DC: American Psychological Association, pp. 15–33.

DeWall, C. N., Anderson, C. A., & Bushman, B. J. (2012). Aggression. Suls, J. (Ed.), *Handbook of Psychology*. New York: Wiley.

DeWall, C. N., Anderson, C. A., & Bushman, B. J. (2011). The general aggression model: Theoretical extensions to violence. *Psychology of Violence*, *1*(13), 245–258.

Dexter, H. R., Penrod, S. D., Linz, D., & Saunders, D. (1997). Attributing responsibility to female victims after exposure to sexually violent films. *Journal of Applied Social Psychology*, *27*, 2149.

Dill, J., & Anderson, C. A. (1995). Effects of justified and unjustified frustration on aggression. *Aggressive Behavior*, *21*, 359–369.

Dill, K. E., & Dill, J. (1998). Video game violence: A review of the empirical literature. *Aggression and Violent Behavior*, *3*, 407–428.

Dollard, J., Doob, L., Miller, N., Mowrer, O., & Sears, R. (1939). *Frustration and Aggression*. New Haven, CT: Yale University Press.

Dominick, J. R. (1984). Videogames, television violence, and aggression in teenagers. *Journal of Communication*, *34*, 136–147.

Donnerstein, E., & Berkowitz, L. (1981). Victim reactions in aggressive erotic films as a factor in violence against women. *Journal of Personality and Social Psychology*, *41*, 710–724.

Elliot, D. S., Huizinga, D., & Ageton, S. S. (1985). *Explaining delinquency and drug use*. Beverly Hills, CA: Sage.

Eron, L. D., Huesmann, L. R., Lefkowitz, M. M., & Walder, L. O. (1972). Does television violence cause aggression? *American Psychologist*, *27*, 253–263.

Ferguson, C. J., & Kilburn, J. (2010). Much ado about nothing: The misestimation and overinterpretation of violent video game effects in Eastern and Western nations: Comment on Anderson et al. (2010). *Psychological Bulletin*, *136*.

Ferguson, C. J., Olson, C. K., Kutner, L. A., & Warner, D. E. (2010). Violent video games, catharsis-seeking, bullying and delinquency: A multivariate analysis of effects. *Crime and Delinquency*.

Ferguson, C. J., San Miguel, C., & Hartley, R. D. (2009). A multivariate analysis of youth violence and aggression: The influence of family, peers, depression and media violence. *Journal of Pediatrics*, *155*(6), 904–908.

Finch, S., Cumming, G., Williams, J., Palmer, L., Griffith, E., Alders, C., et al. (2004). Reform of statistical inference in **(p. 134)** psychology: The case of Memory & Cognition. *Behavior Research Methods, Instruments, & Computers*, *36*, 312–324.

Fischer, P., Greitemeyer, T., Kastenmüller, A., Vogrincic, C., & Sauer, A. (2011). The effects of risk-glorifying media

exposure on risk-positive cognitions, emotions, and behaviors: A meta-analytic review. *Psychological Bulletin*. Advance online publication.

Fleming, M. J., & Rickwood, D. J. (2001). Effects of violent versus nonviolent video games on children's arousal, aggressive mood and positive mood. *Journal of Applied Social Psychology*, *31*, 2047–2071.

Fling, S., Smith, L., Rodriguez, T., Thornton, D., Atkins, E., & Nixon, K. (1992). Videogames, aggression, and self-esteem: A survey. *Social Behavior and Personality*, *20*, 39–46.

Funk, J. B., Baldacci, H. B., Pasold, T., & Baumgardner, J. (2004). Violence exposure in real-life, video games, television, movies, and the Internet: is there desensitization? *Journal of Adolescence*, *27*(1), 23–39.

Funk, J. B., Buchman D. D., Jenks, J., & Bechtoldt, H. (2003) Playing violent video games, desensitization, and moral evaluation in children. *Applied Developmental Psychology*, *24*, 413–436.

Funk, J. B., Elliott, R., Urman, M. L., Flores, G. T. & Mock, R. M. (1999). The Attitudes Towards Violence scale: A measure for adolescents. *Journal of Interpersonal Violence*, *14*, 1123–1136.

Funk, J. B., Fox, C. M., Chan, M., & Curtiss, K. (2008). The development of the Children's Empathic Attitudes Scale using classical and Rasch analyses. *Journal of Applied Developmental Psychology*, *29*, 187–198.

Gaebelein, J. W., & Taylor, S. P. (1971). The effects of competition and attack on physical aggression. *Psychonomic Science*, *24*, 65–67.

Geen, R. G. (2001). *Human Aggression*, 2nd ed. Philadelphia: Open University Press.

Geen, R. G., & O'Neal, E. C. (1969). Activation of cue-elicited aggression by general arousal. *Journal of Personality and Social Psychology*, *11*, 289–292.

Geen, R. G., & Quanty, M. B. (1977). The catharsis of aggression: An evaluation of a hypothesis. In Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*. New York: Academic Press, pp. 1–37.

Geen, R. G., Stonner, D., & Shope, G. L. (1975). The facilitation of aggression by aggression: Evidence against the catharsis hypothesis. *Journal of Personality and Social Psychology*, *31*, 721–726.

Gentile, D. A., Anderson, C. A., Yukawa, N., Saleem, M., Lim, K. M., Shibuya, A., Liau, A. K., Khoo, A., Bushman, B. J., Huesmann, L. R., & Sakamoto, A. (2009). The effects of prosocial video games on prosocial behaviors: International evidence from correlational, longitudinal, and experimental studies. *Personality and Social Psychology Bulletin*, *35*, 752–763.

Gentile, D. A., Lynch, P. J., Linder, J. R., & Walsh, D. A. (2004). The effects of violent video game habits on adolescent aggressive attitudes and behaviors. *Journal of Adolescence*, *27*, 5–22.

Gentile, D. A., & Sesma, A. (2003). Developmental approaches to understanding media effects on individuals. In D. A. Gentile (Ed.), *Media violence and children*. Westport, CT: Greenwood Publishing.

Gerbner, G., Gross, L., Jackson-Beeck, M., Jeffries-Fox, S., & Signorielli, N. (1978). Cultural indicators: Violence profile no. 9. *Journal of Communication*, *28*, 176–207.

Gerbner, G., Gross, L., Morgan, M., & Signorielli, N. (1980). The "mainstreaming" of America: Violence profile no. 11. *Journal of Communication*, *30*, 10–29.

Giancola, P. R., & Chermack, S. T. (1998). Construct validity of laboratory aggression paradigms: A response to Tedeschi and Quigley (1996). *Aggression and Violent Behavior*, *3*, 237–253.

Giancola, P. R., & Parrott, D. J. (2008). Further evidence for the validity of the Taylor aggression paradigm. *Aggressive Behavior*, *34*, 214–229.

Glantz, M. D., & Johnson, J. L. (Eds.) (1999). *Resilience and Development: Positive Life Adaptations*. New York: Kluwer.

Greitemeyer, T. (2009). Effects of songs with prosocial lyrics on prosocial behavior: further evidence and a mediating mechanism. *Personality and Social Psychology Bulletin*, *35*, 1500–1511.

Greitemeyer, T. (2011). Exposure to music with prosocial lyrics reduces aggression: First evidence and test of the underlying mechanism. *Journal of Experimental Social Psychology*, *47*, 28–36.

Greitemeyer, T., & Osswald, S. (2010): Effects of Prosocial video games on prosocial behavior. *Journal of Personality and Social Psychology*, *98*(2), 211–221.

Griffiths, M. (1999). Violent video games and aggression: A review of the literature. *Aggression and Violent Behavior*, *4*, 203–212.

Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist*, *58*, 78–79.

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

Hopf, W. H., Huber, G. L., & Weiß, R. H. (2008). Media violence and youth violence: A 2-year longitudinal study. *Journal of Media Psychology: Theories, Methods, and Applications*, *20*(3), 79–96.

Huesmann, L. R. (1986). Psychological processes promoting the relation between exposure to media violence and aggressive behavior by the viewer. *Journal of Social Issues*, *42*, 125–139.

Huesmann, L. R. (1988). An information processing model for the development of aggression. *Aggressive Behavior*, *14*, 13–24.

Huesmann, L. R. (1998). The role of social information processing and cognitive schema in the acquisition and maintenance of habitual aggressive behavior. In Geen, R., & Donnerstein, E. (Eds.) *Human Aggression: Theories, Research, and Implications for Policy* . New York: Academic Press, pp. 73–109.

Huesmann, L. R., Eron, L. D., Klein, R., Brice, P., & Fischer, P. (1983). Mitigating the imitation of aggressive behaviors by changing children's attitudes about media violence. *Journal of Personality and Social Psychology*, *44*, 899–910.

Huesmann, L. R., Guerra, N. G., Miller, L., & Zelli, A. (1992). The role of social norms in the development of aggression. In Zumkley, H., & Fraczek, A. (Eds.), *Socialization and Aggression* . New York: Springer, pp. 139–151.

Huesmann, L. R., Moise-Titus, J., Podolski, C. L., & Eron, L. (2003). Longitudinal relations between children's exposure to TV violence and their aggressive and violent behavior in young adulthood: 1977–1992. *Developmental Psychology*, *39*, 201–221.

Huesmann, L. R., & Taylor, L. D. (2003). The case against the case against media violence. In Gentile, D. (Ed.) *Media Violence and Children*. Westport, CT: Greenwood Press, pp. 107–130.

(p. 135) Irwin, A. R., & Gross, A. M. (1995). Cognitive tempo, violent video games, and aggressive behavior in young boys. *Journal of Family Violence*, *10*, 337–350.

Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence*, *29*, 589–611.

Josephson, W. L. (1987). Television violence and children's aggression: Testing the priming, social script, and disinhibition predictions. *Journal of Personality and Social Psychology*, *53*, 882–890.

Kenny, D. A. (1984). The NBC study and television violence: A review. *Journal of Communication*, *34*, 176–188.

Kirk, R. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement*, *56*, 746–759.

Kirsh, S. J. (1998). Seeing the world through Mortal Kombat-colored glasses: Violent video games and the development of a short-term hostile attribution bias. *Childhood*, *5*, 177–184.

Koehler, J. J. (1993). The influence of prior beliefs on scientific judgments of evidence quality. *Organizational Behavior and Human Decision Processes*, *56*, 28–55.

Konijn, E. A., Bijvank, N. M., & Bushman, B. J. (2007). I wish I were a warrior: The role of wishful identification in effects of violent video games on aggression in adolescent boys. *Developmental Psychology*, *43*, 1038–1044.

Kronenberger, W. G., Matthews, V. P., Dunn, D. W., Wang, Y., Wood, E. A., Giauque, A. L., et al. (2005). Media violence exposure and executive functioning in aggressive and control adolescents. *Journal of Clinical Psychology*, *61*, 725–737.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*, 480–498.

Lewin, K. (1951). Problems of research in social psychology. In Cartwright, D. (Ed.), *Field Theory in Social Science* (pp. 155–169). New York: Harper & Row.

Lieberman, J. D., Solomon, S., Greenberg, J., & McGregor, H. A. (1999), A hot new way to measure aggression: Hot sauce allocation. *Aggressive Behavior*, *25*, 331–348.

Lin, S., & Lepper, M. R. (1987). Correlates of children's usage of video games and computers. *Journal of Applied Social Psychology*, *17*, 72–93.

Lindsay, J. J., & Anderson, C. A. (2000). From antecedent conditions to violent actions: A general affective aggression model. *Personality and Social Psychology Bulletin*, *26*, 533–547.

Linz, D., Donnerstein, E., & Adams, S. M. (1989). Physiological desensitization and judgments about female victims of violence. *Human Communication Research*, *15*, 509–522.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*, 2098–2109.

Mallick, S. K., & McCandless, B. R. (1966). A study of catharsis of aggression. *Journal of Personality and Social Psychology*, *4*, 591–596.

Mathews, V. P., Kronenberger, W. G., Wang, Y., Lurito, J. T., Lowe, M. J., & Dunn, D. W. (2005). Media violence exposure and frontal lobe activation measured by functional magnetic resonance imaging in aggressive and nonaggressive adolescents. *Journal of Computer Assisted Tomography*, *29*, 287–292.

McLeod, J. M., Atkin, C. K., & Chaffee, S. H. (1972). Adolescents, parents, and television use: Adolescent self-report measures from Maryland and Wisconsin samples. In Comstock, G. A., & Rubinstein, E. A. (Eds.), *Television and Social Behavior: A Technical Report to the Surgeon General's Scientific Advisory Committee on Television and Social Behavior: Vol. 3. Television and Adolescent Aggressiveness* (DHEW Publication No. HSM 72–9058, pp. 173–238). Washington, DC: U.S. Government Printing Office.

Meltzoff, A. N., & Moore, K. M. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *109*, 77–78.

Milgram, S. (1974). *Obedience to Authority*. New York: Harper & Row.

Milavsky, J. R., Kessler, R. C., Stipp, H. H., & Rubens, W. S. (1982). *Television and aggression: A panel study*. New York: Academic Press.

Möller, I., & Krahé, B. (2009). Exposure to violent video games and aggression in German adolescents: A longitudinal analysis. *Aggressive Behavior*, *35*, 75–89.

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

Mullin, C. R., & Linz, D. (1995). Desensitization and resensitization to violence against women: Effects of exposure to sexually violent films on judgments of domestic violence victims. *Journal of Personality and Social Psychology*,

69, 449–459.

Münsterberg, H. (1899). Psychology and history. *Psychological Review*, *6*, 1–31.

Nickerson, R. S. (2000). Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods*, *5*(2), 241–301.

Ohbuchi, K.-I., Kameda, M., & Agarie, N. (1989). Apology as aggression control: Its role in mediating appraisal of and response to harm. *Journal of Personality and Social Psychology*, *56*(2), 219–227.

Prentice, D. A., & Miller, D. T. (1992). When small effects are impressive. *Psychological Bulletin*, *112*, 160–164.

Richard, F. D., Bond, C. F., Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, *7*, 331–363.

Rideout, V. J., Foehr, U. G., & Roberts, D. F. (2010). *Generation M2—Media in the Lives of 8- to 18-Year Olds*. Menlo Park, CA: Kaiser Family Foundation.

Robinson, T. N., Wilde, M. L., Navracruz, L. C., Haydel, K. F., & Varady, A. (2001). Effects of reducing children's television and video game use on aggressive behavior: A randomized controlled trial. *Archives of Pediatric Adolescent Medicine*, *155*, 17–23.

Roediger, H. L., Weldon, M. S., Stadler, M. L., & Riegler, G. L. (1992). Direct comparison of two implicit memory tests: Word fragment and word stem completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1251–1269.

Rosenthal, R. (1986). Media violence, antisocial behavior, and the social consequences of small effects. *Journal of Social Issues*, *42*, 141–154.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin*, *118*(2), 183–192.

Rozeboom, W. W. (1960). The fallacy of the null hypothesis significance test. *Psychological Bulletin*, *57*, 416–428.

Ryan, R., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion*, *30*, 344–360.

(p. 136) Saleem, M., & Anderson, C. A. (in press). Arabs as terrorists: Effects of stereotypes within violent contexts on attitudes, perceptions and affect. *Psychology of Violence*.

Schmidt, D. F., & Sherman, R. C. (1984). Memory for persuasive messages: A Test of a Schema-Copy-Plus-Tag Model. *Journal of Personality and Social Psychology*, *47*, 17–25.

Schutte, N. S., Malouff, J. M., Post-Gorden, J. C., & Rodasta, A. L. (1988). Effects of playing video games on children's aggressive and other behaviors. *Journal of Applied Social Psychology*, *18*, 454–460.

Sharif, I., & Sargent, J. D. (2006). Association between television, movie, and video game exposure and school performance. *Pediatrics*, *118*(4), 1061–1070.

Sheese, B. E., & Graziano, W. G. (2005). Deciding to defect. The effects of video-game violence on cooperative behavior. *Psychological Science*, *16*, 354–357.

Shibuya, A., Sakamoto, A., Ihori, N., & Yukawa, S. (2004). Media bouryoku heno sesshoku, sesshoku kankyo ga kougekisei ni oyobosu choukiteki eikyou: Shougakusei heno paneru chosa [The long-term effects of media violence and its situational variables on aggression: A panel study to elementary school children]. *Proceeding of the 45th convention of the Japanese Society of Social Psychology*, pp. 248–249.

Silvern, S. B., & Williamson, P. A. (1987). The effects of video game play on young children's aggression, fantasy and prosocial behavior. *Journal of Applied Developmental Psychology*, *8*, 453–462.

Spielberger, C. D. (1988). *State-Trait Anger Expression Inventory*. Odessa, FL: Psychological Assessment Resources.

Swing, E. L., & Anderson, C. A. (2010). Media violence and the development of aggressive behavior. In DeLisi, M., & Beaver, K. M. (Eds.), *Criminological Theory: A Life-Course Approach*. Sudbury, MA: Jones & Bartlett, pp. 87–108.

Swing, E. L., Gentile, D. A., & Anderson, C. A. (2008). Violent video games: Learning processes and outcomes. In R. E. Ferdig (Ed.), *Handbook of research on effective electronic gaming in education* (pp. 876–892). Hershey, PA: Information Science Reference.

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

Thomas, M. H., Horton, R. W., Lippincott, E. C., & Drabman, R. S. (1977). Desensitization to portrayals of real life aggression as a function of television violence. *Journal of Personality and Social Psychology*, *35*, 450–458.

Thompson, B. (1998). In praise of brilliance: Where that praise really belongs. *American Psychologist*, *53*, 799–800.

Uhlmann, E., & Swanson, J. (2004). Exposure to violent video games increases automatic aggressiveness. *Journal of Adolescence*, *27*, 41–52.

U.S. Department of Health and Human Services (2001). *Youth violence: A report of the Surgeon General* Rockville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. National Center for Injury Prevention and Control; Substance Abuse and Mental Health Services Administration, Center for Mental Health Services; and National institutes of Health, National Institute of Mental Health. Retrieved April 2, 2010 from **http://www.ncbi.nlm.nih.gov/books/NBK44293/**.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063–1070.

Weber, R., Ritterfeld, U., & Mathiak, K. (2006). Does playing violent video games induce aggression? Empirical evidence of a functional magnetic resonance imaging study. *Media Psychology*, *8*, 39–60.

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

Wilkinson, L., & the Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54*, 594–604.

Williams, D., & Skoric, M. (2005). Internet fantasy violence: A test of aggression in an online game. *Communication Monographs*, *72*, 217–233.

Wolpe, J. (1982). *The Practice of Behavior Therapy*, 3rd ed. New York: Pergamon Press.

Zillmann, D. (1971). Excitation transfer in communication-mediated aggressive behavior. *Journal of Experimental Social Psychology*, *7*, 419–434.

Zillmann, D. (1983). Arousal and aggression. In Geen, R., & Donnerstein, E. (Eds.), *Aggression: Theoretical and Empirical Reviews*, vol.1. New York: Academic Press, pp. 75–102.

Zuckerman, M. (1960). The development of an affect adjective checklist for the measurement of anxiety. *Journal of Consulting Psychology*, *24*, 457–462.

Zuckerman, M., Lubin, B., Vogel, L., & Valerius, E. (1964). Measurement of experimentally induced affects. *Journal of Consulting Psychology*, *28*, 418–425.

Notes:

(1.) We find it ironic that the lead author of that study, Dmitri Williams, in 2005 criticized the experimental study reported in Anderson and Dill (2000) for selecting a violent and a nonviolent game based on pilot testing of several games that included self-reported ratings on a variety of dimensions and physiological measures of arousal.

Williams apparently didn't like the two games chosen because they didn't fit his intuitions about excitement levels induced by the games. What he fails to note is that: (1) Anderson and Dill reported that there were differences in self-reported excitement; (2) there were not differences in heart rate or blood pressure; (3) excitement was statistically controlled in the main experiment; (4) the excitement did not influence the results of the main experiment. Furthermore, in science when intuition conflicts with empirical data, it is intuition that has to yield. In fact, the Anderson and Dill studies set the methodological standard for later video game studies (both experimental and correlational), and their basic findings have been replicated numerous times by numerous research teams from many countries around the world. We are not saying that this early experimental study was perfect; no single study is perfect. In fact, several more recent studies from our and other labs are, in our view, stronger methodologically; they built on the insights and knowledge gained from the earlier study.

Sara Prot

Sara Prot, Center for the Study of Violence, Department of Psychology, Iowa State University, Ames, IA

Craig A. Anderson

Craig A. Anderson, Center for the Study of Violence , Department of Psychology, Iowa State University, Ames, IA