Psychological Bulletin
2010, Vol. 136, No. 2, 151–173

0033-2909/10/$12.00 DOI: 10.1037/a0018251

Violent Video Game Effects on Aggression, Empathy, and Prosocial Behavior in Eastern and Western Countries: A Meta-Analytic Review

Craig A. Anderson
Iowa State University

Akiko Shibuya
Keio University

Nobuko Ihori
Ochanomizu University

Edward L. Swing
Iowa State University

Brad J. Bushman
VU University and University of Michigan

Akira Sakamoto
Ochanomizu University

Hannah R. Rothstein
City University of New York

Muniba Saleem
Iowa State University

Meta-analytic procedures were used to test the effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, empathy/desensitization, and prosocial behavior. Unique features of this meta-analytic review include (a) more restrictive methodological quality inclusion criteria than in past meta-analyses; (b) cross-cultural comparisons; (c) longitudinal studies for all outcomes except physiological arousal; (d) conservative statistical controls; (e) multiple moderator analyses; and (f) sensitivity analyses. Social–cognitive models and cultural differences between Japan and Western countries were used to generate theory-based predictions. Meta-analyses yielded significant effects for all 6 outcome variables. The pattern of results for different outcomes and research designs (experimental, cross-sectional, longitudinal) fit theoretical predictions well. The evidence strongly suggests that exposure to violent video games is a causal risk factor for increased aggressive behavior, aggressive cognition, and aggressive affect and for decreased empathy and prosocial behavior. Moderator analyses revealed significant research design effects, weak evidence of cultural differences in susceptibility and type of measurement effects, and no evidence of sex differences in susceptibility. Results of various sensitivity analyses revealed these effects to be robust, with little evidence of selection (publication) bias.

Keywords: media violence, aggression, video games, empathy and desensitization, prosocial behavior

Supplemental materials: http://dx.doi.org/10.1037/a0018251.supp

> You know what's really exciting about video games is you don't just interact with the game physically—you're not just moving your hand on a joystick, but you're asked to interact with the game psychologically and emotionally as well. You're not just watching the characters on screen; you're becoming those characters.
>
> —Nina Huntemann, *Game Over*

People of all ages in most modern countries get a heavy dose of violent media, especially in TV programs, films, and video games (e.g., Comstock & Scharrer, 2007; Gentile, 2003; Gentile, Saleem, & Anderson, 2007; Kirsh, 2006; Singer & Singer, 2001). Potential harmful effects of media violence have been scrutinized for over six decades, and considerable consensus has been reached on several of the most important issues. As stated by a recent panel of experts assembled by the U.S. Surgeon General, "Research on violent television and films, video games, and music reveals unequivocal evidence that media violence increases the likelihood of aggressive and violent behavior in both immediate and long-term contexts" (Anderson et al., 2003, p. 81). Numerous reports by professional health associations (e.g., American Academy of Pediatrics, American Psychological Association, Australian College of Paediatrics, Canadian Paediatric Society) and government health agencies (e.g., U.S. Office of the Surgeon General, U.S. Department of Health and Human Services) have reached the same conclusion after reviewing the available scientific evidence (Gentile et al., 2007; Ontario Office for Victims of Crime, 2004).

Craig A. Anderson, Edward L. Swing, and Muniba Saleem, Center for the Study of Violence, Department of Psychology, Iowa State University; Akiko Shibuya, Institute for Media and Communications Research, Keio University, Tokyo, Japan; Nobuko Ihori and Akira Sakamoto, Department of Psychology, Ochanomizu University, Tokyo, Japan; Brad J. Bushman, Vrije Universiteit, Amsterdam, the Netherlands, and Institute for Social Research, University of Michigan; Hannah R. Rothstein, Department of Management, Baruch College, City University of New York.

Correspondence concerning this article should be addressed to Craig A. Anderson, Department of Psychology, W112 Lagomarcino Hall, Iowa State University, Ames, IA 50011. E-mail: caa@iastate.edu

The majority of media violence studies have focused on violent television and film effects, and most have been conducted in Western countries, especially the United States. There are theoretical reasons to expect that type of media (e.g., newspapers, literature, comic books, graphic novels, television, film, video games, music) and culture will moderate violent media effects. For example, watching the *Lord of the Rings* films should increase aggressive tendencies more than reading the books because of the higher concentration and glorification of violence in the films.

Similarly, cultural factors may either exacerbate or reduce violent media effects for both statistical and psychological reasons. For example, the context of violence on Japanese television is very different from that on U.S. television, even though the total amount of violence shown is similar (Kodaira, 1998). Japanese TV tends to portray violent actions and their consequences much more vividly, with a particular emphasis on the suffering of the victims. This might explain why the effects of TV violence on aggression sometimes appear smaller in Japan than in the United States.

Other multinational research has found considerable variation in access to and content of “violent television” and a few differences in observed effects (Huesmann & Eron, 1986; Huesmann, Lagerspetz, & Eron, 1984). For example, within Israel there were significant correlations between TV violence viewing and children’s aggression for urban children but not for rural children being raised on a kibbutz, where socialization is conducted in a communal manner (Bachrach, 1986). What is currently unclear is the extent to which the occasional cross-cultural differences in media violence effects result from cross-cultural differences in the content of their violent media (a type of artifact), true differences in media violence effects (perhaps communal, collectivist, or politically unstable countries are less susceptible), or a combination of the two.

Video Game Violence

Past Findings

Video game violence is the new kid on the media violence block, having emerged in the late 1980s and early 1990s. Currently, one can play video games on computers, consoles (e.g., Xbox 360, PlayStation, Wii), handhelds (e.g., Nintendo DS), computers, iPods, personal digital assistants, and mobile telephones. Because video game technology is relatively new, there are fewer empirical studies on video game violence than on TV and film violence. Nonetheless, several meta-analytic reviews have reported significant harmful effects of exposure to violent video games, both in short-term experimental studies and in cross-sectional correlational studies (Anderson, 2004; Anderson & Bushman, 2001; Anderson et al., 2004; Sherry, 2001). Briefly, these reviews found that across these two research designs, exposure to violent video games is associated with higher levels of aggressive behavior, aggressive cognition, aggressive affect, and physiological arousal and with lower levels of prosocial behavior. The earliest meta-analyses reported average effects on aggressive behavior of $r+ = .15$ ($K = 25$, $N = 2{,}722$; Sherry, 2001) and $r+ = .19$ ($K = 33$, $N = 3{,}033$; Anderson & Bushman, 2001). Anderson (2004) found an average effect size of $r+ = .20$ ($K = 32$, $N = 5{,}240$) when all relevant studies were included and a larger effect when more stringent methodological criteria were applied, $r+ = .26$ ($K = 17$, $N = 2{,}731$).[1] In general, the violent video game research mirrors findings from the violent TV and film research, with some evidence that the violent video game effects may be somewhat larger (Anderson, Gentile, & Buckley, 2007; Polman, Orobio de Castro, & Van Aken, 2008).

Recent Skepticism

However, three recent meta-analyses by the same author, each using a very small set of available studies, have suggested that the effects of violent video games on aggression have been substantially overestimated because of publication bias (Ferguson, 2007a, 2007b; Ferguson & Kilburn, 2009) and that therefore there is little-to-no evidence of a violent video game effect on aggression. However, these three meta-analyses have numerous problems that call into question their results and conclusions. For example, counter to widely accepted procedures for reducing the impact of publication bias, only published articles were included in the analyses and then procedures for addressing publication bias were misinterpreted. Also, studies published prior to 1995 were ignored and a large number of studies published since that time apparently were missed.

The text on publication bias cited by Ferguson (2007a; Rothstein, Sutton, & Borenstein, 2005) specifically recommends that the primary way to assure that meta-analytic results will not be affected by publication bias is to conduct a search for relevant studies that is thorough, systematic, unbiased, transparent, and clearly documented. Authors are told to include book chapters, dissertations, conference papers, and unpublished manuscripts that meet the inclusion criteria for the meta-analysis, because this is widely viewed as the best way to ameliorate publication bias.

Ferguson (2007a, 2007b; Ferguson & Kilburn, 2009) used the trim and fill method to estimate the “true” effect size corrected for publication bias. The originators of the trim and fill method (Duval, 2005; Duval & Tweedie, 2000a) have cautioned that the “adjusted” estimate of an effect using imputed studies provided by trim and fill should not be taken as the “true” effect, because it is based on imputed data points (that do not really exist). Trim and fill provides a useful sensitivity analysis that assesses the potential impact of missing studies on the results of a meta-analysis by examining the degree of divergence between the original effect-size estimate and the trim and fill adjusted effect-size estimate.

It has also been widely cautioned that because trim and fill and some other techniques for assessing publication bias are based on an association between effect size and sample size, other explanations of this association should be considered. For example, effect sizes in experimental studies may be larger than those in cross-sectional or longitudinal studies due to the reduced error variance that results from tight experimental controls; researchers may know this and therefore may intentionally plan to use larger sample sizes when conducting nonexperimental studies. Similarly, in some research contexts with very large sample sizes (e.g., national surveys) a researcher may have to use less precise measures (e.g., fewer items) that result in smaller effect sizes. In sum,

[1] The Anderson et al. (2004) analysis differed only slightly from Anderson (2004) and yielded an almost identical effect for the methodologically better studies, $r+ = .27$ ($K = 18$, $N = 3{,}537$).

it is possible that the effects in the studies with small samples really are larger than those in the studies with large samples (cf. Sterne and Egger, 2005).

In addition, the meta-analyses published by Ferguson are not independent of each other because they use highly overlapping subsets of the same small sample of studies, which includes at least one study that does not even have a valid measure of aggressive behavior (i.e., Williams & Skoric, 2005). For example, the Ferguson (2007b) meta-analysis used data from 17 articles, 14 of which were used in Ferguson (2007a), making the two meta-analyses largely redundant.[2] The average effect-size estimates computed by Ferguson (rs = .29, .14, and .15 for Ferguson 2007a, 2007b, and Ferguson & Kilburn, 2009, respectively) before "correcting" for publication bias are very similar in magnitude to those computed by other researchers.

Need for a New Meta-Analysis

Thus, there is some inconsistency between the recent meta-analyses conducted by Ferguson (2007a, 2007b; Ferguson & Kilburn, 2009) and most of the published research and earlier, more comprehensive meta-analyses on media violence effects. Clearly, all agree that prior meta-analyses have not answered all relevant questions about violent video game effects. Furthermore, there has been an explosion of research on violent video game effects since the last comprehensive meta-analysis was published in 2004. For example, none of the prior meta-analytic reviews of violent video game effects included longitudinal studies because none existed, but now several such studies are available. Past meta-analyses also frequently included cross-sectional studies in which sex differences were not statistically controlled. Although there are both theoretical and methodological reasons for not using partial correlations, it certainly is of interest to know whether the average effect size is reliably different from zero when sex has been controlled. A sufficient number of studies now exists to allow meaningful tests of this question.

Other important questions could not be tested in prior meta-analyses because of the small number of available studies. For example, does player perspective (first person vs. third person) influence the magnitude of violent video game effects? Does killing human targets yield larger effects than killing nonhuman targets? Are younger game players more affected than older ones?

Furthermore, almost all of the studies reviewed in prior meta-analyses came from U.S. samples or from similar Western individualistic-culture samples (e.g., Australia, Germany, the Netherlands, United Kingdom). Thus, the possibility that video game violence effects might be smaller (or larger) in collectivistic societies than in individualistic societies has never been explored. Indeed, it was the combination of availability of Japanese studies (following a visit to Japan in 2003 by Craig A. Anderson), the explosion of research in this domain, and the publication of several longitudinal studies that inspired us to begin the present meta-analysis.

Cultural Differences in Aggression

Aggression rates differ greatly across countries and cultures; cross-national comparisons have implicated various cultural variables as possible contributors to these differences. For example, an analysis of peer-directed aggression in 28 countries found that "in general, cultures characterized by collectivistic values, high moral discipline, a high level of egalitarian commitment, low uncertainty avoidance, and which emphasize values that are heavily Confucian showed lower levels of aggression than their counterparts" (Bergeron & Schneider, 2005, p. 116).[3]

However, the rank order of countries by aggression rates varies from one measure of aggression to another. The United States has a higher homicide rate than do many industrialized countries in Europe and Asia but similar or lower rates of other forms of violent crime, such as assault. For example, the average annual homicide rate per hundred thousand for 1998–2000 was almost 400% larger in the United States than in England and Wales (5.87 vs. 1.50; Barclay & Tavares, 2002). But in this time period, the rate for all violent crimes was almost 250% greater in England and Wales than in the United States (1,295 vs. 536; computed from data reported in Barclay & Tavares, 2002).[4]

Japan is generally considered to be a relatively peaceful society. It has lower rates of homicide (1.06) and violent crime (39) than does the United States or most Western countries (Barclay & Tavares, 2002). Japan is also a more collectivistic society, emphasizing high moral discipline, egalitarian values, and Confucian values of peace and nonviolence.

One argument frequently offered by those who claim that media violence doesn't increase aggressive tendencies is that Japan has high levels of media violence but low overall levels of violent crime. If media violence is truly a causal risk factor for violence and aggression, so the argument goes, Japan should have a high violent crime rate. There are multiple problems with this argument, of course. Perhaps the most obvious problem is that exposure to violent media is not the only important risk factor (DeLisi, 2005). Japan differs from the United States and other Western nations on many known causal risk factors for aggression and violence, such as easy access to firearms.

There are at least five reasons to expect smaller media violence effect sizes in Japan (and similar Eastern societies) than in Western societies. First, a relatively smaller effect size may result from differences in how violence is contextualized in Japanese versus U.S. media. Today, global boundaries do not exist when it comes to video games. The most popular video games are played in many countries, under different titles and with different languages. Nonetheless, the contexts of violence in video games played most frequently in Japan can be different from the contexts in video games played most frequently in the United States. Whereas action and sports games are the most popular genre in the United States and Western countries, role-playing games are the most popular

[2] We thank Christopher Ferguson for providing the list of articles used in his three meta-analyses.

[3] Note that of the violent video game studies we located from Eastern cultures, the vast majority came from Japan. Indeed, there were only two studies from other Eastern cultures, one from Singapore and the other from China.

[4] Differences in crime definitions and reporting method may account for some portion of this reversal, but most scholars in the field agree that violence rates in the United Kingdom are higher than in the United States, with the exception of that for homicide, which is considerably higher in the United States. A common explanation for the high homicide rate in the United States is the easy availability of guns, especially handguns.

genre in Japan (Yahiro, 2005). Japanese role-playing games often involve text reading, patience, and cooperative fights against computer-controlled characters, and the contexts of the violent video games that children and adolescents are exposed to in Japan are not the same as those in the West. Second, people in Japan are more likely to pay attention to situational contexts than are people in Western countries (e.g., Masuda & Nisbett, 2001; Nisbett, Peng, Choi, & Norenzayan, 2001). A third reason concerns cultural differences in the meaning, experience, and processing of emotions and their emotion–action linkages. As noted by Mesquita and Leu (2007), "Whereas people in independent contexts view emotional situations mainly from their own perspective . . . people in interdependent contexts assess the emotional meaning from the perspective of other people or a generalized other" (p. 739). One result of this difference in perspective is that people from Japan report being less likely to respond aggressively to an offense or insult than do people in Western cultures. Other research similarly suggests that Japanese culture tends to foster socially engaging emotions, whereas Western culture tends to foster socially disengaging emotions (e.g., Kitayama, Mesquita, & Karasawa, 2006). Similarly, research on ideal affect (what one typically would like to feel) suggests that easterners are more likely to have adjustment goals, whereas westerners are more likely to have influence goals (Tsai, Knutson, & Fung, 2006). A fourth difference concerns the context in which video games are played. One unpublished study (Kodomo no taiken katsudo kenkyukai, 2000) found that considerably fewer Japanese and South Korean fifth graders had their own TV sets (14% and 11%, respectively) than did American, British, and German fifth graders (39%, 69%, and 29%, respectively). Similar results were obtained for eighth graders (28% and 10% vs. 63%, 68%, and 62%). This suggests that Japanese youths may be more likely than Western youths to play their video games in public space, where parents can watch and monitor what they play. Research shows that parental involvement may reduce violent video game effects (e.g., Anderson et al., 2007). Research has also found that the number of friends was not different for frequent versus infrequent gamers in Japan, but that in the United Kingdom frequent gamers had fewer friends than did infrequent gamers (Colwell & Kato, 2003). Again, this suggests important context differences between East and West that might moderate video game effect sizes. Thus, violent video game effects on aggression and related outcome variables may be smaller and more complex in Japan than in Western countries.

On the other hand, most basic emotion and behavior processes are universal. For example, Frijda, Markam, Sato, and Wiers (1995) studied the action readiness after emotional experiences in Dutch, Indonesian, and Japanese participants and found that five factors were quite similar across cultures, including a factor labeled *moving against*. (There were some nonuniversal factors as well.) Similarly, there are numerous cross-cultural differences in average Big Five personality traits, some of which suggest that Eastern collectivist cultures might be more susceptible to media violence effects, others of which suggest the opposite (Schmitt, Allik, McCrae, & Benet-Martinez, 2007). Thus, there also are reasons to believe that media violence effects may well be fairly similar across cultures or even larger in Japan and Eastern cultures. In the present meta-analysis, we investigated the possibility that effect sizes might differ between Western cultures (primarily the United States) and Eastern cultures (primarily Japan).

Additional Theoretical Considerations

Over the last 45 years, an array of social–cognitive models of aggression has systematically improved the field's understanding of the processes involved in the instigation of aggressive behavior and the development of individuals prone to aggression and violence (e.g., Bandura, 1973; Berkowitz, 1984, 1993; Geen, 2001; Huesmann, 1988, 1998). The two most detailed current models are Huesmann's (1998) script model and Crick and Dodge's (1994) social information processing model. Recently, the general aggression model was developed to provide a simplified overview of the common elements among prior models of the development and expression of aggressive behavior (Anderson & Bushman, 2002; Anderson & Carnagey, 2004; Anderson & Huesmann, 2003).

Explicating and comparing these various models lies well outside the scope of this article, but these social–cognitive models allow several important predictions concerning the likely short-term and long-term effects of exposure to violent video games. In general, both short-term and long-term effects of environmental variables (e.g., insult, physical pain, violent media) on aggressive behavior operate by affecting cognitive, emotion, and/or arousal systems.

Aggression Facilitation and Inhibition

Social–cognitive models of aggression distinguish between factors that facilitate the emergence of aggression from those that inhibit it (e.g., Anderson & Huesmann, 2003; Bandura, Barbaranelli, Caprara, & Pastorelli, 1996; Berkowitz, 1984). Common facilitating factors in the immediate situational context include aggression cues (e.g., weapons, violent media) and unpleasant situational events that put people in a bad mood (e.g., provocation, frustration, hot temperatures, loud noises, unpleasant odors, pain). Inhibiting factors include fear of retaliation, negative emotional reactions to images and thoughts of violence, moral beliefs opposing violence, and pleasant situational events that put people in a good mood.

Short-Term Versus Long-Term Effects

Short-term effects are those in which a person plays a video game for a brief time (e.g., 15 min) before relevant measures are obtained. Usually, short-term effects are assessed in experimental studies conducted in labs or in schools. Long-term effects are those that accrue from repeated exposures over a relatively long period of time, such as months or years. Long-term effects typically are assessed in cross-sectional and longitudinal studies.

For theoretical reasons, the effects of video game violence might differ as a function of whether one is discussing short- or long-term effects. This is because the same stimulus can have multiple effects on several factors that facilitate or inhibit aggression. For example, playing a video game with sanitized violence versus a bloody version of the same game may lead to similar levels of aggressive behavior in the immediate situation, whereas repeated exposure to one or to the other version may lead the bloody version to have larger long-term effects. This could happen if both versions equally prime aggressive scripts while being played but lead to differential changes in more stable, long-term factors, such as emotional desensitization to violence, after the game has been turned off. On the other hand, if the bloody version not only primes aggressive behavioral scripts but also increases arousal, it might lead to more aggression in the immediate situation than does the sanitized version.

Immediate, short-term effects are mainly the result of priming existing knowledge structures, such as various types of schemata and scripts (see Bushman & Huesmann, 2006). Priming processes require only (a) a person who already has at least a few well-developed aggression scripts and (b) brief exposure to a video game that requires violent action. There need be no surface-level similarity between the violence in the video game and the aggression measure, as long as the person's aggression scripts have been activated. That is, the game characters do not need to be similar to the player or the player's later real-world target, and the violence in the game does not need to be similar to the player's real-world aggression options. Once aggressive scripts have been activated, additional exposure to the violent video game is unlikely to have more than a minimal impact on later aggressive behavior. If priming of existing knowledge structures is the main process underlying an observed increase in aggression following video game play, playing the randomly assigned games for 15 min versus 30 min should make little difference, all else being equal.

Short-term effects might also reflect mimicry or observational learning of new behaviors and of new beliefs about their likely success. If the main process underlying an observed short-term violent video game effect is such mimicry/observational learning, greater exposure to the violent game (e.g., 30 vs. 15 min) should lead to better learning of the new aggression script and, in the right circumstance, to larger increases in aggression. The context most likely to favor this type of short-term effect is when the participants do not already have well-learned aggression scripts (e.g., very young children); when the aggressive behavior being modeled in the game is novel; and when the aggressive behavior test situation closely resembles the video game in terms of the characters, the provocation, and the possible aggressive action that is available to the participants. Such conditions are rarely (or never) encountered in the existing violent video game experiments, which is why most video game violence researchers believe that the existing short-term effects are mainly the result of priming effects (e.g., Anderson et al., 2003, 2007; Bushman & Huesmann, 2006; Kirsh, 2006; Krahé, 2001).

Long-term effects mainly result from relatively permanent changes in beliefs, expectancies, scripts, attitudes, and other related person factors that are brought about by repeated exposure to video game violence. Because these person factors are relatively stable, repeated exposure to video game violence (or to other environmental risk factors) is required to create significant change. Playing a violent video game one time for 20 min will not change a well-adjusted adolescent into a potential school shooter, with all of the anger, hostile beliefs, expectations, and personality traits that go along with such extreme behavior. But repeated exposure to violent media is expected to lead to measurable changes in the chronic accessibility of aggression-related knowledge structures (e.g., aggression scripts, attitudes and beliefs that support aggressive action) and in relatively automatic reactions to scenes or thoughts of violence (e.g., lack of empathy, physiological desensitization). Another factor important in understanding long-term effects of exposure to violent media is whether the person's environment encourages or discourages aggression. For example, some cultures are relatively supportive of certain types of violence, whereas other cultures condemn them. Similarly, different families within a culture may respond differently. This may be why it appears that having parents who are very involved in one's media usage sometimes acts as a protective factor (e.g., Anderson et al., 2007). Of course, if the highly involved parents actively encourage violent behavior, they are likely to exacerbate the media violence effect.

Aggressive Cognition Versus Affect and Arousal

Video games can be exciting, fun, frustrating, exhilarating, and boring. Being the target of potential harm, even in the virtual world of video games, is likely to prime aggressive cognitions and emotions and to increase physiological arousal. The aggressive cognition aspect is of particular interest for two reasons. First, many situational factors can increase arousal and anger, even certain nonviolent video games. For instance, race driving video games, sports video games, and even perceptual/motor skills games that require intense concentration and rapid responses (e.g., Tetris, Bejeweled) can increase heart rate and blood pressure. Similarly, video games that are too fast paced or too difficult for the player are likely to increase frustration and anger, which in turn might activate aggressive thoughts. But violent video games, by their nature, *require* the activation of aggressive thoughts, whereas nonviolent games do not require it. Second, the repeated activation of aggressive thoughts, both novel ones (especially in children) and well-practiced ones, is the most likely route to relatively permanent changes in the person, because the activation of aggression-related knowledge structures becomes more automatic and chronic with repetition and eventually becomes part of the person's personality (Strack & Deutsch, 2004; Wegner & Bargh, 1998). The negative affect and physiological arousal instigated by a video game (violent or nonviolent) likely dissipate fairly quickly and are less likely to leave long-term traces in the brain than are the cognitive learning and overlearning of aggression-related perceptual and social schemata (including aggressive behavioral scripts) that are rehearsed constantly while playing violent games.

Predictions

Before spelling out our specific predictions, we want to raise two key points. First, predicting the pattern of all the possible combinations of variables in video game studies requires a thorough knowledge of which processes are engaged by the video game. Will a third-person shooter have a different impact on immediate aggression than a first-person shooter? Will gorier games have a bigger impact than less gory games? Without knowledge of how well each specific game activates aggressive thoughts, feelings, and physiological arousal, any prediction is risky at best. Second, the importance of assessing a host of potential short-term and long-term effects of different types of violent video games becomes obvious. Nonetheless, a number of broader scale predictions are possible. For example, all else being equal, participants randomly assigned to play a violent video game should tend to behave more aggressively for a short period of time afterward than those randomly assigned to play an equally fun and equally challenging nonviolent video game. We next offer additional predictions, grouped by outcome variable.

Aggressive behavior. We expected to find that playing a violent video game would increase aggressive behavior in a short-term experimental context, relative to playing a nonviolent video game that is equally exciting, arousing, and enjoyable. We expected similar effects in long-term contexts, that is, in cross-sectional correlation studies and in longitudinal studies. We expected the largest effects in short-term experimental studies and the smallest effects in longitudinal studies, once sex has been controlled. This is because experimental studies generally are better at controlling for effects of extraneous variables that increase the error variance and therefore decrease the effect-size magnitude.

Aggressive cognition. The predictions for aggressive cognition were the same as for aggressive behavior.

Aggressive affect and physiological arousal. Brief exposure to violent video games should, on average, increase physiological arousal and aggressive affect. An important methodological caveat is warranted, however. Studies based on violent and nonviolent video games that have been preselected to be equally arousing obviously are not appropriate tests of the short-term arousal- and affect-inducing effects of violent video games. Thus, they should be excluded from the analyses designed to test this specific hypothesis. The same is true when comparison games have been preselected to create equivalent affective states.

It is less clear what to expect in long-term contexts, but the temporary nature of moods and of physiological arousal leads us to expect either very weak long-term effects or none at all. Weak long-term effects might occur on aggressive affect indirectly through habitual increases in aggressive thinking or through problems engendered by habitual aggressive behavior. Weak long-term effects on arousal might occur in young people through changes to brain regions that control cardiovascular and other arousal-related functions. Unfortunately, there are no long-term studies of physiological arousal with which to test this hypothesis.

Empathy/desensitization. It is unclear whether playing a violent video game for a brief period of time should have a detectable impact on measures of desensitization to violence or of empathy for violence victims. Systematic desensitization therapies suggest that repeated exposures to gory scenes of violence and to pain and suffering of others will have some impact on a person's physiological reactions to new scenes of violence (desensitization) and on empathetic responses to victims, but such therapies typically take place over a period of days or even weeks. Thus, we expected brief exposure to a violent video game would have a relatively small impact on desensitization and empathy. However, we expected larger effects in long-term studies and in experimental studies that involve longer desensitization procedures. Unfortunately, there are no long-duration experimental desensitization studies of violent video game effects.

Prosocial behavior. Social–cognitive models of social behavior suggest that briefly playing a violent video game should reduce prosocial or helping behavior in the immediate situation. The temporary increase in aggressive cognition and affect might be incompatible with, or might interfere with, empathic thoughts and emotions that frequently underlie helping behavior. Similarly, short-term desensitization effects could reduce helping victims of violence in several ways (Bushman & Anderson, 2009; Carnagey, Anderson, & Bushman, 2007). We therefore expected that video game violence would produce short-term decreases in some forms of prosocial behavior.

On the other hand, many violent video games involve the use of violence to help others, such as saving the princess, one's teammates, or all of humanity from enemies that need to be killed. Thus, it is possible that playing certain types of violent video games might prime a type of "hero" script and thereby lead to an increased likelihood of certain limited types of helping behavior. No studies covered by our search period tested this hypothesis, though we are aware of several such studies currently in progress.

We did not expect to find strong long-term decreases in prosocial behavior, because the types of situations that inspire helping behavior are relatively unlikely to be of the ambiguous kind that allow spontaneous priming of aggressive thoughts and feelings. One exception to this latter prediction concerns helping victims of violence or injury. Because emotional desensitization to injury-related cues (expressions of pain, presence of blood) reduces the perceived need for help by violence victims, repeated exposure to violent video games should yield long-term declines in this specific type of prosocial behavior (Carnagey et al., 2007). Unfortunately, there have not been enough direct tests of helping victims of violence, and there are no longitudinal studies testing this specific hypothesis.

In sum, theory suggests that violent video game effects on prosocial behavior should be very context specific. However, the specific contexts used in existing prosocial behavior studies are of the type that lead to predictions of a significant decrease in short-term experimental studies and a small effect (or no effect) in long-term studies.

The Present Meta-Analysis

We undertook the present meta-analysis for four related reasons. First, the video game violence research literature is expanding rapidly, with new studies being reported almost monthly. An updated meta-analysis is badly needed to take into account the new research. Second, many of the newer studies are of better methodological quality than some of the earlier studies (see the meta-analysis in Anderson et al., 2004). With this larger sample of higher quality studies, one can use stricter inclusion criteria for the main analyses of potential moderators and still have a sufficient sample of studies to yield meaningful results. In essence this larger set of high-quality studies allows tests of theoretical propositions that could not be tested in prior years. Third, there is a growing body of research using Japanese samples, a literature that has gone largely unnoticed by scholars in the West. This body of research not only adds to the total body of studies available for an updated meta-analysis but also allows examination of whether video game violence effects occur in a low violence society that differs from the West in so many important ways. Fourth, the larger body of studies in this domain allows tests of a number of potentially important moderator variables. For example, in experimental studies the video game violence effect size may differ as function of whether the violent game is played from a first-person or third-person perspective.

Method

Literature Search Procedures

Outcome variables. We focused on six outcome variables, the first five of which have been used in prior meta-analyses. The outcome variables were physically aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, prosocial (helping) behavior, and a combined empathy/desensitization variable. All are described more fully in the Results section.

Western studies literature search. We searched PsycINFO and MEDLINE for all entries through 2008 using the following terms: (*video** or *computer* or *arcade*) and (*game**) and (*attack** or *fight** or *aggress** or *violen** or *hostil** or *ang** or *arous** or *prosocial* or *help** or *desens** or *empathy*). In addition, we searched the reference sections of prior meta-analytic and narrative reviews. We included dissertations, book chapters, and unpublished papers.

Japan studies literature search. There is no search engine comparable to PsycINFO for psychological research in Japan. Therefore, we searched CiNii (NII Scholarly and Academic Information Navigator) and Magazine Plus (Nichigai Associates, Inc.) for all entries through 2008 using the following terms: (*terebigemu* [TV game] or *konpuutaagemu* [computer game]) or *bideogemu* [video game]).

From these two searches we retrieved over 130 research reports that contained some potentilly relevant original data, with over 380 effect-size estimates based on over 130,000 participants. As shown in Table 1, this is a huge increase since the last comprehensive meta-analysis (Anderson et al., 2004) as well as the most recent meta-analyses by Ferguson (2007a, 2007b; Ferguson & Kilburn, 2009).

Outcome Variable Details

Aggressive behavior. High-quality experimental studies typically measure aggressive behavior using noise blasts, electric shocks, or hot sauce given to an ostensible partner (in the last case, the partner is known to hate spicy food; for discussions and studies of validity, see Anderson, Lindsay, & Bushman, 1999; Bushman & Anderson, 1998; Carlson, Marcus-Newhall, & Miller, 1989; Giancola & Parrott, 2008). High-quality nonexperimental studies typically measure aggressive behavior using standardized questionnaires (e.g., Buss & Perry, 1992), self-reports, peer reports, teacher reports, or parent reports. Whenever possible, we used measures of physical aggression, because that is the type of aggression most frequently modeled and rewarded in violent video games. In many of the nonexperimental studies, the aggression measure was a composite of physical and verbal aggression.

Table 1
A Comparison of the Sizes of Recent Meta-Analyses of Violent Video Game Effects to That of the Current Meta-Analysis

	Violent video game studies		
Meta-analysis	No. papers	*K*	*N*
Anderson et al., 2004	44	97	16,534
Ferguson, 2007a	24	25	4,205
Ferguson, 2007b	17	21	3,602
Ferguson & Kilburn, 2009	14	15	Unknown
Present article	136	381	130,296

Note. It was not possible to derive *N* for violent video game effects for Ferguson and Kilburn (2009) because the reported *N*s included studies that had TV and film effects confounded with video game effects. $K =$ number of effects sizes; $N =$ number of participants.

Aggressive cognition. Aggressive cognition has been assessed in numerous ways. Short-term experimental studies have used reading reaction time, story completion, word fragment completion, Stroop interference, speed to recognize facial emotions, and hostile attribution bias measures. Occasionally, more traitlike measures of aggressive cognition (such as attitudes toward violence) have been used in short-term experimental studies; these are inappropriate because they measure stable thoughts and beliefs that develop over a lifetime and should not be influenced by playing a video game for a few minutes. Nonexperimental studies have used measures of trait hostility, hostile attribution bias, attitudes toward violence, hypothetical aggression statements, aggression vignettes, implicit association tests, and normative beliefs about aggression. A few measures, such as variants of the Implicit Association Test, have been found to be sensitive to short-term experimental manipulations as well as to reflect longer term attitudes and so have properly been used in both short-term and long-term studies (Lane, Banaji, Nosek, & Greenwald, 2007).

Aggressive affect. Aggressive affect measures used appropriately in short-term experimental studies include self-report measures of state hostility, state anger, and feelings of revenge. One experimental study assessed brain function in regions of the brain known to be affected by anger. Most measures for nonexperimental studies were self-reported trait anger scales.

Physiological arousal. Physiological arousal was assessed with measures of heart rate, blood pressure, or skin conductance.

Empathy/desensitization. Empathy refers to the degree to which a person subjectively identifies and commiserates with a victim and feels emotional distress. Empathy measures are almost always based on self-report scales in which participants indicate the extent to which they empathize with, feel sympathy for, or feel sorry for a particular person or group of people. In high-quality studies, state measures are used in short-term experimental contexts, whereas trait measures are used in nonexperimental contexts.

The term *desensitization* has been used to cover a wide range of measures, including shorter recommended jail terms for persons convicted of a violent crime to longer latency to intervene in a violence episode (e.g., Carnagey et al., 2007). Theoretically, however, desensitization refers to a reduction in negative emotional response to scenes of violence. The best measures of such effects are negative emotion-related measures, such as heart rate, skin conductance, or other physiological indicators of emotion-related reactions to scenes of violence. In the present article, desensitization specifically refers to a reduction in physiological reactivity to scenes of violence. Most other measures that have been called desensitization are actually theoretical sequelae of reduced negative emotional reactions.

Empathy and desensitization are similar in that both refer to automatic emotional reactions to harm befalling someone else. They differ in directionality and in type of measurement (physiological vs. self-report). We combined these two outcome variables into one category because of their conceptual similarity and because there were too few studies to warrant separate meta-analyses. We reverse scored the desensitization effects, so that negative effects indicated that high exposure to video game violence was associated with high desensitization or low empathy. In other words, theory predicts negative effect sizes.

Prosocial behavior. Experimental studies used donating of jelly beans or money, helping someone succeed at a task, or helping a victim of a staged violent episode. Nonexperimental studies used self- and other reports of helping behavior.

Methodological Criteria Assessment

Many of the effect-size estimates are from high-quality studies that used well-established and theoretically appropriate measures or manipulations of exposure to violent video games and well-established, theoretically appropriate outcome measures. However, other studies suffer from one or more serious weaknesses relative to the specific hypothesis. For example, some experimental studies used violent and nonviolent video games that were chosen on the basis of pilot testing because they yielded equal states of arousal; obviously, such studies do not provide appropriate tests of the effect of violent content on arousal. Usually, this type of piloting procedure was done by the original authors to demonstrate that the selected video games did indeed yield similar arousal states, so that other hypotheses could be more accurately tested in the main study (e.g., Anderson et al., 2004; Anderson & Dill, 2000).[5]

Other studies used weak or inappropriate measures of exposure to video game violence, such as the amount of time spent playing any type of video game rather than time spent playing violent video games. Indeed, many studies report the correlations of both the time on violent games and the time on all games with physical aggression in order to test whether the more theoretically appropriate measure yields larger effects (e.g., Anderson & Dill, 2000; Anderson et al., 2007).

Of course, meta-analysis researchers always face the dilemma of dealing with studies of widely varying quality and characteristics. The common solution is to establish a set of methodological criteria and then exclude studies that fail to meet these criteria. In some domains this works well, but in more controversial domains the inclusion/exclusion decisions themselves become the focus of extended debate, thus decreasing the value of the meta-analysis itself. We dealt with this issue in multiple ways. First, we divided studies into two broad categories, those whose methods reflected the best practices in the manipulation and measurement of theoretically appropriate independent and dependent variables versus those that did not.[6] The main analyses (including the moderator analyses) were performed on this set of high-quality studies. Second, we contacted the authors of reports that appeared to have additional unpublished data that could be used to compute effect-size estimates that met the best practice criteria. In this way, we were able to obtain several best practice effect-size estimates that were not in the original reports.[7] Third, we conducted several types of sensitivity analyses. As has been done in other recent meta-analyses (e.g., Chida & Hamer, 2008), the average effect was estimated for each outcome variable on both the full sample and on the best practice sample. If both types of effect-size estimates could be computed from the same study, we kept only the one based on best practices. The full sample analysis reduces the plausibility of claims of selection bias, because all potentially relevant studies are included. The full sample may either underestimate or overestimate the true effect sizes, because it includes studies whose methods might artifactually inflate or deflate the reported effect size. We therefore also report a comparison of best practice versus other studies. As a final type of sensitivity analysis, we used the trim and fill procedure to see how much various effect-size estimates changed as a result of potential selection (publication) bias.

Best practice inclusion criteria. The inclusion criteria for best practice studies are listed in Table 2, which also gives examples of criteria violations. A more detailed listing of the specific violations for specific studies is too lengthy for inclusion in this article but can be downloaded at the following web page: http://www.psychology.iastate.edu/faculty/caa/abstracts/2010-2014/NotBestViolations.pdf. Two independent raters examined each effect and judged whether it met all six criteria. Initial agreement was over 93%. Discrepancies were examined and discussed with a third judge until consensus was reached.

Correlated data. For the longitudinal studies we included both a longitudinal effect size and a cross-sectional effect size. The latter was the average of the two cross-sectional effects, one measured at Time 1 and the other at Time 2.

For studies that reported multiple effects on the same conceptual outcome variable, we took one of two actions. In those cases where one measure was clearly better than the others, based on theoretical relevance (e.g., physical aggression is more relevant to violent video game effects than is verbal aggression), established validity (e.g., use of a well-validated multiple item measure of trait physical aggression vs. a new single item measure of trait aggression), or other empirical evidence offered in the study, we used the best measure. For example, if a study reported two new outcome measures of aggressive behavior, and only one of them correlated significantly with a third variable known to be related to physical aggression (e.g., trait irritability), we used that measure (e.g., Anderson & Dill, 2000, Study 2). In those cases in which there was not a clear best measure, we used the average effect size (Bartholow, Bushman, & Sestir, 2006). Note that we also repeated the main analyses, always using the average effect, and found essentially the same results. For all analyses, we used fixed effects models, although random effects models yielded very similar results.

Coding frame: Moderator variables.

All studies. We coded the following information for each effect size: research design (experimental, cross-sectional, longitudinal); av-

[5] Of course, the quality of that study relative to tests of violent video game effects on aggressive cognition would be very high.

[6] Study quality also varies within each of these two broad categories. Although one might attempt a more fine-grained, multilevel assessment of quality, such an attempt would require more studies than presently exist in this domain.

[7] As should be clear, many of the not best practices effects were never intended by the original authors as tests of the specific hypothesis for which they earned not best practices status. In many cases, the not best practices effects were the result of high-quality procedures used to improve the precision of the main hypothesis test. In other cases, the not best practices effect was not part of the main study at all but was simply reported in a correlation matrix that included other variables that were the main focus of the article. Thus, neither readers nor authors of original reports should interpret a not best practices listing as a negative judgment about the author's methodological skills or the overall value of the study that included the not best practices effect.

Table 2
Inclusion Criteria for Determining Whether a Study Qualifies as a "Best Practices" Study

Criteria	Examples of inclusion criterion violations
1. The compared levels of the independent variable were appropriate for testing the hypothesis.	In a short-term experiment, participants in the "nonviolent" condition played a video game that contained considerable violence.
2. The independent variable was properly operationalized.	In a nonexperimental study, total video game play rather than violent video game play was used as the predictor variable.
3. The study had sufficient internal validity in all other respects.	Participant retention was substantially lower in one experimental condition than another, indicating potential self-selection of participants.
4. The outcome measure used was appropriate for testing the hypothesis.	The hypothesis specifies an effect on aggressive behavior, but the outcome measure assessed behavior directed toward an inanimate object rather than a person.
5. The outcome measure could reasonably be expected to be influenced by the independent variable if the hypothesis was true.	A measure of personality trait aggression was used as the measure of aggressive behavior in a short-term experimental study.
6. The outcome variable was properly computed.	Pre- and postmanipulation scores were averaged but were not reported separately.

erage age (when only "college students" was reported, we assigned an age of 20); culture (East [Japan, China, Singapore] vs. West); and sex of participants. We also coded a number of other characteristics specific to a research design and/or an outcome variable.

Experimental studies. For experimental studies we coded the following features: violent game player's perspective (first or third person); violent game player's role (hero, criminal, neither); violent game targets (human, nonhuman, both); and duration of time spent playing the assigned video game immediately prior to assessment of the dependent variable. There are no obvious theory-based predictions for these four moderators for most dependent variables in short-term experimental studies. For example, if the short-term effect of playing a violent game on aggressive behavior is a priming phenomenon, playing the game for 30 min is unlikely to have a greater impact than playing it for 15 min, unless the content or difficulty changes a lot in the last 15 min.

Many experimental studies of aggressive behavior used some version of the competitive reaction time (CRT) task. In this task participants are told that they are competing against another person on a series of reaction time trials, that the loser of each trial will receive a punishment immediately after losing the trial, and that before each trial participants set the punishment level that their opponents will receive. The punishment settings are used to assess physical aggression. Original versions used electric shocks (Taylor, 1967), but in more recent years the punishments usually involve blasts of white noise. The CRT is one of the most widely used laboratory techniques for measuring physical aggression and has been shown to have good external validity (Carlson et al., 1989; Giancola & Chermack, 1998; Giancola & Parrott, 2008). Thus, for aggressive behavior studies we also coded whether or not the CRT task was used.

For experimental studies of aggressive cognition, we coded whether or not the outcome variable was some type of rapid automatic cognitive response task (e.g., reaction times) or some type of more thoughtful measure, such as hostile attributional style.

Nonexperimental studies. Nonexperimental studies included several different ways of measuring exposure to violent video games. We created a dichotomous code that distinguished between studies in which the amount of time spent playing violent games was specific to each game or game type (and then summed or averaged across games or game type), and studies that used some other measures of exposure to violent video games. An example of the former type of measure is the one reported by Anderson and Dill (2000). Their video game violence (VGV) exposure measure has participants list their five most frequently played games. Participants then indicate for each game how frequently they play that game, how violent the graphics are, and how violent the content is. The two violence ratings are averaged and then multiplied by the frequency. This is done for each game listed, and then these five scores are averaged. We refer to this as the VGV-specific type of measure. The second type of measure was used in most of the Japanese studies. Participants rate how frequently they see violent scenes in the games they play. In some studies, this violent scenes rating was then multiplied by a measure of how many hours per week the participant played video games of any type. We refer to this as the VGV-general measure.

For nonexperimental studies of aggressive behavior we coded whether the measure was of physical aggression versus a mixture of physical and some other type (most commonly verbal). We also coded whether or not the measure was primarily composed of more extreme physical aggression that is illegal (violence, such as assault).

For nonexperimental studies of aggressive cognition we coded whether the measure was of trait hostility versus some other type of aggressive cognition, such as attitudes toward violence or hostile attributional style. We did this because a large number of studies used the Hostility subscale of the Aggression Questionnaire (Buss & Perry, 1992) and because that measure is very similar to trait anger, which is not theoretically expected to yield strong cross-sectional or longitudinal effects.

For longitudinal studies we coded the length of time between the initial and the final assessment period. This ranged from 3 to 30 months.

Partial correlations. Normally, partial correlations are not used in meta-analyses because the statistical theory underlying meta-analytic procedures assumes that one is working with raw (zero-order) correlations. This was not an issue for experimental studies, even when separate effects were not reported by sex, because random assignment removes any unwanted correlation between the independent variable and sex.

As noted earlier, however, a number of nonexperimental studies in this domain either ignored sex or reported finding no Sex × Video Game Violence (VGV) interaction and then reported an overall VGV effect that combined across sex. Theoretically this could inflate the VGV effect, because males tend to play more violent video games and tend to be more physically aggressive than females. Of course, if violent video games actually do increase physical aggression (and other aggression-related outcome variables), controlling for sex could lead to artificially low effect-size estimates. Our solution to these issues was to contact researchers with a request for additional data that would allow us either to get separate estimates for males and females or to statistically partial out the effect of sex. We then created two overlapping data sets. One, labeled the "best raw" data set, consisted of all best practices studies with effects in their rawest form.[8] The other, labeled the "best partials" data set, contained only effects that had been corrected for sex, either by separate estimates for males and females or by use of partial correlations.

For longitudinal studies, the best raw data set contained the correlations between Time 1 VGV exposure and Time 2 outcomes. For the best partials data set, the effect sizes were partial correlations with both sex and Time 1 outcomes partialed out. The main analyses and all of the moderator analyses were carried out on the best partials data set. Thus, results from the best partials data set are very conservative estimates and may well underestimate the true video game effects. However, comparing these effects to the corresponding best raw effects gives another indication of the strength (or weakness) of the overall effects of violent video games.

Our final sample consisted of 381 effect-size estimates based on 130,296 participants. Of these, over half (221 effects) met the best practice inclusion criteria. Table 3 illustrates the breakdown by outcome variable and, for the best partials data set (the one on which subsequent analyses focus), by research design and culture. Appendix A lists the sources that contributed at least one best practice effect. Appendix B lists additional studies that contributed effects that did not meet the best practice criteria. (The Appendices can be found in the online supplemental materials.)

Table 3
Characteristics of the Samples

Variable	Full sample *K*	Full sample *N*	Best raw *K*	Best raw *N*	Best partials *K*	Best partials *N*
Category						
Aggressive behavior	140	68,313	79	21,681	75	18,751
Aggressive cognition	95	24,534	59	16,271	53	12,598
Aggressive affect	62	17,370	37	9,191	35	7,543
Prosocial behavior	23	9,645	16	6,906	16	6,905
Empathy/desensitization	32	8,528	15	6,580	14	6,268
Physiological arousal	29	1,906	15	969	15	969
Total	381	130,296	221	61,598	208	53,034
Culture						
East					64	32,436
West					144	20,598
Research design						
Experimental					92	8,705
Cross-sectional					82	28,788
Longitudinal					34	15,541

Note. K = number of effects; N = total sample size.

About one third of the best practice effects and over half of the participants were from Eastern cultures, mainly Japan. The majority of the remaining best practice effects were from U.S. samples, but samples also came from Australia, Germany, Italy, the Netherlands, Portugal, and the United Kingdom.[9]

Meta-analytic procedures. All effects sizes were converted to the correlation coefficient, denoted by r. We used the software program Comprehensive Meta-Analysis and used a fixed effects model so that we could assess the heterogeneity in various subsets of studies. For each outcome variable we first computed the overall average effect size and then computed a moderator analysis based on type of study. This was done for both the best raw and the best partials data sets. We then conducted more detailed moderator analyses on the best partials data.

We also conducted analyses to address the possibility that results might be affected by selection bias (also called publication bias) in the sample of studies included in the meta-analyses. The concern is based on the premise that studies that fail to "work" are less likely to be published, which might bias the results of a meta-analysis. To assess the possibility that publication bias affected our results, we used the trim and fill procedure (Duval & Tweedie, 2000a). Trim and fill is based on the assumption that in the absence of publication bias, the studies will be distributed symmetrically about the mean effect size (plotted on the x-axis) relative to standard error (plotted on the y-axis), because the sampling error is random. In the presence of publication bias, studies are expected to be systematically missing in a manner that can be identified by the trim and fill analysis. In the case of positive effect data, if low-effect nonsignificant studies are missing, we would expect a gap on the left-bottom quadrant in the plot, where the nonsignificant studies would have been if we had located them. If, based on other selection mechanisms, high-effect studies are selectively missing, the gap would be on the right side of the mean effect. If asymmetry is detected, trim and fill uses an iterative procedure to remove the most extreme small studies from the specified side of the funnel plot, re-computing the effect size at each iteration until the funnel plot is symmetric about the (new) effect size. In theory, this will provide an unbiased estimate of the effect size. Although trimming yields an adjusted effect size, it also reduces the variance of the effects, resulting in a too-narrow confidence interval. Therefore, the algorithm then adds the original studies back into the analysis and imputes a mirror image for each. This fill has no impact on the point estimate but produces a better estimate of the variance (Duval and Tweedie, 1998, 2000a, 2000b). The major advantage of this approach is that it addresses an important question, What is the best estimate of the unbiased effect size? But, as noted earlier, this estimate should not be interpreted as the true effect size, because it is based on imputed data. Furthermore, if there is a true, theoretically meaningful relationship between effect size and sample size (e.g., researchers use larger samples when conducting longitudinal studies because

[8] For several studies, the only estimates possible were some type of partial correlation. Rather than discard potentially useful information, we kept these in both of the best practices analyses.

[9] The totals across outcome variables in the table are not all independent samples, because many studies reported multiple outcome variables for the same sample.

they know longitudinal effects are smaller than cross-sectional ones), the trim and fill procedure can erroneously adjust the average effect sizes. Ideally, the trim and fill procedure is used on appropriate subsets of studies, and the difference between the original effect size and the trim and fill adjusted effect size should be used to get a feel for the possible biasing effect of publication or selection bias, not as an estimate of the true effect size.

Results

Analysis Plan

Main and moderator analyses were done separately for each outcome variable. In other words, we conducted six independent meta-analyses. We first present results of analyses on the best practice effects, for both the best raw and the best partials data sets. These are the effects that met the inclusion criteria. We then present more detailed moderator analyses on the best partials data. Finally, we present results from several sensitivity analyses for all outcome variables, including analyses of methodological quality as a moderator of effect size. Note that in this section, "study" refers to a sample on which an effect size was computed.

Violent Video Game Effects on Aggressive Behavior

Main analyses. Table 4 presents the results of the main analyses on aggressive behavior for both the best raw and the best partials data.[10] Figure 1 illustrates several main points from these analyses. First, regardless of research design and regardless of whether the standard zero-order correlation approach or the much more conservative partial correlation approach was used, VGV exposure was significantly related to higher levels of aggressive behavior. Most notably, in longitudinal studies even when sex and Time 1 aggressive behavior were controlled, amount of violent video game play at Time 1 significantly predicted an increase in aggressive behavior at Time 2.

Second, partialing out sex effects (in cross-sectional and longitudinal studies) and Time 1 aggressive behavior effects (in longitudinal studies) greatly reduced the average effect size of VGV. Third, in the best partials data, experimental studies yielded the largest effects whereas longitudinal studies yielded the smallest effects. Fourth, ignoring research design led to very large heterogeneity effects. Fifth, when research design, sex, and Time 1 effects were controlled, none of the heterogeneity effects were significant. In sum, this much larger meta-analysis, with over 70 independent effects involving over 18,000 participants from multiple countries, ages, and culture types, yielded strong evidence that playing violent video games is a significant risk factor for both short-term and long-term increases in physically aggressive behavior.

Additional moderator tests—best partials data. We conducted additional moderator tests within each research design, even though the heterogeneity test results yielded little evidence that the studies within design type came from different populations. There were two reasons for doing these additional tests. First, there are several specific comparisons that are of special interest for theoretical (e.g., culture), methodological (e.g., how to measure VGV exposure), or public policy reasons (e.g., player perspective). Second, the heterogeneity tests are omnibus tests; it is possible that more focused tests will yield significant differences.

Culture. The effect of culture (Eastern vs. Western) was not significant in any of the research designs. The average effect in experimental studies was slightly larger in Eastern than in Western studies but not significantly so, $Q(1) = 0.28$, $p > .50$. In cross-sectional studies, the VGV effect was slightly larger in Western than in Eastern studies but not significantly so, $Q(1) = 1.08$, $p > .20$. In longitudinal studies the VGV effect size was somewhat larger in Western ($r+ = .126$, $K = 5$, $N = 1{,}037$) than in Eastern studies ($r+ = .059$, $K = 7$, $N = 3{,}392$). This effect was marginally significant, $Q(1) = 3.52$, $p < .07$.

Sex. There was no evidence that the VGV effect on aggressive behavior differed for males and females ($ps > .10$). The VGV effect was slightly larger for females in experimental and longitudinal studies and was slightly larger for males in the cross-sectional studies, but none of these differences were significant.

Age. Average age of the participants was not significantly related to the VGV effect sizes in experimental or longitudinal studies ($ps > .50$). However, it is important to note that there were no longitudinal studies on participants older than 16. For cross-sectional studies there was a marginally significant effect of age ($b = -.005$, $Z = -1.82$, $p < .07$). Studies with older participants tended to yield slightly smaller effect sizes than did those with younger participants.

Moderators specific to experiments. Of the moderators specific to experiments, the only one with at least a marginally significant effect was the CRT variable, $Q(1) = 2.90$, $p < .09$. Experimental studies that used some version of the CRT task ($r+ = .188$, $K = 15$, $N = 1{,}724$) yielded slightly smaller effects than did those with some other measure of aggression ($r+ = .259$, $K = 12$, $N = 789$). None of the other moderators associated with experimental design approached significance (i.e., player perspective, player role, target type, time on game).

Moderators specific to nonexperimental studies. In the cross-sectional studies, one of the additional moderators was marginally significant. Studies with a pure physical aggression measure ($r+ = .184$, $K = 28$, $N = 7{,}137$) yielded slightly larger VGV effects than did studies that used some mixed aggression measure ($r+ = .153$, $K = 8$, $N = 4{,}672$), $Q(1) = 2.82$, $p < .10$. The method of measuring VGV exposure did not approach significance ($p > .20$), but the average effect was slightly larger when VGV-specific measures were used than when VGV-general measures were used. The violence moderator (violent behavior vs. aggressive but not violent behavior) also did not approach significance ($p > .30$).

For longitudinal studies, the method of measuring violent video game exposure was significantly related to the magnitude of the effect size, $Q(1) = 6.81$, $p < .01$. Studies that used VGV-specific type measures ($r+ = .152$, $K = 4$, $N = 902$) yielded significantly larger VGV effects than did those that used VGV-general type measures ($r+ = .055$, $K = 8$, $N = 3{,}527$). However, this moderator was somewhat confounded with the East/West mod-

[10] For completeness, we also included results from the full sample, but these results are not discussed in any detail.

Table 4
Aggressive Behavior: Average Effect of Violent Video Game Exposure by Study Design for Best Raw, Best Partials, and Full Sample Data Sets

Design	*N*	*K*	Effect size and 95% CI: Point estimate	*LL*	*UL*	Test of null (two-tailed): *z*	*p*	Heterogeneity: *Q*	*df*(*Q*)	*p*
Best raw										
Experimental	2,513	27	.210	0.172	0.248	10.512	.000	19.41	26	.819
Longitudinal	4,526	12	.203	0.175	0.231	13.787	.000	40.72	11	.000
Cross-sectional	14,642	40	.262	0.247	0.277	32.291	.000	207.99	39	.000
Total within								268.12	76	.000
Total between								16.74	2	.000
Overall	21,681	79	.244	0.231	0.256	36.422	.000	284.86	78	.000
Best partials										
Experimental	2,513	27	.210	0.172	0.248	10.512	.000	19.41	26	.819
Longitudinal	4,429	12	.075	0.045	0.104	4.974	.000	9.22	11	.601
Cross-sectional	11,809	36	.171	0.154	0.189	18.732	.000	49.55	35	.052
Total within								78.19	72	.289
Total between								40.17	2	.000
Overall	18,751	75	.154	0.140	0.168	21.118	.000	118.36	74	.001
Full sample										
Experimental	3,464	45	.181	0.148	0.213	10.538	.000	79.08	44	.001
Longitudinal	5,513	14	.198	0.172	0.223	14.812	.000	41.53	13	.000
Cross-sectional	59,336	81	.189	0.181	0.196	46.412	.000	771.85	80	.000
Total within								892.46	137	.000
Total between								0.70	2	.706
Overall	68,313	140	.189	0.182	0.196	49.838	.000	893.16	139	.000

Note. Effect sizes measured as *r*. CI = confidence interval; *LL* = lower limit; *UL* = upper limit.

erator, and this clouded interpretation. Three of the four VGV-specific studies were from the West, but only two of the eight VGV-general studies were from the West. Both types of measures, however, yielded significant longitudinal effects ($ps <$ .001 and .005, respectively). Neither the time between measurements nor the physical versus mixed aggression moderators approached significance.

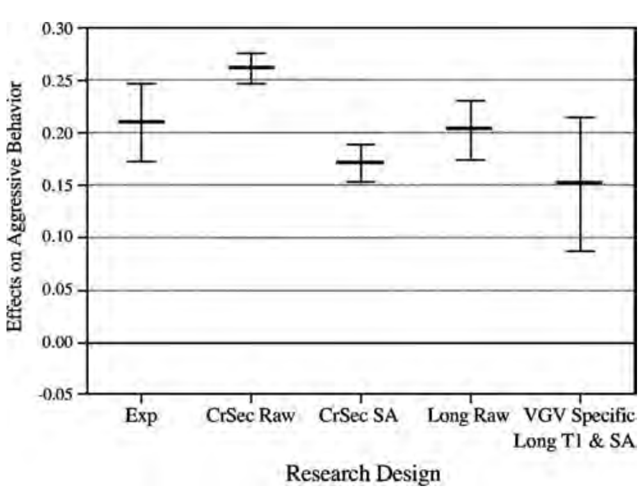

Figure 1. Effects of playing violent video games on aggressive behavior: Averages and 95% confidence intervals by research design. Exp = experimental studies (same in best raw and best partials data); CrSec = cross-sectional studies; Raw = data from best raw samples; SA = sex adjusted (data from best partials samples); Long = longitudinal studies; VGV Specific = studies that used the more specific type of video game violence exposure measure; T1 & SA = Time 1 and sex adjusted.

Summary of main findings. Regardless of research design or conservativeness of analysis, exposure to violent video games was significantly related to higher levels of aggressive behavior. For experimental studies, $r+ = .210$. For cross-sectional studies the best raw and best partials analyses yielded average effect sizes of $r+ = .262$ and .171, respectively. For longitudinal studies the best raw and best partials analyses yielded average effect sizes of $r+ =$ .203 and .075, respectively. Neither culture nor sex yielded any significant moderator effects.

The fact that significant positive effect sizes were obtained in short- and long-term contexts confirms our main theoretical hypotheses. Of particular importance is the finding of a significant longitudinal effect. This shows that playing violent video games can increase aggression over time. Thus, the present findings, especially the longitudinal ones, fill the main gap in the empirical literature on violent video game effects. Furthermore, these effects appear to generalize across culture.

The marginally significant age effect suggests that children might be more susceptible than young adults to violent video game effects, but more research specifically targeted to this question is needed. The lack of a time on game effect in experimental studies suggests that these effects are largely based on priming processes that are triggered in the first few minutes of game play. Experimental studies that vary time on game within the same study are needed to provide a more precise look at this question.

Violent Video Game Effects on Aggressive Cognition

Main analyses. Table 5 presents the results of the main analyses on aggressive cognition. As with aggressive behavior, VGV exposure was significantly related to increased levels of aggressive cognition, regardless of research design and regardless of whether zero-order correlations or the more conservative partial correlation approach was used. Furthermore, even when sex and Time 1 aggressive cognition were controlled, amount of violent video game play at Time 1 predicted a significant increase in aggressive cognition at Time 2.

Also as expected, partialing out sex and Time 1 aggressive cognition effects reduced the average VGV effect size in nonexperimental studies. In the best partials data, experimental studies yielded the largest effects, whereas longitudinal studies yielded the smallest effects. Once again, ignoring research design led to very large heterogeneity effects. Finally, when research design, sex, and Time 1 effects were controlled, none of the heterogeneity effects were significant. In sum, this much larger meta-analysis, with over 50 independent effects involving over 12,000 participants from multiple countries, ages, and cultures, yielded strong evidence that playing violent video games increases aggressive cognition in both short- and long-term contexts.

Additional moderator tests—best partials data.

Culture. Culture (Eastern vs. Western) was not a significant moderator in either the experimental or the cross-sectional studies. However, in longitudinal studies the VGV effect was significantly larger in Western ($r+ = .137$, $K = 3$, $N = 710$) than in Eastern studies ($r+ = .038$, $K = 5$, $N = 2{,}602$), $Q(1) = 5.50$, $p < .02$.

Sex. There were insufficient data to test the sex moderator effect in experimental and longitudinal studies. In cross-sectional studies, the VGV effect was slightly larger for females but not significantly so.

Age. There were too few longitudinal studies to test the moderating effect of age. Age was not a significant moderator in either experimental or cross-sectional studies.

Moderators specific to experiments. None of the moderators specific to experiments (i.e., player perspective, player role, target type, time on game, type of aggressive cognition measure) approached significance.

Moderators specific to nonexperimental studies. For cross-sectional studies neither of the additional moderators approached significance (VGV exposure measure, type of aggressive cognition measure). However, for longitudinal studies, the VGV measure of video game exposure moderator was marginally significant, $Q(1) = 3.52$, $p < .07$. Studies that used VGV-specific measures ($r+ = .113$, $K = 4$, $N = 891$) yielded larger VGV effects than did those that used VGV-general type measures ($r+ = .040$, $K = 4$, $N = 2{,}421$). Also of importance was the finding that type of aggressive cognition measure (trait hostility vs. other) was completely confounded with culture and therefore yielded exactly the same moderation effect as reported earlier for culture. The five longitudinal studies from Japan all used a trait hostility measure, whereas the three studies from Western cultures used hostile

Table 5
Aggressive Cognition: Average Effect of Violent Video Game Exposure by Study Design for Best Raw, Best Partials, and Full Sample Data Sets

Design	*N*	*K*	Effect size and 95% CI: Point estimate	*LL*	*UL*	Test of null (two-tailed): *z*	*p*	Heterogeneity: *Q*	*df*(*Q*)	*p*
Best raw										
Experimental	2,887	24	.217	0.181	0.252	11.695	.000	35.11	23	.051
Longitudinal	3,408	8	.115	0.082	0.148	6.728	.000	13.08	7	.070
Cross-sectional	9,976	27	.183	0.164	0.202	18.445	.000	185.56	26	.000
Total within								233.75	56	.000
Total between								18.74	2	.000
Overall	16,271	59	.175	0.160	0.190	22.440	.000	252.49	58	.000
Best partials										
Experimental	2,887	24	.217	0.181	0.252	11.695	.000	35.11	23	.051
Longitudinal	3,312	8	.059	0.025	0.093	3.400	.001	7.81	7	.349
Cross-sectional	6,399	21	.114	0.090	0.138	9.128	.000	19.84	20	.468
Total within								62.76	50	.106
Total between								40.49	2	.000
Overall	12,598	53	.123	0.106	0.141	13.826	.000	103.25	52	.000
Full sample										
Experimental	4,289.5	48	.207	0.177	0.236	13.496	.000	90.00	47	.000
Longitudinal	4,178.5	9	.110	0.080	0.140	7.142	.000	13.50	8	.096
Cross-sectional	16,066	38	.164	0.149	0.179	20.951	.000	269.06	37	.000
Total within								372.56	92	.000
Total between								20.41	2	.000
Overall	24,534	95	.162	0.150	0.175	25.528	.000	392.98	94	.000

Note. Effect sizes measured as *r*. CI = confidence interval; *LL* = lower limit; *UL* = upper limit.

attribution bias or attitude/belief measures. Thus, it is impossible to know whether this moderation effect results from culture differences in the VGV longitudinal effect or from measurement instrument differences.

Summary. Exposure to violent video games was significantly related to higher levels of aggressive cognition, regardless of research design or conservativeness of analysis. For experimental studies, $r+ = .217$. For cross-sectional studies the best raw and best partials analyses yielded average effect sizes of $r+ = .183$ and .114, respectively. For longitudinal studies, the best raw and best partials analyses yielded average effect sizes of $r+ = .115$ and .059, respectively. Culture significantly moderated the longitudinal VGV effect, but the VGV effect was significantly greater than zero in both cases. Furthermore, in this small set of longitudinal studies culture was perfectly confounded with type of cognition measure. In addition, studies that used a VGV-specific measure yielded larger effects than those that used a VGV-general measure. Three of the four VGV-specific studies and none of the VGV-general studies were from the West. Therefore, it is unclear whether the smaller longitudinal effect in studies from Japan was the result of culture or of either or both of two measurement instrument differences. Additional research could easily resolve this.

As with the aggressive behavior results, perhaps the most important finding relative to prior meta-analyses is the significant longitudinal effect of VGV on aggressive cognition. In combination with the experimental and the cross-sectional findings, the data provide strong evidence that playing violent video games is a significant causal risk factor for both short- and long-term increases in aggressive thinking.

Violent Video Game Effects on Aggressive Affect

Main analyses. Table 6 presents the results of the main analyses on aggressive affect. As with aggressive behavior and aggressive cognition, VGV exposure was significantly related to higher levels of aggressive affect regardless of research design and regardless of whether zero-order correlations or the more conservative partial correlations were used. Furthermore, even when sex and Time 1 aggressive affect were controlled, amount of violent video game play at Time 1 predicted a significant increase in aggressive affect at Time 2.

As with aggressive behavior and aggressive cognition, research design was a significant moderator of the VGV effect on aggressive affect. Experimental studies yielded the largest effects and longitudinal studies the smallest.

Additional moderator tests—best partials data. There were too few longitudinal studies for us to do any additional moderator analyses. Furthermore, none of the moderator variables yielded a significant effect in experimental or cross-sectional studies, even though there was evidence of significant heterogeneity within experimental studies.

Violent Video Game Effects on Prosocial Behavior

Main analyses. Table 7 presents the main results on prosocial behavior. VGV exposure was significantly related to lower levels

Table 6
Aggressive Affect: Average Effect of Violent Video Game Exposure by Study Design for Best Raw, Best Partials, and Full Sample Data Sets

			Effect size and 95% CI			Test of null (two-tailed)		Heterogeneity		
Design	*N*	*K*	Point estimate	*LL*	*UL*	*z*	*p*	*Q*	*df*(*Q*)	*p*
				Best raw						
Experimental	1,454	21	.294	0.245	0.341	11.289	.000	49.15	20	.000
Longitudinal	2,602	5	.075	0.037	0.113	3.836	.000	13.19	4	.010
Cross-sectional	5,135	11	.101	0.074	0.128	7.227	.000	16.56	10	.085
Total within								78.91	34	.000
Total between								53.18	2	.000
Overall	9,191	37	.124	0.104	0.144	11.883	.000	132.08	36	.000
				Best partials						
Experimental	1,454	21	.294	0.245	0.341	11.289	.000	49.15	20	.000
Longitudinal	2,602	5	.039	0.0001	0.077	1.967	.049	9.76	4	.045
Cross-sectional	3,487	9	.110	0.077	0.143	6.509	.000	9.78	8	.281
Total within								68.70	32	.000
Total between								63.83	2	.000
Overall	7,543	35	.121	0.098	0.143	10.481	.000	132.53	34	.000
				Full sample						
Experimental	3,015	37	.181	0.146	0.216	9.863	.000	111.22	36	.000
Longitudinal	3,373	6	.082	0.048	0.116	4.768	.000	13.73	5	.017
Cross-sectional	10,982	19	.145	0.126	0.163	15.215	.000	153.17	18	.000
Total within								278.12	59	.000
Total between								16.87	2	.000
Overall	17,370	62	.139	0.124	0.153	18.293	.000	294.99	61	.000

Note. Effect sizes measured as *r*. CI = confidence interval; *LL* = lower limit; *UL* = upper limit.

Table 7
Prosocial Behavior: Average Effect of Violent Video Game Exposure by Study Design for Best Raw, Best Partials, and Full Sample Data Sets

			Effect size and 95% CI			Test of null (two-tailed)		Heterogeneity		
Design	*N*	*K*	Point estimate	*LL*	*UL*	*z*	*p*	*Q*	*df*(*Q*)	*p*
				Best raw						
Experimental	633	4	−.182	−0.257	−0.106	−4.599	.000	3.79	3	.285
Longitudinal	2,778	5	−.114	−0.151	−0.077	−6.022	.000	15.91	4	.003
Cross-sectional	3,495	7	−.093	−0.126	−0.060	−5.506	.000	19.07	6	.004
Total within								38.77	13	.000
Total between								4.46	2	.107
Overall	6,906	16	−.110	−0.133	−0.086	−9.125	.000	43.23	15	.000
				Best partials						
Experimental	633	4	−.182	−0.257	−0.106	−4.599	.000	3.79	3	.285
Longitudinal	2,777	5	−.062	−0.099	−0.025	−3.268	.001	2.32	4	.677
Cross-sectional	3,495	7	−.094	−0.127	−0.061	−5.544	.000	19.25	6	.004
Total within								25.36	13	.021
Total between								7.74	2	.021
Overall	6,905	16	−.089	−0.113	−0.066	−7.404	.000	33.10	15	.004
				Full sample						
Experimental	875	8	−.161	−0.226	−0.095	−4.748	.000	5.26	7	.629
Longitudinal	2,778	5	−.114	−0.151	−0.077	−6.022	.000	15.91	4	.003
Cross-sectional	5,992	10	−.086	−0.111	−0.061	−6.659	.000	29.29	9	.001
Total within								50.46	20	.000
Total between								5.05	2	.080
Overall	9,645	23	−.101	−0.121	−0.081	−9.904	.000	55.51	22	.000

Note. Effect sizes measured as *r*. CI = confidence interval; *LL* = lower limit; *UL* = upper limit.

of prosocial behavior regardless of research design and regardless of whether zero-order or the partial correlations were used. Even when sex and Time 1 prosocial behavior were controlled, amount of violent video game play at Time 1 predicted a significant decrease in prosocial behavior at Time 2 in longitudinal studies.

Yet again, research design was a significant moderator of the VGV effect on prosocial behavior, with experimental studies yielding the largest (negative) effects and longitudinal studies the smallest. Once sex and Time 1 effects were partialed out, there was no evidence of heterogeneity in the experimental or longitudinal effects, but there was in the cross-sectional studies.

Additional moderator tests—best partials data. There were too few experimental and longitudinal studies to do any additional moderator analyses on them. For cross-sectional studies we were able to test for culture, sex, and VGV type of measure effects. On average, the VGV effect on prosocial behavior was larger in the Western studies ($r+ = -.225$, $K = 2$, $N = 347$) than in Eastern studies ($r+ = -.079$, $K = 5$, $N = 3{,}148$), $Q(1) = 6.83$, $p < .01$. There also was a significant effect of violent video game exposure measure, $Q(1) = 13.69$, $p < .001$. Studies that used the VGV-specific type of measure yielded larger (negative) effects ($r+ = -.186$, $K = 3$, $N = 1{,}074$) than those that used general measures ($r+ = -.052$, $K = 4$, $N = 2421$). Furthermore, culture and VGV measure were confounded; two of the three VGV-specific studies came from a Western culture, whereas all of the VGV-general studies came from Eastern cultures. Finally, there was no evidence of sex differences in the effect of violent games on prosocial behavior.

Violent Video Game Effects on Empathy/Desensitization

Main analyses. Table 8 presents the main results on empathy/desensitization. VGV exposure was significantly related to less empathy (and more desensitization) regardless of research design and regardless of whether the zero-order or partial correlations were used.

When sex and Time 1 effects were controlled, research design was a significant moderator variable. Of course, because there was only one experimental study, comparisons across designs should be made with caution.

Additional moderator tests—best partials data. There were too few experimental and longitudinal studies to do any additional moderator analyses. For cross-sectional studies, we were able to test the moderating effects of culture and video game exposure measure. On average, effect sizes were larger in Western studies ($r+ = -.294$, $K = 4$, $N = 450$) than in Eastern studies ($r+ = -.144$, $K = 5$, $N = 3{,}148$), $Q(1) = 9.53$, $p < .01$. There also was a significant effect for video game exposure measure, $Q(1) = 4.36$, $p < .05$. Studies that used the VGV-specific type of measure yielded larger (negative) effects ($r+ = -.211$, $K = 5$, $N = 1{,}177$) than did those using the VGV-general measure ($r+ = -.139$, $K = 4$, $N = 2{,}421$). Furthermore, culture and video game exposure measure were confounded; four of the five VGV-specific studies came from a Western culture, whereas all of the VGV-general studies came from Eastern cultures.

Table 8
Empathy/Desensitization: Average Effect of Violent Video Game Exposure by Study Design for Best Raw, Best Partials, and Full Sample Data Sets

			Effect size and 95% CI			Test of null (two-tailed)		Heterogeneity		
Design	*N*	*K*	Point estimate	*LL*	*UL*	*z*	*p*	*Q*	*df*(*Q*)	*p*
				Best raw						
Experimental	249	1	−0.138	−0.258	−0.014	−2.175	.030	0.00	0	1.000
Longitudinal	2,421	4	−0.184	−0.223	−0.145	−9.147	.000	24.88	3	0.000
Cross-sectional	3,910	10	−0.203	−0.233	−0.173	−12.845	.000	24.46	9	0.004
Total within								49.34	12	0.000
Total between								1.44	2	0.488
Overall	6,580	15	−0.194	−0.217	−0.170	−15.873	.000	50.78	14	0.000
				Best partials						
Experimental	249	1	−0.138	−0.258	−0.014	−2.175	.030	0.00	0	1.000
Longitudinal	2,421	4	−0.070	−0.109	−0.030	−3.427	.001	12.82	3	0.005
Cross-sectional	3,598	9	−0.163	−0.195	−0.131	−9.817	.000	23.10	8	0.003
Total within								35.92	11	0.000
Total between								12.87	2	0.002
Overall	6,268	14	−0.126	−0.150	−0.102	−9.999	.000	48.79	13	0.000
				Full sample						
Experimental	537	11	−0.148	−0.232	−0.062	−3.351	.001	24.20	10	0.007
Longitudinal	2,796	6	−0.160	−0.196	−0.123	−8.501	.000	37.23	5	0.000
Cross-sectional	5,195	15	−0.188	−0.214	−0.162	−13.671	.000	55.96	14	0.000
Total within								117.39	29	0.000
Total between								2.00	2	0.369
Overall	8,528	32	−0.177	−0.197	−0.156	−16.383	.000	119.38	31	0.000

Note. Effect sizes measured as *r*. CI = confidence interval; *LL* = lower limit; *UL* = upper limit.

Violent Video Game Effects on Physiological Arousal

Main analyses. All of the physiological arousal studies were experiments. Overall, playing a violent video game increased physiological arousal ($r+ = .184$, $p < .001$, $K = 15$, $N = 969$). The heterogeneity test was significant, $Q(14) = 30.43$, $p < .01$. We were able to conduct moderator analyses for player perspective, player role, game violence target, average age, and game-playing time. However, none of the moderator tests approached statistical significance.

Sensitivity Analyses

Full sample. Recall that numerous studies did not meet the inclusion criteria listed in Table 2. What happens if these methodologically weak studies are included in the main analyses? Are there systematic differences in average effect size? Table 9 displays the results. Three main points emerge. First, for each of the six outcome variables, the violent video game effect size was still significant even when methodologically weaker studies were included (*p*s $< .001$). Second, for five of the outcome variables the methodologically weak studies yielded smaller effect sizes than did the methodologically strong studies; in four cases the difference was statistically significant. The exception was aggressive affect, for which the methodologically weak studies yielded a significantly larger average effect size. Third, for each outcome variable even the methodologically weak studies yielded a significant overall effect.

Trim and fill analyses. Table 10 presents the results of the trim and fill analyses, as a further check on possible selection/publication bias. For each outcome variable, we first applied the trim and fill procedure to the full sample (which included both methodologically weak and strong studies) and the best raw sample, ignoring research design.

Next, because research design was such a strong and consistent moderator variable for the best partials samples and because these samples are the main focus of this article, we also applied the trim and fill procedure to the best partials data by research design. Recall that sample size, research design, culture, and specific research instruments are somewhat confounded in these studies. Therefore, within each outcome and research design we broke down the studies into smaller subgroups using the following decision rule: If there was a significant moderator effect, separate trim and fill analyses were done on the different levels of that moderator; otherwise, separate analyses were done by culture. Of course, in several cases there were too few studies for these breakdowns, in which case the trim and fill procedure was applied at a higher level of studies. For example, all of the best practices physiological studies used an experimental design, and only one of these studies was from an Eastern culture.

The summarized results at the bottom of Table 10 suggest that if there has been selection or publication bias in favor of theoretical hypotheses in these samples, the bias has been weak and has had relatively little impact on average effect-size estimates. For example, experimental studies were weakened by .017 by the trim and fill imputation procedure. The cross-sectional studies were weakened by an even smaller amount, .005. Conversely, the longitudinal studies were strengthened by an average of .008. In sum,

Table 9
Effects of Methodologically Weak Versus Strong Studies

Methodological quality	*N*	*K*	Effect size and 95% CI: Point estimate	*LL*	*UL*	Test of null (two-tailed): *z*	*p*	Heterogeneity: *Q*	*df*(*Q*)	*p*
Aggressive behavior										
Weak	46,632	61	0.163	0.154	0.172	35.506	.000	504.84	60	.000
Strong	21,681	79	0.244	0.231	0.256	36.422	.000	284.86	78	.000
Total within								789.70	138	.000
Total between								103.46	1	.000
Overall	68,313	140	0.189	0.182	0.196	49.838	.000	893.16	139	.000
Aggressive cognition										
Weak	8,263	36	0.138	0.116	0.159	12.497	.000	132.46	35	.000
Strong	16,271	59	0.175	0.160	0.190	22.440	.000	252.49	58	.000
Total within								384.95	93	.000
Total between								8.02	1	.005
Overall	24,534	95	0.162	0.150	0.175	25.528	.000	392.98	94	.000
Aggressive affect										
Weak	8,179	25	0.155	0.134	0.176	14.060	.000	158.67	24	.000
Strong	9,191	37	0.124	0.104	0.144	11.883	.000	132.08	36	.000
Total within								290.75	60	.000
Total between								4.24	1	.039
Overall	17,370	62	0.139	0.124	0.153	18.293	.000	294.99	61	.000
Prosocial behavior										
Weak	2,739	7	−0.078	−0.116	−0.041	−4.095	.000	10.33	6	.111
Strong	6,906	16	−0.110	−0.133	−0.086	−9.125	.000	43.23	15	.000
Total within								53.56	21	.000
Total between								1.95	1	.163
Overall	9,645	23	−0.101	−0.121	−0.081	−9.904	.000	55.51	22	.000
Empathy/desensitization										
Weak	1,948	17	−0.116	−0.160	−0.071	−5.078	.000	59.26	16	.000
Strong	6,580	15	−0.194	−0.217	−0.170	−15.873	.000	50.78	14	.000
Total within								110.03	30	.000
Total between								9.35	1	.002
Overall	8,528	32	−0.177	−0.197	−0.156	−16.383	.000	119.38	31	.000
Physiological arousal										
Weak	937	14	0.085	0.020	0.150	2.552	.011	10.47	13	.655
Strong	969	15	0.184	0.121	0.245	5.647	.000	30.43	14	.007
Total within								40.89	27	.042
Total between								4.59	1	.032
Overall	1,906	29	0.135	0.090	0.180	5.814	.000	45.48	28	.020

Note. Effect sizes measured as *r*. CI = confidence interval; *LL* = lower limit; *UL* = upper limit.

there is no evidence that publication or selection bias had an important influence on the results.

Discussion

Main Findings

Although the meta-analyses in this article revealed numerous findings about the short- and long-term effects of playing violent video games, six findings are particularly important. First, social–cognitive models and other theoretical considerations predicted the broad pattern of results quite well. As expected, VGV exposure was positively associated with aggressive behavior, aggressive cognition, and aggressive affect. These effects were statistically reliable in experimental, cross-sectional, and longitudinal studies, even when unusually conservative statistical procedures were used. Also as expected, VGV exposure was related to desensitization and lack of empathy and to lack of prosocial behavior. Furthermore, the relative magnitudes of effects for different outcome variables and moderator variables were mainly consistent with theory. For example, the longitudinal effect of VGV was somewhat smaller on aggressive affect than on aggressive cognition or behavior. This suggests that the processes underlying media violence effects are well understood (cf. Anderson et al., 2003).

Second, the VGV effects are significant in Eastern as well as Western cultures. There are hints that VGV effects may be larger in Western than Eastern cultures, but these occur only in nonex-

Table 10
Results of the Trim and Fill Selection/Publication Bias Analysis

Sample	*K*	Imputed studies *N*	Imputed studies Direction	Obs. *r*+	Adj. *r*+	Change	Strength change
Aggressive behavior							
Full	140	10	Right	0.189	0.192	0.003	0.003
Best raw	79	1	Right	0.244	0.244	0.000	0.000
Best partials: Experimental	27						
East	5	0		0.245	0.245	0.000	0.000
West	22	8	Left	0.207	0.178	−0.029	−0.029
Best partials: Cross-sectional	36						
East	8	1	Left	0.163	0.160	−0.003	−0.003
West	28	3	Left	0.182	0.176	−0.006	−0.006
Best partials: Longitudinal	12						
VGV-specific	4	1	Left	0.152	0.143	−0.009	−0.009
VGV-general	8	1	Left	0.055	0.054	−0.001	−0.001
Aggressive cognition							
Full	95	12	Right	0.163	0.170	0.008	0.008
Best raw	59	0		0.175	0.175	0.000	0.000
Best partials: Experimental	24	5	Left	0.217	0.199	−0.018	−0.018
Best partials: Cross-sectional	21						
East	6	1	Left	0.102	0.100	−0.002	−0.002
West	15	4	Left	0.127	0.106	−0.021	−0.021
Best partials: Longitudinal	8						
East	5	0		0.038	0.038	0.000	0.000
West	3	2	Right	0.137	0.182	0.045	0.045
Aggressive affect							
Full	62	17	Left	0.139	0.100	−0.039	−0.039
Best raw	37	15	Left	0.124	0.102	−0.022	−0.022
Best partials: Experimental	21	0		0.294	0.294	0.000	0.000
Best partials: Cross-sectional	9						
East	5	2	Left	0.097	0.062	−0.035	−0.035
West	4	2	Left	0.150	0.137	−0.013	−0.013
Best partials: Longitudinal	5	1	Right	0.039	0.049	0.011	0.011
Prosocial behavior							
Full	23	10	Right	−0.101	−0.064	0.037	−0.037
Best raw	16	4	Right	−0.110	−0.089	0.021	−0.021
Best partials: Experimental	4	2	Right	−0.182	−0.125	0.057	−0.057
Best partials: Cross-sectional	7						
VGV-specific	3	2	Right	−0.186	−0.168	0.018	−0.018
VGV-general	4	2	Right	−0.052	−0.033	0.019	−0.019
Best partials: Longitudinal	5	2	Left	−0.062	−0.070	−0.008	0.008
Empathy/desensitization							
Full	32	2	Left	−0.177	−0.179	−0.002	0.002
Best raw	15	0		−0.194	−0.194	0.000	0.000
Best partials: Cross-sectional	9						
East	5	2	Left	−0.144	−0.167	−0.023	0.023
West	4	0		−0.294	−0.294	0.000	0.000
VGV-specific	5	0		−0.211	−0.211	0.000	0.000
VGV-general	4	2	Left	−0.139	−0.171	−0.032	0.032
Best partials: Longitudinal	4	0		−0.070	−0.070	0.000	0.000
Physiological arousal							
Full	29	0		0.135	0.135	0.000	0.000
Best raw and partials	15	0		0.184	0.184	0.000	0.000
Overall average							−0.006
Full average							−0.011
Best raw average							−0.007

(*table continues*)

Table 10 (*continued*)

Sample	*K*	Imputed studies *N*	Imputed studies Direction	Obs. *r*+	Adj. *r*+	Change	Strength change
Best partials: Experimental average							−0.017
Best partials: Cross-sectional average							−0.017
Best partials: Longitudinal average							0.008

Note. Strength change is the difference between the observed and adjusted, taking into account the hypothesized effect direction. VGV = video game violence; Obs. *r*+ = observed average effect size; Adj. *r*+ = Adjusted average effect size.

perimental studies. Indeed, in experimental studies all three outcome variables for which there were sufficient studies (aggressive behavior, aggressive cognition, and aggressive affect) yielded slightly larger effects in Eastern studies, though none approached statistical significance. In the few nonexperimental cases in which culture yielded a significant moderator effect, it was unclear whether the difference should be attributed to cultural differences in vulnerability or to the use of different measures.

Third, there is evidence that the magnitude of the effect size obtained in a study is influenced by the types of measures one uses. How one measures exposure to violent video games in nonexperimental contexts influences the magnitude of effects. Measurement procedures, similar to those used by Anderson and Dill (2000), that elicit violence and time ratings for each specific game played by the participant appear to yield larger effect sizes than do other methods. However, to some extent this finding is confounded with culture. What is needed to clarify this issue are studies in which multiple methods of assessing violent video game exposure are used in the same sample of participants, so that they can be directly compared. Similarly, it appears that trait hostility measures may not be the best way to assess aggressive cognition.

Fourth, and perhaps most important, the newly available longitudinal studies provide further confirmation that playing violent video games is a causal risk factor for long-term harmful outcomes. This is especially clear for aggressive behavior, aggressive cognition, and empathy/desensitization. But significant longitudinal VGV effects also were found for aggressive affect and for prosocial behavior.

Fifth, the failures of several specific moderators to yield significant effects are also important for theoretical and practical reasons. In experimental studies, the lack of player perspective (first or third person), player role (hero, criminal), time on game, and target (human, nonhuman) all suggest that the short-term effects of violent video games on aggression are largely the result of priming processes (see Bushman & Huesmann, 2006). The fact that the CRT task yielded effect sizes similar to those from other measures of aggressive behavior in experimental studies strongly contradicts recent claims that it is only CRT studies that find such effects; indeed, the average CRT effect was slightly (nonsignificantly) smaller than for the other measures. The fact that studies using violent behavior measures did not yield significantly smaller effects than comparable studies using less extreme aggression measures, and that the violent behavior studies by themselves yielded a significant VGV effect, is important. The facts that sex did not significantly moderate the findings and that age yielded only one marginally significant moderation effect suggest that large portions of the population (at least through college age) are susceptible to the harmful effects of violent video game play.

Sixth, our results confirm that partialing out sex effects and (in longitudinal studies) Time 1 outcome variable effects reduces average effect size. This is not surprising, of course, but it does lead to interesting questions about what is the "best" estimate of the true effect sizes. Partialing out sex effects certainly is a safe, conservative procedure in that it guarantees that the estimated average effect size will not be upwardly biased. However, if violent video game exposure truly plays a causal role in aggression and in the other outcomes assessed in these meta-analyses (and our conservative procedures do demonstrate this), partialing out sex effects yields underestimates of the true effect size. Perhaps the best solution is to view sex-controlled estimates as the lower boundary of the true effects and the sex-ignored effects as the upper boundary.[11]

Two additional findings warrant highlighting, especially in view of all the attention that has paid recently to claims of publication bias and other weaknesses in the violent video game literature. First, there was no evidence that publication (or study selection bias) is responsible for the observed relations between exposure to violent video games and aggressive behavior or the other five outcome variables. Differences between the observed average effect sizes and the corresponding trim and fill adjusted averages were small in magnitude and did not substantially change any overall effects or conclusions. Second, our other sensitivity analyses (see Table 9) showed that even including methodologically weak studies (studies that would not be analyzed in noncontroversial fields of study) yields relatively small changes in average effect sizes. Even the weak studies yielded significant overall effects on all six outcome variables. In sum, the only way one can "demonstrate" that the existing literature on violent video game effects does not show multiple causal harmful effects is to use an incredibly small subset of the existing literature, include some of the methodologically poorest studies, exclude many of the methodologically strongest studies, and misuse standard meta-analytic techniques.

Finally, it is important to note that our main analyses were conducted on the most conservative sample of studies, in which all explained variance that is confounded with sex or Time 1 outcome measures is statistically removed from the estimates of the VGV effects. Thus, the finding of significant VGV effects on all outcome measures and in all research designs speaks to the power of video games.

[11] Of course, there also are measurement reliability issues that will tend to make empirical estimates smaller than true effects.

Magnitude of Average Effect Sizes

We have no doubt that the import of the average effect sizes reported in this article will be greatly debated. There are at least two major issues, one concerning which estimates are most appropriate and the other concerning the importance of the average effect sizes.

Consider the results for physically aggressive behavior, shown in Figure 1. Which are the most appropriate estimates of the true violent video game effect, the raw estimates for cross-sectional and longitudinal studies or those that partialed out sex (and Time 1) effects? We believe that there is no single correct answer. On the one hand, the two facts that males spend more time than females playing violent video games and that males generally (but not always) are more physically aggressive than females suggest that an appropriately conservative test of violent video game effects should partial out sex effects. On the other hand, once it has been established that violent video game play is a causal risk factor for physical aggression in both males and females, it becomes clear that partialing out sex effects yields underestimates of the effect of violent video games.

Concerning the importance issue, are the effect sizes large enough to be considered important? From a theoretical standpoint, the answer is pretty clear. Prevailing social–cognitive theories predict statistically significant effects but do not predict absolute magnitude; finding the predicted effects to be significant therefore lends support to the theories.

From a practical or applied standpoint, the answer is less clear. From a strict "percentage of variance explained" perspective, the longitudinal effect size with sex and Time 1 aggression partialed out might seem small (i.e., $0.152^2 = 2.31\%$). However, as numerous authors have pointed out, even small effect sizes can be of major practical significance. When effects accumulate across time, or when large portions of the population are exposed to the risk factor, or when consequences are severe, statistically small effects become much more important (Abelson, 1985; Rosenthal, 1986, 1990). All three of these conditions apply to violent video game effects.

Furthermore, when dealing with a multicausal phenomenon such as aggression, one should not expect any single factor to explain much of the variance. There are dozens of known risk factors for both short-term aggression and the development of aggression-prone individuals. To expect any one factor to account for more than a small fraction of variance is unrealistic. This suggests that a better way to assess the practical importance of violent video game effects is to compare it to some other known risk factors for youth violence. In fact, even the overly conservative sex and Time 1 adjusted VGV effect size estimate ($r+ = .152$) compares favorably to such risk factors as substance use, abusive parents, and poverty (U.S. Department of Health and Human Services, 2001).[12]

Learning From the Past

Learning from our mistakes. Additional findings of importance come not from the meta-analytic results but from the review and selection process involved in identifying studies of sufficient quality to merit inclusion in the focal analyses. There are a lot of methodological pitfalls that the wary researcher must avoid in this domain, there are a lot of researchers who are not avoiding them, and there are a lot of editors who are publishing them. The quality criteria described in Table 2 and the more specific examples outlined on our website could (and should) prove useful to many people who are involved in media violence research, including faculty supervisors of undergraduate and graduate students, journal editors and their manuscript reviewers, and some media violence scholars themselves.

We do not intend to imply that those studies that did not meet all of our quality criteria are useless; some of them yielded findings of some value on other research questions. Similarly, we do not intend to imply that those studies that met our best practices criteria are without shortcomings. There is no such thing as a perfect study. Several of the most frequent weaknesses of studies that did pass the present quality check are use of single-item outcome measures; use of longitudinal time periods that may be too short for the phenomenon of interest; use of weak measures of exposure to violent video games; and sample sizes that are too small given the expected effect size. The first three tend to lead to underestimates of the true effect size, whereas the third tends to yield nonsignificant statistical test results.

Contrary to claims by video game industry representatives, some gamers, and a few researchers, in general it is not the methodologically poor studies that tend to yield big effects. Rather, methodologically superior studies tend to yield larger effects.

Gaps and future research. This review also revealed a number of gaps in the research literature. One involves the lack of age effects, especially in longitudinal studies. Most scholars in this area believe that younger children are likely to be more vulnerable than older adolescents and young adults to long-term media violence effects. But there is little evidence of larger effect sizes for younger than for older participants. Indeed, a recent television study suggests that even late adolescents are vulnerable (Johnson, Cohen, Smailes, Kasen, & Brook, 2002).

Our cross-sectional analysis yielded a marginally significant negative relationship between age and the effect of violent games on aggressive behavior, but the longitudinal studies did not find an age effect. Part of the problem may be that meta-analytic procedures are not ideally suited to testing such effects. Different studies typically use different measures of aggression and video game habits, include different participant populations, and take place during different years. All of the differences add noise to the meta-analysis, noise that may well hide real age effects. New research is needed in which different-age participants are sampled from the same population, the same measures are used (or as similar as is reasonable given the age differences), and all of the data collection takes place in the same years. Such studies should yield cleaner tests of whether or not long-term effects of video game violence are larger for younger than older participants.

[12] It is unclear from this Surgeon General report whether all of the reported longitudinal effects were similarly based on sex and Time 1 adjusted partial correlations. Also note that one criterion for inclusion in the Surgeon General report was violent behavior. The present longitudinal studies of the VGV/physical aggression include some violent behavior but also include somewhat less severe forms of physical aggression. Thus, the effects are not strictly comparable.

Similarly, we believe that many moderator variables of interest can be more precisely examined within the same study than they can in meta-analyses. For example, as a test for short-term player perspective effects, it would be best to conduct experiments in which participants are randomly assigned to play the same violent game but from different perspectives. Similar studies are needed to more closely examine player role effects, target effects, and time effects.

In the aggression domain, pre–post and repeated measures designs are generally problematic because of suspicion, practice, and carryover effects. Longitudinal designs solve these problems to some extent, but such designs do not allow experimental tests of the immediate, short-term consequences of playing violent video games. Some types of dependent variables are less susceptible to these problems, such as reaction time tasks that presumably assess effects without the participant's awareness (e.g., lexical decision, reading reaction time, and implicit association tasks). Cognitive neuroscience approaches also are likely to prove especially valuable in more precisely assessing the immediate effects of playing violent and nonviolent video games on a host of cognitive and affective processes (e.g., Bailey, West, & Anderson, in press; Weber, Ritterfeld, & Mathiak, 2006). For example, previous research has mapped specific brain areas that are especially active during violent and nonviolent actions within the same game, using functional magnetic resonance imaging techniques (Weber et al., 2006).

Conclusions

The present findings show that the social–cognitive theoretical view fits the existing data on video game violence effects quite well. This has important implications for public policy debates, for further development and testing of basic theory, and for development and testing of potential intervention strategies designed to reduce harmful effects of playing violent video games. Concerning basic theory, additional research of all three types (but especially experimental and longitudinal) is needed, especially on VGV effects on empathy, desensitization, and prosocial behavior. Additional longitudinal studies with longer intervals are needed for aggressive behavior and aggressive cognition. Furthermore, longitudinal studies with very large samples and very long time spans between the first time period and the last are needed so we can assess the impact of violent video games on very serious forms of physical aggression (i.e., violence). Concerning interventions, there have been a few studies with findings that suggest that specific programs involving schoolchildren and their parents can reduce exposure to violent media and the frequency of unwarranted aggressive behavior (e.g., Huesmann, Eron, Klein, Brice, & Fischer, 1983; Robinson, Wilde, Navracruz, Haydel, & Varady, 2001).

Concerning public policy, we believe that debates can and should finally move beyond the simple question of whether violent video game play is a causal risk factor for aggressive behavior; the scientific literature has effectively and clearly shown the answer to be "yes." Instead, we believe the public policy debate should move to questions concerning how best to deal with this risk factor. Public education about this risk factor—and about how parents, schools, and society at large can deal with it—could be very useful.

It is true that as a player you are "not just moving your hand on a joystick" but are indeed interacting "with the game psychologically and emotionally." It is not surprising that when the game involves rehearsing aggressive and violent thoughts and actions, such deep game involvement results in antisocial effects on the player. Of course, the same basic social–cognitive processes should also yield prosocial effects when game content is primarily prosocial. Unfortunately, there has been relatively little research on purely prosocial game effects, largely because there are few games that have the main characters modeling helpful behavior in the complete absence of violent behavior. However, some recent studies have found that prosocial games can increase cooperation and helping (Gentile et al., 2009; Greitemeyer & Osswald, 2009). Video games are neither inherently good nor inherently bad. But people learn. And content matters.

References

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

Anderson, C. A. (2004). An update on the effects of violent video games. *Journal of Adolescence, 27,* 113–122.

Anderson, C. A., Berkowitz, L., Donnerstein, E., Huesmann, L. R., Johnson, J., Linz, D., . . . 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. (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. (2002). Human aggression. *Annual Review of Psychology, 53,* 27–51.

Anderson, C. A., & Carnagey, N. L. (2004). Violent evil and the general aggression model. In A. Miller (Ed.), *The social psychology of good and evil* (pp. 168–192). New York, NY: Guilford.

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., & 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, NY: Oxford University Press.

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

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.

Bachrach, R. (1986). The differential effect of observation of violence on kibbutz and city children in Israel. In L. R. Huesmann & L. D. Eron (Eds.), *Television and the aggressive child: A cross-national comparison* (pp. 210–238). Hillsdale, NJ: Erlbaum.

Bailey, K., West, R., & Anderson, C. A. (in press). The influence of video games on social, cognitive, and affective information processing. In J. Decety & J. Cacioppo (Eds.), *Handbook of social neuroscience.* New York, NY: Oxford University Press.

Bandura, A. (1973). *Aggression: A social learning theory analysis.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996).

Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology, 71,* 364–374.

Barclay, G., & Tavares, C. (2002). *International comparisons of criminal justice statistics 2000.* Retrieved from United Kingdom Home Office website: http://www.homeoffice.gov.uk/rds/pdfs2/hosb502.pdf

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,* 532–539.

Bergeron, N., & Schneider, B. H. (2005). Explaining cross-national differences in peer-directed aggression: A quantitative synthesis. *Aggressive Behavior, 31,* 116–137.

Berkowitz, L. (1984). Some effects of thoughts on anti- and prosocial influences of media events: A cognitive–neoassociation analysis. *Psychological Bulletin, 95,* 410–427.

Berkowitz, L. (1993). *Aggression: Its causes, consequences, and control.* New York, NY: McGraw-Hill.

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

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

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

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., & 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.

Chida, Y., & Hamer, M. (2008). Chronic psychosocial factors and acute physiological responses to laboratory-induced stress in healthy populations: A quantitative review of 30 years of investigations. *Psychological Bulletin, 134,* 829–885.

Colwell, J., & Kato, M. (2003). Investigation of the relationship between social isolation, self-esteem, aggression and computer game play in Japanese adolescents. *Asian Journal of Social Psychology, 6,* 149–158.

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

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.

DeLisi, M. (2005). *Career criminals in society.* Thousand Oaks, CA: Sage.

Duval, S. (2005). The "Trim and Fill" method. In H. Rothstein, A. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 127–144). Chichester, England: Wiley.

Duval, S. J., & Tweedie, R. L. (1998). Practical estimates of the effect of publication bias in meta-analysis. *Australasian Epidemiologist, 5*(4), 14–17.

Duval, S., & Tweedie, R. (2000a). A nonparametric "Trim and Fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95,* 89–98.

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56,* 455–463.

Ferguson, C. J. (2007a). Evidence for publication bias in video game violence effects literature: A meta-analytic review. *Aggression and Violent Behavior, 12,* 470–482.

Ferguson, C. J. (2007b). The good, the bad and the ugly: A meta-analytic review of positive and negative effects of violent video games. *Psychiatric Quarterly, 78,* 309–316.

Ferguson, C. J., & Kilburn, J. (2009). The public health risks of media violence: A meta-analytic review. *Journal of Pediatrics, 154,* 759–763.

Frijda, N. H., Markam, S., Sato, K., & Wiers, R. (1995). Emotion and emotion words. In J. A. Russell (Ed.), *Everyday conceptions of emotion* (pp. 121–143). Dordrecht, the Netherlands: Kluwer Academic.

Geen, R. G. (2001). *Human aggression.* Philadelphia, PA: Open University Press.

Gentile, D. A. (Ed.). (2003). *Media violence and children.* Westport, CT: Praeger.

Gentile, D. A., Anderson, C. A., Yukawa, S., Ihori, N., Saleem, M., Ming, L. K., . . . Sakamoto, A. (2009). The effects of prosocial video games on prosocial behaviors: International evidence from correlational, experimental, and longitudinal studies. *Personality and Social Psychology Bulletin, 35,* 752–763.

Gentile, D. A., Saleem, M., & Anderson, C. A. (2007). Public policy and the effects of media violence on children. *Social Issues and Policy Review, 1,* 15–61.

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.

Greitemeyer, T., & Osswald, S. (2009). Prosocial video games reduce aggressive cognitions. *Journal of Experimental Social Psychology, 45,* 896–900.

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 R. G. Geen & E. Donnerstein (Eds.), *Human aggression: Theories, research, and implications for social policy* (pp. 73–109). New York, NY: Academic Press.

Huesmann, L. R., & Eron, L. D. (Eds.). (1986). *Television and the aggressive child: A cross-national comparison.* Hillsdale, NJ: Erlbaum.

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., Lagerspetz, K., & Eron, L. D. (1984). Intervening variables in the TV violence–aggression relation: Evidence from two countries. *Developmental Psychology, 20,* 746–775.

Johnson, J. G., Cohen, P., Smailes, E. M., Kasen, S., & Brook, J. S. (2002, July 5). Television viewing and aggressive behavior during adolescence and adulthood. *Science, 295,* 2468–2471.

Kirsh, S. J. (2006). *Children, adolescents, and media violence: A critical look at the research.* Thousand Oaks, CA: Sage.

Kitayama, S., Mesquita, B., & Karasawa, M. (2006). Cultural affordances and emotional experience: Socially engaging and disengaging emotions in Japan and the United States. *Journal of Personality and Social Psychology, 91,* 890–903.

Kodaira, S. I. (1998). A review of research on media violence in Japan. In U. Carlsson & C. von Feilitzen (Eds.), *Children and media violence* (pp. 81–105). Göteborg, Sweden: UNESCO International Clearinghouse on Children and Violence on the Screen.

Kodomo no taiken katsudo kenkyukai. (2000). *Kodomo no taiken katsudo tou ni kansuru kokusai hikaku chosa* [A cross-cultural study about children's experiences and activities]. Nagano, Japan: Author.

Krahé, B. (2001). *The social psychology of aggression.* East Sussex, England: Psychology Press.

Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). Understanding and using the Implicit Association Test: IV. What we know (so far) about the method. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes* (pp. 59–102). New York, NY: Guilford Press.

Masuda, T., & Nisbett, R. E. (2001). Attending holistically versus analytically: Comparing the context sensitivity of Japanese and Americans. *Journal of Personality and Social Psychology, 81,* 922–934.

Mesquita, B., & Leu, J. (2007). The cultural psychology of emotion. In S. Kitayama & D. Cohen (Eds.), *Handbook of cultural psychology* (pp. 734–759). New York, NY: Guilford Press.

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review, 108,* 291–310.

Ontario Office for Victims of Crime. (2004). *Action agenda: A strategic blueprint for reducing exposure to media violence in Canada.* Retrieved from http://www.fradical.ca/Action_Agenda_November_2004.pdf

Polman, H., Orobio de Castro, B., & Van Aken, M. A. G. (2008). Experimental study of the differential effects of playing versus watching violent video games on children's aggressive behavior. *Aggressive Behavior, 34,* 256–264.

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.

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

Rosenthal, R. (1990). How are we doing in soft psychology? *American Psychologist, 45,* 775–777.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005). Publication bias in meta-analysis: Prevention, assessment and adjustments: Chichester, England: Wiley.

Schmitt, D. P., Allik, J., McCrae, R. R., & Benet-Martinez, V. (2007). The geographic distribution of Big Five personality traits: Patterns and profiles of human self-description across 56 nations. *Journal of Cross-Cultural Psychology, 38,* 173–212.

Sherry, J. L. (2001). The effects of violent video games on aggression: A meta-analysis. *Human Communication Research, 27,* 409–431.

Singer, D. G., & Singer, J. L. (Eds.). (2001). *Handbook of children and the media.* Thousand Oaks, CA: Sage.

Sterne, J. A. C., & Egger, M. (2005). Regression methods to detect publication and other bias in meta-analysis. In H. Rothstein, A. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 99–110). Chichester, England: Wiley.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8,* 220–247.

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

Tsai, J. L., Knutson, B., & Fung, H. H. (2006). Cultural variation in affect valuation. *Journal of Personality and Social Psychology, 90,* 288–307.

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.

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.

Wegner, D. M., & Bargh, J. A. (1998). Control and automaticity in social life. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 446–496). New York, NY: McGraw-Hill.

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

Yahiro, S. (2005). *Terebigemu kaishakuron josetsu: Assanburaaju* [Introduction of interpreting theory of video games: Assemblage]. Tokyo, Japan: Gendaishokan.

Received July 6, 2009
Revision received October 17, 2009
Accepted October 21, 2009 ■