Much Ado About Something: Violent Video Game Effects and a School of Red Herring: Reply to Ferguson and Kilburn (2010)

Brad J. Bushman
University of Michigan and VU University

Hannah R. Rothstein
Baruch College, City University of New York

Craig A. Anderson
Iowa State University

In this article we reply to C. J. Ferguson and J. Kilburn's (2010) critique of our meta-analysis on violent video game effects (C. A. Anderson et al., 2010). We rely on well-established methodological and statistical theory and on empirical data to show that claims of bias and misinterpretation on our part are simply wrong. One should not systematically exclude unpublished studies from meta-analytic reviews. There is no evidence of publication or selection bias in our data. We did not purposely exclude certain studies; we included all studies that met our inclusion criteria. Although C. J. Ferguson and J. Kilburn believe that the effects we obtained are trivial in size, they are larger than many effects that are deemed sufficiently large to warrant action in medical and violence domains. The claim that we (and other media violence scholars) are attempting to create a false crisis is a red herring.

Keywords: meta-analysis, violent video games, aggression

We appreciate the opportunity to reply to the Ferguson and Kilburn (2010) critique of our meta-analysis on violent video game effects (Anderson et al., 2010). Healthy debate about such issues is how scientific knowledge progresses. In this reply we address the criticisms Ferguson and Kilburn have raised about our meta-analysis.

Author Expertise in Violent Media Research and Meta-Analysis

The three authors who wrote this reply have considerable expertise in conducting violent media research, in meta-analysis, or in both (as do the other authors on our meta-analysis). Two of us (Anderson and Bushman) have been conducting research on violent media (including violent video games) for at least 20 years (e.g., Anderson & Ford, 1986; Bushman & Geen, 1990). Two of us (Bushman and Rothstein) teach graduate-level courses on meta-analysis, have written meta-analysis books (Borenstein, Hedges, Higgins, & Rothstein, 2009; Wang & Bushman, 1999), have contributed chapters to reference books on meta-analysis (Bushman & Wang, 2009; Rothstein, 2003; Rothstein & Hopewell, 2009; Rothstein, McDaniel, & Borenstein, 2001), and have written peer-reviewed articles that advance meta-analytic theory and methods (Bushman & Wang, 1995, 1996; Hedges, Cooper, & Bushman, 1992; Ioannidis, Patsopoulos, & Rothstein, 2008; Rothstein, 2008b; Rothstein & McDaniel, 1989; Schmidt et al., 1993; Valentine, Pigott, & Rothstein, in press; Wade, Turner, Rothstein, & Lavenberg, 2006; Wang & Bushman, 1998). One of us (Rothstein) is an expert on publication bias in meta-analysis (McDaniel, Rothstein, & Whetzel, 2006; Rothstein, 2004, 2008a, 2008b; Rothstein & Hopewell, 2009; Rothstein, Sutton, & Borenstein, 2005).

Excluding Unpublished Studies From Meta-Analytic Reviews

The term *unpublished study* means that the study was not published in a peer-reviewed journal, although it could have been published in another outlet (e.g., book). In their comment on our meta-analysis, Ferguson and Kilburn (2010) stated that "Anderson et al. failed to note that many scholars have been critical of the inclusion of unpublished studies in meta-analyses" (p. 175). This is simply false, at least when considering the writings of meta-analytic scholars. Consider the following statements from individuals who have written books on how to conduct meta-analytic reviews. Lipsey and Wilson (2001) stated that including only published material because it is refereed and represents "higher quality research" is "generally not very convincing" (p. 19). Petticrew and Roberts (2005) recommended searching for journal articles, books and book chapters, conference proceedings, dissertations, and other "gray" literature. Cooper (2009) specifically pointed out the limitations of relying only on peer-reviewed journal articles, stating that

> bias against null findings and confirmatory bias means that quality-controlled journal articles (and conference presentations) should not be used as the sole source of information for a research synthesis unless you can convincingly argue that these biases do not exist in the specific topic area. (p. 63, italics in the original)

Brad J. Bushman, Institute for Social Research, University of Michigan, and VU University, Amsterdam, the Netherlands; Hannah R. Rothstein, Department of Management, Baruch College, City University of New York; Craig A. Anderson, Center for the Study of Violence, Department of Psychology, Iowa State University.

Correspondence concerning this article should be addressed to Brad J. Bushman, Institute for Social Research, University of Michigan, 426 Thompson Street, Ann Arbor, MI 48106. E-mail: bbushman@umich.edu

Borenstein et al. (2009) stated that "publication status cannot be used as a proxy for quality; and in our opinion should not be used as a basis for inclusion or exclusion of studies" (p. 279). Littell, Corcoran, and Pillai (2008) urged individuals who conduct a meta-analysis to invest the extra effort needed to obtain gray or unpublished studies.

The view advanced by Egger and Smith, who are cited by Ferguson and Kilburn (2010) as arguing that inclusion of unpublished studies increases bias, was taken out of context. In fact, what Egger and Smith said is that inclusion of data from unpublished studies can, under some conditions, introduce bias, but they did not recommend limiting meta-analyses to peer-reviewed journal articles (see also Egger, Dickersin, & Smith, 2001). The Cook et al. article (1993) cited by Ferguson and Kilburn is the report of an opinion survey conducted almost 20 years ago among journal editors. This survey is out of date, and more recent surveys indicate that opinions have changed (e.g., Tetzlaff, Moher, Pham, & Altman 2006).

In summary, the current consensus among meta-analysis experts is that publication status is not a good proxy for methodological rigor and that any study that (otherwise) meets the inclusion criteria for a meta-analysis should not be excluded because it was not published in a peer-reviewed journal. There is absolutely no support for Ferguson and Kilburn's position that unpublished studies should not be included in a meta-analysis.

Publication Bias

There is ample evidence from multiple sources that publication bias is pervasive. That is why meta-analysts are urged to try to track down unpublished studies. Even when a researcher sets out to locate all potentially eligible studies, and unpublished articles such as dissertations and conference proceedings are included in a review, it is possible that some studies meeting the inclusion criteria were not found, and that these studies differed in some systematic way from those that were found. The purpose of conducting publication bias analyses is to assess the likelihood that, if such studies exist, they would threaten the validity of the results obtained by meta-analyzing only the retrieved studies.

In their meta-analytic reviews (Ferguson, 2007a, 2007b; Ferguson & Kilburn, 2009), Ferguson and his colleagues claimed that the trim and fill technique produces a "corrected" coefficient; it does not. In fact, the trim and fill technique produces an estimate of the effect adjusted for imputed missing studies. Both the originators of trim and fill technique (cf. Duval, 2005) and other meta-analysis experts who advocate its use have stated unequivocally that one should not view the adjusted estimate as a corrected or more accurate estimate of the effect, because it is based on imputed data points. Trim and fill is most appropriately considered a useful sensitivity analysis that assesses the potential impact of missing studies on the meta-analysis. It does this by examining the degree of divergence between the original effect-size estimate and the trim and fill adjusted effect-size estimate. This point is made numerous times in the key reference source for publication bias in meta-analysis (Rothstein et al., 2005), including in chapters cited by Ferguson and his colleagues.

Additionally, the key assumption of trim and fill is that the observed asymmetry in effects is due to publication bias rather than to real differences between effects found in small- versus large-sample studies. Sterne and Egger (2001) noted that it is possible that studies with smaller samples actually do have larger effects, perhaps because the smaller studies used different populations or designs than did the larger ones. Sterne and Egger coined the term *small study effect* to denote this alternative explanation for the results of the trim and fill and other publication bias procedures (e.g., Begg and Mazumder and Egger tests). Ferguson and his colleagues do not mention this critical caveat, even though these are the procedures they are relying upon.

Finally, it has been established that under conditions of heterogeneity, trim and fill may "impute" missing studies that do not actually exist in file drawers or anywhere else (Peters, Sutton, Jones, Abrams, & Rushton, 2007; Terrin, Schmid, Lau, & Olkin, 2003). The results of both our and Ferguson's work show that the effect sizes are quite heterogeneous. This is yet another reason to interpret the trim and fill results as a test of the robustness of the observed effects to the threat of publication bias, rather than as the correct effects.

We endorse Ferguson and Kilburn's (2010) observation that the politicization of this research area increases the risk for bias. Unlike the typical scenario in which publication bias is created by censoring on the basis of statistical significance, in politicized areas of research, there is at least the possibility that data are censored on the basis of political or other personal interests of researchers, reviewers, or editors. Thus, unlike typical publication bias, where one is concerned that the small effect size, nonsignificant results are missing (i.e., the ones that show that violent video games have no or minimal effects), in cases such as the current one, there is equal cause for concern that some large effect size results could be missing due to deliberate suppression. Because we considered both possibilities, we used the trim and fill method to look for putatively missing studies higher than the mean effect, as well as to look for putatively missing studies lower than the mean effect. We conducted these analyses on relatively homogeneous subgroups, in an attempt to avoid the problems that can occur when trim and fill is used when there is a lot of between study heterogeneity. The results, noted as sensitivity analyses and reported in Table 10 of Anderson et al. (2010), show that for some outcomes it appeared that low-effect studies were missing (the trim and fill adjusted correlation was lower than the observed correlation), whereas for other outcomes it appeared that high-effect studies were missing (the trim and fill adjusted correlation was higher than the observed correlation). We therefore do not understand Ferguson and Kilburn's (2010) objection to our conclusion that, overall "there is no evidence that publication or selection bias had an important influence on the results" (p. 167).

Inclusion Criteria and Classification of Studies as "Best Practice"

As stated in our article, unpublished studies were retrieved from PsycINFO and MEDLINE databases in the United States and from proceedings and annual reports in Japan. In addition, there were a number of "unpublished" Japanese studies from proceedings compilations. The publication bias analyses we conducted confirm that if we missed any unpublished studies, they would not have significantly influenced our findings.

Ferguson and Kilburn disagreed with our classification of some studies as "best practice." Agreement among coders was 93% for

best practice studies. More important, the pattern of results was the same for best practice studies and for all studies. Ferguson and Kilburn (2010) stated that we were "disinclined toward Williams and Skoric (2005), despite the fact that this study does indeed (contrary to Anderson et al.'s assertions) include a measure of verbal aggression at least as ecologically valid, if not more so, than many of those nominated as best practices." This study did not meet our inclusion criteria because it measured verbal rather than physical aggression.

There are a host of other problems with Ferguson and Kilburn's (2010) claims about what was (or was not) included in our meta-analysis. Indeed, detailing all of them would take more space than is allocated for such replies. None of the studies that they claimed we missed were in fact missed. Several studies that are now available were not available at the time of the cutoff for the meta-analysis (i.e., Ferguson & Rueda, in press; Ferguson, San Miguel, & Hartley, in press; Olson et al., 2009; Przybylski, Weinstein, Ryan, & Rigby, 2009). We could redo all the meta-analyses again, including these and other recent studies, but by the time we finished there would be still more studies. Besides, adding all of the newly available studies would not change the results of our meta-analysis in even a minor way, for two reasons: (a) the effect sizes are similar in size to the ones in our meta-analysis and (b) their sample sizes are not large enough to change the average effect size much, even if the new studies had effect sizes around zero (which they do not).

Posters, such as Barnett, Coulson, and Foreman (2008), are not included in PsycINFO or MEDLINE, so there could be no bias in our selection of posters. Furthermore, although Ferguson and Kilburn claimed that these authors had a published report in 2008, they failed to provide a reference for it and there is no record of it in PsycINFO or MEDLINE. Also, we did not ask any research groups for unpublished studies or posters.

It is unclear why Ferguson and Kilburn think that work by Ryan and his colleagues (Przybylski et al., 2009; Ryan, Rigby, & Przybylski, 2006) contradicts our meta-analysis findings. They studied why people are attracted to video games, not the effects of violent video games on aggression. The relevant data from all of the remaining research groups that "arguably, have presented research not in line with Anderson et al.'s hypotheses" (Ferguson & Kilburn, 2010, p. 175) were in fact included in our meta-analysis. Furthermore, even though each of these remaining studies failed to meet one or more best practice inclusion criteria, their effects were similar in size to those obtained in other studies ($r+ = .184$, $K = 7$, $N = 2,080$, $Z = 8.45$, $p < .001$). In summary, Ferguson and Kilburn failed to identify any biased search processes, any biased search outcomes, or any studies that should have been but were not included in our meta-analysis.

Magnitude of Effect of Violent Video Games on Aggressive Behavior

Ferguson and Kilburn (2010) stated, "Our analyses agree that the uncorrected estimate for violent video game effects is quite small ($r = .15$ in both analyses)." We are not sure where Ferguson and Kilburn came up with the $r = .15$ value. Perhaps they used the "best partials" estimate for all study designs, an estimate that actually does "correct" for gender differences in all studies and initial aggression levels in longitudinal studies (see Anderson et al., 2010, Table 4). The overall estimate of the effect of violent video games on aggression was $r = .19$ for all studies and $r = .24$ for studies of higher methodological quality.

Ferguson and Kilburn claimed that the .15 estimate is too liberal because it does not control for other risk factors, such as depression, peer group influence, and family environment. There are at least four problems with this claim. First, it is irrelevant to experimental studies in which participants are randomly assigned to groups. Second, the point estimates for cross-sectional studies were all larger than $r = .15$ (rs = .26, .17, and .19 for best raw, best partials, and full sample, respectively). Third, one cannot combine correlations from studies unless all studies controlled for exactly the same variables. Fourth, Ferguson and Kilburn do not mention that some well-known cross-sectional studies controlled for several individual-difference risk factors and still found significant violent video game effects (e.g., Anderson et al., 2004).

Ferguson and Kilburn (2010) considered the effects we obtained to be so small that they are not worth worrying about. Other meta-analyses cited by Ferguson and Kilburn as supposedly refuting the effect of violent video games on aggressive and violent behavior have found correlations in the same range (e.g., Sherry, 2001). What differs is not the magnitude of the obtained effects but rather how the effects are interpreted.

By conventional standards (Cohen, 1988), our correlations are between "small" ($r = .1$) and "medium" ($r = .3$) in size. However, this is the range of effects most commonly observed in social psychology. For example, one meta-analysis examined the magnitude of effects obtained in social psychology studies during the past century. The average effect obtained from 322 meta-analyses of more than 25,000 social psychology studies involving over 8 million participants was about $r = .2$ (Richard, Bond, & Stokes-Zoota, 2003). This not surprising, because human behavior is extremely complex and has multiple causes. For this reason Hemphill (2003) recommended a reconceptualization of effect size, in which $r = .1$ is small, $r = .2$ is medium, and $r = .3$ is large. Similarly, Lipsey (1990) recommended a reconceptualization of effect-size conventions, based on reviews of effects of social science interventions, in which $r = .07$ is small, $r = .22$ is medium, and $r = .41$ is large.

The effects we obtained for violent video games are similar in size to the effects of risk factors for physical health, such as exposure to lead, asbestos, or secondhand smoke (Bushman & Anderson, 2001). They are also similar in size to other risk factors for violent and aggressive behavior, such as poverty, substance abuse, and low IQ (U.S. Department of Health and Human Services, 2001). We do not consider the magnitude of these effects to be trivial. Neither do professional physical and mental health organizations, which issued the *Joint Statement on the Impact of Entertainment Violence on Children.* According to the statement, "Entertainment violence can lead to increases in aggressive attitudes, values, and behavior, particularly in children" (Congressional Public Health Summit, 2000, p. 1). The six organizations that signed the statement were the American Academy of Child and Adolescent Psychiatry, American Academy of Family Physicians, American Academy of Pediatrics, American Medical Association, American Psychiatric Association, and American Psychological Association. More recently, the American Psychological Association (2005) issued a similar statement on violent video game effects.

Finally, there are circumstances in which small effect sizes warrant serious concern: "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" (Anderson et al., 2010, p. 170).[1]

Effects of Violent Video Games on Serious Acts of Aggression or Violence

Ferguson and Kilburn (2010) probably are correct in noting that violent video games have a weaker effect on serious acts of aggression and violence than on less serious acts.[2] This is no surprise. Because serious acts of aggression and violence are relatively rare, they are difficult to predict using violent video game exposure or any other single risk factor. Violent crimes typically result from a combination of multiple risk factors. No single risk factor accounts for a large proportion of variance, but that does not mean that the risk factors are trivial and should be ignored.[3]

Unstandardized Aggression Measures

Ferguson and Kilburn (2010) raised a potentially valid point about the use of unstandardized aggression measures. Variations of the competitive reaction time task developed by Taylor (1967) have been used by aggression researchers for over 40 years. Different researchers have used different measures of aggression from this task, and this practice could increase the probability of a Type I error if researchers were systematically choosing a measure on the basis of the size of the media violence effect. If the overall meta-analytic experimental effect size is inflated by such a reporting bias in competitive reaction time studies, these studies should yield systematically larger effect sizes than experimental studies using other aggressive behavior measures, but they did not. This is not surprising. A previous meta-analysis found that different laboratory measures of aggressive behavior produce similar results and are highly correlated (Carlson, Marcus-Newhall, & Miller, 1989). For example, the correlation between physical punishment intensity and duration was .76 across 92 experimental studies.

Is Psychology Inventing a Phantom Youth Violence Crisis?

There are at least five problems with Ferguson and Kilburn's (2010) claims in this section of their comment. First, we have never claimed that national violent crime data are a good test of media violence effects. Because violent crime is influenced by so many risk factors, simple studies of national crime rate changes are difficult to interpret (Anderson & Bushman, 2002). Second, there is evidence that certain types of youth violence rates are increasing (Escobar-Chaves & Anderson, 2008). The data are not definitive yet and are not relevant to testing media violence effects. They do, however, suggest that there may be some selectivity in Ferguson and Kilburn's choice of violence indicators. Third, the only studies we have seen in which controls for a variety of "third" variables wiped out the video game violence effect have done so by "controlling" for variables that themselves could be conceptualized as

additional outcomes of high exposure to violent video games, such as trait aggressiveness. Fourth, the experimental studies yielded significant effects, and they used random assignment to control for the types of third variables that Ferguson and Kilburn claimed explain video game violence effects. Fifth, media violence researchers do not claim that violent media are the most important risk factor for aggressive and violent behavior. However, of all the risk factors linked to aggressive and violent behavior, exposure to violent media may be the easiest factor for parents to control. In summary, claims that the mainstream media violence research community is trying to invent a phantom youth violence crisis have no basis in fact.

Overlapping Meta-Analyses

This is not the place to go through a long list of concerns we have about the violent video game meta-analyses conducted by Ferguson and his colleagues, but we would like to mention briefly one concern. Ferguson and his colleagues provide very little information about the studies included in their three meta-analyses (Ferguson, 2007a, 2007b; Ferguson & Kilburn, 2009). This too goes against current practice. Most important, no list of included studies is provided for any of the meta-analyses. We asked Ferguson for a list of the studies included in each of his meta-analyses, and he graciously provided us these lists. The percentages of overlapping studies in the meta-analyses ranged from 54% to 100%. Thus, their meta-analytic results are not independent.

Summary

In summary, we conducted a state-of-the art meta-analysis on violent video game effects, one that includes data from more that 10 times as many participants as in meta-analyses conducted by Ferguson and his colleagues. We included unpublished studies, as recommended by virtually all meta-analysis experts. We created and tested stringent inclusion criteria. We conducted appropriate analyses to assess the impact of publication bias and found minimal bias. One could still argue that the magnitude of effects we observed was so small that it is trivial, but most meta-analysis experts, physicians, psychologists, and psychiatrists would disagree with Ferguson and Kilburn on this point as well. Our results suggest that violent video games increase aggressive thoughts, angry feelings, and aggressive behaviors and decrease empathic feelings and prosocial behaviors. Moreover, we obtained similar

---

[1] For example, Rosenthal (1990) noted the case involving the effects of taking a daily aspirin (vs. a placebo) on the occurrence of a heart attack. The original, double-blind placebo-randomized experiment was stopped early because the preliminary results were so strong that it was deemed unethical to continue giving placebos. The effect size was $r = .034$.

[2] But seriousness of aggression was not a significant moderator in our meta-analysis.

[3] For example, in a longitudinal study (Anderson, Gentile, & Buckley, 2007, Chapter 7), a fairly extreme behavior (getting into a physical fight at school) was affected by a risk factor of small size. After controlling for Time 1 fighting behavior, sex, and hostile attribution bias, those who played a lot of violent video games early in the school year were about 20% more likely to be involved in a subsequent physical fight.

effects in Western and Eastern countries. Violent video game exposure is a causal risk factor for later aggression.

References

American Psychological Association. (2005, August 17). *Resolution on violence in video games and interactive media.* Retrieved from http://www.apa.org/pubs/journals/releases/resolutiononvideoviolence.pdf

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

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., & Ford, C. M. (1986). Affect of the game player: Short-term effects of highly and mildly aggressive video games. *Personality and Social Psychology Bulletin, 12,* 390–402.

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

Barnett, J., Coulson, M., & Foreman, N. (2008, April). *The WoW! factor: Reduced levels of anger after violent on-line play.* Poster session presented at the annual meeting of the British Psychological Society, Dublin, Ireland.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* New York, NY: Wiley.

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

Bushman, B. J., & Geen, R. G. (1990). The role of cognitive–emotional mediators and individual differences in the effects of media violence on aggression. *Journal of Personality and Social Psychology, 58,* 156–163.

Bushman, B. J., & Wang, M. C. (1995). A procedure for combining sample correlations and vote counts to obtain an estimate and a confidence interval for the population correlation coefficient. *Psychological Bulletin, 117,* 530–546.

Bushman, B. J., & Wang, M. C. (1996). A procedure for combining sample standardized mean differences and vote counts to estimate the population standardized mean difference in fixed effects models. *Psychological Methods, 1,* 66–80.

Bushman, B. J., & Wang, M. C. (2009). Vote counting methods in meta-analysis. In H. M. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *Handbook of research synthesis* (2nd ed., pp. 207–220). New York, NY: Russell Sage Foundation.

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

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York, NY: Academic Press.

Congressional Public Health Summit. (2000, July 26). *Joint statement on the impact of entertainment violence on children.* Retrieved from www.aap.org/advocacy/releases/jstmtevc.htm

Cooper, H. (2009). *Research synthesis and meta-analysis: A step-by-step approach* (4th ed.). Thousand Oaks, CA: Sage.

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

Egger, M., Dickersin, K., & Smith, G. (2001). Problems and limitations in conducting systematic reviews. In M. Egger, G. Davey-Smith, & D. Altman (Eds.), *Systematic reviews in health care: Meta-analysis in context* (pp. 43–68). London, England: BMJ Books.

Escobar-Chaves, S. L., & Anderson, c. A. (2008). Media and risky behaviors. *The Future of Children, 18,* 147–180.

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.

Ferguson, C. J., & Kilburn, J. (2010). Much ado about nothing: The misestimation and overinterpretation of violent video game effects in Eastern and Western nations: Comment on Anderson et al. (2010). *Psychological Bulletin, 136,* 174–178.

Ferguson, C. J., & Rueda, S. M. (in press). The Hitman study: Violent video game exposure effects on aggressive behavior, hostile feelings and depression. *European Psychologist.*

Ferguson, C. J., San Miguel, C., & Hartley, R. D. (in press). A multivariate analysis of youth violence and aggression: the influence of family, peers, depression, and media violence. *Journal of Pediatrics.*

Hedges, L. V., Cooper, H. M., & Bushman, B. J. (1992). Testing the null hypothesis in meta-analysis. A comparison of combined probability and confidence interval procedures. *Psychological Bulletin, 111,* 188–194.

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

Ioannidis, J. P. A., Patsopoulos, N. A., & Rothstein, H. R. (2008). Reasons or excuses for avoiding meta-analysis in forest plots. *British Medical Journal, 336,* 1413–1415.

Lipsey, M. W. (1990). *Design sensitivity: Statistical power for experimental research.* Newbury Park, CA: Sage.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

Littell, J. H., Corcoran, J., & Pillai, V. (2008). *Systematic reviews and meta-analysis.* Oxford, England: Oxford University Press.

McDaniel, M. A., Rothstein, H. R., & Whetzel, D. (2006). Publication bias: A case study of four test vendor manuals. *Personnel Psychology, 59,* 927–953.

Olson, C., Kutner, L., Baer, L., Beresin, E., Warner, D., & Nicholi, A. (2009). M-rated video games and aggressive or problem behavior among young adolescents. *Applied Developmental Science, 13*(4), 1–11.

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2007). Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. *Statistics in Medicine, 26,* 4544–4562.

Petticrew, M., & Roberts, H. (2005). *Systematic reviews in the social sciences: A practical guide.* Chichester, England: Wiley-Blackwell.

Przybylski, A., Weinstein, N., Ryan, R., & Rigby, C. (2009). Having to versus wanting to play: Background and consequences of harmonious versus obsessive engagement in video games. *CyberPsychology & Behavior, 12,* 485–492.

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

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

Rothstein, H. R. (1990). Interrater reliability of job performance ratings: Growth to asymptote level with increasing opportunity to observe. *Journal of Applied Psychology, 75,* 322–327.

Rothstein, H. R. (2003). Progress is our most important product: Contributions of validity generalization and meta-analysis to the development and communication of knowledge in I/O psychology. In K. R. Murphy (Ed.), *Validity generalization: A critical review* (pp. 115–154). Mahwah, NJ: Erlbaum.

Rothstein, H. R. (2004). File drawer analysis. In M. Lewis-Beck (Ed.), *The*

encyclopedia of research methods for the social sciences. Thousand Oaks, CA: Sage.

Rothstein, H. R. (2008a). Publication bias. In *Wiley encyclopedia of clinical trials.* Chichester, England: Wiley.

Rothstein, H. R. (2008b). Publication bias as a threat to the validity of meta-analytic results. *Journal of Experimental Criminology, 4,* 61–81.

Rothstein, H. R., & Hopewell, S. (2009). Grey literature. In H. M. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *Handbook of research synthesis* (2nd ed., pp. 103–125). New York, NY: Russell Sage Foundation.

Rothstein, H. R., & McDaniel, M. A. (1989). Guidelines for conducting and reporting meta-analyses. *Psychological Reports, 65,* 759–770.

Rothstein, H. R., McDaniel, M. A., & Borenstein, M. (2001). Meta-analysis: A review of quantitative cumulation methods. In N. Schmitt & F. Drasgow (Eds.), *Advances in measurement and data analysis* (pp. 534–570). San Francisco, CA: Jossey-Bass.

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

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

Schmidt, F. L., Law, K., Hunter, J. E., Rothstein, H. R., Pearlman, K., & McDaniel, M. (1993). Refinements in validity generalization methods: Implications for the situational specificity hypothesis. *Journal of Applied Psychology, 78,* 3–12.

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

Sterne, J. A. C., & Egger, M. (2001). Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *Journal of Clinical Epidemiology, 54,* 1046–1055.

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.

Terrin, N., Schmid, C. H., Lau, J., & Olkin, I. (2003). Adjusting for publication bias in the presence of heterogeneity. *Statistics in Medicine, 22,* 2113–2126.

Tetzlaff, J., Moher, D., Pham, B., & Altman, D (2006, October). *Survey of views on including grey literature in systematic reviews.* Paper presented at the 16th Cochrane Colloquium, Dublin, Ireland.

U.S. Department of Health and Human Services. (2001). *Youth violence: A report of the Surgeon General.* Rockville, MD: U.S. Government Printing Office.

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (in press). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics.*

Wade, C. A., Turner, H. M., Rothstein, H. R., & Lavenberg, J. (2006). Information retrieval and the role of the information specialist in producing high-quality systematic reviews in the social, behavioral, and education sciences. *Evidence & Policy: A Journal of Research, Debate and Practice, 2,* 89–108.

Wang, M. C., & Bushman, B. J. (1998). Using normal quantile plots to explore meta-analytic data sets. *Psychological Methods, 3,* 46–54.

Wang, M. C., & Bushman, B. J. (1999). *Integrating results through meta-analytic review using SAS software.* Cary, NC: SAS Institute.