Directional is the new null?

Psychology of Popular Media

© 2023 American Psychological Association
ISSN: 2689-6567

2023, Vol. 12, No. 3, 364–372
https://doi.org/10.1037/ppm0000447

COMMENTARY

Directional Is the New Null? A Comment on Bushman and Anderson (2021)

Grant J. Devilly[1], Aaron Drummond[2], James D. Sauer[3], Allen Copenhaver[4], Julia Kneer[5], and Christopher J. Ferguson[6]

[1] School of Applied Psychology and Griffith Criminology Institute, Griffith University
[2] School of Psychology, College of Humanities and Social Sciences, Massey University
[3] School of Psychological Sciences, College of Health and Medicine, University of Tasmania
[4] Criminal Justice, Lindsey Wilson College
[5] Department of Media and Communication, ESHCC, Erasmus University Rotterdam
[6] Department of Psychology, Stetson University

Bushman and Anderson (2021) have recently argued that the evidence of harm after playing violent video games is so strong that this effect should be our starting point for future research. They base this claim on an argument that: (a) many professional bodies agree with this opinion; (b) strong theories, such as their General Aggression Model (GAM), predict such outcomes; (c) experimental and meta-analytic studies back such a claim; and (d) people who do not obtain this effect are in the minority and their studies have methodological shortcomings or they misanalyze their data. It is argued here that this is not consistent with the precepts of falsificationism and that: (a) their argument from authority is problematic; (b) they appear unconcerned with, or unaware of, increasing null studies, particularly missing out international research or preregistered studies; (c) the majority of research groups outside of Bushman/Anderson and their coauthors do not concur with their results; and (d) there are theories (e.g., the Immersive Media Prediction model) which better account for the data than their GAM. It is also argued that when theories and data collide it is the theories that need revision not the data, demand effects of researchers are strong in gaming research, and Bushman and Anderson's results may themselves have been influenced by their methodology rather than accurately describing a naturally occurring weakness in the human condition. It is also argued that we need a more collegial approach to gaming research and make suggestions to facilitate this shift.

Public Policy Relevance Statement

In a recent article by Bushman and Anderson (2021), violent video gameplay (VVG) has been referenced as a cause of community violence—believing that the scientific evidence for this is so overwhelming that we should reposition the null hypothesis to reflect that specific interpretation of the data. We argue here that Bushman and Anderson are in a minority of researchers who continue to find harmful effects following VVG and explain how and why their conclusions have likely come from straw man arguments. We offer suggestions to rectify these errors and propose methods to consolidate a more collegial approach to the subject matter.

Keywords: violent video games, aggression, scientific method, effect sizes, argument from authority

"Cherish those who seek the truth but beware of those who find it."
(Proverb traditionally attributed to André Gide or Voltaire)

Bushman and Anderson (2021) wish to solve "the puzzle of null violent media effects." Wishing to address nonreplication is admirable and we appreciate all attempts to move the field forward. However, this proposal is problematic because a null effect is usually seen as the *terminus a quo* and nonreplication usually focuses on why a select group of researchers obtain a unique directional effect. To better explain this issue, it is first necessary to have a common understanding of falsificationism.

Grant J. Devilly https://orcid.org/0000-0001-6740-186X

Correspondence concerning this article should be addressed to Grant J. Devilly, School of Applied Psychology and Griffith Criminology Institute, Griffith University, Mt Gravatt, QLD 4122, Australia. Email: grant@devilly.org

Falsificationism and Science

Social scientists have traditionally performed research in the spirit of Popper's falsification principle, whereby the purpose of research is to objectively quantify a testable hypothesis (arrived at through induction or deduction) and to put this hypothesis to the test, rejecting where necessary rather than affirming. There is some debate regarding different types of falsificationism. Popper saw falsificationism as a way of demarking between science and pseudoscience (and, in particular, psychoanalysis and astrology; Popper, 1962). However, it is not just that a theory can be falsified, some saw Popper as arguing that (a) there should be some effort applied to falsify it, and (b) when it is falsified the result should be accepted—presumably by the experimenters and by the community (Cioffi, 1985).

There is some argument regarding the exact nature of Popperian falsificationism, with different commentators seeing various numbers of versions of falsificationism in Popper's writings—works, which spanned some 60 years. Lakatos (1970) sees that there are three main types of falsificationism in Popper's work: dogmatic falsificationism (a reaction to inductive reasoning, it is the argument there are empirical reasons, and empirical reasons only, that can be trusted to refute hypotheses); naïve falsificationism (where one can see statements as observations and these can be used in the falsificationism process); and sophisticated falsificationism (where one is not considering just a specific theory or hypothesis, but running a comparative falsificationism on other available theories and seeing which best approximates the data, is most falsifiable and leads us to new predictions—moving the field forward). We are promoting here a sophisticated falsificationism whereby new data leads, at least, to the proposition of a revised theory. But we argue that this theory must be overtly distinguishable from previous theories and that any new terms, provisions, or expectations are clearly labeled as such. Throughout, though, Popper believed that in evaluating a scientific theory we should attempt to falsify and not confirm, with all hypotheses seen as hypothetical explanations of the world.

This methodological approach appears to be at odds with what Bushman and Anderson (2021) are proposing in their manuscript. In this paper, we argue why we see it as problematic that a specific view of human nature—one of human frailty, emotional contagion, affective lability, and negative cognitive consequences from video games, and one that is held by a small group of academics (see Ferguson & Colwell, 2017; Quandt et al., 2015)—is being promoted as our baseline for investigation ("our new null"?).

The Argument From Authority

The argument for the direction of investigation in the current case relies on an argument from authority in that multiple "professional and scientific organizations have issued statements about the potentially harmful effects of exposure to violent media." Such a baseline is problematic on many counts—not least because authority is the least valid source of knowledge under the scientific model (see Pelham & Blanton, 2018). For example, Bushman and Anderson (2021) note the American Psychological Association and the Australian Psychological Society as examples of authorities who warn about the risks associated with violent media. However, there are reasons to be concerned about appeals to these authorities. First, such calls to authority may turn out to be incorrect. For example, upon checking, the Australian Psychological Society have confirmed via email that while they have "some resources that reference gaming, the APS has not made any such statement" (Personal Communication, October 6, 2021).

Second, the authority themselves may be mistaken. For example, when Ferguson et al. (2020) attempted to replicate the research of the American Psychological Association's (APA) 2015 Taskforce on Violent Media, they found associations much weaker than those reported by the taskforce and also found "some studies were included that did not appear to have relevant data, and many other available studies were not included" (p. 1423). To our knowledge, the results of Ferguson et al.'s replication have not been shown to be invalid, yet the American Psychological Association has not updated its position. Furthermore, sometimes professional organizations appear to have greater certitude in the alleged effects of violent video games on aggression than the data can support (Copenhaver & Ferguson, 2018). For example, it is our understanding that, even within the APA, not all subcommittees supported the 2015 Resolution on Violent Video Games. Thus, even among bodies that might be considered authorities, the degree of consensus and the strength of evidence supporting any consensus remains questionable.

Predominantly citing their own research, Bushman and Anderson go on to claim that "the overwhelming majority of paediatricians, media researchers, and parents believe that exposure to violent media can increase aggression in children." We would note that as most of the organizations they cite have, at one time or another, been contracted or advised by Bushman or one of his coauthors, making such claims is somewhat circular. Thus, citing professional bodies may amount to "echo attribution" in which some scholars directly influenced or wrote those professional body statements and then cite them to bolster their own worldviews. It is also worrying that, to our knowledge, such organizations do not consult with international researchers, increasing the likelihood of common method variance errors in their decision-making.

Further, such appeals to authority, and reliance on the representativeness heuristic (the heuristic that organizations making recommendations ostensibly based on the available evidence must draw conclusions that represent the relevant scientific data in its entirety), contain many dangers—some obvious, some subtle. Indeed, claims of consensus are not even supported by Bushman et al.'s (2015) data whose sample showed that 39.0% of media psychologists did not believe there to be a causal link between violent games and aggression (leaving those who did far from a consensus view), and more than half of those surveyed (57.6%) did not believe violent games to be a major factor in real-life violence. Moreover, other surveys of scholars (e.g., Ferguson & Colwell, 2017; Quandt et al., 2015) have failed to replicate Bushman et al.'s (2015) results and, in fact, found the "harm" view to be in the minority. Lastly, we note that most US Government reviews (including by the School Safety Commission in the United States, as well as the Brown v. Entertainment Merchants Association, 2011 US Supreme Court decision), and Government reviews by the UK, Australia, and Sweden, all came to the conclusion that the research evidence was not able to conclusively support links between violent games and aggression or violent crime.

During this process, Bushman and Anderson argue that their position is also predicted by various theories and, therefore, violent games "should" increase aggression. However, although overlooked by Bushman and Anderson, there are other theories that do not hold the same predictions as their General Aggression Model (Anderson and Bushman, 2002). These theories have arisen out of empirical data (data generating theory) rather than being theories looking for data. Such data-driven theories include, for example, the Immersive Media Prediction model (Unsworth et al., 2007), a motivational hypothesis of self-determination theory (Przybylski et al., 2014), an evolutionary model on prosocial behavior (Devilly et al., 2017), and the Catalyst Model (Ferguson et al., 2008).

Situational or Dispositional Associations With Violence

The authors also erroneously cite Social Learning Theory as predictive of violence following violent game exposure. Bandura renamed this Social Cognitive Theory precisely because of his conviction that cognitive control can stop the behavior, seeing that "people are self-organizing, proactive, self-reflecting, and self-regulating, not just reactive organisms shaped and shepherded by environmental events or inner forces" (Bandura, 2001, p. 266). This also speaks to the problem of imitation versus intention and the lack of a clear pathway explaining behavior. In fact, one paper that investigated whether Bandura's theory or the GAM was a better model for media violence effects revealed that the context for in-game aggression was a more powerful determinant of post-game aggression than the in-game violence itself—suggesting that the contextual factors included in Bandura's theory (but absent from the GAM) were a crucial factor (Sauer et al., 2015).

In effect, there are theories that see (a) humans as organisms that are, on the whole, resilient to the trials and tribulations inherent to life; (b) individual and personological variables as predictors of reactivity and refraction following emotionally laden life events; and (c) do not see any short-term reactivities as the result of a general vulnerability of the human spirit—one that can be vaccinated by confiscation and prohibition. Anderson and Bushman (2002, 2018), on the other hand, see the etiological genesis of any violence or aggression in society as specifically situational rather than dispositional (in the entirety in 2002 and predominantly in 2018), and primarily preventable through legislation.

Anderson and Bushman's (2002) episodic model of reactivity does allow room for a "thoughtful action" in the short term; however, this luxury is not afforded the player in the long-term process of the GAM. In effect, the GAM argued very clearly in 2002 that violent video gaming increases aggressive personality (which we take to mean "disposition") while the person may have a "thoughtful action" during an individual scenario. This is because the long-term increase in aggression from the model has no escape clause. As Anderson says himself, after likening gaming to smoking cigarettes and getting cancer, "scientists have not been able to find any group of people who consistently appear immune to the negative effects of media violence or video game violence" (#5, Anderson, 2009).

By 2003, it was even plainly stated by Anderson and colleagues that there was "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). However, by 2007 this issue of intentional behavior had become "media violence exposure can increase the likelihood of aggressive and violent behavior in both short- and long-term contexts" (emphasis added; Anderson et al., 2007). Aside from these statements exemplifying just one aspect of the micromutations of GAM predictions over time, this approach fails the sophisticated falsificationism perspective of societal knowledge acquisition.

One thing for certain, though, is that this perspective of gaming, metaphorically, appears to be blaming the gun for a shooting but not a person, their upbringing or their particular set of circumstances—but, of course, only if the gun is an "assault rifle" and not a "hunting rifle", and irrespective of "whether the person holding the gun is a 'good guy' or a 'bad guy'" (as they claim on p. 7), and summary, the fact that some organizations have been advised to warn against violent gaming and some theories predict violent media effects makes no difference to how we should view actual data. When a theory is incompatible with empirical data, the scientific method demands that we must either edit or dispose of the theory, not the data—particularly after replication. Alternatively, one may look to extant theories that *do* accommodate the data, such as those theories of aggressive behavior outlined above. In contrast, Bushman and Anderson seem to be implying in their paper that when data do not align with a subset of theories that predict harmful violent media effects, then we should look for errors in the methods of the studies returning these null results, with a view to discarding the data as anomalous. Such ad hoc theorizing would constitute a form of the confirmation bias (Oswald & Grosjean, 2004), increase publication bias (Begg, 1994), be one of the practices which the open-science movement and preregistration was formed to prevent (Simmons et al., 2021; van't Veer & Giner-Sorolla, 2016), and, if followed through by a researcher, would be considered by many to constitute a form of questionable research practice (John et al., 2012) or, in extreme cases, a form of data manipulation (Haverford College, 2015).

Errors in a Test

Bushman and Anderson (2021) also provide examples of phenomena, and methodological and analytical practices, which may increase the error in a test of the gaming-violence nexus. We appreciate the effort applied to these lists and agree with most of them. However, we have two main points to raise regarding these lists.

First, we are grateful to see this list in print as it formalizes the item-by-item approach we have seen by negative reviewers in submissions (predominantly to APA journals) where null effects papers are rejected. It is our contention, however, that these items should be seen as discussion points rather than exclusion criteria. Placing the method and various controls into perspective is the role of the discussion section of a manuscript. The fact that one particular study did not include one particular type of control (or, even, controlled too much) is not a reason to reject the study, but rather a reason to place this study into context with other research and place relevant caveats on the conclusion drawn from the data in that study. We are sure that Bushman and Anderson intended this, but we would like this point to be explicit.

Saturated Models

The issue of controlling for "everything but the kitchen sink" (p. 3) is more complex than Bushman and Anderson present here. They express the concern that including irrelevant predictor variables can decrease the variance explained by a critical predictor of interest. In effect, they argue against saturating a model to such a degree that truly relevant relationships are missed (frequently referred to by the phrase "death by a thousand cuts"). It is possible that truly irrelevant control variables may either decrease the variance explained by an important predictor or, conversely, result in spurious confidence in a predictor that, in fact, is not as critical as the researchers believe. However, this concern does not apply to theoretically relevant control variables. Many of the factors Bushman and Anderson criticize researchers for including as unnecessary have been identified in the psychological and criminological research (independent of gaming research) as associated with actual violent crime—and not just aggression. Virtually all introductory criminology/criminal justice textbooks inform students that the factors predictive of criminal behavior, violent crime included, are myriad. Thus, controlling for theoretically relevant predictors of violence and aggression is, in fact, good scientific practice to generate more accurate and comprehensive models of human aggression. As Bushman and Anderson argue that theory should direct analysis (and, even, the interpretation of results), we would expect them to be in favor of controlling for known predictors unrelated to gaming. In contrast, to conflate the effects of more video gaming with, say, lack of schooling or absent parenting, in the prediction of aggression or violence would be akin to argue that children who eat more ice cream have more teeth cavities because ice cream spoons have been known to chip teeth (i.e., Type III error).

Exposure Time

Second, it is important to point out that these error tables apply equally in all directions. For example, one of the items listed is "insufficient exposure to violent media." They are correct—for instance, researchers have shown that short periods of 20-min exposures to violent gaming increase scores on anger measures far more than 60-min exposures (Devilly et al., 2012). Compare this data to Anderson et al.'s (2010) prima facie reasoning on the topic of gaming length: "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" (p. 159). This is a good example of why we should not rely solely upon theory—theory works right up *until it no longer explains the empirical data*, and then it must be replaced by a new theory that better explains that data. In this case, Anderson et al. assume a direction of effect exactly opposite to the results of experimental manipulation.

Moreover, it is difficult to know what constitutes "insufficient" exposure—particularly if gaming time as brief as 15 min (as argued above by Anderson et al., 2010) might be considered sufficient? Consumer research has shown that gamers play an average of 1 hr and 4 min per session (Limelight Networks, 2019). Gentile et al. argued that adolescents played for an average of 77 min/day (Gentile et al., 2004). In many, if not all, of Anderson and Bushman's experimental studies, the exposure to violent gaming is less than 30 min, with 15 min being common (e.g., Lull & Bushman, 2016). In such a case, their directional effect may well be the result of a Type III error—failing to measure frustration with the experimental procedure that occurs when gameplay ceases just as the participant gains mastery of the controls—and then misinterpreting that as the effect of the media exposure (e.g., see Breuer et al., 2015; Devilly et al., 2012, 2021). It could even be that one is measuring the frustration with a lack of mastery in knowing how to play a game—one which dissipates over time as participants become frequented with the controls (Przybylski et al., 2009).

Unfortunately, in peer review, it has been our experience that the "insufficient exposure to violent media" argument is interpreted as meaning that the game used is not contemporary enough and null effect research is rejected following advice from a reviewer. The implication, sometimes made explicitly, is that the realism of the graphics of this older game (one that 18 months ago may have been held aloft as the pinnacle of aggression and violence) is now no longer able to stimulate the aggressive reaction that Bushman and Anderson expect. This does "beg the question" of the hypothesized long-term effects of violent gameplay as argued by those who see violent video games as causing personal and societal harm. The somewhat contradictory argument here seems to be that people quickly desensitize to new stimuli and the effects are not obvious in experimentation, yet these stimuli have caused long-term harm.

This is also a no-win publishing scenario in which a game greater than 18 months old is seen as unremarkable—even though ethics, testing, write-up, submission, and first peer review nearly always take longer than 18 months. It is also difficult to see why effects of this type would be specific to studies showing null effects as studies reporting differences between conditions would surely be vulnerable to the same experimental time frame and the corresponding cost to the "contemporariness" of game stimuli.

Peer Review and Demand Effects

Reasons to Reject Are Reasons to Accept

This raises a couple of important factors—not listed in Bushman and Anderson's tables—that may explain why some scholars continue to erroneously believe there are more published demonstrations of violent media effects than published null effects. First among these is the peer-review process. In submitting an article to a journal (and, in particular, to an *APA* journal), there is a strong probability that at least one reviewer will be from the "harm" perspective. There is now the risk that, referencing Bushman and Anderson's paper, one of the items listed in their tables will be used to reject the study. If, and when, this were to occur one would be (and frequently is) left with the impression that it is the result which is being rejected and not the experimental data. For example, most of the items raised in these lists can be used to disqualify a study either because it did or did not generate an effect. To give a concrete example of this in action, we recently had a paper rejected on the advice of the reviewer who, after arguing they had no allegiance to Bushman and Anderson or the GAM theory, stated "the experiment was a failure. There were no differences in anger and aggression as a function of the four media types."

Bushman and Anderson (2021) also spend a substantial proportion of their paper on what they see as the misuse of covariation. Indeed, researchers have been discussing the issue of covariation in all areas of research (e.g., Miller & Chapman, 2001). Not accounting for co-varying effects can be used to reject a study, usually accompanied by the phrase "this study contains a fatal flaw…." Yet accounting for co-varying effects can also be used to discount a study by virtue of multicollinearity as they represent in their Figure 1. Under such circumstances, we fear for an independent reviewing process, one where an action editor is placed in the invidious position of rejecting a paper because it did or did not account for other variables in the measure of aggression. It is our experience that this can leave the editor in an awkward position, with a reviewer encouraging the rejection of the paper or, as some of the present authors have directly experienced, requesting a full replication after two rounds of reviewing—effectively stopping publication of the data.

There is a risk that a peer-review process that draws heavily on the (potentially biased) application of Bushman and Anderson's factors of concern may lead to the over-representation of the "harm" perspective upon which Bushman and Anderson rely in their previously described argument from authority. Note, any bias in a reviewer's interpretation and application of Bushman and Anderson's factors of concern may be entirely nonconscious. The fact remains that many of these factors can contribute to Type I or Type II error. Moreover, Bushman and Anderson note that some factors listed in their Table 1 might be threats to validity in some studies and "theoretically interesting variable(s)" in another. If the threats in Table 1 are selectively applied to account for null effects, they will contribute to a bias in the published literature—one that could be based on the study conclusion rather than the study data.

The Good Researcher

We agree with Bushman and Anderson (2021) that suspicion of study aims is well worth measuring: There is a long history in psychology on the effects that a "good participant" can have on study results (e.g., Orne, 1962) and they are right to be alert to this bias. Indeed, "demand effects" while being a *researcher* on one of Anderson's or Bushman's experiments may also be strong. With both academics having devoted almost their entire career to violent gaming and only ever found negative and aggressive effects, their faculty webpages/links leave little to the imagination of enquiring students as to the laboratory zeitgeist on the effects of violent video games. We would like to build upon Bushman and Anderson's point by also cautioning against the bias that "good research assistants" can bring to study outcomes.

This "Rosenthal effect" has been shown to affect a chief investigator's assistants even when the chief investigator does not verbally communicate the direction or nature of the bias (Rosenthal et al., 1963). Bushman and Anderson argue for a suspicion measure of the experiment's goals for participants, but it is also important to keep research assistants and supervised students neutral to outcome. Preregistration of data analysis plans, including explicit preregistration of relevant exclusions, controls, and all moderation/mediation analyses will help to limit unidentified researcher degrees of freedom. Otherwise, to alter the analysis strategy based on expected outcomes greatly increases the chances of Type-1 error rates (Simmons et al., 2011). It is noteworthy that unidentified researcher degrees of freedom have been shown to be associated with false-positive rates of more than 60% (Simmons et al., 2011), and the vast majority of preregistered studies in the area of violent video game effects appear to yield nonsignificant effects (e.g., Hilgard et al., 2019; Przybylski & Weinstein, 2019).

The Retracted Studies

We wonder whether these demand factors might be contributing to the growing number of published retractions of work published by those who see violent media as harmful (e.g., Benjamin & Bushman, 2018; Çetin et al., 2016; Whitaker & Bushman, 2017; Zhang, Espelage, & Rost, 2018; Zhang, Espelage, & Zhang, 2018), as well as editorial concerns for the integrity of conclusions (e.g., Benjamin et al., 2018). This is worrying because even replications of retracted papers conducted by the original authors do not seem to replicate the initial results (see Markey, 2018, for an informed discussion of such an example). There are many reasons why a paper gets retracted, but demand effects and the "good researcher" role, as well as the requirement for many top journals only to publish papers with significant effects (publication bias), are among the most likely. However, it may also be procedural, such as using nontransparent procedures, materials, or data. This leads on to our suggestions later in this paper.

If we cannot replicate our own studies (which is occurring for all academics in this area on both sides of the argument—although to differing degrees—see Devilly et al., 2021) then it is likely that the effects are slight, easily disturbed by third variables and, likely as not, open to change. Linking these retractions to Bushman and Anderson's later claim that "It is difficult to interpret the magnitude of any effect size in the absence of theory. Fortunately, there are many theoretical reasons why exposure to violent media should increase aggression" it should be self-evident that this argument is unstable and tautological.

Effect Sizes

There is a further problem with the logic of the effect size argument made by Bushman and Anderson. We agree that effect size estimates are, in many ways, more important than simple probability levels. Probability can be affected by having too much power (where trivial differences become significant due to a very large sample size, leading to the need for equivalence testing) or not enough power (where large differences are not seen as significant due to a small sample size). This can render point estimates unstable and result in both Type-1 and Type-2 errors. However, Bushman and Anderson make the assumption that the distribution of effect sizes in violent video gaming research on aggression is likely to be normal. However, the claimed schism in the literature suggests that one could conceivably see a bimodal distribution in effects.

Meta-Analytic Problems

Considering the likely problems in the peer-review process and researcher demand effects, both outlined above, we are not confident that there is a normal distribution around a true effect size that is represented in the literature. The meta-analysis that Bushman and Anderson rely upon to derive their effect size of $d = 0.4$ was the one conducted by Anderson et al. (2010). This has been criticized for the omission of studies that found effects opposite to their team and the inclusion of studies that misused the competitive reaction time task (Ferguson & Kilburn, 2010; see also Hilgard et al., 2017 for a largely failed replication of this meta-analysis). Further, most other meta-analyses find much smaller effects (e.g., Ferguson, 2007, 2015; Furuya-Kanamori & Doi, 2016; Hilgard et al., 2017; Prescott et al., 2018; Sherry, 2001), with best-practice longitudinal studies, for instance, finding effects statistically no different from zero (e.g., Drummond et al., 2021).

Contrary to Bushman and Anderson's implication that questionable research practices might account for null effects, recent evidence (Drummond et al., 2021; Ferguson & Hartley, 2022) suggests a positive relationship between the presence of questionable research practices and the size of reported effects. Such a result is also more parsimonious with the history of scientific research—errors tend to cause effects, not stability.

Further, there is a large "file draw" problem in this domain due to the peer-review problems outlined above. This problem has already been well demonstrated for the Anderson et al. (2010) meta-analysis (Hilgard et al., 2017), so this concern is no longer in doubt. There may also be unpublished studies with catharsis effects from gaming that have fallen foul of the peer-review process and, if so, these would lead to even fewer studies needed to reduce the average effect.

Suggestions

We offer several basic proposals in our conclusion to help move this field forward toward better science.

First, we believe that from now on journals should support the use of preregistration of research designs, hypotheses, methods, materials, and statistical analyses for all studies. We understand that this is more difficult for student research and nonfunded research and suggest that latitude is applied to these cases. However, our observation is that preregistered studies are far less likely to find effects than non-preregistered studies, suggesting that researcher publication bias or expectancy effects may contribute to false-positive results in this literature (see for instance, Drummond et al. 2021; Hilgard et al., 2019; Przybylski & Weinstein, 2019). This mirrors findings throughout psychology's "replication crisis" more broadly, with one study finding preregistered trial median effect sizes significantly lower than in trials that were not preregistered ($d = 0.16$ vs. $d = 0.36$; Schäfer & Schwarz, 2019). This approach of trying to control for investigator degrees of freedom could also include preplanned multiverse analyses (e.g., Steegen et al., 2016), particularly when study analyses have not been registered first.

Second, there is an urgent need for a clear smallest effect size of interest to be established for interpreting results as hypothesis supportive (e.g., Anvari & Lakens, 2021). Current evidence suggests that effect sizes below $r = .10$ and possibly even below $r = .20$ (note the difference above between effect sizes in Schäfer & Schwarz, 2019) are very often due to methodological noise rather than real effects. Effect sizes in this range should no longer be interpreted as hypothesis supportive (Ferguson & Heene, 2021). Our argument here is not that there should be a specific effect size for publication but, quite the opposite, null or small effects should be published where these occur and they should be interpreted in light of other interfering factors, and in comparison to the size of other relevant and irrelevant effects (see for instance Orben & Przybylski, 2019).

Third, the use of covariates should be guided by clear theory. Covariates such as family environment, trait aggressive personality, mental health, gender, and peer influences are well-known important covariates throughout aggression research (Savage, 2004) and should be employed wherever applicable and possible in video game research. Bivariate correlations or effect sizes should not necessarily be interpreted as meaningful on their own as they are likely to be spuriously inflated. This is because uncontrolled confounders can inflate effects (e.g., households with poor parenting might have increased aggression as well as fewer controls on playtime). Alternatively, when covariates are not necessarily systematically related to the independent variable, they may reduce the validity/precision of the point estimates (i.e., act as statistical noise).

Fourth, journals should consider adopting an open peer-review process where possible. This would require nonblind and signed peer reviews from reviewers, as well as the publication of the peer-review reports alongside the article if there is major disagreement. This system would increase the transparency of the reviewing process, and the accountability of reviewers to ensure that they were prepared to publicly stand by their comments solicited within the peer-review process.

Studies into the effectiveness of open peer review have found that open peer reviews have similar quality (van Rooyen et al., 1999) or modest benefits to the quality of the paper compared to when reviews are closed (see Bingham et al., 1998). As noted by Smith (1999), the main argument for open peer review is an ethical one: "a court with an unidentified judge makes us think immediately of totalitarian states and the world of Franz Kafka" (p. 4) while, in contrast:

> …identifying the reviewer links privilege and duty, by reminding the reviewer that with power comes responsibility: that the scientist invested with the mantle of the judge cannot be arbitrary in his or her judgment and must be a constructive critic.… All editors have seen curt, abusive, destructive reviews and assumed that the reviewer would not have written in that way if he or she were identifiable. (Smith, 1999, p. 5)

DeCoursey further notes that "Although manuscripts are rarely reviewed by a single reviewer, anonymous review does offer unscrupulous reviewers more opportunities for blocking publication without repercussion." The gaming research field is, unfortunately, characterized—in both scientific and non-academic communities—by rhetoric and emotionally laden discourse. Examples include likening playing violent video games to smoking cigarettes (e.g., Anderson, 2003, 2004) and the "no harm" researchers to tobacco industry scientists denying that smoking causes lung cancer (Anderson, 2004, 2019). We feel it is now necessary that open reviews be conducted to improve the validity and temper of the publication process. The overriding arguments against this process are that: Obtaining reviewers willing to self-identify is difficult; reviewers may be less critical and the science will suffer accordingly, and the approach has been unsuccessfully tried in the past. However, we argue that much has changed since the 1990s when this approach was last seriously pursued. Recent analysis of open peer-review systems within one system (*F1000Research*) has identified that:

> some weak evidence that being based in the same country as an author may influence a reviewer's decision, while there was insufficient evidence to conclude that being able to read an existing published review prior to submitting a review encourages conformity. (Thelwall et al., 2021, p. 809)

We argue that the open-source (and open-science) movement has grown since those days and younger, more open researchers are attaining positions of influence and responsibility. Indeed, open peer review is now predominantly seen by researchers as an umbrella term that may encompass many different approaches to the review process, such as named reviewers, published reviews, and representative reviews (see Ross-Hellauer, 2017). It is our hypothesis that these emerging leaders of the field are less territorial, have grown up with self-identified social media and, with an increase in support for and participation in open-science practices, we might expect to see an increase in open peer-reviewing. One could even imagine a metric where open and blind reviews receive differential weighting (as long as this does not build upon any existing academic hegemony), or where opposed blind reviews that diverge dramatically from open peer reviews are published alongside the article. At the very least, bringing to the editor's awareness that the area is contentious may help create the correct environment for a more impartial (and, possibly, signed) review.

We also need to distinguish the specific claims of any one theory. For example, in creating the current research we found that our team had very different understandings of the General Aggression Model

(GAM). It seems to us that, following various beta versions of the GAM in the late 1990s, there is the GAM Version 1.0 (GAM 1.0; Anderson & Bushman, 2002) and this does not postulate personality and individual differences as being *significantly* involved in the gaming/violence nexus. Following a few "point upgrades" across the literature, there is the GAM Version 2.0 (GAM 2.0; Anderson & Bushman, 2018) which does accept that some aspects of personality and individuality *can* contribute to aggression following gaming. However, the GAM 2.0 also seems to claim that "all things can contribute in all possible ways." Specificity in predictions is the key to falsificationism and is the underlying principle of science. We, therefore, welcome in the reply to our comment here any indications as to what evidence (exactly) would invalidate the GAM 2.0.

Perhaps the greatest gesture of scholarship is when two or more academics with differing results visit each other's laboratory and conduct joint research. We believe it is time that this occurs within the gaming research arena, incorporating Anderson and/or Bushman. The lead, second and third authors here suggest that a collaborative research funding application with cross-national research arms be attempted. This should help clarify whether the differences between the two sets of results are differences in method, sample, or approach. This, coupled with a triple-blind research protocol with incremental data being held by a third party, would be a state-of-the-art approach to understanding why we are obtaining such wildly different interpretations of our natural world. One could even include a third arm with a scientist who has no historical baggage in the area and is widely respected and seen as independent.

Conclusion

In conclusion, we argue here that appeals to authority for any scientific position are weak, there are multiple theories that explain the data more specifically than any interpretation of the various GAM releases, and there are multiple ways that methodological and analytic issues can be intentionally or unintentionally misused to stifle dissent to a given position. We believe that this is important because we have large discrepancies between many researchers and the authors of the original article in their study outcomes and interpretation. We have suggested open review, specificity in predictions, clarification around effect sizes of interest, preregistration of studies, and a joint research protocol to clarify methodologies.

References

Anderson, C. A. (2003). *Violent video games: Myths, facts, and unanswered questions*. American Psychological Association Science Brief. http://www.apa.org/science/about/psa/2003/10/anderson

Anderson, C. A. (2004). An update on the effects of playing violent video games. *Journal of Adolescence*, *27*(1), 113–122. https://doi.org/10.1016/j.adolescence.2003.10.009

Anderson, C. A. (2009). *FAQs on violent video games and other media violence*. http://www.craiganderson.org/wp-content/uploads/caa/Video_Game_FAQs.html

Anderson, C. A. (2019). *Do violent video games have negative effects on teenagers?* Metafact. https://metafact.io/factcheck_answers/1832

Anderson, C. A., Berkowitz, L., Donnerstein, E., Huesmann, L. R., Johnson, J. D., Linz, D., Malamuth, N. M., & Wartella, E. (2003). The influence of media violence on youth. *Psychological Science in the Public Interest*, *4*(3), 81–110. https://doi.org/10.1111/j.1529-1006.2003.pspi_1433.x

Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual Review of Psychology*, *53*(1), 27–51. https://doi.org/10.1146/annurev.psych.53.100901.135231

Anderson, C. A., & Bushman, B. J. (2018). Media violence and the general aggression model. *Journal of Social Issues*, *74*(2), 386–413. https://doi.org/10.1111/josi.12275

Anderson, C. A., Gentile, D. A., & Buckley, K. E. (2007). *Violent video game effects on children and adolescents: Theory, research, and public policy*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195309836.001.0001

Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., Rothstein, H. R., & Saleem, M. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in eastern and western countries: A meta-analytic review. *Psychological Bulletin*, *136*(2), 151–173. https://doi.org/10.1037/a0018251

Anvari, F., & Lakens, D. (2021). Using anchor-based methods to determine the smallest effect size of interest. *Journal of Experimental Social Psychology*, *96*,104159. https://doi.org/10.1016/j.jesp.2021.104159

Bandura, A. (2001). Social cognitive theory of mass communication. *Media Psychology*, *3*(3), 265–299. https://doi.org/10.1207/S1532785XMEP0303_03

Begg, C. B. (1994). Publication bias. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 399–409). Russell Sage Foundation.

Benjamin, A. J., & Bushman, B. J. (2018). Retraction notice to 'The weapons effect' [*Curr. Opin. Psychol.* 19 (2018) 93–97]. *Current Opinion in Psychology*, *22*, 96. https://doi.org/10.1016/j.copsyc.2018.08.005

Benjamin, A. J., Kepes, S., & Bushman, B. J. (2018). Effects of weapons on aggressive thoughts, angry feelings, hostile appraisals, and aggressive behavior: A meta-analytic review of the weapons effect literature. *Personality and Social Psychology Review*, *22*(4), 347–377. https://doi.org/10.1177/1088868317725419

Bingham, C. M., Higgins, G., Coleman, R., & Van Der Weyden, M. B. (1998). The Medical Journal of Australia internet peer-review study. *Lancet*, *352*(9126), 441–445. https://doi.org/10.1016/S0140-6736(97)11510-0

Breuer, J., Scharkow, M., & Quandt, M. (2015). Sore losers? A reexamination of the frustration-aggression hypothesis for colocated video game play. *Psychology of Popular Media Culture*, *4*(2), 126–137. https://doi.org/10.1037/ppm0000020

Brown v. Entertainment Merchants Association. (2011). 08 U.S. 1148. Retrieved from http://www.supremecourt.gov/opinions/10pdf/08-1448.pdf

Bushman, B. J., & Anderson, C. A. (2021). Solving the puzzle of null violent media effects. *Psychology of Popular Media*. Advance online publication. https://doi.org/10.1037/ppm0000361

Bushman, B. J., Gollwitzer, M., & Cruz, C. (2015). There is broad consensus: Media researchers agree that violent media increase aggression in children, and pediatricians and parents concur. *Psychology of Popular Media Culture*, *4*(3), 200.

Çetin, Y., Wai, J., Altay, C., & Bushman, B. J. (2016). Effects of violent media on verbal task performance in gifted and general cohort children. *Gifted Child Quarterly*, *60*(4), 279–287. https://doi.org/10.1177/0016986216660382

Cioffi, F. (1985). Psychoanalysis, pseudoscience and testability. In G. Currie, & A. Musgrave (Eds.), *Popper and the human sciences* (pp. 13–44). Martinus Nijhoff Publishers.

Copenhaver, A., & Ferguson, C. J. (2018). Selling violent video game solutions: A look inside the APA's Internal notes leading to the creation of the APA's 2005 Resolution on violence in video games and interactive media. *International Journal of Law and Psychiatry*, *57*(March-April), 77–84. https://doi.org/10.1016/j.ijlp.2018.01.004

Devilly, G. J., Brown, K., Pickert, I., & O'Donohue, R. (2017). An evolutionary perspective on cooperative behavior in gamers. *Psychology of Popular Media Culture*, *6*(3), 208–221. https://doi.org/10.1037/ppm0000097

Devilly, G. J., Callahan, P., & Armitage, G. (2012). The effect of violent videogame playtime on anger. *Australian Psychologist*, *47*(2), 98–107. https://doi.org/10.1111/j.1742-9544.2010.00008.x

Devilly, G. J., O'Donohue, R. P., & Brown, K. (2021). Personality and frustration predict aggression and anger following violent media. *Psychology, Crime and Law*, 1–37. https://doi.org/10.1080/1068316X.2021.1999949

Drummond, A., Sauer, J. D., Ferguson, C. J., Cannon, P. R., & Hall, L. C. (2021). Violent and non-violent virtual reality video games: Influences on affect, aggressive cognition, and aggressive behavior. Two pre-registered experiments. *Journal of Experimental Social Psychology*, *95*,104119. https://doi.org/10.1016/j.jesp.2021.104119

Ferguson, C. J. (2007). Evidence for publication bias in video game violence effects literature: A meta-analytic review. *Aggression and Violent Behavior*, *12*(4), 470–482. https://doi.org/10.1016/j.avb.2007.01.001

Ferguson, C. J. (2015). Do angry birds make for angry children? A meta-analysis of video game influences on children's and adolescents' aggression, mental health, prosocial behavior, and academic performance. *Perspectives on Psychological Science*, *10*(5), 646–666. https://doi.org/10.1177/1745691615592234

Ferguson, C. J., & Colwell, J. (2017). Understanding why scholars hold different views on the influences of video games on public health. *Journal of Communication*, *67*(3), 305–327. https://doi.org/10.1111/jcom.12293

Ferguson, C. J., Copenhaver, A., & Markey, P. (2020). Reexamining the findings of the American Psychological Association's 2015 Task force on violent media: A meta-analysis. *Perspectives on Psychological Science*, *15*(6), 1423–1443. https://doi.org/10.1177/1745691620927666

Ferguson, C. J., & Hartley, R. D. (2022). Pornography and sexual aggression: Can meta-analysis find a link? *Trauma, Violence and Abuse*, *23*(1), 278–287. https://doi.org/10.1177/1524838020942754

Ferguson, C. J., & Heene, M. (2021). Providing a lower-bound estimate for psychology's 'Crud factor': The case of aggression. *Professional Psychology, Research and Practice*, *52*(6), 620–626. https://doi.org/10.1037/pro0000386

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*(2), 174–178. https://doi.org/10.1037/a0018566

Ferguson, C. J., Rueda, S. M., Cruz, A. M., Ferguson, D. E., Fritz, S., & Smith, S. M. (2008). Violent video games and aggression: Causal relationship or byproduct of family violence and intrinsic violence motivation? *Criminal Justice and Behavior*, *35*(3), 311–332. https://doi.org/10.1177/0093854807311719

Furuya-Kanamori, L., & Doi, S. A. (2016). Angry birds, angry children, and angry meta-analysts: A reanalysis. *Perspectives on Psychological Science*, *11*(3), 408–414. https://doi.org/10.1177/1745691616635599

Gentile, D. A., Lynch, P. J., Linder, J. R., & Walsh, D. A. (2004). The effects of violent video game habits on adolescent hostility, aggressive behaviors, and school performance. *Journal of Adolescence*, *27*(1), 5–22. https://doi.org/10.1016/j.adolescence.2003.10.002

Haverford College. (2015). *Plagiarism and data manipulation*. http://ds-wordpress.haverford.edu/psych2015/projects/chapter/plagiarism-and-data-manipulation/

Hilgard, J., Engelhardt, C. R., & Rouder, J. N. (2017). Overstated evidence for short-term effects of violent games on affect and behavior: A reanalysis of Anderson et al. (2010). *Psychological Bulletin*, *143*(7), 757–774. https://doi.org/10.1037/bul0000074

Hilgard, J., Engelhardt, C. R., Rouder, J. N., Segert, I. L., & Bartholow, B. D. (2019). Null effects of game violence, game difficulty, and 2D: 4D digit ratio on aggressive behavior. *Psychological Science*, *30*(4), 606–616. https://doi.org/10.1177/0956797619829688

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, *23*(5), 524–532. https://doi.org/10.1177/0956797611430953

Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos, & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91–195). Cambridge University Press. https://doi.org/10.1017/CBO9781139171434.009

Limelight Networks. (2019). *The state of online gaming—2019*. Retrieved from https://www.limelight.com/resources/white-paper/state-of-online-gaming-2019/

Lull, R. B., & Bushman, B. J. (2016). Immersed in violence: Presence mediates the effect of 3D violent video gameplay on angry feelings. *Psychology of Popular Media Culture*, *5*(2), 133–144. https://doi.org/10.1037/ppm0000062

Markey, P. (2018). *Web post: The failed replication of a retracted study*. The 100% CI. http://www.the100.ci/2018/09/27/the-failed-replication-of-a-retracted-study/

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, *110*(1), 40–48. https://doi.org/10.1037/0021-843X.110.1.40

Orben, A., & Przybylski, A. K. (2019). The association between adolescent well-being and digital technology use. *Nature Human Behaviour*, *3*(2), 173–182. https://doi.org/10.1038/s41562-018-0506-1

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*(11), 776–783. https://doi.org/10.1037/h0043424

Oswald, M. E., & Grosjean, S. (2004). Confirmation bias. In R. F. Pohl (Ed.), *Cognitive illusions. A handbook on fallacies and biases in thinking, judgement and memory* (pp. 79–96). Psychology Press.

Pelham, B. W., & Blanton, H. (2018). *Conducting research in psychology: Measuring the weight of smoke*. Sage Publications.

Popper, K. (1962). *Conjectures and refutations. The growth of scientific knowledge*. Basic Books.

Prescott, A. T., Sargent, J. D., & Hull, J. G. (2018). Metaanalysis of the relationship between violent video game play and physical aggression over time. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(40), 9882–9888. https://doi.org/10.1073/pnas.1611617114

Przybylski, A. K., Deci, E. L., Rigby, C. S., & Ryan, R. M. (2014). Competence-impeding electronic games and players' aggressive feelings, thoughts, and behaviors. *Journal of Personality and Social Psychology*, *106*(3), 441–457. https://doi.org/10.1037/a0034820

Przybylski, A. K., Ryan, R. M., & Rigby, C. S. (2009). The motivating role of violence in video games. *Personality and Social Psychology Bulletin*, *35*(2), 243–259. https://doi.org/10.1177/0146167208327216

Przybylski, A. K., & Weinstein, N. (2019). Violent video game engagement is not associated with adolescents' aggressive behaviour: Evidence from a registered report. *Royal Society Open Science*, *6*(2),171474. https://doi.org/10.1098/rsos.171474

Quandt, T., Van Looy, J., Vogelgesang, J., Elson, M., Ivory, J. D., Consalvo, M., & Mäyrä, F. (2015). Digital games research: A survey study on an emerging field and its prevalent debates. *Journal of Communication*, *65*(6), 975–996. https://doi.org/10.1111/jcom.12182

Rosenthal, R., Persinger, G. W., Kline, L. V., & Mulry, R. C. (1963). The role of the research assistant in the mediation of experimenter bias. *Journal of Personality*, *31*(3), 313–335. https://doi.org/10.1111/j.1467-6494.1963.tb01302.x

Ross-Hellauer, T. (2017). What is open peer review? A systematic review. *F1000Res*, *6*, 588. https://doi.org/10.12688/f1000research.11369.2

Sauer, J. D., Drummond, A., & Nova, N. (2015). Violent video games: The effects of narrative context and reward structure on in-game and postgame aggression. *Journal of Experimental Psychology: Applied*, *21*(3), 205–214. https://doi.org/10.1037/xap0000050

Savage, J. (2004). Does viewing violent media really cause criminal violence? A methodological review. *Aggression and Violent Behavior*, *10*(1), 99–128. https://doi.org/10.1016/j.avb.2003.10.001

Schäfer, T., & Schwarz, M. A. (2019). The meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. *Frontiers in Psychology*, *10*, 813. https://doi.org/10.3389/fpsyg.2019.00813

Sherry, J. L. (2001). The effects of violent video games on aggression: A meta-analysis. *Human Communication Research*, *27*(3), 409–431. https://doi.org/10.1111/j.1468-2958.2001.tb00787.x

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Simmons, J. D., Nelson, L. D., & Simonsohn, U. (2021). Pre-registration: Why and how. *Journal of Consumer Psychology*, *31*(1), 151–162. https://doi.org/10.1002/jcpy.1208

Smith, R. (1999). Opening up BMJ peer review. *BMJ*, *318*(7175), 4–5. https://doi.org/10.1136/bmj.318.7175.4

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*, *11*(5), 702–712. https://doi.org/10.1177/1745691616658637

Thelwall, M., Allen, L., Papas, E. R., Nyakoojo, Z., & Weigert, V. (2021). Does the use of open, non-anonymous peer review in scholarly publishing introduce bias? Evidence from the F1000Research post-publication open peer review publishing model. *Journal of Information Science*, *47*(6), 809–820. https://doi.org/10.1177/0165551520938678

Unsworth, G., Devilly, G. J., & Ward, T. (2007). The effect of playing violent video games on adolescents: Should parents be quaking in their boots. *Psychology, Crime and Law*, *13*(4), 383–394.

van Rooyen, S., Godlee, F., Evans, S., Black, N., & Smith, R. (1999). Effect of open peer review on quality of reviews and on reviewers' recommendations: A randomised trial. *BMJ*, *318*(7175), 23–27. https://doi.org/10.1136/bmj.318.7175.23

van't Veer, A. E., & Giner-Sorolla, R. (2016). Pre-registration in social psychology—A discussion and suggested template. *Journal of Experimental Social Psychology*, *67*(November), 2–12. https://doi.org/10.1016/j.jesp.2016.03.004

Whitaker, J. L., & Bushman, B. J. (2017). 'Boom, headshot!': Effect of video game play and controller type on firing aim and accuracy': Retraction. *Communication Research*, *44*(1), 144. https://doi.org/10.1177/0093650217690274

Zhang, Q., Espelage, D. L., & Rost, D. H. (2018). Retracted: Short-term exposure to movie violence and implicit aggression during adolescence. *Youth and Society*. Advance online publication. https://doi.org/10.1177/0044118X18775846

Zhang, Q., Espelage, D. L., & Zhang, D. J. (2018). Retracted: The priming effect of violent game play on aggression among adolescents. *Youth and Society*. Advance online publication. https://doi.org/10.1177/0044118X18770309

Received November 8, 2021
Revision received October 11, 2022
Accepted October 12, 2022 ■