Dear Author

Please use this PDF proof to check the layout of your article. If you would like any changes to be made to the layout, you can leave instructions in the online proofing interface. Making your changes directly in the online proofing interface is the quickest, easiest way to correct and submit your proof. Please note that changes made to the article in the online proofing interface will be added to the article before publication, but are not reflected in this PDF proof.

If you would prefer to submit your corrections by annotating the PDF proof, please download and submit an annotatable PDF proof by clicking here and you'll be redirected to our PDF Proofing system.
Aggressive video games research emerges from its replication crisis (Sort of)

Christopher J Ferguson

The impact of aggressive video games (AVGs) on aggression and violent behavior among players, particularly youth, has been debated for decades. In recent years, evidence for publication bias, questionable researcher practices, citation bias and poor standardization of many measures and research designs has indicated that the false positive rate among studies of AVGs has been high. Several studies have undergone retraction. A small recent wave of preregistered studies have largely returned null results for outcomes related to youth violence as well as outcomes related to milder aggression. Increasingly, evidence suggests AVGs have little impact on player behavior in the realm of aggression and violence.

Nonetheless, most professional guild policy statements (e.g., American Psychological Association) have failed to reflect these changes in the literature. Such policy statements should be retired or revised lest they misinform the public or do damage to the reputation of these organizations.

Address
Department of Psychology, Stetson University, 421 N. Woodland Blvd., DeLand, FL 32729, United States

Corresponding author: Ferguson, Christopher J (cjfergus@stetson.edu)

Current Opinion in Psychology 2020, 36:xx–yy
This review comes from a themed issue on Cyberpsychology
Edited by Jon Elhai and Dmitri Rozgonjuk

https://doi.org/10.1016/j.copsyc.2020.01.002
2352-250X/© 2020 Elsevier Ltd. All rights reserved.

Introduction

Concerns about aggressive video games (AVGs) are nothing new. They began in the 1970s with cabinet games like Death Race and continue today into the virtual reality age. Policy makers have often sought to regulate these games, ultimately culminating in the Brown v EMA (2011) US Supreme Court decision which forbade such regulation, and which declared research evidence linking AVGs to aggression to be unconvincing. Scientific studies into AVG effects have existed for nearly the same time frame, beginning in the 1980s. Despite producing hundreds of studies, no consensus among scholars ever emerged regarding effects [1]. More recently, psychology has been experiencing a replication crisis and evidence suggests that AVG research has been part of that replication crisis.

AVG’s replication crisis

Recent meta-analyses converge on the observation that effect sizes in AVG research, regarding outcomes related to aggression and reduced prosocial behavior, are only marginally different from zero [2]. However, scholars are not certain whether such tiny effects should be interpreted as meaningful or attributed to systematic issues such as demand characteristics, questionable researcher practices, unstandardized measures, single-responder bias and poor matching of experimental and control conditions [3].

Studies typically fall within two broad types. The first of these are experiments wherein players are randomized to play either violent or non-violent games, then examined on some measure of aggression [4]. Confounds can occur in such experiments through multiple means. Violent and non-violent games may differ in features other than violent content such as difficulty, competitiveness and frustration. Unstandardized aggression measures may create false positives when scholars can select outcomes that best fit their hypotheses [5]. Obvious hypotheses can create demand characteristics, influencing participant behaviors in the direction of the hypothesis.

There have, at this juncture, been approximately a half dozen preregistered experimental studies of AVG effects. In addition to preregistration used to reduce questionable researcher practices, most such used standardized outcomes and sought to match video game conditions more closely than had been the case in the past. To date, no preregistered experiment of video game effects has produced findings supportive of the hypothesized effects [6. 7. 8. 9. 10]. To be clear, this is not to suggest there are no studies at all that link AVGs to aggression, but rather that when studies are preregistered, they tend not to find effects. This does suggest that many reported experimental findings are false positive results. Indeed, recent meta-analyses have confirmed that publication bias explains most of the variance in experimental studies [11]. This is consistent with research across social science which finds that preregistered studies produce weaker effect sizes than non-preregistered studies, suggesting a fairly dramatic impact of researcher expectancy effects and questionable researcher practices on research outcomes [12].
The second type of study includes correlational and longitudinal studies, wherein individuals, often youth, are tracked over time. Such studies examine whether playing AVGs at time 1 can predict aggression at time 2 (months or years apart). The more rigorous studies control for other variables including gender, time 1 aggression, mental health, trait aggression and family environment. Such studies are weakened by poor controls, questionable researcher practices, unstandardized measures and demand characteristics.

Correlational studies are numerous in number with a wide variety of outcomes. There are, at present, only two cross-sectional studies that are preregistered. One found no evidence for a relationship between AVGs and aggression in youth [13]. The other did find evidence for a cross-sectional association between AVGs and some (but not all) aggression outcomes in college students [14]. This latter study is, in fact, the only preregistered study to find any evidence for a relationship between AVGs and aggression.

With longitudinal studies, effects are more definitive. Most longitudinal studies find trivial to no relationships between early AVG exposure and later aggression or decreased empathy [15,16,17,18,19]. There are currently only two preregistered longitudinal analyses, both of preexisting datasets. Neither of these found evidence that AVGs predict aggression or decreased prosocial behavior [20,21]. However, more preregistered longitudinal studies would definitely be welcome.

Thus, we can see although many earlier studies did find relationships between AVGs and aggression, such outcomes are largely explained by publication bias and researcher degrees of freedom including questionable researcher practices. Several other issues point to the problem of the replication crisis in AVG research. One is that questionable researcher practices, such as recomputing AVGs variables differently across different publications using the same dataset, have been well documented [13]. The other is that some studies have specifically been retracted [22,23]. The intent here is not to be unduly critical of those studies. Good-faith mistakes can always happen, and authors should be commended for responsibly correcting the scientific record. However, it is plausible that a pressure to produce volumes of significant findings may create perverse incentive structures that make false positive results more likely [24].

Sexualization in games

From the AVG literature we can see several patterns. First, statements of effects by some scholars and professional guilds such as the American Psychological Association (APA) and American Academy of Pediatrics (AAP) often do not match the actual available evidence. Second, there are a number of methodological and science culture issues that elevate the risk of false positive results. Third, preregistered studies are less likely to produce hypothesis supportive findings.

A parallel area of research has developed regarding the potential impact of sexualized video games (SVGs) on aggression or hostility toward woman. What constitutes a ‘sexualized’ video game is notoriously ill-defined (as is ‘violent’ video game) but can be broadly construed as games wherein characters are presented in hypersexual, commodified ways, in which their role in the game is reduced to their sexual value. Most concerns are regarding female characters and impacts on male aggression or female body dissatisfaction.

Compared to the AVG field, the SVG field is much smaller and the involvement of preregistration much more limited. Some early studies did not consider actual gameplay at all, but simply showed slides of game characters [25]. Others are difficult to interpret. For instance, one study suggested women playing a sexualized avatar with their own face increased rape myth acceptance whereas playing a sexualized avatar with a fictional character’s face decreased rape myth acceptance compared to a control group [26]. Methodological issues have also plagued this field. For instance, one study suggested playing sexualized games such as Grand Theft Auto might decrease male players’ empathy toward women [27].

However, a reanalysis of this study found not only were these results mistaken, but the claimed randomization to game condition had not taken place and game condition was irretrievably conflated with participant age (younger players mainly played Grand Theft Auto (GTA)) [28].

Other recent studies have, for the most part, also not found evidence for effects for SVGs on player aggression toward women correlational [29], longitudinally [30] or experimentally [31,32]. To be fair, these findings are not universal, and some studies do find negative effects [33]. Further, some studies are simply difficult to interpret. For instance, one study [34] found that playing sexualized video games was unrelated to hostile sexism (e.g., traditional misogyny) but was correlated with benign sexism (i.e., believing women are morally superior to men or deserving of male protection). However examining the benevolent sexism items, they appeared unassignable such that answering them in the direction of ‘benign sexism’ respondents might simply have been trying to avoid hostile sexism (for instance, ‘Women are more moral men’ an affirmative response could be benign sexism, or it could be simply trying to avoid implying women are less moral than men by saying ‘no’).

Unfortunately, to date, there is only a single preregistered study in this field [35]. This study considered the effects of sexualized games on female players’ body dissatisfaction and aggression toward other women. No effects for
sexualized game play were found. Although this finding replicates similar outcomes for preregistered studies in the AVG realm, more preregistered studies would obviously be welcome.

The way forward
To summarize, early bodies of literature were mixed, but often nonetheless misinterpreted as suggesting clear links between AVGs and aggression. More recent preregistered work has largely clarified that earlier findings were likely spurious and exaggerated the impact of AVGs via publication bias, citation bias and researcher expectancy effects. I close by suggesting some approaches that may help ensure that AVG research is more completely extricated from psychology’s replication crisis.

Preregistration
As indicated, studies which are preregistered likely offer a more rigorous test of hypotheses and tend to produce outcomes that differ from non-preregistered studies. Authors should consider preregistration, including analyses for secondary data [36]. At minimum, preregistration should indicate expected sample size, all hypotheses, all materials and all planned statistical analyses. Any deviations from the preregistration should be openly discussed in the final paper.

Standardization
The use of unstandardized aggression measures has resulted in unreliable results and undoubtedly contributed to AVG’s replication crisis [37]. Future research should focus on standardized laboratory measures [e.g., 38] and clinically validated measures such as the Child Behavior Checklist for survey studies.

Game matching
In experimental studies, most control games have substantially differed from experimental games on qualities other than violent content which introduces significant confounds [39]. Providing perfect matching between games can be difficult, but researchers need to be more diligent in their matching on story, characterization, difficulty, competitiveness and potential for frustration. Simplistic games such as Tetris should never be used as control games.

Multivariate controls
Survey studies should control for theoretically relevant third variables including gender, trait aggressiveness, family environment and mental health. Mor aspirationally, genetics could be controlled where available [40]. Bivariate correlations are upwardly biased and tend to capture spurious variance that has little to do with AVGs.

Betas rule metas
One weakness of meta-analyses is their common reliance on bivariate correlations which, as noted above, makes them spurious evaluators of research fields. Reliance on bivariate correlations results in significant overconfidence in meta-analytic results supporting a particular hypothesis. The common rationale for use of bivariate correlations is that they are more homogeneous than standardized regression coefficients, although an examination of bivariate data in AVG research finds this, in fact, is not true [41]. Many scholars now advocate for the use of standardized regression coefficients (which adjust for variable explained by other theoretically relevant variables . . . for instance boys play more AVGs and are more aggressive, thus it is important to control for gender) in meta-analysis as these tend to provide a clearer picture of whether hypotheses are truly being supported in research fields than do bivariate correlations [42,43,44].

The meaningless of mean effect sizes
Many meta-analyses tend to interpret weighted mean effect sizes as indicative of population effect sizes, in effect declaring ‘the average effect size wins!’ This is a poor and unjustified use of meta-analysis. Mean effect sizes from meta-analyses are now known to significantly inflate effect sizes and do not approximate population-level effects [45]. Because of publication bias, systematic methodological flaws and questionable researcher practices, weighted mean effect sizes tend to be upwardly inflated, for example, the Garbage-In, Garbage-Out (GIGO) phenomenon of meta-analysis. Meta-analyses are more useful in examining how methodological issues such as standardization, attempts to reduce demand characteristics, citation bias, and so on, can influence effect sizes. In the future, meta-analyses could compare preregistered versus non-preregistered AVG studies.

Retirement and moratorium on professional guild statements
Evidence has now clarified that professional guild policy statements such as by the APA and AAP were based on selective interpretation of evidence, citation bias, and grossly misrepresent the AVG and other Media effects fields [46]. These statements should no longer be cited as credible. Current statements such as the APA’s 2015 video game task force should be retracted or retired, and professional guilds should refrain from further statements until more preregistered studies become available.

Conclusions
Arguably, interest in the notion that AVGs contribute to significant violence or aggression in society is waning, as is typical for Media-based moral panics [47]. Although some scholars certainly continue to argue for the potential harmfulness of AVGs [e.g., 48] this increasingly appears to be a minority view [1]. AVG research provides a cautionary tale in how the moral valence of a topic and lax scientific standards can create significant misinformation. Further, AVG research also illustrates how principles
4 Cyberspsychology

of preregistration and open science can extricate a research field from its replication crisis.

Funding source
There is no funding source to report for this article.

Conflict of interest statement
Nothing declared.

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

• of special interest
• of outstanding interest


This preregistered study found that neither randomized exposure to game content, nor purposefully chosen exposure to game content was associated with aggression. This study attempted to assess whether a contrast would be found for randomized play versus freely chosen play, but no games influenced aggressive behavior.


This preregistered study of teens found no evidence that video game content increases aggression. However, teen girls found more violent games to be stressful.


This preregistered study finds that activity choices including violent game play can cathartically reduce stress and hostility after a stressful task. This article challenges the notion that cathartic activities lead to greater aggression.


This meta-analysis suggests that previous experimental findings related to AVGs were largely the product of publication bias, not real effects. This calls into question the results of a previous meta-analysis.


This article compared effect sizes from numerous non-preregistered and preregistered studies in social science. Effect sizes for non-preregistered studies on average were far higher ($r = .38$) than for preregistered studies ($r = .16$) suggesting a dramatic impact for questionable researcher practices on outcomes in social science research.


This registered report cross-sectional study found no evidence to link AVGs to aggression in youth. This was the first registered correlational study to examine relationships between AVGs and youth aggression.


This preregistered study did find associations between AVGs and aggression in college students. Results were not always consistent, but did indicate some correlations between both overall video game play and AVG game play and some but not all outcomes related to aggression.


Longitudinal study finds no evidence early video playing is associated with later aggressive behaviors or reduced prosocial behaviors among youth.


Please cite this article in press as: Ferguson CJ: Aggressive video games research emerges from its replication crisis (Sort of). Curr Opin Psychol (2020). https://doi.org/10.1016/j.copsyc.2020.01.002
Aggressive games and aggression in youth Ferguson

Longitudinal study finds no link between early AVG playing and later aggression across multiple measures. This study also found little evidence that playing sexualized games was related to sexual aggression. 18. Etchells PJ, Gage SH, Rutherford AD, Munafò MR: Prospective investigation of video game use in children and subsequent conduct disorder and depression using data from the Avon Longitudinal Study of Parents and Children. PLoS One 2016, 11(10). This longitudinal study found only trivial relationships between earlier shooter game exposure and later conduct disorder. This game suggests that shooter game play has little impact on the development of clinically significant aggression.


20. Ferguson CJ: A preregistered longitudinal analysis of aggressive video games and aggressive behavior in Chinese youth. (in press) Psychol Pub Media Cult 2020. Preregistered longitudinal analysis finds: evidence early exposure to AVG predicts later aggression among Chinese youth. 1340 youth were involved. Aggression was consistent over time and predicted by moral disengagement, but not AVG use.


28. Ferguson CJ, Donnellan MB: Are associations between “sexist” video games and decreased empathy toward women robust? A reanalysis of Gabbardini et al. 2016. J Youth Adolesc 2017, 46(12):446-2459 https://doi.org/10.1007/s10964-017-0700-x. This reanalysis of a quasi-experimental study found no relationship between playing Grand Theft Auto and player empathy toward women. Serious mistakes including a failed experimental randomization were also found in the dataset.


Correlational study finds no relationship between exposure to sexualized video games and sexist attitudes or decreased empathy toward rape victims. Using sexualized games was, if anything, associated with more positive attitudes among users higher in trait aggressiveness.


Experimental study finds that playing sexualized games reduces rape myth acceptance.


Experimental study finds playing sexualized video games decreases rape myth acceptance over time.


This study considered sexualized and non-sexualized variants of the Tomb Raider game to examine how sexualization impacted women player’s body dissatisfaction and aggression toward other women. All elements of the study were preregistered. Results indicated that sexualization had no impact on either women’s body dissatisfaction or aggression.


This article discusses the importance of open science practices, including preregistration. In particular, this article discusses how these practices can be applied to secondary datasets, that is, studies already conducted with data generally available to all researchers.


41. Ferguson CJ: Pay no attention to that data behind the curtain: on angry birds, happy children, scholarly squabbles, publication bias and wet betas rule metas. Perspect Psychol Sci 2015, 10:683-691.

This paper examined many common standards to the use of standard regression coefficients in meta-analyses for AVGs. Analyses of several data sources indicated that the assumptions that the variance...
Correlations are better suited for meta-analysis than standardized regression coefficients are false. It is argued that standardized regression coefficients are superior to bivariate correlations for use in meta-analysis.


This meta-analysis of AVG studies confirms that standardized regression coefficients are superior to bivariate correlations when meta-analyzing studies. Overall effects of AVGs were trivial in size.


This article finds that meta-analyses tend to significantly overestimate population effect sizes due to significant methodological issues within the studies included in the meta-analysis. This is an illustration of the GIGO (garbage in, garbage out) phenomenon.

