Does Doing Media Violence Research Make One Aggressive?

The Ideological Rigidity of Social-Cognitive Theories of Media Violence and a Response to Bushman and Huesmann (2013), Krahé (2013), and Warburton (2013)

Malte Elson and Christopher J. Ferguson

Department of Communication, University of Münster, Germany, Department of Psychology, Stetson University, DeLand, FL, USA

We thank Bushman and Huesmann (2013), Krahé (2013), and Warburton (2013) for taking the time to comment on our review of digital game violence research. This is obviously a field where considerable controversy continues to exist and an opportunity for cordial debate could help resolve differences in the field. The current crop of comments run the gamut from keeping to reasonable points of disagreement (Warburton), to relying on sometimes snide comments (“The art of omission,” Krahé), to the comment by Bushman and Huesmann which makes use of ad hominem attacks. They also tend to restate old arguments that have been discredited either in our review or other past reviews (e.g., Adachi & Willoughby, 2011; Hall, Day, & Hall, 2011a, 2011b). Some of these statements confirm our initial concerns that the debate on violent media effects has shifted from science to ideology.

In the following, we examine some of the statements by our colleagues that we find to be problematic or misleading.

Theoretical Perspectives on the Effects of Violent Games

Bushman and Huesmann start out by explaining that there is abundant evidence for the harmful effects of observing violence in the home, at school, in the community, or in the culture on children. They refer to studies investigating the development of aggression through observation of violence in the Israeli-Palestinian conflict (Boxer et al., 2013), or the effects of exposure to rocket attacks on distress and violence (Henrich & Shahar, 2013). Then they turn to the audience by asking a rhetorical question (p. 3): “How, then, could viewing violence in the mass media not be harmful to children?” They ask for a psychological theory that explains how the risk of violence is increased by observing violence in the home, school, community, or culture, but not by observing it in the media.

We, in turn, might ask what one thing has to do with the other. Bushman and Huesmann imply two things here: (1) The experience or observation of real acts of violence is qualitatively similar to those of fictional violence. This is quite a stretch given that children aged five or younger are already able to distinguish between real and fictional television (Wright, Huston, Reitz, & Piemiat, 1994). (2) Observing proximal acts of violence (e.g., in the family) has similar psychological effects as watching virtual displays of violence (e.g., in a digital game). This argument is certainly not supported by the clinical practice of this comment’s second author, in which he has regularly observed the devastating influence of family violence on children, but cannot think of a single case in which watching Woody Woodpecker or playing Call of Duty was the root of a child’s mental health suffering. This argument can be dismissed by considering a simple, but obviously unethical and illegal hypothetical experiment. Take 200 children and randomize 100 to watch their parents viciously...
attack one another for a hour a day, the other 100 to watch a violent television program an hour a day, then assess their mental health after one month is over. Although this experiment will obviously always remain in the realm of the hypothetical (one hopes at least), to suggest the mental health outcomes for these children would be even remotely identical is absurd.

Bushman and Huesmann then respond to our review of explanatory models and theories. Apparently, their biggest criticism of the Catalyst Model that we discuss as one theoretical perspective is that it is not “new.” This claim is simply nonfactual. One might take a look at the reference list of our review article to find well over 30 publications in which research questions and results are explicitly explained with the General Aggression Model (GAM; Anderson & Bushman, 2002). Conversely, neither of them cites any existing evidence or perspectives opposing their own views in their comments (other than a select few to criticize). As one example, readers might compare Dr. Bushman’s section of a recent National Science Foundation (NSF) report on gun violence (Subcommittee of Youth Violence, 2013) to a review of media violence almost simultaneously released by Common Sense Media (CSM, 2013), a traditionally anti-media watchdog group. Although we do not agree with the conclusions of CSM, we admire the latter as a model for making an honest and balanced argument for the “harm” perspective. We invite readers to consider these two contemporary reports side-by-side, and examine their differences.

They reject our criticism of the GAM mainly on three grounds: First, that the GAM is a social-cognitive model intended to be used by aggression researchers, not clinicians in the area of pathological aggression. This is in stark contrast with the degree to which these authors themselves generalize the GAM to criminal behavior (Huesmann, 2013) and clinical realms (see, once again, Subcommittee of Youth Violence, 2013). Bushman and Huesmann underline the historical importance of the GAM for the field of aggression research. We are certainly well aware of its impact (although we do note it is seldom used in fields outside of media effects such as criminology), which is why we presented it as one theoretical approach to media violence effects research in our literature review. However, we are uncertain where exactly (and why) the GAM draws the line between the general kind of aggression that researchers are supposedly interested in, and the pathological aggressiveness only clinicians are concerned with. It also does not reflect recent developments in clinical psychology and diagnostics (such as the DSM-5; American Psychiatric Association, 2013), which tend to understand more and more disorders as a spectrum, and not as distinctively “pathological” or “non-pathological.” Proponents of the GAM should explicate the subgroups of the population and the kinds of behaviors covered by the GAM, and those “too pathological.” We believe, however, that this could severely limit the significance and relevance of the GAM, as the general public is concerned with antisocial, violent, or criminal behaviors, rather than higher-order cognitive processes.

We endorse Warburton’s call for empirical research that examines the boundaries of media effects, and would like to take this one step further to theoretical and conceptual work. We believe that the strength of every theory or model lies within the differentiated and accurate description of its boundaries and limitations. For the GAM, however, we are currently observing the opposite, as there are attempts to extend it to other areas, such as suicide and global warming (DeWall, Anderson, & Bushman, 2011). Nevertheless, we welcome Warburton’s suggestion and see this as a potential starting point for a fruitful debate: What kind of processes can be explained by the GAM, and for which behaviors are other theories or models (such as the Catalyst Model) more suited? Or, indeed, can we advance our understanding of media effects research by reconsidering theoretical models altogether? Is it time to replace social-cognitive theories of media effects such as the GAM with more user-driven media theories such as Uses and Gratifications (Sherry, Lucas, Greenberg, & Lachlan, 2006)?

The second argument is that there are strong effects of social-cognitive processes and priming on behavior, for which Krahé cites the famous study by Bargh, Chen, and Burrows (1996) as one example. While discussing the mechanisms of such cognitive activations is beyond the scope of this response, we point to recent failed direct and conceptual replications of the Bargh et al. study (Doyen, Klein, Pichon, & Cleeremans, 2012) and others (Pashler, Coburn, & Harris, 2012; Shanks et al., 2013), which put the robustness of priming effects on behavior into question. Thus, we see a different similarity between the fields of social priming and media violence than Krahé: Both have been considered “received wisdom” and are now experiencing replication crises against which their proponents appear to be extremely defensive. Far from “easily refuting” (p. 5) our comments, Krahé’s observation reinforces our concerns about the overstatement of mixed research results.

Third, Bushman and Huesmann argue that social learning through observation of violence in the school, home, or community is a strong predictor of aggressive and violent behavior. It is wrong to assume that we are denying the effectiveness of social learning, which is also why, in fact, family and peer violence is an important risk factor in the Catalyst Model. This misconception probably stems from the fact that we consider violent games (and other media) to be rather limited as “teachers” or models compared to other influences (e.g., parents). The GAM certainly does not discern between different modes of observation (virtual vs. real) nor different types of observed violence (fictional vs. nonfictional). It fails to explain how observations of fictional violence in virtual worlds (e.g., soldiers fighting)
Empirical Evidence and Methodological Issues

The perspectives on empirical evidence by the commenters can be divided into two sections. The first regards the appropriateness and validity of common methodological procedures in the research. They defend the Competitive Reaction Time Task (CRTT) in particular on the ground that it supposedly has a high experimental or psychological realism. However, neither of them cites any study supporting this argument for the CRTT. Conversely, they do avoid mentioning the study by Mitchell (2012) which revisits the “truth or triviality” issue and presents a gap between the results of laboratory and field studies in aggression research (and other areas of psychological science). There is no evidence for Warburton’s claim that participants actually believe they can hurt another with the CRTT, which is certainly a requirement for its validity. We remain skeptical whether this proves true. The CRTT noises are certainly unpleasant, but they are far from being harmful. Moreover, actually harmful noise blasts would severely limit CRTT’s further use in laboratory experiments for ethical reasons.

Krahé claims that we quote from one unpublished paper to support our criticism of the CRTT, when in truth we are actually citing at least four publications which discuss its lack of standardization and issues in reliability and validity (Ferguson & Rueda, 2009; Ferguson, Smith, et al., 2008; Savage, 2004; Tedeschi & Quigley, 1996), plus one study showing that there is no link between habitual violent media exposure and CRTT scores (Krahé et al., 2011). The study by Breuer, Elson, Mohseni, and Scharkow (2012) that Krahé refers to supports concerns about CRTT’s psychometric objectivity. Instead of providing evidence for the CRTT’s external validity, Bushman and Huesmann refer to the example of prisoners being punished with loud music at Guantanamo Bay. To us, however, this torture scenario, involving non-consenting prisoners exposed to hours upon hours of sleep depriving noise, only vaguely resembles the CRTT with its brief exposure and ostensibly consented opponents in university laboratories. Equating the CRTT to torture seems, to us, one more example of the irresponsible overreach to which this field has become accustomed.

Krahé also criticizes us for avoiding the study by Giancola and Parrott (2008) in support of its validity. In fact, we did not cite this study because it is largely irrelevant to the argument. First, Giancola and Parrott used electrical shocks, not any of the numerous noise-burst variants which are the norm in media effects research. To the best of our knowledge, the electroshock version of the CRTT has never been used in experiments on violent game effects. Instead, media effects researchers usually rely on non-painful noise-bursts instead. We remain skeptical as to whether these two very different stimuli should be equated. Second, the study does in no way provide any evidence for external or construct validity, but only for its group discrimination validity (intoxicated vs. sober participants with a high self-reported propensity toward aggression), which as Tedeschi and Quigley (2000) state is not only weak evidence for validity, but also might easily constitute a logical fallacy. Giancola and Parrott also do not address the issue of lacking standardization, which is certainly an issue for the CRTT’s psychometric objectivity. None of the comments cited any evidence to alleviate our concern that there is a lack of validity data for the CRTT. Therefore, the CRTT should not be generalized to significant real-world aggression. Unfortunately, all too often, scholars do exactly this.

Krahé rejects our call for rather using clinical measures of aggression because they would be inappropriate for community samples of children, adolescents, and adults. We are puzzled by this statement given that such measures are, in fact, normed on community samples and perfectly valid for use with all children. The avoidance of measures that actually document harm, and the reliance on unstandardized, non-validated laboratory procedures such as the CRTT, is particularly problematic when, once again, a field appears very willing to generalize its results to clinical outcomes of “harm to children.”

Krahé underlines the importance of “experimental research in artificial settings (. . .) with a tighter control of confounding or distorting factors that would affect behavior under natural conditions” (p. 7), and we wholeheartedly agree: Proper experimental laboratory research is key to understanding basic relationships between variables. Consequently, however, one should only generalize these findings to less artificial and more naturalistic situations with due care. This is something we see repeatedly disregarded in effects research on violent games. This observation also assumes that those confounding factors in game effects experiments have been properly controlled, which
has been questioned as well (Adachi & Willoughby, 2011; Elson, Breuer, Van Looy, Kneer, & Quandt, in press).

Warburton’s major argument is to put the findings on violence in games into context of other media effects research, with the effects of advertisement on food choice as an example. Warburton states that those comparisons should only be made unless there is “a valid reason why different psychological mechanisms would underlie those effects” (p. 3). Given the title of his comment, “Apples, oranges, and the burden of proof,” we must point to the many differences between the psychological mechanisms underlying the role of advertising and fictional media.

Briefly, to work, advertising need only nudge behavior slightly, from one product to another. Advertisement works particularly well when most people are already motivated to purchase the advertised good, such as food, and not in generating desire for something they have no reason to buy in the first place. Choosing one brand over another is a low-impact choice, while choosing to behave aggressively is not. It does not require the sorts of fundamental changes to motivation or personality often suggested as the outcome of media violence. Second, advertising always relates to the everyday life of the consumer, and thus purports to be “true” (although it often is not). Fictional media rarely resemble everyday experiences of their users, and they only seldom attempt to appear non-fictional. Unlike fictional media, advertisement also addresses the consumer directly by promising a “better life” (e.g., faster cars, healthier food). It is quite reasonable to speculate that advertising works hard to circumnavigate “fiction detectors” in ways fictional media does not. Indeed, equating advertising to fictional media involves the problematic assumption that our brains are incapable of distinguishing fictional from non-fictional sources of media. Indeed, the lab of the second author, applying usual skepticism, has found evidence for advertising effects (Ferguson, Muñoz, & Medrano, 2012) but has been unable to find evidence for many fictional media effects. Thus, calls to equate advertising and fictional media should be rejected as too simplistic.

One reason for Warburton’s suggestion of an overall consistency in scientific findings on violent media effects might be his misconception of the meaning of p-values in null hypothesis significance testing (NHST). p-Values do certainly not determine whether an effect is “real” or not, as Warburton claims (p. 11). His description implies several common misinterpretations, such as p being the probability that the H0 is false, and that statistically significant differences are always relevant. NHST and p-values are not a test of “reality”? We clarify: The p-value is the probability of the observed result or more extreme results under the assumption that the null hypothesis is true. Given that problems with p-values and NHST (and common misconceptions thereof) have been debated for more than 70 years in psychology (and other sciences), and are beyond the scope of this comment, we recommend Cohen (1994) for a further look.

Meta-Analyses

The second argument regards previously published reviews and meta-analyses of the empirical evidence. Bushman and Huesmann criticize us for not conducting a meta-analysis to test the Catalyst Model, and instead supposedly using an “informal vote-counting” approach. We are not sure where they got that impression. The structure in each section of our empirical review is similar: We first present empirical studies investigating the simple relationship between violence in games and aggressive outcomes (thoughts, emotions, and behaviors). Then, we put these findings into the context of other studies looking at other factors potentially influencing relevant (in-)dependent variables. We are, hence, not “missing the point of moderation” (see Krahé’s comment, p. 6), but making it: Researchers must take these third variables (or “distorting factors,” see above) into account to avoid issues in validity when studying the effects of violence in games.

We would like to point out several issues with the meta-analysis by Anderson et al. (2010) that Bushman and Huesmann and Warburton discuss in their comments. First, the “average effect size wins!” approach potentially conceals failed replications. Second, the “best practices” defined by Anderson et al. (2010) are rather ambiguous and nontransparent. For example, studies were rated on whether “the outcome measure was appropriate for testing the hypotheses” (p. 9), without providing a clear definition of what can be considered “appropriate” (such as sufficient standardization and validation). Moreover, due to the common use of unstandardized and unvalidated measures (as discussed earlier), a meta-analysis certainly has the same limitations as the studies it includes, an issue Anderson et al. (2010) failed to consider in their “best practices.” Also, the authors make much mention of Dr. Rothstein’s comments on meta-analysis without specifically noting that she was a coauthor of the Anderson et al. (2010) meta-analysis, and thus not an independent commentator.

Further, and in some ways surprising to us, is the way in which Bushman and Huesmann describe how unpublished material was gathered for the Anderson et al. meta-analysis. It appears that they only asked a limited number of authors for unpublished data supplementary to publications they already identified, and that they did not solicit unpublished studies. In another paper by Rothstein and Bushman (2012), it is explicitly recommended to “include unpublished studies whenever it is possible” (p. 135) and that usually “authors of meta-analyses try to contact every author who has ever published an article on the topic of interest (…))” (p. 131) for unpublished material. We do not wish to speculate why, then, in the case of the Anderson et al. meta-analysis they did not fully heed their own advice. And while it might be true that only 2 out of 88 included papers that reported aggressive behavior were unpublished (although this is not mentioned in the meta-analysis itself), we would like to point out that Appendix A of Anderson
et al. (2010) includes at least 16 unpublished studies, and Appendix B 27 unpublished studies. Many of these are non-English technical reports or conference presentations that would be difficult to locate, making replication of the Anderson et al. meta-analysis unlikely. In an earlier comment on their meta-analysis, it was noted that the majority of these unpublished studies came from the authors of the meta-analysis or their collaborators (Ferguson & Kilburn, 2010). At the time, the second author happened to be in touch with the authors of Anderson et al. (2010) on other matters and at no time was asked for unpublished data, nor informed they were conducting a meta-analysis. Since many of these then unpublished data have subsequently been published, their existence and the failure of Anderson et al. to secure them are irrefutable.

We appreciate that Bushman and Huesmann are forthright in acknowledging that they made little effort to secure unpublished studies from a wider range of authors in the Anderson et al. meta-analysis, despite Rothstein and Bushman’s (2012) advice to do exactly this. This reinforces our concerns about the selection bias against scholarly groups questioning the “harm” belief in that study, which may have influenced its conclusions. Even in that meta-analysis, however, effect sizes tend to be truncated with even simple controls (longitudinal effects drop to a meager $r = .075$, for instance, controlling only for gender and time 1 aggression), a fact the authors often fail to note.

No Consensus on a Consensus

Bushman and Huesmann present some preliminary data from a survey, showing that surveyed members of the American Psychological Association’s (APA) Media Psychology and Technology Division, the International Communication Association’s (ICA) Mass Communication Division, and the American Academy of Pediatrics (AAP) largely agree that violent games increase aggression in children. This is somewhat contrasted by a survey on digital game researchers conducted by the ICA’s Game Studies Interest Group, the European Communication Research and Education Association’s (ECREA) Digital Games Research Temporary Working Group, and the Digital Games Research Association (DiGRA) (Van Looy et al., 2013). Their sample comprised 544 games researchers, of whom only 1.3% strongly agreed and 8.8% agreed that the effects of digital games on aggressive behavior are a problem for society (27% were undecided, 35.5% disagreed, and 27.6% strongly disagreed). The 64 respondents that indicated psychology as their research tradition did not significantly diverge from that overall results (1.6%; 9.4%; 31.3%; 29.7%; 28.1%, respectively). Thus, the results presented by Bushman and Huesmann were not replicated, and we find support for our initial statement that there is, in fact, disagreement among the research community. This has also been expressed recently in an open letter signed by more than 200 scholars that was sent to the APA, urging its task force on violent media to repeal strong claims made in previous policy statements, and to acknowledge the diverse opinions and perspectives that exist on media violence effects.

Despite our criticism of this kind of rhetoric, Bushman and Huesmann repeat the comparisons between the effect sizes of violent media on aggression and smoking on lung cancer, and justify it by stating that “calculations don’t lie” (p. 18). Or perhaps they do. The problems with the calculations made to support such conclusions have, by now, been well documented (Block & Crain, 2007; Ferguson, 2009) which they fail to mention. Even ignoring the problematic statistics underlying these comparisons, methodologies of media effects research and oncology are so drastically different that a comparison of the resulting effect sizes is invalid. If cancer studies would consist of participants smoking cigarettes for 5–10 min and then rating their cancer severity on a 5-point Likert scale, then yes, such analogies would be eligible. But fortunately, cancer research does not have the methodology or validity issues that media effects studies do. Ironically, tests for cancer have everything that currently employed aggression tests do not. They are standardized, they are clinically validated (according to the results, one either has cancer or not), and they have a high reliability and external validity (someone who has cancer in a laboratory also has it outside the laboratory). Unfortunately, the same cannot be said about measurements of aggression.

What is, perhaps, most disappointing about the comments of Bushman and Huesmann is that they spend so much time disparaging those who disagree with them. They attempt to resurrect the now-discredited Pollard Sacks, Bushman, and Anderson (2011), despite it had been debunked by scholars uninvested in either side of the debate (Hall, Day, & Hall, 2011b). Put simply, that Bushman and his colleagues should nominate themselves as the “true experts” is neither surprising, nor illuminating, and certainly not part of credible science. But let us imagine that their claims of publishing more than skeptics are true (despite Hall et al., 2011b). So what? The fact that at one point in time a certain belief is expressed in a majority of publications is hardly any proof for the validity of this opinion. Looking back at the history of psychology (and other sciences), there are many paradigms that once were particularly popular and influential, and then later regarded as insufficient or simply wrong. One might think of theories such as phrenology or humorism, or to give a more recent (and appropriate) example from psychological science, the cognitive revolution as a response to the once dominant radical behaviorism. All these theories have gone through a paradigm shift in which younger scholars overthrow the ideas of older scholars. We believe that violent media effects research is currently facing the same process.

References

1 Unpublished study meaning, according to Bushman, Rothstein, and Anderson (2010), “not published in a peer-reviewed journal, although it could have been published in another outlet.” (p. 182)

It appears to us that Bushman and Huesmann make claims of consensus simply by discounting anyone whom they disagree with. Indeed, toward the end of that piece that appears to include anyone who consumes “large amounts” (undefined, of course) of violent media, or even the authorship of novels (although the second author certainly appreciates the plug by Bushman and Huesmann). We invite readers to consider what a broad brush this is to paint with. They also discount the opinions of the US Supreme Court (Brown v. Entertainment Merchants Association, 2011) and presumably numerous lower courts as well as government reviews by Australia (2010) and Sweden (Statens Medieråd, 2011). Bushman and Huesmann incorrectly imply that scientific evidence had little to do with the US Supreme Court’s decision, despite that the majority decision made clear their (rightful) skepticism of the application of this research to a public health issue. They imply the Supreme Court accepted “industry arguments” ignoring that numerous amicus briefs were filed against the “harm” position, by scholars, attorney generals, legal scholars, and youth advocacy groups. Warburton also implies the Australian report was influenced by pressure from the “gaming lobby.” Thus, both comments blame any differences in opinion on the gaming community or media industry, instead of arguing what might have been wrong scientifically in these two reports.

It Is Time to Change the Culture of Media Violence Research

We wish to be clear that we are not against scholars making an argument linking media violence with aggression. Our concern remains that the culture of this field has evolved to tolerate sweeping statements equating weak research with public health crises and stifling any form of dissent. Indeed, while functioning as reviewers, we have seen examples of reviews which viciously attack findings which disagree with (Gentile, 2013) by demanding that “naysayers” to media violence effects should not be given “valuable (and undeserved)” public attention (Strasburger & Donnerstein, 2013, p. 3).

We find all these observations to be deeply troubling for the credibility of this field. Once again, to be clear, we believe that many scholars on all sides of this debate are doing good work and are dedicated to an open exchange of views, an openness that is at the heart of the scholarly enterprise. But we also observe that some scholars actively and aggressively attempt to quell dissenting views, disparage skeptics, question the motives of those who disagree with them, and enforce a highly ideological view of this field. We believe these efforts have done considerable damage to the scholarly enterprise and the reputation of this field (Hall, Day, & Hall, 2011a). This leads us to wonder what it is about doing aggression research that seems to make some scholars so aggressive. We hope that the majority of scholars will join with us, whatever their personal views may be, in rejecting such a hardline ideological approach to this field and allow it to return to a proper atmosphere of respectful exchange of ideas.

We also express some concern with what appears, to us, to be an overly mechanistic perspective on human learning. This is exemplified by Warburton’s comments on human learning, often expressed in language of rigid certitude (i.e., “It is known that. . .”). We certainly do not deny that humans learn and often learn socially, but we express concern that Warburton’s language has converted social learning from something humans can do to something they must do. Warburton also relies on the problematic area of neuroimaging (Vul, Harris, Winkielman, & Pashler, 2009) to support his conclusions in this regard, but overall we find this approach to human learning unsatisfying. To us it is just as important to understand when humans do not learn as when they do, how they make decisions about when to learn, and when to ignore a learning opportunity. Neglecting this in favor of a mechanistic “monkey see/monkey do” model (a metaphor actually used by Orue et al., 2011 to describe their results), to us, does not remotely begin to capture the subtlety and sophistication of human learning and human existence.

More fundamentally, it may be time for this field to consider serious changes in both theory and in communicating to the public. Several prominent media scholars recently headlined a panel entitled “Why don’t they believe us?” at the International Communication Association conference (Donnerstein, Strasburger, Viner, & Gentile, 2013). The most parsimonious answer to this question is, in fact, “Because the data are not convincing.” Much like psychoanalysts of ages past, media scholars have taken to constructing elaborate theories for why people have not accepted their theories, or even personally attacking those who disagree with them. We contend that the traditional media effects paradigm has failed for the simple reason that it does not comport its own predictions of societal developments. Current theories are arguably too mechanistic, assume viewers are passive receptacles of learning, rather than active shapers and processors of media culture. We do not believe data support the traditional paradigm. We do no less than call on scholars to move past the traditional media effects paradigm, and to an understanding of the interaction between media, behavior, and culture, that is shaped by media users, not media content.

Acknowledgment

The research leading to these results has received funding from the European Union’s Seventh Framework Programme (FP7/2007-2013) under Grant Agreement No. 240864 (SOFOGA).

References


Krahé, B., Möller, I., Huesmann, L. R., Kirwil, L., Felber, J., & Berger, A. (2011). Desensitization to media violence: Links with habitual media violence exposure, aggressive cogni-


