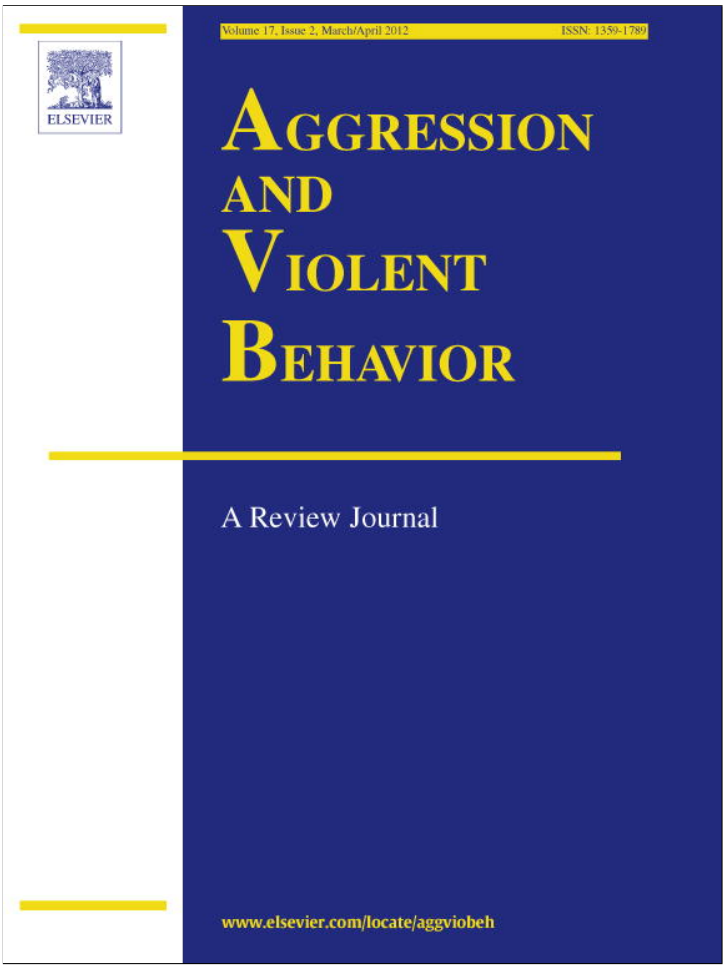
Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Aggression and Violent Behavior 17 (2012) 129–139

Contents lists available at SciVerse ScienceDirect

Aggression and Violent Behavior

Have recent studies addressed methodological issues raised by five decades of television violence research? A critical review

Christopher J. Ferguson [a,*], Joanne Savage [b]

[a] Department of Behavioral, Applied Sciences & Criminal Justice, Texas A&M International University, Laredo, TX 78045, United States
[b] American University, United States

ARTICLE INFO

Article history:
Received 14 October 2011
Received in revised form 20 November 2011
Accepted 29 November 2011
Available online 6 December 2011

Keywords:
Television
Mass media
Aggression
Violence

ABSTRACT

Media violence continues to be a concern to parents, policy makers and researchers. In spite of confidence expressed by some that exposure to television violence causes serious aggression and violent crime, critics hold that serious flaws in research methodology limits or nullifies conclusions drawn from widely-cited studies. In this paper, we will examine a series of classic studies for lessons learned about conducting media violence research, and assess whether recent publications adhere to those standards. We conclude that empirical reports with serious flaws continue to be published, compromising our ability to understand this phenomenon.

© 2011 Elsevier Ltd. All rights reserved.

Contents

1. Introduction

Television, as an entertainment medium, became widely available in the years after World War II. Almost immediately, programs with violent content such as westerns and police dramas became popular. For instance, in the peak year for westerns, 1959, 26 westerns were on the air, with 8 in the top 10 (Time, 1959). This rise in the popularity of television violence appeared to coincide with a rise in violent crimes in the United States lasting from the early 1960s through approximately 1993 (Federal Bureau of Investigation, 1951–2010). Around the same time, Albert Bandura published a series of experiments (e.g., Bandura, 1965; Bandura, Ross, & Ross, 1961, 1963) which suggested that young children model aggressive behavior after watching adults engage in such behavior on videotapes, ushering in the dominance of the social learning paradigm in the field of psychology. Arguably, these three elements, the exploding popularity

* Corresponding author. Tel.: +1 956 326 2636.
E-mail address: CJFerguson1111@Aol.com (C.J. Ferguson).

1359-1789/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.avb.2011.11.001

Author's personal copy

of television, the rise in violent crime beginning in the 1960s, and the increased influence of the social learning paradigm, led to strong concerns among social scientists that TV violence causes serious aggression.

The belief that television causes aggression and violence has been strongly held by many for a very long time. Scholarly reviews that draw this conclusion are too numerous to list here. In addition, health and policy organizations routinely issue statements supporting the conclusion. For example, as recently as 2000, the American Psychological Association co-signed a joint declaration of health organizations (American Academy of Pediatrics, 2000) which stated that research studies on television and other media violence "point overwhelmingly to a causal connection between media violence and aggressive behavior in some children." Nonetheless, some skeptical scholars (e.g., Felson, 1996; Fischoff, 1999; Freedman, 2002; Gauntlett, 1995; Grimes, Anderson, & Bergen, 2008; Moeller, 2005; Olson, 2004; Savage, 2004; Stipp & Milavsky, 1988; Trend, 2007) have stated that the research on television violence was never as strong or consistent as proponents have claimed, and that discussions of television violence research, including those by professional organizations such as the American Psychological Association and the American Academy of Pediatrics have often misled (although probably not intentionally) the public rather than informed it. Claims that the effects for television violence approximate those for smoking on lung cancer (Bushman & Anderson, 2001; Huesmann, 2007), and that television violence was a key factor in the 1970s–1993 violent crime increase (Bushman & Anderson, 2001; Centerwall, 1989), contrast wildly with the conclusions made by other reviewers who suggest that the effect is either nonexistent, or so small as to have no practical significance (Block & Crain, 2007; Ferguson & Kilburn, 2009; Gauntlett, 1995; Grimes et al., 2008; Olson, 2004; Savage, 2004; Savage & Yancey, 2008). For example, Savage and Yancey (2008) conclude, instead, that "Proposals to address the 'media violence problem' under the guise of reducing violent crime in our society are not likely to succeed. We would recommend exploring other policies for reducing violent crime" (p. 788). These critiques have become particularly pronounced as youth and adult violence rates have returned to 1960s levels without any abatement in violence on television.

Given these sharp differences in the interpretation of the findings of the past five decades of research, it would be helpful to elucidate what the principal methodological issues are for estimating the effects of media violence. In the present paper, we will review a select set of widely cited studies and highlight methodological concerns and standards emanating from each. Then we will turn to recent published studies to see if these lessons learned are currently being applied by reviewers and editors.

2. Experimental research

2.1. Demand characteristics

It is probably safe to say that few students pass through introductory psychology classes without learning of the Bandura studies and the purported importance of these studies to social learning theory and the belief that aggression in children is learned. The Bandura investigations, of course, were not remotely the first studies of media violence, as such studies are found at least as far back as the Payne Fund studies of movie violence (Blummer, 1933). Nor are the Bandura studies (e.g., Bandura, 1965; Bandura et al., 1961, 1963) technically media violence studies at all, as the depictions of adults hitting bo-bo dolls shown to children were removed of plot and narrative. However, in showing filmed depictions of adults engaged in person–object aggression, and demonstrating that children mimic these acts under some circumstances, it is clear that the Bandura studies had an enormous impact on general opinions and spawned further television violence research.

Bandura randomly assigned young children to watch a video of an adult engaged in aggressive behavior toward a non-human object (an inflated bo-bo doll, which is intended to be hit) or to a control condition. All children were then purposefully frustrated by having them view toys they could not play with before being placed in the room with the bo-bo doll. Children who saw the model behave aggressively toward the bo-bo doll in the film were more likely to act aggressively toward the bo-bo doll, and even imitated specific behaviors. In a later study, Bandura (1965) added a reward or punishment ending to the video, with the model being rewarded for engaging in aggression in one condition, and punished (ironically using person-on-person violence) in another condition. Children who viewed the punishment condition were far less likely to mimic aggression than those who viewed the reward condition.

This series of experiments is arguably one of the best known and most influential in psychology. However, some scholars (Ferguson, 2010a; Gauntlett, 1995; Tedeschi & Quigley, 1996) have commented that the bo-bo doll studies may not be as conclusive as is commonly believed among psychologists. One major problem is the demand character of the design. Aggressiveness is provoked through an artificial laboratory manipulation, and it is unclear whether the children's intent was to be aggressive. In the series of experiments, children are shown the video of the adult model acting aggressively. Then they are placed in the room with the bo-bo doll. In the absence of any further instructions, the children may have believed that the video provided them with instructions on what the adult experimenter wished them to do, particularly when placed near the very same bo-bo doll.

Further, the sole purpose of a bo-bo doll is for it to be hit in such a way as it will return, resulting in fun rough-and-tumble play. Given that a bo-bo doll is not only an object, but an object that is a salient prime for aggressive play (since it is not designed for cuddling or affectionate play), generalization of the bo-bo doll studies even to other person-on-object aggression is dubious, and generalization to person-on-person aggression is unjustifiable. Naturally, ethical restrictions preclude replacing the bo-bo doll with a live animal or another child. Yet, in choosing a bo-bo doll rather than a more neutral object, the bo-bo doll studies may have produced biased results by strongly priming aggressive acts.

In short, the bo-bo doll studies have not adequately distinguished rough-and-tumble play from true aggression (Tedeschi & Quigley, 1996). Given that most definitions of aggression (e.g., Baron & Richardson, 1994) focus on causing harm to another entity who wishes to avoid such harm, it is unlikely that a bo-bo doll paradigm is best for emulating the type of antisocial aggression we most wish to understand. Kniveton and Stephenson (1975) replicated the study and found that the modeling effects of bo-bo doll aggression did not generalize to interpersonal aggression.

Those who have not read the original Bandura reports may be unaware that the later bo-bo doll studies (Bandura, 1965) include videos involving not only person-on-bo-bo doll aggression, but an actual person-on-person assault. Some children were randomized to a punishment condition in which aggression toward the bo-bo doll was punished in the video through physical assault by the experimenter against the adult model. As described in the original text (Bandura, 1965. p. 591):

> For children in the model-punished condition, the reinforcing agent appeared on the scene [this occurs after the children watched the model hit the bo-bo doll] shaking his finger menacingly and commenting reprovingly, "Hey there you big bully. You quit picking on that clown. I won't tolerate it." As the model drew back he tripped and fell, and the other adult sat on the model and spanked him with a rolled up magazine while reminding him of his aggressive behavior. As the model ran off, cowering,

Author's personal copy

the agent forewarned him, "If I catch you doing that again, you big bully, I'll give you a hard spanking. You quit acting that way."

As is well known, children who viewed this condition, behaved less aggressive. This has two implications. First, the vicarious punishment of aggression is effective. This supports the criticism that experimental demand may be at work as it suggests that children are looking for information on how to interact with the toys. Second, the modeling effect is limited; if it were not, the children who saw the aggression against the bo-bo doll and the punitive assault against the model should have been doubly aggressive.

Very commonly textbooks of psychology, in discussing the Bandura studies, fail to note the limitations of these studies and often suggest that these studies "prove" aggression is or can be learned through imitation (e.g., Lilienfeld, Lynn, Namy, & Woolf, 2009; Myers, 2009). In spite of these serious criticisms, rather than portraying the Bandura studies as foundational rather than conclusive, they are still accepted as evidence of media violence effects.

After criticisms of demand characteristics of studies surfaced, researchers turned to clever empirical designs to avoid the problem. For example, Josephson (1987) arranged a game of floor hockey, rather than setting the subjects up for a contrived aggressive paradigm. Others turned to correlational research, which circumvents the problem entirely.

2.2. Operationalizing exposure to violent TV

Operationalizing exposure to violent TV in the laboratory seems simple enough: randomly assign half the subjects to view a violent film and half to view a nonviolent one of similar length (to control for the act of watching a film at all for a period of time) and compare the groups later. However, Zillmann and Geen pointed out that violent films have a special characteristic that might also be associated with aggression: they tend to be arousing or exciting. Zillmann (1991) points out that arousal is a force that energizes or intensifies behavior and that television can be highly arousing. Studies where experimental subjects are exposed to violence, and control subjects are exposed to something calm or boring, may report statistically significant differences between groups due to the differences in excitement or arousal elicited by the material, rather than the violent content itself (e.g., Geen, 1976). One example of a study criticized in this regard is that conducted by Friedrich and Stein (1973). They compared children who had watched Batman and Superman to those who watched Mr. Rogers. Another study compared children who watched violent action to a group who watched French comedies. Huston-Stein, Fox, Greer, Watkins, and Whitaker (1981) tested whether media portrayals of high action – high violence would elicit significantly more aggression than those with high action – low violence. While they their findings suggest that high-action was associated with aggression, there was no effect of violent content.

Findings by Feshbach and Singer (1971) had an interesting, unexpected outcome. They conducted a field experiment of boys living in several residential facilities. Within each facility, boys were randomly assigned to watch either a diet of nonviolent or violent programs over a course of six weeks. Aggressive behavior was rated by an adult guardian and included physically and verbally aggressive acts toward other people. The importance of this study, compared to many others, is that it was conducted in the field and actual, spontaneous aggressive behavior was observed, removing the demand characteristics of a laboratory study. The authors found that boys who watched violent programming were *less aggressive* than those who watched nonviolent programming. There were twice as many incidents of fighting, hitting, and kicking among controls than among the treatment group. This finding has been severely criticized on several grounds (Huesmann, Eron, Berkowitz, & Chaffee, 1992). Foremost is that it may be possible that boys in the nonviolent group became aggressive out of frustration at not being allowed to watch their favored shows, although the authors made significant attempts to reduce external sources of frustration, such as allowing the boys to drop out of the study. It might also be the case that the boys watching calm, boring shows became restive and were more prone to energetic (aggressive) play after the program was over. The criticism we highlight here, however, which was applied in many subsequent studies, is the fact that the nonviolent programs viewed by control subjects were less arousing than the violent programs. Matching the level of excitement of the violent film and the control condition became de rigueur due to this problem.

Finally, it is clear that some shows are more likely to be imitated, for fun, than others. For example Leyens, Parke, Camino, and Berkowitz (1975) showed children the violent film Zorro; other studies have used the Three Stooges. It has been argued if children in such studies then imitate the behaviors seen in these shows, it should not be counted as "aggression" given that the children may simply be imitating the adventuresome or silly aspects of these films, particularly when compared to "control" conditions that lack these aspects. In essence, some films may be imitated not because they are violent but because their narrative is more compelling to the participants in these studies than the films chosen as controls. As occurred in Bandura's Bo-Bo doll studies, the stimulus may have primed a specific behavior without any particular aggressive motivation or intent.

3. Interaction effects

Josephson (1987) presented an interesting experimental study of violent television exposure on young boys. In this study, 396 boys in second and third grade were randomly assigned to watch violent or nonviolent television. Afterward, the boys played floor hockey and observers took note of all instances of aggression during the game (e.g., pushing, hitting with a stick, tripping, and name calling.) The boys who watched the violent television were overall less aggressive than the other boys. However, the author also tested an interaction effect and found that boys high in trait aggression behaved more aggressively after watching the violent program than those who watched the control program. Further, the effect was enhanced if they were frustrated during the procedure. These results suggest that response to television violence may not be consistent across viewers, but rather depends upon the internal characteristics of the viewer. Subsequent studies would be remiss if they did not consider this possibility.

4. Correlational studies

4.1. Operationalizing exposure to media violence

Correlational research was undertaken to address the limited generalizability of experimental research. Many published articles report simple correlations between measures of violent television exposure and aggression. Because such correlational studies do not include manipulation of the independent variable or random assignment, it is possible that correlations might indicate that persons with an aggressive trait are more likely to choose to watch violent television, rather than television causing their aggression. Similarly, even gender effects may account for bivariate correlations, with boys both acting more aggressive and selecting more violent television than girls. This problem is exacerbated by a common problem in the early studies: the operationalization of exposure to media violence. Eron (1963) collected data for 367 boys and 322 girls in New York beginning in 1960. Subjects were asked what their three favorite shows were and these were rated as to whether or not they emphasized antisocial aggression. Correlations between this measure and indicators of aggression have been reported in numerous reports (e.g., Eron, 1963; Eron, Huesmann, Lefkowitz, & Walder, 1972; Huesmann, 1986).

Author's personal copy

Robinson and Bachman (1972) reported a large-scale survey study of young males (n = 2200). They asked subjects what their four favorite programs were and report a linear trend with individuals with low scores having the lowest average aggression scores and those reporting strongly favoring violent TV having the highest (no correlation coefficient was reported or statistical significance). McIntyre, Teevan, and Hartnagel (1972) used a very similar measure. They reported a "very small positive relationship" (p. 404), though a close look at their table suggests inconsistencies with the text. Although objective raters of violent content were usually used, so the authors could examine "preference for violent programs," the "preference for violent TV" or "favorite programs" measure has been criticized because it is thought to strongly reflect the *personality* of the subject. The amount of viewing is not taken into account, so the measure cannot be said to reflect the quantity of exposure to violent television (compare two children reported enjoying Bonanza; one child might watch it every week and the other child might have only seen it once).

This problem was remedied considerably in other studies which employed a measure of exposure to media violence that approximates the state of the art today. McLeod, Atkin, and Chaffee (1972) provided a checklist of prime time television shows and asked subjects to indicate how often they watched (e.g., almost always, often, sometimes, never). This removes the personality aspect of the "favorite programs" measure. Violent content was rated independently and was multiplied by the frequency of viewing to compute an overall indicator of violence viewing. Other correlational studies turned to this type of measure as well (e.g., Belson, 1978; Huesmann & Eron, 1986; Milavsky, Kessler, Stipp, & Rubens, 1982).

4.2. Establishing temporal order

The early correlational studies earlier mentioned were consistent in reporting significant correlations between indicators of exposure to TV violence and aggression (Savage, 2004). However, each additional study using indicators of similar constructs is no longer instructive, as temporal order is not established. It is clear that aggressive individuals prefer to watch violent programs more so than nonaggressive individuals, so a correlation only reiterates this knowledge. Interest in learning more about a causal connection between violent media exposure and aggression led some authors to undertake ambitious longitudinal studies. There have been only a few widely-cited ones to date (the studies published by Eron et al.; those published in the Huesmann & Eron, 1986 volume; Milavsky, Kessler, et al., 1982). These studies all used a wave-to-wave design which accomplished two things. First, it allowed them to use aggression from a later wave as a dependent variable and exposure to TV violence at an earlier wave as the independent variable, improving claims of temporal ordering. Because aggression and exposure to television violence is highly correlated over time, though, it is not enough to look at correlations from wave to wave. In addition, the new design allowed researchers to control for early wave trait aggression. The multivariate design, using a measure of aggression as the dependent variable, a good measure of exposure to violent media in an earlier wave as the independent variable, and controlling for a series of potential confounds, including trait aggression, continues to be the state of the art in the field for this type of research. As discussed later, these designs produce far less consistent evidence to support television violence/aggression effects.

4.2.1. Measuring aggression

In both laboratory and correlational studies, the appropriate measurement of aggression has been debated. That the concept of aggression is difficult both the define satisfactorily and to measure validly has been understood for some time (Benjamin, 1985; Hawley & Vaughn, 2003; Ritter & Eslea, 2005; Smith, 2007; Tedeschi & Quigley, 1996;2000; Van der Dennen, 1980). Van der Dennen (1980) provides a comprehensive discussion of definitions of aggression and notes that definitions of aggression tend to vary depending upon the theoretical perspective of the individual providing the definition. Arguably the definition provided by Baron and Richardson (1994) is reasonably representative of the social science understanding of aggression in which aggressive behavior is defined as intentional behavior produced to cause physical harm or humiliation to another person who wishes to avoid the harm. Nonetheless, this form of definition has been criticized as value-laden, specifically in the implicit understanding of aggression as inherently bad and nonadaptive at any level (Ferguson & Beaver, 2009; Hawley & Vaughn, 2003; Smith, 2007). In order to reduce the inherent value statements in common definitions of aggression, Ferguson and Beaver (2009, p. 387) proposed one in which implicit desirability was removed "behavior which is intended to increase the social dominance of the organism relative to the dominance position of other organisms." This definition reflects that aggression may have adaptive value, particularly at moderate levels.

Most definitions of aggression agree that the recipient of aggression must not be willing or consenting. As such, aggressive play, such as "cowboys and Indians," or sports activities, such as football, would not typically count as aggressive behavior, so long as both parties "play by the rules" as the parties involved are all consenting to the play whether or not it has aggressive overtones. Similarly, imitative aggressive play has been excluded in some studies. With this in mind, studies, such as the famous (or infamous) Milgram studies (Milgram, 1963), would appear as a remarkable approximation of measured aggression in a laboratory setting, as part of the experimental design involved communicating to participants that they were causing real physical injuries to the "learner" who repeatedly asked to be released from the electric shocks. However, ethical issues arising from the Milgram studies, at least in part (see Baumrind, 1964), have made the use of similar techniques prohibited. As a result, in most laboratory paradigms of aggression, measurement procedures, while ethically appropriate, but by necessity are distant from common definitions of aggression.

The validity of aggression measures used in television research has developed into its own contentious debate (e.g., Anderson & Bushman, 1997; Ferguson & Rueda, 2009; Giancola & Chermack, 1998; Giancola & Parrott, 2008; Ritter & Eslea, 2005; Tedeschi & Quigley, 1996, 2000). Meta-analyses of television violence reveal that most studies do not involve anything resembling physical aggression at all. For instance, measures of aggression in television violence studies have included outcome variables such as asking children if they wanted to pop a balloon (Mussen & Rutherford, 1961), asking college students if they would like to have a graduate student confederate (who had just insulted them) as an instructor in a course (Berkowitz, 1965), puffing up a blood-pressure cuff on a person's arm (Zillmann & Weaver, 2007), and applications of Tabasco sauce and chilly water to learners and opponents (see Ritter & Eslea, 2005; Tedeschi & Quigley, 1996 for further descriptions).

Supporters of laboratory aggression paradigms in particular have tended to argue that various aggression measures correlate with third variables that also correlate with aggression, or correlate with each other (e.g., Anderson & Bushman, 1997). However correlations between these laboratory measures of aggression and real-world behavior are wanting. The real question is whether these measures are highly correlated with the types of behavior to which the authors (and reviewers) wish to generalize their findings. For example, Anderson and Bushman (1997) cite sources and correlations between written and physical response measures in the lab, but when responding to questions about the association between lab measures and physical violence, they only comment that the "lack of surface similarity … leads to the overly pessimistic critiques" (p. 25). They emphasize that laboratory aggression measures are indicative of the participant's intention to harm the confederate and this can help us

Author's personal copy

understand the causes of violent crime. Giancola and Parrott (2008) report that the original electric shock measure may have some weak validity. However the validity depended upon self-report measures of anger and trait aggression, rather than against real-world violent acts. It is not clear that validating one set of questionable measures against a second set of measures of limited validity advances the arguments about validity very far.

What many media scholars fail to acknowledge is that serious, violent aggression in the real world is *very* different from laboratory aggression. In the real world, one is not told to harm other people — "press this button to shock the other person," "press this to punish the competitor with a noise blast." Instead, we are subject to frequent messages that hurting others is wrong. People are punished for hurting other people in the real world. There are laws against it and some penalties are severe. In the real world, when you hurt someone, it is likely that person or his associates will try to hurt you back. Finally, while laboratory aggression is designed to cause minimal harm to subjects, actually hitting someone with hand or object, or attacking a person with a weapon has unpredictable consequences. Our moral values deter us from behavior that might really harm another person. Even under highly provocative circumstances, people can almost always refrain from committing violence, yet in a laboratory, subjects are almost universally willing to obey directives to perform "aggressive" acts — a fact that demonstrates the distance between the two behaviors. Thus, intense scrutiny of the generalizability of these measures *are* justified when generalizations to serious aggression and violence are made, as they frequently are (Bushman & Anderson, 2001; Cantor & Wilson, 2003; Huesmann, Moise, & Podolski, 1997).

Survey instruments used in correlational research also do not correlate highly with real-world aggressive acts (Ferguson, 2009a; Gothelf, Apter, & van Praag, 1997; Henry & Metropolitan Area Child Study Research Group, 2006). For example, peer, teacher, and parent reports of aggression largely correlated less than $r = 0.2$ with actual observed physical aggression (Henry & Metropolitan Area Child Study Research Group, 2006). Ferguson and Rueda (2009) noted that these problems extend to often used surveys of "trait aggression." Surveys of "intention" to be aggressive or "violent attitudes" have obvious limitations. A popular but hotly debated measure is the peer-nominated measure of aggression. The measure enjoys very good reliability yet it is unclear whether it adequately taps the concept or aggression (or is susceptible to 'popularity contest' effects). Reviewers have questioned its generalizability to serious aggression or violence. The measure includes ten items of which only two clearly represent interpersonal aggression (see Table 1). Most of the items are rather benign forms of aggression (e.g., "Who often says 'Give me that'"; "Who often says mean things"). It is also possible that a "peer-nominated" measure of any negative outcome may function more as a popularity contest than a valid measure of aggression. The measure is used widely, in highly respected studies and in many publications, Huesmann et al. have reported that this measure is associated with exposure to television violence.

While the measure has its virtues, it is not a good proxy for serious aggression. The peer-nominated aggression has been found to have weak validity coefficients (Henry & Metropolitan Area Child Study Research Group, 2006). Huesmann, Eron, Lefkowitz, and Walder (1984), for example, report significant but low simple correlations between peer-nominated aggression and later spouse abuse (0.27), self-rated physical aggression (0.25) and other indicators of criminality based on their longitudinal data from New York State. Outcomes reported by Henry and Metropolitan Area Child Study Research Group (2006) were far more modest with peer nominated aggression displaying validity coefficients between 0.03 and 0.17 against observed aggressive behavior. Teacher reports and self-reports fared little better.

Finally, it is worth noting that, despite the claims of some scholars that laboratory measures of aggression are highly valid (e.g., Anderson & Bushman, 1997; Giancola & Chermack, 1998), we have looked repeatedly for evidence, tracking back from several citations, and found no correlations high enough to justify any of them as a proxy measure for truly violent behavior. Ferguson and Rueda (2009) suggested that aggression measures need to achieve consistent predictive validity results of $r = 0.3$ or preferably $r = 0.4$, a level which is reasonably consistent with recommendations for validity coefficients more broadly (Anastasi & Urbina, 1996). We are not aware of any laboratory measures of aggression (and few survey measures of aggression) which have consistently done so. Arguably even this standard is rather low. It is noteworthy that psychiatrists have not adopted these measures to predict violence risk in clinical or forensic settings despite the fact that many of these measures are less expensive and easier to administer than other clinical/forensic measures of violence risk currently in use. Yet, these measures are being presented to the scholarly community and general public as if they are equivalent to these very same clinically and forensically relevant violent outcomes.

Table 1
Items from the Lefkowitz, Eron, Walder and Huesmann measure of aggression.

1. Who does not obey the teacher?
2. Who often says, "Give me that?"
3. Who gives dirty looks or sticks out their tongue at other children?
4. Who makes up stories and lies to get other children into trouble?
5. Who does things that bother others?
6. Who starts a fight over nothing?
7. Who pushes or shoves other children?
8. Who is always getting into trouble?
9. Who says mean things?
10. Who takes other children's things without asking?

4.3. Controlling for aggressive trait and other "third factors"

An important issue for many correlational studies is the failure of these studies to control for third factors which are likely to be correlated with both exposure to violent media and aggression. For example, if children with violent parents are, as a result, exposed to lots of violent TV (because violent people favor violent fare), and also display highly aggressive behavior (due to exposure to actual violence), we would find a significant correlation between exposure to violent TV and aggression, unless we controlled for parental violence. Without controls for confounding factors, the coefficient for TV violence effect is said to be *biased* — larger or smaller than it should be. Savage and Yancey (2008) suggest that controls for SES, parent education, parent violence, neglect, and intelligence should be applied. All are likely to lead to differential television exposure and all are empirically associated with violent behavior.

Due to the lack of computer technology, prior to the late 1980s, reliance on bivariate correlations was common (Freedman, 2002; Gauntlett, 1995; Savage, 2004). Without control variables, a correlation between violent TV exposure and aggression only teaches us that aggressive people watch more violent TV. It does not establish causal order and it does not remove the possibility that the correlation might be due to the fact that aggressive persons, those in poverty, those who are neglected by their parents etc. are more likely to watch violent fare and to behave aggressively. The advent of computer technology enabled researchers to conduct multivariate analyses more easily and numerous analyses were published.

Lefkowitz, Eron, Walder, and Huesmann (1977) (and Huesmann, Moise-Titus, Podolski, & Eron, 2003), controlling for variables one by one, reported a significant association between preference for violent TV and aggression, although it is unclear why the variables were entered one by one, rather than together as would have been preferable. Milavsky, Kessler, et al. (1982) reported nonsignificant partial effects of exposure to media violence on aggression. Controlling for other factors, Huesmann and Eron (1986) found that exposure to television

Author's personal copy

violence was associated with later aggression for girls but not boys and Sheehan (1986) and Wiegman, Kuttschreuter, and Baarda (1985) found no association between TV violence exposure and aggression. In short, the relationship between exposure to media violence and aggression has been dramatically attenuated when other factors have been controlled. Savage and Yancey (2008) summarize effect sizes for models controlling for aggressive trait and SES; in the small number of studies where both are controlled simultaneously, effect sizes are generally small and not significant. Multivariate analyses controlling for confounding factors became the state of the art and its importance in producing unbiased results cannot be understated.

5. Model specification

Some cases studies report using good measurement procedures, and multivariate analyses, but employ poor statistical models. Two problems have been prominent. The first is the substitution of inferior measures for good ones, when the desirable measures were available. The second is the selection of control variables.

Some view the ambitious project led by Huesmann and Eron (1986) to be the most important study on the effects of television violence exposure on aggression, and it probably is. The study was conducted in six countries, including boys and girls in the United States, Australia, Finland, Israel (both a city sample and a kibbutz sample), Poland, and the Netherlands. The Dutch group published their results in a separate manuscript (Wiegman & Kuttschreuter, 1992) due to differences in opinion regarding how to analyze the data. This large-scale multinational study addressed several previous problems. First, its wave-to-wave design allowed the authors to both address the problem of temporal order and control for trait aggression. They also carefully measured actual exposure to media violence by asking subjects which programs they watched the most and having an independent rater rate the violence level of those shows. Having read the chapter in their book on methodology, the reader anticipates a model where later wave aggression is regressed onto earlier wave television violence exposure, controlling for earlier wave trait aggression, plus a host of other measured factors such as SES and parenting measures.

Unexpectedly, out of all six countries (including the Dutch), the authors reported the expected model for only U.S. boys and girls, Australian boys and girls, and children in the Netherlands. The Finnish study reported that the original models were not significant and then reported tables from a post hoc analysis (Lagerspetz & Viemero, 1986). The authors of the Israeli (Bachrach, 1986) and the Polish study (Fraczek, 1986) do not report the original model and only report models using alternative measures. For the Finnish sample and U.S. boys, the authors substituted a measure multiplying TV violence exposure by "identification with aggressive characters." For the U.S. boys sample, this was done as a follow-up, presumably to elucidate the nonsignificant findings for the original model. It is unclear why the authors for the Finnish sample report only the findings for this measure, not the more desirable direct measure of exposure to media violence. The Polish study reported only findings using "violence of preferred shows" rather than the measure weighted by frequency (Fraczek, 1986). These alternatives are much less desirable than the original because identification with aggressive characters, and preferences are really derivative of a personality characteristic that is highly correlated with aggressive behavior and as a result the influence of aggressive personality on aggression can no longer be teased out from the effect of television viewing habits in these reports (Savage, 2004). Coefficients are likely to be biased strongly in an upward direction when the indicator of exposure to TV violence is confounded with a measure of personality. Another unexplained change was the substitution of the outcome "ratio of aggression to avoidance of aggression" for the peer-nominated aggression measure which was used in the Israeli study (Bachrach, 1986).

The authors report statistically significant associations between media violence exposure and aggression, using these substitutions, and the reader is placed in a dilemma regarding the determination of what s/he has learned. The book editors accept the substitutions as equal to the original analyses, and proclaim in the conclusions chapter of the book that the results overall are supportive of the link between television violence exposure and aggression (Huesmann, 1986). The Dutch authors came to the opposite conclusion (Wiegman & Kuttschreuter, 1992) as did Savage (Savage, 2004; Savage & Yancey, 2008). Savage and Yancey (2008) compared effect size estimates of the "original" models and the "post hoc" models and find dramatic differences. Overall, the effect size for longitudinal studies where the multivariate model includes a control for earlier aggressive trait is only 0.038 (Savage & Yancey, 2008). They conclude that "there is not one study that reports the comparison we would really like to see" (p. 786) and that their estimates of average effect sizes, low as they are, are likely to be biased in an upward direction. Based on the evidence, they conclude that the association between media violence and serious aggression has not been established.

The importance of designing a study that adheres to the lessons of previous research, and following through by reporting an analysis that lives up to the plan, cannot be understated. Substituting less-valid measures for better ones makes little sense and failing to include control variables that were measured during the data collection process is unjustified.

6. The debate about the practical significance of television violence effects

Effect size estimates for television violence on aggression range from a low of approximately $r = 0.038$ (Savage & Yancey, 2008) or $r = 0.08$ (Ferguson & Kilburn, 2009) to a high of $r = 0.31$ (Paik & Comstock, 1994). Most other meta-analyses have produced estimates below $r = 0.2$ (Bushman & Anderson, 2001; Hogben, 1998; Wood, Wong, & Chachere, 1991). The high estimate produced by Paik and Comstock includes studies with any outcome of "antisocial behavior" which ranged from aggressive behavior against toys to criminal violence. Forty-five percent of the studies included self reports of behavior intention or performance on an aggression machine, 25% used interpersonal aggression during play or social interaction, and less than 10% involved illegal behavior (Paik & Comstock, 1994). Meta-analyses which limited their samples to criminal aggression, including a sub-analysis by Paik and Comstock themselves, report much lower effect sizes. Paik and Comstock's own estimate of the association between exposure to television violence and criminal violence is 0.10. Although the American Psychological Association has been clear in highlighting the importance of effect size interpretation in the analysis of statistical findings (Wilkinson & Task Force on Statistical Inference, 1999), much debate persists regarding how to interpret them. Cohen (1988) suggested $r = 0.1$ as a cut-off for "small effects" below which effects would be considered trivial, whereas Ferguson (2009b) suggested a more conservative $r = 0.2$, arguing that effects below this range were highly prone to type I error. Bushman and Anderson (2001) make the point that "small is big" arguing that if 1% of the population responds more aggressively after viewing a popular program, this translates to hundreds of thousands of people, despite that this is a rather dubious interpretation of effect size. It would be more correct to say that this 1% represents a 1% change in aggression within individuals, not a dramatic effect on 1% of the population. Further, this does not take into account the quality of the aggression measurement.

While Bushman and Anderson (2001) call to mind mass homicides in Jonesboro and Littleton, they have no evidence that demonstrates that exposure to media violence causes extreme violence of that nature. In fact, in decades of media violence research, where thousands of children and young adults have been shown violent

Author's personal copy

programming and, more recently, asked to play violent video games, we have seen no direct evidence that even one truly violent act has been caused by television violence exposure, either alone or in part. We have our own example of a low correlation. If you have a sample of 100 persons, and divide it into two groups of 50 who each get a score of 10 on some outcome, it takes only *one subject* in the experimental group to get a score of 11 (a 10% increase) to result in a correlation of $r = 1$. Given that a 10% increase in the aggression level of sitting on the couch watching TV is likely to be minimal, it is hard for us to envision a public health problem.

Most meta-analyses agree that television violence rates are in the range of trivial to small if one uses Cohen's recommendations and clearly trivial if one uses Ferguson's recommendations. For those interested in generalization to violent crime, Table 2 displays the effect size estimates from numerous meta-analyses related to various possible causes of violent crime. The estimate for television violence effects is among the smallest in the published literature.

The small effect size for television violence research, particularly with outcomes related to violence has been used to support strong claims for the importance of this variable. No scholar claims television violence is the only variable of importance, related to societal violence, of course, but some scholars certainly argue it is a critically important variable. In the abstract of one paper it is claimed that "Conservative estimates are that media violence may be causing 10% of real-life violence—not the leading cause by any means, but an unhealthy chunk that we could do something about if we chose to" (Strasburger, 2009). This 10% estimate is apparently based on squaring $r = 0.31$ from Paik and Comstock (1994). Note that this most extreme of all the meta-analytic estimates is seen by Strasburger as a "conservative estimate" and that this estimate is based on the widest array of measures of aggression, not the estimates related to criminal violence which are closer to $r = 0.1$ in the same study (Paik & Comstock, 1994). Similarly, Huesmann (2007, p. S6) recently stated "For better or worse the mass media are having an enormous impact on our children's values, beliefs, and behaviors." Bushman and Anderson (2001, p. 477) state "Modern society is exposed to a massive dose of violent media. What effect, if any, does this exposure have on people? In the 2Oth century, two major explosions occurred: a mass media explosion and a violent crime explosion" and later link the advent of television with the rise in violent crime beginning in 1965. Even if one were to accept the "small but meaningful" argument regarding the effect sizes in television violence research, the claims of "enormous" or "explosive" effects by many leading television researchers are lamentable. Few of these researchers have taken care to note the sudden plummet in violent crime which began in 1993 and continues to the time of this writing.

It is often claimed that the effect of media violence on aggression is comparable in to that for well-known medical effects such as smoking and lung cancer (Bushman & Anderson, 2001; Council on Communications and Media, 2009; Huesmann, 2007). As noted, this is done by selectively focusing on the $r = 0.31$ outcome from Paik and Comstock (1994) and ignoring all other meta-analytic results, particularly those which more narrowly focus on violent crime as an outcome. The comparison is otherwise flawed both conceptually and statistically. The conceptual flaws are most evident. Even taken on face value, comparisons between psychological research on media effects, and medical research on disease is a comparison of apples to oranges. Many of the important medical research studies commonly cited in such comparisons, such as Salk Vaccine Trials, or the Physician's Aspirin Study, were well-designed randomized controlled outcome studies. Other research fields such as smoking and lung cancer do not benefit from such trials due to ethical concerns. However, many of these medical studies do have one benefit not experienced by psychological studies: perfectly valid outcomes. Put bluntly, death is a perfect measurement of death. One need not worry about measurement validity when considering death as an outcome. By contrast, near-complete reliance on proxy measures of serious aggression raises serious validity concerns.

Statistically speaking, most medical researchers report effect sizes in terms of Odds Ratios (OR) or Relative Risks (RR). It is known that direct comparisons of OR/RR to psychological effect size estimates such as *r* or *d* are fraught (Rosenthal & DiMatteo, 2001) with measurement error. Researchers have tried to adjust for this by using phi and Cramer's V formulas to calculate *r* from medical research. Unfortunately it has been known for some time that these formulas do not work well with medical studies (Block & Crain, 2007; Crow, 1991; Ferguson, 2009b, 2009c; Hsu, 2004; Kraemer, 2005). Specifically these formulas grossly underestimate the effects of many medical studies, particularly those with very large samples, because the effect size estimate is divided by sample size. For studies with very large samples (such as for smoking and lung cancer research), the resultant effect size is thus vastly underestimated when calculated in terms of "*r*" rather than OR/RR. That the comparison between television violence effects and smoking and lung cancer as well as other medical effects is based both upon these flawed estimates of medical effects, as well as selective use of the $r = 0.31$ estimate for television violence has been pointed out several times (Block & Crain, 2007; Ferguson, 2009c), yet it continues to be repeated. The 2009 AAP policy statement repeated these claims without either noting that they had been challenged in the published literature, or that the $r = 0.31$ was the highest among several meta-analytic results.

The debate about the practical significance of effect sizes is both heated and unresolved. Nonetheless, authors who attempt to make an interpretation of the effect sizes they report can contribute to the ongoing interpretation of the body of findings.

Table 2
Effect sizes in criminal justice research.

Relationship	Effect size (*r*)
Video game sales and youth violence rates in United States[a]	−0.95
Genetic influences on antisocial behavior[b,c]	0.64 to 0.75
Self control and perceptions of criminal opportunity on crime[d]	0.58
Protective effect of community institutions on neighborhood crime[d]	0.39
Cruelty to animals and later violence toward people[e]	0.37
Firearms ownership on crime[d]	0.35
Incarceration use as a deterrent on crime[d]	0.33
Religiosity and crime[d]	−0.29
Psychopathy on criminal recidivism in juvenlies[f]	0.26
Negative family environment on crime[d]	0.26
Peer group on crime[d]	0.25
Poverty on crime[d]	0.25
Exposure to child abuse and aggressive behavior in young men[g]	0.23
Television violence on violent crime	0.02 to 0.10

[a] Ferguson (2010a).
[b] Rhee and Waldman (2002).
[c] Ferguson (2010b).
[d] Pratt and Cullen (2005).
[e] Merz-Perez, Heide, and Silverman (2001).
[f] Douglas, Epstein, and Poythress (2008).
[g] Nicholas and Rasmussen (2006).

7. New studies: Are we making progress?

In this paper we have highlighted several milestones in research methodology for studying the effects of violent television on aggressive behavior. Our question at present is: have recent published studies applied this wisdom? We ask this question because we have been concerned about trend in recent research on TV and video game violence, in the published literature and in manuscripts we have been asked to review. We will discuss five recent studies in TV violence.

Ybarra et al. — Ybarra et al. (2008) recently reported a study of the effects of media violence on self-reported violent behavior in the

Author's personal copy

pages of *Pediatrics*. The data were based on a cross-sectional survey of more than 155 youths. The dependent variable included self-reported violence including aggravated assaults, robberies and sexual assaults (0 = none, 1 = at least one). The major contribution of this study was the analysis of the relationship between *serious violent behavior* and a variety of types of exposure to violent media including music, games, television and Internet Web sites. The authors collected data on many important potentially confounding factors including age, sex, race, anger propensity, substance use, family income family criminality, emotional closeness, parental monitoring, coercive discipline, spousal abuse, delinquent peers, academic performance and community violence. Of course a major weakness of the paper is the inability to establish temporal order. Another weakness is the measure of exposure to violent media. Subjects were grouped so that individuals who reported that "many, most, or all" of the games they played, music they listened to, Internet sites they visited, or television shows they watched were compared to the other subjects. The use of such a measure makes it important to control for trait characteristics. Nonetheless, the methodological design has the potential to make an important contribution to the literature.

In their multivariate analysis, Ybarra et al. (2008) found that only subjects heavily exposed to violent web sites were significantly more likely to self-report serious violence: "All other exposures to violent media were not significantly related to seriously violent behavior when underlying differences in individual, family, school, peer and community characteristics were taken into account" (p. 932). Thus one would expect that the authors would emphasize the fact that exposure to TV violence and video game violence were not significantly associated with serious aggression, which is a major advance in the literature. Instead, the authors focus on the simple correlations, highlighted in their Table 3, which shows that the simple odds ratios (essentially bivariate correlations, which as we have noted, are highly misleading) for reporting serious violence were significantly higher for those with high exposure to video games, violent music and violent Internet sites — the odds ratio for exposure to violent television or movies was nearly statistically significant. Their abstract fails to draw a hard line between the simple correlations and the multivariate findings: "Exposures to violence in the media, both online and off-line, were associated with significantly elevated odds for concurrently reported seriously violent behavior" (p. 929). This is true, except of course, when the proper statistical controls are employed in a multivariate analysis. Then this statement is false. Due to the authors' mis-emphasis on the bivariate correlational findings, rather than the multivariate findings, reviewers are likely to cite this study as further support for a link between media violence and serious aggression, rather than as a study which suggests that the association between video game violence, television violence and music violence and serious aggression is not significantly different from zero, which would be the appropriate conclusion from the paper as presented.

Boxer, Huesmann, Bushman, O'Brien, and Moceri (2009) were interested in the role of violent media in "shaping violent and other serious antisocial behavior" (p. 417). They used a sample of 820 youth, including 390 juvenile delinquents and 430 high school students. Because the authors hoped to assess the effect of childhood exposure to media violence on present-day juvenile violence and aggression, they used a retrospective measure of "favorite programs" which were rated for violent content independently. As discussed earlier, the "favorite programs" measure has been criticized and the authors do not report attempting to assess actual frequency or duration of exposure. While the guided recall procedure they used for the retrospective data is likely to reduce confounding exposure with present-day preferences and exposure, it does not establish temporal order because media preferences are likely to be enduring. One way to strengthen the arguments would be to control for present day preferences, but they did not elect to do so.

The most concerning problem in this study is the model specification. Rather than control for potentially confounding factors, such as callousness, academic skills, depression, psychoticism, neighborhood violence, and witnessing violence (which they measured, presumably to use as control factors), the authors create a "cumulative risk" variable, by dummy coding risk factors and adding them together. While cumulative risk is an interesting construct on its own, here the independent variable is preference for media violence so the model should be designed to eliminate rival hypotheses by controlling for these factors individually using continuous, not dummy-coded categorical variables. Controlling for the number of risks does not accomplish this end. In essence the authors greatly reduce the variance of their control variables, reducing the effectiveness of the controls. The authors employ a stepwise procedure and report that media violence risk is significantly associated with childhood preferences for violent television, above and beyond the effects of "other risk" (the cumulative risk indicator). They conclude that violent media preferences were significantly associated with violence and general aggression, but the reader must wonder why they did not simply control for the individual control variables and whether the coefficient would be statistically significant had this been done.

Johnson, Cohen, Smailes, Kasen, and Brook (2002) published a relatively brief report that garnered a lot of attention after its publication in the journal *Science*. The authors include some control variables such as parental neglect, family income and mental illness, making this one of the better studies regarding controlling for "third" variables. The outcome variable specifically involved person-on-person acts of aggression, also demonstrating superiority over many earlier studies. They reported that television viewing in early adolescence (age 14) was weakly related to person-on-person aggression (OR 1.57–1.58) in later adolescence or early adulthood, and later at age 30 (OR 1.26–2.62). A relatively inconsistent pattern of significant and non-significant results was found by gender at the earlier outcome period. At age 30, results were only significant for females, not males. A major limitation that was ignored by a number of reviewers is that Johnson et al. examined *total television viewing*, not violent media exposure per se. While this is one of the best studies examining potential harmful effects of television, the findings cannot be generalized to violent media exposure, in particular since there were no controls for past aggressiveness, neglect or abuse, or an array of other factors that might be related to television viewing and aggression such as neglect, peer rejection, intelligence, and so on. Although the sample appears to have been taken from a larger study of personality disorders, personality disorder diagnosis was not controlled, despite that this would have been a critical control variable.

Ferguson et al. — Ferguson, San Miguel, and Hartley (2009) present research on 603 mostly Hispanic youth between the ages of 10–14. Their measure of exposure to television violence was comprised from the amount of time children spent watching their three favorite television shows and the amount of violence in these shows. This improves upon the "favorites" measure described earlier by including measurement of exposure time. It was also felt by the authors that the massively expanded television offerings available currently, make a standardized list of shows both impractical and potentially biasing. However, this measure is still not a perfect measure of television violence exposure, highlighting how difficult it can be to adequately measure this construct. The authors use an array of well-validated instruments of youth violence including the Child Behavior Checklist aggression and rule-breaking scales (both parent and youth self-report), the Negative Life Events scale, which measures violent criminal acts, and the Olweus Bullying Scale. Control variables included family violence, antisocial personality, peer, school and community variables and depressive symptoms. Their results suggested that the best predictors of youth violence were depression and peer delinquency as well as antisocial personality and psychological

Author's personal copy

abusiveness in the family. Television violence exposure was not predictive of youth violence on these well-validated outcome measures. Although the study suffers from some limitations, such as the failure to establish temporal order, it contributes to the literature by providing an analysis of data, including a measure of serious violence, controlling for various confounds, including antisocial personality.

Coyne et al. — Several recent papers have examined the impact of television violence on domestic violence in adult relationships using measures such as the Conflict Tactics Scale. One such recent paper considered the moderating role of television violence exposure on the relationship between psychopathy and aggressive behavior (Coyne, Nelson, Graham-Kevan, Keister, & Grant, 2010). The authors studied 337 undergraduate students who had been involved in a serious relationship. They measured psychopathy, relationship aggression (using the Conflict Tactics Scale), and exposure to television violence by having subjects list their favorite shows and how often they viewed those shows, with independent raters assessing violence content. They did not find that exposure to TV violence mediated the effects of primary psychopathy on physical aggression or romantic relationship aggression. They do report a more direct effect of viewing physical violence on TV on physical aggression for men, and viewing relational violence on TV on violence in romantic relationships for both men and women. Also, they report that male "secondary" psychopaths are significantly more likely to choose to watch programs with relational violence. Importantly, their structural model does not control for potentially confounding factors as it directly links TV violence measures with aggressive outcomes, so we do not learn much from this study except that the correlations we find between aggression and exposure to TV violence extend to the specific case of domestic violence; we do not discover whether watching relationship violence on TV causes partner violence, or whether batterers are simply more likely to watch shows with relationship violence in them.

8. Conclusion

In this essay we have reviewed some of the evidence related to the effects of media violence on aggressive behavior. We have highlighted several methodological issues and several ongoing debates. Finally, we assessed several recent studies to determine whether lessons learned from decades of research are being applied. We are concerned that some of these lessons are not being applied. Several recently published studies emphasize simple versus multivariate findings or fail to control for important third factors. We express concern that little effort appears to be made to employ more careful research designs, particularly when well-controlled multivariate designs with well-validated outcome measures appear to produce different outcomes than less-controlled studies. We are also concerned that these common problems in the study of television violence may be repeated in the study of video game violence. Although we cannot cite unpublished papers, both authors of this paper have peer-reviewed papers submitted for publication with these flaws on numerous occasions. For example, in papers where authors report measuring important covariates, they fail to control for these factors in the actual statistical model. Very recently, authors argue that "hierarchical" regression is a worthy substitute for multiple regression, when the variable of interest is, at times, entered into models before the supposed control factors. Some authors in the video game violence area have exhumed the "favorite show" measure, asking subjects only what their favorite games are, without a measure of the actual number of hours of exposure. The problem with ignoring what we have learned from decades of research is that the new research reifies the same, flawed conclusions by introducing new studies with the same biases. In other words, the research doesn't move forward and does not self-correct.

The evidence is clear both from skeptical scholars (Ferguson & Kilburn, 2009; Savage & Yancey, 2008) as well as non-skeptical scholars (Paik & Comstock, 1994) that the best studies, particular those which use outcome measures for the most serious aggression, produce the weakest effects. If authors are convinced that media violence is causing violent crime, they must apply better methodological paradigms to demonstrate this. Were media scholars more cautious in the generalization of their research to outcomes related to violent crime (as Coyne et al., 2010 are to their credit), this would not be an issue. But as long as scholars and other reviewers continue to make claims about effects on violent crime, the size of which are inconsistent with published research, and important health organizations continue to push for policy that curbs media violence as a remedy to real-life violence, the quality of the evidence will remain centrally important.

References

American Academy of Pediatrics (2000). Joint statement on the impact of entertainment violence on children. Retrieved 5/10/10 from: http://www.aap.org/advocacy/releases/jstmtevc.htm

Anastasi, A., & Urbina, S. (1996). *Psychological testing.* New York: Prentice Hall.

Anderson, C., & Bushman, B. (1997). External validity of 'trivial' experiments: The case of laboratory aggression. *Review of General Psychology*, *1*(1), 19–41. doi:10.1037/1089-2680.1.1.19.

Bachrach, Riva S. (1986). The differential effect of observation of violence on Kibbutz and city children in Israel. In L. R. Huesmann, & L. D. Eron (Eds.), *Television and the aggressive child: A cross-national comparison* (pp. 201–238). Hillsdale, NJ: Lawrence Erlbaum.

Bandura, A. (1965). Influence of models' reinforcement contingencies on the acquisition of imitative response. *Journal of Personality and Social Psychology*, *1*, 589–595.

Bandura, A., Ross, D., & Ross, S. A. (1961). Transmission of aggression through imitation of aggressive models. *Journal of Abnormal and Social Psychology*, *63*, 575–582.

Bandura, A., Ross, D., & Ross, S. A. (1963). Imitation of film-mediated aggressive models. *Journal of Abnormal and Social Psychology*, *66*, 3–11.

Baron, R., & Richardson, D. (1994). *Human aggression.* New York: Plenum Press.

Baumrind, D. (1964). Some thoughts on ethics of research: After reading Milgram's 'Behavioral Study of Obedience'. *American Psychologist*, *19*(6), 421–423. doi:10.1037/h0040128.

Belson, W. (1978). *Television violence and the adolescent boy.* Westmead, England: Saxon House.

Benjamin, L. (1985). Defining aggression: An exercise for classroom discussion. *Teaching of Psychology*, *12*(1), 40–42. doi:10.1207/s15328023top1201_11.

Berkowitz, L. (1965). Some aspects of observed aggression. *Journal of Personality and Social Psychology*, *2*, 359–369.

Block, J., & Crain, B. (2007). Omissions and errors in 'Media violence and the American public'.. *American Psychologist*, *62*, 252–253.

Blummer, H. (1933). *Movies and conduct.* New York: MacMillan.

Boxer, P., Huesmann, L. R., Bushman, B. J., O'Brien, M., & Moceri, D. (2009). The role of violent media preference in cumulative developmental risk for violence and general aggression. *Journal of Youth and Adolescence*, *38*, 417–428.

Bushman, B., & Anderson, C. (2001). Media violence and the American public. *American Psychologist*, *56*, 477–489.

Cantor, J., & Wilson, B. (2003). Media and violence: Intervention strategies for reducing aggression. *Media Psychology*, *5*(4), 363–403. doi:10.1207/S1532785XMEP0504_03.

Centerwall, B. (1989). Exposure to television as a risk factor for violence. *American Journal of Epidemiology*, *129*, 643–652.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Council on Communications and Media (2009). Media violence. *Pediatrics*, *124*, 1495–1503.

Coyne, S., Nelson, D., Graham-Kevan, N., Keister, E., & Grant, D. (2010). Mean on the screen: Psychopathy, relationship aggression, and aggression in the media. *Personality and Individual Differences*, *48*(3), 288–293. doi:10.1016/j.paid.2009.10.018.

Crow, E. (1991). Response to Rosenthal's comment, "How are we doing in soft psychology?". *American Psychologist*, *46*(10), 1083. doi:10.1037/0003-066X.46.10.1083.a.

Douglas, K., Epstein, M., & Poythress, N. (2008). Criminal recidivism among juvenile offenders: Testing the incremental and predictive validity of three measures of psychopathic features. *Law and Human Behavior*, *32*(5), 423–438. doi:10.1007/s10979-007-9114-8.

Eron, L. D. (1963). Relationship of TV viewing habits and aggressive behavior in children. *Journal of Abnormal and Social Psychology*, *67*(2), 193–196.

Eron, L. D., Huesmann, L. R., Lefkowitz, M. M., & Walder, L. O. (1972). Does television violence cause aggression? *American Psychologist*, *27*, 253–263.

Federal Bureau of Investigation (1951–2010). *Uniform crime reports.* Washington, DC: Government Printing Office.

Felson, R. B. (1996). Mass media effects on violent behavior. *Annual Review of Sociology*, *22*, 103–128.

Ferguson, C. J. (2009a). Media violence effects: Confirmed truth, or just another X-File? *Journal of Forensic Psychology Practice*, *9*(2), 103–126.

Ferguson, C. J. (2009b). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, *40*(5), 532–538.

Author's personal copy

Ferguson, C. J. (2009c). Is psychological research really as good as medical research? Effect size comparisons between psychology and medicine. *Review of General Psychology*, *13*(2), 130–136.

Ferguson, C. J. (2010a). Blazing Angels or Resident Evil? Can violent video games be a force for good? *Review of General Psychology*, *14*(2), 61–82.

Ferguson, C. J. (2010b). Genetic contributions to antisocial personality and behavior (APB): A meta-analytic review from an evolutionary perspective. *Journal of Social Psychology*, *150*(2), 160–180.

Ferguson, C. J., & Beaver, K. M. (2009). Natural born killers: The genetic origins of extreme violence. *Aggression and Violent Behavior*, *14*(5), 286–294.

Ferguson, C. J., & Kilburn, J. (2009). The public health risks of media violence: A meta-analytic review. *Journal of Pediatrics*, *154*(5), 759–763.

Ferguson, C. J., & Rueda, S. M. (2009). Examining the validity of the Modified Taylor Competitive Reaction Time Test of aggression. *Journal of Experimental Criminology*, *5*(2), 121–137.

Ferguson, C. J., San Miguel, C., & Hartley, R. D. (2009). A multivariate analysis of youth violence and aggression: The influence of family, peers, depression and media violence. *Journal of Pediatrics*, *155*(6), 904–908.

Feshbach, S., & Singer, R. (1971). *Television and aggression.* San Francisco: Jossey-Boss.

Fischoff, S. (1999). Psychology's quixotic quest for the media-violence connection. *An invited address at the Annual Convention of the American Psychological Association, Boston, August 21, 1999* Retrieved 5/11/10 from: http://www.calstatela.edu/faculty/sfischo/violence.html

Fraczek, Adam (1986). Socio-cultural environment, television viewing, and the development of aggression among children in Poland. In L. R. Huesmann, & L. D. Eron (Eds.), *Television and the aggressive child: A cross-national comparison* (pp. 119–160). Hillsdale, NJ: Lawrence Erlbaum.

Freedman, J. (2002). *Media violence and its effect on aggression: Assessing the scientific evidence.* Toronto: University of Toronto Press.

Friedrich, L., & Stein, A. (1973). Aggressive and prosocial television programs and the natural behavior of preschool children. *Monographs of the Society for Research in Child Development*, *38*, 63.

Gauntlett, D. (1995). *Moving experiences: Understanding television's influences and effects.* Luton: John Libbey.

Geen, R. G. (1976). Observing violence in the mass media: Implications of basic research. In R. G. Geen, & E. C. O'Neal (Eds.), *Perspectives on aggression* (pp. 193–234). New York: Academic Press.

Giancola, P., & Chermack, S. (1998). Construct validity of laboratory aggression paradigms: A response to Tedeschi and Quigley (1996). *Aggression and Violent Behavior*, *3*(3), 237–253. doi:10.1016/S1359-1789(97)00004-9.

Giancola, P., & Parrott, D. (2008). Further evidence for the validity of the Taylor aggression paradigm. *Aggressive Behavior*, *34*(2), 214–229. doi:10.1002/ab.20235.

Gothelf, D., Apter, A., & van Praag, H. (1997). Measurement of aggression in psychiatric patients. *Psychiatry Research*, *71*(2), 83–95. doi:10.1016/S0165-1781(97)00047-4.

Grimes, T., Anderson, J., & Bergen, L. (2008). *Media violence and aggression: Science and ideology.* Thousand Oaks, CA: Sage.

Hawley, P., & Vaughn, B. (2003). Aggression and adaptive function: The bright side to bad behavior. *Merrill-Palmer Quarterly*, *49*, 239–242.

Henry, D.Metropolitan Area Child Study Research Group. (2006). Associations between peer nominations, teacher ratings, self-reports, and observations of malicious and disruptive behavior. *Assessment*, *13*, 241–252.

Hogben, M. (1998). Factors moderating the effect of television aggression on viewer behavior. *Communication Research*, *25*, 220–247.

Hsu, L. M. (2004). Biases of success rate differences shown in binomial effect size displays. *Psychological Bulletin*, *9*(2), 183–197. doi:10.1037/1082-989X.9.2.183.

Huesmann, L. R. (1986). Psychological processes promoting the relation between exposure to media violence and aggressive behavior by the viewer. *Journal of Social Issues*, *423*, 125–139.

Huesmann, L. R. (2007). The impact of electronic media violence: Scientific theory and research. *Journal of Adolescent Health*, *41*, S6–S13.

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

Huesmann, L., Eron, L., Berkowitz, L., & Chaffee, S. (1992). *The effects of television violence on aggression: A reply to a skeptic.* Washington, DC US: Hemisphere Publishing Corp.

Huesmann, L. R., Eron, L. D., Lefkowitz, M. M., & Walder, L. O. (1984). Stability of aggression over time and generations. *Developmental Psychology*, *20*, 1120–1134.

Huesmann, L., Moise, J., & Podolski, C. (1997). The effects of media violence on the development of antisocial behavior. *Handbook of antisocial behavior* (pp. 181–193). Hoboken, NJ US: John Wiley & Sons Inc.

Huesmann, L., Moise-Titus, J., Podolski, C., & Eron, L. (2003). Longitudinal relations between children's exposure to TV violence and their aggressive and violent behavior in young adulthood: 1977–1992. *Developmental Psychology*, *39*(2), 201–221. doi:10.1037/0012-1649.39.2.201.

Huston-Stein, A., Fox, S., Greer, D., Watkins, B. A., & Whitaker, J. (1981). The effects of TV action and violence on children's social behavior. *Journal of Genetic Psychology*, *138*, 183–191.

Johnson, J. G., Cohen, P., Smailes, E. M., Kasen, S., & Brook, J. S. (2002, March 29). Television viewing and aggressive behavior during adolescence and adulthood. *Science*, *295*, 2468–2472.

Josephson, W. (1987). Television violence and children's aggression: Testing the priming, social script, and disinhibition predictions. *Journal of Personality and Social Psychology*, *53*(5), 882–890. doi:10.1037/0022-3514.53.5.882.

Kniveton, B., & Stephenson, G. (1975). The effects of an aggressive film model on social interaction in groups of middle-class and working-class boys. *Journal of Child Psychology and Psychiatry*, *16*(4), 301–313. doi:10.1111/j.1469-7610.1975.tb00364.x.

Kraemer, H. C. (2005). A simple effect size indicator for two-group comparisons?: A comment on $r_{equivalent}$. *Psychological Methods*, *10*(4), 413–419. doi:10.1037/1082-989X.10.4.413.

Lagerspetz, K., & Viemero, V. (1986). Television and aggressive behavior among Finnish children. In L. R. Huesmann, & L. D. Eron (Eds.), *Television and the aggressive child: A crossnational comparison*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Lefkowitz, M. M., Eron, L. D., Walder, L. O., & Huesmann, L. R. (1977). *Growing up to be violent: A longitudinal study of the development of aggression.* New York: Pergamon.

Leyens, J. P., Parke, R. D., Camino, L., & Berkowitz, L. (1975). Effects of movie violence on aggression in a field setting as a function of group dominance and cohesion. *Journal of Personality and Social Psychology*, *32*, 346–360.

Lilienfeld, S., Lynn, S., Namy, L., & Woolf, N. (2009). *Psychology: A framework for everyday thinking.* : Allyn & Bacon.

McIntyre, J. J., Teevan, J. J., & Hartnagel, T. (1972). Television violence and deviant behavior. In G. A. Comstock, & E. A. Rubinstein (Eds.), *Television and social behavior. Television and adolescent aggressiveness*, *Vol. 3*. (pp. 383–435)Washington, DC: U.S. Government Printing Office.

McLeod, J. M., Atkin, C. K., & Chaffee, S. H. (1972). Adolescents, parents, and television use. Adolescent self-report measures from Maryland and Wisconsin samples. In G. A. Comstock, & E. G. Rubinstein (Eds.), *Television and social behavior. Television and adolescent aggressiveness*, *Vol. 3*. (pp. 35–135)Washington, DC: U.S. Government Printing Office.

Merz-Perez, L., Heide, K., & Silverman, I. (2001). Childhood cruelty to animals and subsequent violence against humans. *International Journal of Offender Therapy and Comparative Criminology*, *45*(5), 556–573. doi:10.1177/0306624X01455003.

Milavsky, J. R., Kessler, R., Stipp, H., & Rubens, W. S. (1982). Television and aggression: Results of a panel study. In D. Pearl, L. Bouthilet, & J. Lazar (Eds.), *Television and behavior: Ten years of scientific progress and implications for the 80s. Technical reviews*, *Vol. 2*. (pp. 138–157) Washington, DC: U.S. Government Printing Office.

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, *67*(4), 371–378. doi:10.1037/h0040525.

Moeller, T. (2005). How 'unequivocal' is the evidence regarding television violence and children's aggression? *APS Observer*, *18*, 6.

Mussen, P., & Rutherford, E. (1961). Effects of aggressive cartoons on children's aggressive play. *Journal of Abnormal and Social Psychology*, *62*, 461–464.

Myers, D. (2009). *Psychology.* : Worth Publishers.

Nicholas, K., & Rasmussen, E. (2006). Childhood abusive and supportive experiences, inter-parental violence, and parental alcohol use: Prediction of young adult depressive symptoms and aggression. *Journal of Family Violence*, *21*(1), 43–61. doi:10.1007/s10896-005-9001-3.

Olson, C. (2004). Media violence research and youth violence data: Why do they conflict? *Academic Psychiatry*, *28*, 144–150.

Paik, H., & Comstock, G. (1994). The effects of television violence on anti-social behavior: A meta-analysis. *Communication Research*, *21*, 516–546.

Pratt, T., & Cullen, C. (2005). Assessing macro-level predictors and theories of crime: A meta-analysis. In Michael Tomry (Ed.), *Crime and justice: A review of research*, *Vol. 32*. (pp. 373–450)Chicago: University of Chicago Press.

Rhee, S., & Waldman, I. (2002). Genetic and environmental influences on antisocial behavior: A meta-analysis of twin and adoption studies. *Psychological Bulletin*, *128*, 490–529.

Ritter, D., & Eslea, M. (2005). Hot sauce, toy guns and graffiti: A critical account of current laboratory aggression paradigms. *Aggressive Behavior*, *31*, 407–419.

Robinson, J. P., & Bachman, J. G. (1972). Television viewing habits and aggression. In G. A. Comstock, & E. A. Rubinstein (Eds.), *Television and social behavior: Television and adolescent aggressiveness*, *Vol. 3*. (pp. 372–382)Washington, DC: U.S. Government Printing Office.

Rosenthal, R., & DiMatteo, M. (2001). Meta analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology*, *52*, 59–82. doi:10.1146/annurev.psych.52.1.59.

Savage, J. (2004). Does viewing violent media really cause criminal violence? A methodological review. *Aggression and Violent Behavior*, *10*, 99–128.

Savage, J., & Yancey, C. (2008). The effects of media violence exposure on criminal aggression: A meta-analysis. *Criminal Justice and Behavior*, *35*, 1123–1136.

Sheehan, Peter W. (1986). Television viewing and its relation to aggression among children in Australia. In L. R. Huesmann, & L. D. Eron (Eds.), *Television and the aggressive child: A cross-national comparison* (pp. 161–200). Hillsdale, NJ: Lawrence Erlbaum.

Smith, P. (2007). Why has aggression been thought of as maladaptive? In P. Hawley, T. Little, & P. Rodkin (Eds.), *Aggression and adaptation: The bright side to bad behavior* (pp. 65–83). Mahwah, NJ: Lawrence Erlbaum.

Stipp, H., & Milavsky, J. (1988). U.S. television programming's effects on aggressive behavior of children and adolescents. *Current Psychology: Research & Reviews*, *7*(1), 76–92. doi:10.1007/BF02686665.

Strasburger, V. (2009). Why do adolescent health researchers ignore the impact of the media? *Journal of Adolescent Health*, *44*(3), 203–205. doi:10.1016/j.jadohealth.2008.12.019.

Tedeschi, J., & Quigley, B. (1996). Limitations of laboratory paradigms for studying aggression. *Aggression & Violent Behavior*, *2*, 163–177.

Tedeschi, J., & Quigley, B. (2000). A further comment on the construct validity of laboratory aggression paradigms: A response to Giancola and Chermack. *Aggression & Violent Behavior*, *5*, 127–136.

Time (1959). Westerns: The six-gun Galahad. Time Magazine. Retrieved 5/11/10 from: http://www.time.com/time/magazine/article/0,9171,892441,00.html?internalid=ACA

Trend, D. (2007). *The myth of media violence: A critical introduction.* Malden, MA: Blackwell.

Van der Dennen, J. M. G. (1980). *Problems in the concepts and definitions of aggression, violence, and some related terms.* Groningen: Rijksuniversiteit Polemologisch Instituut.

Author's personal copy

Wiegman, O., & Kuttschreuter, M. (1992). A longitudinal study of the effects of television viewing on aggressive and prosocial behaviors. *British Journal of Social Psychology*, *31*, 147–164.

Wiegman, O., Kuttschreuter, M., & Baarda, B. (1985). *Television viewing related to aggressive and prosocial behavior.* The Hague: Stitching voor Onderzoek van het Onderwijs.

Wilkinson & Task Force on Statistical Inference (1999). Statistical methods in psychological journals: Guidelines and explanations. *American Psychologist*, *54*(8), 594–604. doi:10.1037/0003-066X.54.8.594.

Wood, W., Wong, F., & Chachere, J. (1991). Effects of media violence on viewers' aggression in unconstrained social interaction. *Psychological Bulletin*, *109*(3), 371–383. doi:10.1037/0033-2909.109.3.371.

Ybarra, M., Diener-West, M., Markow, D., Leaf, P., Hamburger, M., & Boxer, P. (2008). Linkages between internet and other media violence with seriously violent behavior by youth. *Pediatrics*, *122*(5), 929–937.

Zillmann, D. (1991). Television viewing and physiological arousal. In J. Bryant, & D. Zillmann (Eds.), *Responding to the screen: Reception and reaction processes* (pp. 103–134). Hillsdale, NJ: Lawrence Erlbaum.

Zillmann, D., & Weaver, J. (2007). Aggressive personality traits in the effects of violent imagery on unprovoked impulsive aggression. *Journal of Research in Personality*, *41*, 753–771.