GAME VIOLENCE, GAME DIFFICULTY, AND 2D:4D DIGIT RATIO AS PREDICTORS OF AGGRESSIVE BEHAVIOR

A Dissertation

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

by

JOSEPH HILGARD

Dr. Bruce Bartholow, Dissertation Supervisor

MAY 2015

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

GAME VIOLENCE, GAME DIFFICULTY,
AND 2D:4D DIGIT RATIO AS PREDICTORS
OF AGGRESSIVE BEHAVIOR

presented by Joseph Hilgard,

a candidate for the degree of doctor of philosophy,

and hereby certify that, in their opinion, it is worthy of acceptance.

Professor Bruce Bartholow

Professor Lissa Behm-Morawitz

Professor Ken Sheldon

Professor Todd Schachtman

Dedicated to the rational updating of beliefs.

ACKNOWLEDGEMENTS

I thank my family for their unwavering support and the reservoir of sanity they provide. I thank my advisor, Dr. Bruce Bartholow, for sponsoring and advising me all these six years and for granting me his trust and respect. I thank Dr. Ines Segert for sending me her brightest students as research assistants, and I thank those research assistants for many hours of volunteer labor. Finally, I thank Dr. Jeffery Rouder for introducing me to many of my most valued tools – R, Git, and Bayesian model comparison.

While studying a textbook on the use of Akaike Information Criterion for model selection, I came across a quote from J. Bronowski regarding Ludwig Eduard Boltzmann, the originator of statistical models of entropy. In it, Boltzmann is described as "an irascible, extraordinary man… quarrelsome and delightful, and everything that a human should be." The quote always reminds me of my equally irascible, extraordinary, quarrelsome, and delightful collaborator and friend, Dr. Rouder.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

GAME VIOLENCE, GAME DIFFICULTY, AND 2D:4D DIGIT RATIO
AS PREDICTORS OF AGGRESSION

Joseph Hilgard

Dr. Bruce Bartholow, Dissertation Supervisor

ABSTRACT

Previous research has sought to determine whether violent video games cause increases in aggressive behavior. Game difficulty is also thought to potentially cause changes in aggression. Finally, prenatal testosterone exposure, as measured by the ratio of lengths of the index and ring fingers (2D:4D digit ratio), is thought to cause dispositional increases in aggressive behavior.

In the present research, game violence and game difficulty were manipulated by making four modified versions of the same game. Participants played a game, were provoked, and then had an opportunity to aggress against another participant in the study. Neither game violence nor game difficulty was found to influence aggressive behavior in a theory-consistent way, and Bayesian model comparison favored the null model over all theory-derived alternative hypotheses. Similarly, 2D:4D ratio did not predict aggressive behavior, either alone or in interaction with the other study variables.

The results suggest that brief exposure to violent or difficult games does not influence aggressive behavior when game stimuli are closely matched, and furthermore, that differences in finger length do not predict aggression. I discuss the implications for brief-session experimental research of causes of aggression.

Introduction

Media shapes the world. In the absence of media, people's perceived worlds would be limited to only that fraction of the world which they personally experience. Through media, however, people develop ideas, attitudes, and schema about people they've never met, experiences they've never had, and concepts they've never personally explored. Because these ideas, attitudes, and schema are the basis of behavior, media messages promise to influence human psychology and behavior, whether for good or for ill.

One of the newest and most popular forms of media is video games. Compared to traditional media such as books, movies, and film, video games are highly interactive; the player controls, to at least some degree, the game character. Moreover, video games are generally highly motivating and engaging and can be played for many hours at a time. These properties have inspired research that investigates whether video games are an effective way to teach skills and behaviors (Green & Bavelier, 2003).

However, not all learned skills and behaviors are desirable. While video games are a broad and heterogeneous form of media, many of the most popular video games contain violent content. Violent content ranges from the mild, fantastic, and cheerful (e.g., *Super Mario Galaxy*) to the graphic, realistic, and depraved (e.g., *Grand Theft Auto 3, Manhunt, Mortal Kombat*). Exposure to this violent content is expected to teach players aggressive behaviors and schema. It is further supposed that media effects of video games are more potent than those of other forms of media because the player is an active participant in the violent content, rather than a passive viewer. Over the past two decades, research has sought to measure and understand the possible relationships between consumption of violent media and changes in aggressive and violent behavior.

Psychological Theories of Aggressive Behavior

Like most psychological and behavioral phenomena, aggressive and violent behaviors are thought to be complex and multiply determined, with no single clear cause. However, a variety of theories have emerged to describe the diverse possible causal pathways from stimuli to aggression.

Social learning theory. Research of media effects on behavior began with Bandura's social learning theory (Bandura & McClelland, 1977). In contrast to behaviorist theories, which proposed that individuals learn through experienced rewards and punishments for behaviors, social learning theory suggested that behavior also could be shaped through observational learning. Instead of having to experience a reward or punishment firsthand, a person could learn behavior through observing others' behaviors and the rewards or punishments those others received.

An early study examined the possibility of such a learning process. Children were randomly assigned to watch a version of a video of a lab assistant in a room full of toys. Among the toys in the room is a "bobo doll", an inflatable, durable doll with a weighted base, such that it springs upright when pushed over. In one version of the video, the lab assistant ignores the doll and plays with the other toys in the room. In another version, the lab assistant repeatedly attacks the doll, hitting it with a mallet, throwing it into the air, or sitting on it and punching it repeatedly. Children who watched this version of the video were more likely to engage in aggressive play with the doll or with other toys, imitating behaviors learned from watching the assistant's behavior (Bandura, Ross, & Ross, 1961).

This theory had alarming implications for the possible effects of violent media. If behavior is easily and readily shaped by observing others, then watching media in which violent behaviors are justified and rewarded could teach people to use violence. Future research attempted to test Social Learning Theory models of violent media and aggressive behavior with more externally valid stimuli (e.g., violent cartoons) and dependent measures (e.g., aggressing against another person rather than hitting an inanimate doll) (e.g., Josephson, 1987; Potts, Huston, & Wright, 1986).

The General Aggression Model. The General Aggression Model (GAM; Anderson & Bushman, 2002; Lindsay & Anderson, 2000) is an attempt to integrate social learning theory and later-developed theories into a single broad model that would still be specific enough to be falsifiable. GAM does this by describing a cycle consisting of person/situation inputs, an internal state of the individual, and outcomes resulting from the process. These outcomes are expected to cycle back to affect the person/situation inputs, leading to lasting changes. The theories integrated in the GAM explicate the theoretically-relevant inputs, states, and outcomes, as well as their relationships.

Causal pathways within the GAM reflect the predictions of many theories. For example, cognitive-neoassociation theory (Berkowitz, 1984) considers learned associations between cues and aggression, suggesting that a conditioned stimulus can later prime associated cognition or affect. Script theory (Huesmann, 1986, 1998) proposes that well-rehearsed sets of concepts are selected and applied for their resemblance to the current context; increased rehearsal of aggressive scripts, then, is expected to increase the likelihood that these scripts are activated and applied. Excitation transfer theory (Zillmann, 1983) argues that that arousal from a previous event can be applied to a later unrelated event, causing inappropriate affective overreactions to interpersonal situations and increased aggressive behavior. Cultivation theory (see Shanahan & Morgan, 1999), which argues that media portrayals influence the receivers' perception of the real world, suggests that persons exposed to violent media develop a distorted worldview, overestimating the frequency and social normativity of aggressive or violent behavior. Desensitization theory (Wolpe, 1958), which proposes that repeated exposure to an affective stimulus causes decreased affective response over time, indicating that violent media may make aggressive or violent behaviors less affectively aversive and reduce others' apparent need for help.

GAM broadly describes internal states leading to behavior as belonging to three categories: cognition, affect, and arousal. Violent media is expected to influence all of these in short-term contexts. Given the theories combined by GAM, increased aggressive thought accessibility, hostile feelings, arousal, rehearsal of aggressive behaviors, and expectations of aggressive behavior from others are all believed to increase aggressive behavior. Many experiments have found associations between violent media, violent behavior, and these hypothesized mediating processes.

GAM is also argued to predict *long-term* changes in behavior. Recall that GAM is a cycle. Its outcomes (e.g., aggressive behavior) are thought to shape the individual's personality and future situations; an aggressive individual is thought to be more likely to find himself in aggressive contexts in the future. In those aggressive contexts, the individual is expected to use previously-exercised aggressive behavior. Repeated exposure to aggressive primes is argued to make the prime *chronically accessible,* causing prolonged priming of aggressive behavior (Anderson & Bushman, 2002).

GAM has recently been further generalized to explain effects of nonviolent media on prosocial behavior in a model called the General Learning Model (GLM; Buckley & Anderson, 2006). This model argues that games can be teaching tools and can teach aggressive or prosocial behaviors. This model is structurally analogous to the GAM, featuring the same series of person/situation inputs, which contribute to a present internal state, leading to outcomes such as appraisals and behaviors. Affect, cognitions, and arousal derived from media are again expected to influence a person's internal states and choices of actions, allowing calming (Whitaker & Bushman, 2012) or prosocial (Greitemeyer & Osswald, 2010) video games to cause increased prosocial behavior.

Evidence for violent game effects on aggressive outcomes

To date, twenty five years of violent video game research indicates a causal effect of violent games on aggressive thoughts, feelings, and behaviors. Researchers have found significant effects of violent game contents on aggressive outcomes whether comparing early arcade games like *Centipede* and *Zaxxon* (Anderson & Ford, 1986) or more modern, realistic video games such as *Grand Theft Auto 4* (e.g., Gabbiadini, Riva, Andrighetto, Volpato, & Bushman, 2013).

In summarizing this literature, meta-analysists have argued that effects are positive and highly statistically significant (Anderson et al., 2010; Greitemeyer & Mügge, 2014), a finding which one researcher saw as "nailing the coffin shut on doubts" (Huesmann, 2010). Effect sizes have been recognized as modest in magnitude (r = .21, Anderson et al., 2010; r = .19, Greitemeyer & Mügge, 2014), but these effect sizes are argued to be practically meaningful based on their putative implications for public health. Accordingly, professional societies have released public statements on the harmful effects of violent media (American Psychological Association, Task Force on Violent Media, 2005; American Academy of Pediatrics, Council on Communications and Media, 2009). It has been argued that there is now “broad consensus” among media researchers, pediatricians, and parents that media violence increases aggression in children (Bushman, Gollwitzer, & Cruz, in press).

The controversy. However, not all researchers have been convinced by these research findings (Adachi et al., 2013). First, critics argue that the evidence has been overstated due to publication bias (Ferguson, 2007), that is, that studies which do not find significant effects are not submitted or not accepted for publication, causing the existing literature to provide a biased overestimate of the effect size. It has also been argued that the construct of “violent games” is lacking in content validity (Hilgard, Engelhardt, & Bartholow, in revision; Ferguson, 2014; Progress and Freedom Foundation & Electronic Frontier Foundation, 2009). For example, many violent game scholars have defined mostly-innocuous games like *Pac-Man* as being “violent” (Thompson & Haninger, 2001; Rushton, 2013) and meta-analyzed research findings accordingly (e.g. “best-practices” criteria from Anderson et al., 2010). Invalid definitions suggest that the cause of increased aggression may be confounding variables rather than violent game content itself (Adachi & Willoughby, 2011a, 2011b). Finally, the internal and external validity of aggression measures is sometimes called into question. It has been argued that the Competitive Reaction Time Task, a common measure of aggressive behavior, does not have a single standardized form of quantification, and so researchers may flexibly analyze several quantifications and selectively report the one that rejects the null or indicates the largest effect size (Elson, Mohseni, Breuer, Scharkow, & Quandt, 2014). Flexible analysis would, like publication bias, overestimate the size of the true effect.

Sample size. Some skeptics of violent-media effects have conducted their own experiments to attempt to test better-controlled violent game manipulations (e.g. Adachi and Willoughby, 2011b; Elson et al., 2014; Valadez & Ferguson, 2012). However, many of these experiments have suffered from insufficient sample size. When sample size is too small, and the hypothesis test underpowered, a nonsignificant test result does not necessarily present positive evidence for the truth of the null hypothesis. In our Bayesian re-analysis of these studies, we find that evidence for the null is mixed, and that some studies reporting nonsignificant results nonetheless find some evidence for the alternative hypothesis of an effect (Hilgard, Engelhardt, Bartholow, & Rouder, submitted). An ideal experiment would include a large sample and consider the strength of evidence as a continuous quantity, perhaps through use of effect sizes and confidence intervals or Bayesian analyses.

Testing Specific Effects of Violent Game Contents

Researchers have attempted to test the specific effects of violent game content, not other potential confounding game features. However, violent and nonviolent games are often very different, usually belonging to very different genres with very different rules of play. For example, violent games are often shooter games, fighting games, or action games, while nonviolent games are often racing games, puzzle games, or sports games. Therefore, while tested games do differ in their *violent content,* they are also different in their controls, strategies, and other gameplay features we call *game mechanics* (Hilgard et al., in revision). It would be possible that these confounding differences in game mechanics, rather than the actual violent content, are responsible for the observed changes in aggressive outcomes.

Researchers have attempted several ways to account for these potential differences. First, one might conduct a pilot test, collecting ratings of some potential confounds, hoping not to observe a significant difference between the games on any confound. This approach is flawed in that retention of the null hypothesis does not provide evidence for the null hypothesis, especially when sample sizes are small, as they often are in pilot tests (Hilgard et al., submitted). Another approach is to apply the potential confounds as covariates. This approach, however, is less than ideal. On the one hand, if the confound does cause aggression, in the case that the confound is measured with error (as is likely, given that these confounds are often measured with single-item covariates), residual variance will remain in the model. Analysis of covariance might mitigate, but not eliminate, influence of the confound, leading to an overestimated effect size. On the other hand, certain apparent confounds might be meaningful outcomes of violent content, mediating the relationship between violent content and aggressive outcomes. Applying these mediators as covariates would eliminate much of the relationship between violent content and aggressive outcome, underestimating the effect size.

Game modification paradigms provide greater experimental control and eliminate the need for post-hoc statistical adjustments of questionable value. Rather than comparing two separate games, or different activities within a single game, modification allows the researcher to exercise control over the game contents. For example, a game can be modified so that the same level is played either with violent or nonviolent contents, but all other game parameters are kept the same (as demonstrated in Carnagey & Anderson, 2005; Elson, Breuer, Van Looy, Kneer, & Quandt, in press; Engelhardt, Hilgard, & Bartholow, in press; Przybylski, Deci, Rigby, & Ryan, 2014). This approach allows for accurate tests of the effects of very specific game features.

2D:4D Ratio

Media is not the only anticipated cause of aggression. Because males are generally more aggressive (see Campbell, 2006), it has been suggested that aggression, being a sexually-influenced trait, is affected by the sex hormone testosterone. Some support for this idea has been found in lizards (Moore & Marler, 1987) and in birds (Wingfield, Ball, Dufty, Hegner, & Ramenofsky, 1987), but effects among humans are less apparent, perhaps because of the role of culture in establishing sexually-dimorphic behavior (see Archer, 2009).

Nevertheless, it has been suggested that prenatal testosterone exposure could influence a variety of physiological and psychological constructs through organizational effects on the developing brain. While ethical reasons forbid the investigation of the effects of prenatal testosterone on psychological development, the measurement of 2D:4D digit ratio has been suggested as an alternative approach to measurement of prenatal testosterone. 2D:4D, the ratio of the lengths of the index and ring finger, is thought to be sexually dimorphic. On average, men have shorter index fingers relative to their ring fingers (2D:4D: ~ 0.95) as compared to women (2D:4D: ~ 1.0; Manning, Scutt, Wilson, & Lewis-Jones, 1998; Phelps, 1952). Within each sex, 2D:4D has been found to be associated with higher prenatal levels of the androgen testosterone and lower levels of the estrogen estradiol (Lutchmaya, Baron-Cohen, Raggatt, Knickmeyer, & Manning, 2004).

Inconsistent effects of 2D:4D on aggressive behavior. The testosterone-aggression hypothesis would predict that 2D:4D ratios indicative of greater developmental androgen exposure would be associated with greater aggression. However, evidence does not seem to support this relationship. Meta-analysis indicates that there is no relationship between 2D:4D and aggression in females, and that the relationship between 2D:4D and aggression in males is quite small (r = -.06, Hönekopp & Watson, 2011). Effect sizes for several studies were not reported other than as "not significant" and imputed as r = .00 (n = 284 out of the total sample N = 1895), so this meta-analysis may provide an overly conservative test.

In an attempt to resolve this inconsistency, it has been proposed that 2D:4D ratio only predicts aggressive behavior in an aggressive context (Millet, 2011). For example, 2D:4D ratio is argued to interact with the effect of an aggressive music video on aggressive intent, with more masculine ratios leading to greater aggressive intent when the music video was aggressive (r = -.46), but not when the music video was not aggressive (r = -.03) (Millet & Dewitte, 2007). Similarly, it is argued that the relationship between 2D:4D ratio and an behavior in an economic dictator game reverses depending on whether participants are in a neutral or aggressive context, e.g., having been previously primed with aggressive words (Millet & Dewitte, 2009). It is possible, however, that these moderation models are overfitting the data, especially if they are attempted post-hoc when the anticipated main effects are not found.

Null results from gene expression data. Recent meta-analytic efforts call into question the validity of 2D:4D ratio as a measurement of prenatal testosterone action. Voracek (submitted) investigated the estimated effect of the gene Xq11.2-12, expected to influence androgen responsivity. Longer variants of this gene are less active, and thus would be expected to lead to reduced response to testosterone, and thus, less masculine 2D:4D ratio. An initial small-sample study did indeed find such a relationship (Manning, Bundred, Newton, & Flanagan, 2003). However, several subsequent studies have found no significant relationship, and Voracek estimates the effect size as r = .02, [-.02, .06]. Thus, it is possible that 2D4D is not a valid measurement of prenatal testosterone activity in typical populations.

Purpose

The proposed study examines the effects of game violence, game difficulty, and 2D:4D ratio on aggressive behavior among college-aged males. The study will thereby test the hypotheses generated by previous research and theory. First, I aim to test the effect of violent game content on aggressive behavior, deriving the first hypothesis:

H1: Violent game content will increase aggressive behavior as measured by duration of coldpressor assignment, even when games are identical in all ways save violent content.

It has also been argued that observed changes in aggressive behavior are not due to the violent content of tested games, but rather, confounded elements of difficult or competitive game content (Adachi & Willoughby, 2011a, 2011b; Przybylski et al., 2013). My previous research suggests that challenging games may also deplete cognitive resources typically employed in the control of behavior (Engelhardt, Hilgard, and Bartholow, in press). Since cognitive control is theorized to be an important element in inhibiting aggressive behavior (Anderson & Bushman, 2002), depletion of cognitive control resources would be expected to increase aggressive behavior. Difficult and competitive gameplay, then, should be expected to lead to increased aggression through frustration, priming of competition, thwarted competency, or mental fatigue. I thereby derive my second hypothesis:

H2: Difficult games will increase aggressive behavior relative to easy games, whether through increased competitive content or the depletion of cognitive resources.

As summarized before, it has been suggested that lower, more masculine 2D:4D ratio predicts increased aggression. Although this effect may be context-dependent, the current context should allow the effect to emerge, as participants are provoked by their partners. From this suggestion, I derive my third hypothesis:

H3: Lower 2D:4D ratio will predict greater aggression.

Finally, theories of aggressive behavior predict that multiple effects should have super-additive effects. For example, violent games may prime and facilitate aggressive thoughts and behaviors, but these effects should be especially potent when cognitive resources are depleted (Anderson & Bushman, 2002). These predictions are mirrored by I^3 Theory (Finkel, 2013), a theory that describes sources of aggression as being instigating, impelling, or (dis)inhibiting, and predicts that combinations of the three yield superadditive effects. Not only are all participants provoked (instigation), but others are hypothetically driven by prenatal testosterone and violent game content (impelled) and/or cognitively depleted by challenging gameplay (disinhibited). Thus, I derive my fourth and last hypothesis:

H4: Combinations of aggression-inducing factors tested in H1, H2, and H3 should lead to superadditive increases in aggression, as predicted by I^3 theory and the General Aggression Model.

Method

Participants

Participants were 335 male undergraduate students at a state university. The target sample size was 450 subjects, anticipating a loss of about 50 subjects due to failures of the experiment or of deception; data collection and entry are still in progress. Participation was restricted to males because 2D:4D effects are thought to apply only to males (McIntyre et al., 2007; but see Millet & Dewitte, 2007). This also had the positive side-effect of eliminating gender as a potential source of variance. Participants were primarily Caucasian (76.7%), with some African-American (8.9%), Asian (7.8%), and Latino (3.6%). On average, participants were 18.9 (SD: 1.9) years old.

Scientific integrity

Hypotheses and sample size were preregistered at https://osf.io/cwenz/. Materials including game files are also available at that URL. Data and analytic code will later be made available at that website. Data and code are currently available upon request at https://collaborate.missouri.edu/jhilgard/vg-dissertation.

Measures

2D:4D ratio. Participants placed their hands on a flatbed scanner, fingers held together and fully extended. The scanner imaged their hands. The distance from tip to basal crease of each index and ring finger was measured using the caliper tool in the GNU Image Manipulation Program (www.gimp.org), a freeware Photoshop-like tool. 2D:4D ratios were created for each hand by taking the ratio of lengths of the index and ring fingers. Data are planned to be double-coded for maximum reliability. As of now, 122 subjects have been double-entered. Among these, inter-rater reliability was excellent (>90%).

Coldpressor task. Participants had an opportunity to aggress against their partner in the experiment by assigning the partner to immerse his fist in a bucket of painfully-cold water for an amount of time. Before making the assignment, the participant first sampled the cold water himself for five seconds to learn that cold-water immersion is unpleasant. The participant then assigned the partner to a duration of cold-water immersion on a 9 point scale, ranging from 0 to 80 seconds in 10-second intervals. This measure can be quantified only in one way (e.g. 1-9 rating), eliminating the concerns about which is the “correct” quantification strategy often associated with the competitive reaction time measure of aggression (see Elson et al., 2014). This measure is attached as Appendix A.

Manipulation checks. Participants completed a questionnaire assessing the efficacy of the various parts of the experimental manipulation. First, participants rated their exchange with their partner for how helpful, pleasant, irritating, etc. their partner’s feedback was. Then, participants rated the video game they played, indicating how violent, enjoyable, exciting, and challenging it was. Participants then rated their degree of experience with video games, first-person shooter video games, and playing video games with a keyboard and mouse. Finally, participants provided demographic information about themselves. This measure is attached as Appendix B.

Probe for suspicion. Participants completed a questionnaire intended to imitate a funneled debriefing. It begins with broad questions about the study and its purpose, and whether anything seemed strange about the study, and then grows increasingly specific, asking the participant about the aggression measure and other participant in the study. This measure is attached as Appendix C.

Materials

Modified video games. Four modified versions of the video game *Doom II* (iD Software, 1994) were created using software modification tools (Judd, 2011; vd Heiden, 2012). These four versions were designed to create a 2 (Difficulty: Easy, Difficulty) x 2 (Violence: Nonviolent, Violent) Latin squares design.

Across the four video games, all gameplay variables are held constant. The player moves at the same speed, the player's abilities have the same effects on enemies, and the enemies have the same abilities and artificial intelligence. A series of unique levels were designed that would be easy for players to navigate. This was done to minimize the amount of time players spent wandering aimlessly or being lost and maximize the amount of time engaged in gameplay and violence, as appropriate. All four versions of the game used the same levels so that level geography and the placement of supplies and enemies were the same across conditions. In the case that the player's health was reduced to zero, he would start again from the most recent of six checkpoints.

Violent content of the games was manipulated by changing the graphical and auditory representation of the player's tools and of the enemies. In the nonviolent version, enemy graphics and sounds were borrowed from *Chex Quest* (Digital Café, 1996), a modified version of *Doom II* that replaces the enemies with silly-looking booger aliens. The players' weapons are similarly replaced with "zorchers", science-fiction tools similar in appearance to remote controls. Participants in this condition are told that the aliens are lost and confused and need to be sent home with the zorcher. Players maintain their health and ammunition by picking up fruits, vegetables, “zorch pellets,” and “zap tapes.” In the violent version, enemy graphics and sounds were borrowed from *Brutal Doom* (Abenante, 2012), a modified form of *Doom II* that increases the degree of violence by making defeated enemies explode into fountains of gore, severed limbs, and scattering teeth. In the violent version of the game, the texture of some map scenery was replaced with more hellish imagery such as rivers of blood, demonic skulls, or bodies chained to walls. The functional aspects of map geometry remained the same across versions, however.

The difficulty of the games was manipulated by changing the enemies’ artificial intelligence. In the difficult version of the game, the enemies fought per their original artificial intelligence. In the violent version of the game, monsters tried to wound the player with guns, claws, teeth, or fireballs. In the nonviolent version of the game, aliens tried to slime the player by throwing boogers. Thus, in the difficult version of the game, it was possible that players would be wounded or slimed too many times and have to restart the level. Players had to attend to the game environment to find supplies such as health, armor, and ammunition. In the easy version of the game, however, enemies had their artificial intelligence changed so that they could not attack the player. Instead, they would walk very slowly towards the player and wait to be killed or zorched. In the easy version of the game, it was impossible for the player to lose health or to have to restart the level. Players were also given infinite ammunition so that they would not have to search the environment for supplies.

The modified games were also programmed to track players’ in-game behavior and performance. Across the gameplay session, the game tracked: the number of times the player had to restart the level, the number of enemies slain or zorched, the number of times the rapid-fire tool was used, the number of times the slow and powerful tool was used, the furthest point reached by the player, and the number of times the player was hit by an enemy.

Procedure

Participants arrived at the lab in pairs and were immediately escorted to separate adjacent rooms. Following consent, participants' hands were photographed with a flatbed scanner for measurement of 2D:4D. Because there was only one scanner, participants were able to see each other as scans were taken, demonstrating the presence of another participant in the study. After scanning, participants returned to their desks.

Participants were then given an envelope, a sheet of loose-leaf paper, and a printed essay prompt. They were informed that the first task is to write a five-minute persuasive essay of their personal views on abortion which would later be judged by the other participant. (To justify this practice, participants were told that participants rate essays just as well as trained research assistants.) At the end of these five minutes, the essays were collected so that they purportedly could be exchanged with the other participant.

Instead of exchanging the essays, each participant received a fake, premade essay designed to oppose their beliefs. Participants who wrote a pro-life essay received a pro-choice essay, while participants who wrote a pro-choice essay received a pro-life essay. With this essay, participants received a form for rating the essay. This form asked participants to rate the organization, originality, writing style, clarity of expression, persuasiveness of arguments, and overall quality of the essay. Participants also could leave comments. Once finished, the participant returned the essay and the evaluation form to the partner's envelope, which was then taken from the room, ostensibly for data entry.

Participants then played their assigned version of the video game. Each received a cover story that explained the story and controls of the game (see attached in Appendix D). In the nonviolent condition, the story explained that the booger aliens are lost and confused, and that when the player has "zorched" them all, he sees a scene of the aliens playing together on their homeworld. By comparison, in the violent condition, the story explains that the aliens must all be slain, and that when the player has killed them all, he sees a scene of the player character posing with his shotgun. The cover story also explained whether enemies would or would not attack the player per the difficulty manipulation.

Participants were then given 15 minutes to play the game. They were monitored for a few minutes to make sure that they successfully completed the first level of the game and moved on to the second level, at which time the participant was left to play alone.

While the participant played the video game, materials were prepared for subsequent provocation and measurement of aggression. An insulting essay evaluation form was placed in the participant's envelope; on it, the partner had rated all dimensions as between -8 and -10 in quality, and commented "This is the stupidest thing I've ever read." To prepare the coldpressor task, a dozen ice cubes were added to the coldpressor pitcher 5 minutes before the end of the game session.

When the game session ended, the research assistant brought the coldpressor pitcher and a towel into the room. A key was pressed on the keyboard to print the game variables, which the assistant then logged. The game was then quit by pressing Alt+F4. The RA then navigated to a folder containing an E-Prime task in preparation for the purported second portion of the experiment.

At this point, the participant was told that the next portion of the experiment involves performing a computer task while distracted by cold-water exposure. The participant was asked to sample the coldpressor by placing his fist in it for five seconds. At the end of five seconds, the participant was allowed to withdraw his hand and towel off. The participant was then asked if he would be okay with the coldpressor. (No participants indicated unwillingness to participate in the coldpressor task.)

The research assistant then brought the participant's original envelope into the room and asked him to read the partner's rating of his essay. The research assistant again left the room to fetch a distraction assignment form and gave it to the participant, explaining that "to avoid experimenter bias," participants were being asked to randomly assign each other to the various levels of distraction. The participant was asked to circle a number on the sheet, thereby assigning the partner to an amount of coldpressor exposure ranging from 0 seconds to 80 seconds in 10 second intervals.

Once this sheet was retrieved, participants were told that the experiment was running out of time and that the distraction task would be skipped. Participants completed post-questionnaires asking them to rate the games, their partner's feedback, and what they suspected was the purpose of the study. Participants were then fully debriefed and dismissed.

Results

For each tested effect, I report an effect size and confidence interval as well as a Bayesian model comparison. In Bayesian model comparison, an alternative hypothesis is explicitly stated for each effect. While in frequentist power analysis it is common to presume a single effect size (e.g. a point hypothesis of $r = .21$), Bayesians can describe effects in probability distributions. Effects are often described as following a Cauchy distribution of certain scale such that large scale parameters reflect large effects and small scale parameters reflect small effects. In the present research, effects are expected to be small, so alternative hypotheses were set to $\delta \sim$ Cauchy(.4). By comparing the probability of the data given the null hypothesis ($d = 0$) and the probability of the data given the alternative hypothesis ($\delta \sim$ Cauchy(.4)), a Bayes factor is obtained. The Bayes factor describes how many times more likely the data are given one hypothesis than the other. This Bayes factor also has a natural interpretation as the multiplicative change in beliefs. If one believed an effect was more likely than no effect with 10-to-1 odds, but the data favor the null hypothesis with 1-to-15 odds, one should update beliefs to 1-to1.5 against the effect.

Quality Control

Of the 335 participants, 86 indicated on the debriefing form that the purpose of the experiment was to study the effects of violent games on aggressive behavior without selecting any of the other offered purposes. A further 2 subjects had gameplay data indicating that they had been wounded or slain in the easy game condition. A further 24 subjects were excluded because the research assistants indicated some failure of deception or of methodology. The effective sample size was 223. Of these, digit ratios are available for only 152 at the present moment.

We note that our failure of deception rate of 25.7% is considerably higher than our anticipated 11% rate or of rates reported in previous work. It is possible that our funneled debriefing process is more sensitive than that used in previous work. It is also possible that our hypothesis was too obvious to the participants given the study design. A final possibility is that repeated press releases on the study of violent game effects is reaching broader awareness, reducing the proportion of naïve participants in the population. To be conservative, I report analyses with hypothesis-aware participants removed.

Manipulation Check

Participant ratings on the post-questionnaires were submitted to 2 (Violence) x 2 (Difficulty) ANOVA. The manipulation was highly effective: participants indicated that the violent game (M = 5.2; SD = 1.27) was much more violent than the nonviolent game (M = 2.2, SD = 1.49; *d* = 2.2 (1.87, 2.54)).

Mean evaluations of the participants' interactions with the partner were also assessed. Participants generally indicated that they were irritated (*M* = 4.92, *SD* = 1.71), angered (*M* = 4.22, *SD* = 1.75), and annoyed (*M* = 4.92, *SD* = 1.80) by their partner. Furthermore, they were not happy (*M* = 2.45, *SD* = 1.41) or pleased (*M* = 2.18, *SD* = 1.37) with their partner and found the feedback unhelpful (*M* = 1.78, *SD* = 1.23).

To determine whether the coldpressor dependent variable was a sensitive measure of aggression, I tested whether these participant evaluations were related to coldpressor assignments. First, a principal component was extracted from participants' six ratings of the interaction, described above. The first component accounted for 57% of the variance and had the expected pattern of loadings: .51, .45, and .50 for irritation, anger, and annoyance, -.35, -.22, and -.33 for happiness, helpfulness, and pleasure. This component, hereafter referred to as composite irritation, was then used as a linear predictor of coldpressor assignment. The relationship was moderately strong, $t(196) = 5.43$, $r = .36$ (.22, .46), suggesting that the coldpressor measure was indeed influenced by participants' intent to aggress. A scatterplot and loess regression line are provided in Figure 1.

A 2 (Violence) x 2 (Difficulty) ANOVA was conducted to determine whether the game played influenced participants' ratings of the interaction. Effects were small and not statistically significant, suggesting that the game played had a minimal influence on participants' composite irritation. See Table 1 for this ANOVA output.

Primary Outcome

Coldpressor assignments were found to be non-normally distributed. Distributions appeared to resemble a mixture of a uniform and a point such that participants either followed directions and assigned a random value between 1 and 9 or they decided to aggress against their partner and assigned a 9. See histograms in Figure 2 and means and SDs in Table 2.

Because of this non-normal distribution, I attempted to model the data in several ways. First, I treated the data as normally distributed for a typical ANOVA, generating effect sizes, confidence intervals, and Bayes factors. Next, I treated the data as being censored from above, attempting to model possible coldpressor assignments above the maximum. Finally, I treated coldpressor assignment as a categorical outcome with 1-8 representing a single nonaggressive response category and 9 representing an aggressive response category. This categorized variable was analyzed with logistic regression.

Conventional ANOVA. Beginning with the full 2 (Violence) x Submitting the data to ANOVA, effects were found to be very small. Estimates of the main effects depended considerably on the treatment of the 2 (Violence) x 2 (Difficulty) interaction, which was statistically significant ($t(219) = -2.21$, $r = -.14$ (-.27, -.01)) but negative, such that violent content increased aggressive behavior among players of the easy game ($r = .20$, (.01, .37)) but decreased aggressive behavior among players of the difficult game ($r = -.10$, (-.28, .09)). This interaction would seem at odds with the previous literature on violent game effects, which almost exclusively uses video games in their default, challenging parameters (e.g. my difficult-game condition). Suffice it to say that this interaction does not support the hypothesis of super-additive effects (H4, above) and is not interpretable under the theories outlined previously.

If this uninterpretable interaction is included in the ANOVA, the main effects of Violence and of Difficulty are small, positive, and statistically significant (Violence: $t(219) = 2.04$, $r = .14$ (.00, .26); Difficulty: $t(219) = 2.19$, $r = .15$ (.01, .27)). Because this interaction is negative, representing a cross-over, removing it from the model causes a dramatic decrease in the main effects (Violence: $t(220) = 0.67$, $r = .05$ (-.09, .18); Difficulty: $t(220) = 0.89$, $r = .06$ (-.07, .19)). These estimated effects are dramatically smaller than those reported in meta-analyses of previous violent-games research ($r = .21$, Anderson et al., 2010; $r = .19$, Greitemeyer & Mugge, 2014). A frequentist might even say that they are *statistically significantly* smaller than the previously-reported effect sizes.

Main effects of left and right 2D:4D were negligible ($t(151) = -0.19$, $r = -.02$ (-.17, .14); $t(151) = .129$, $r = .01$ (-.15, .17). Two- and three-way interactions of 2D:4D with violence and difficulty were also small and negligible (all $|t| < 1.3$).

Because the earlier manipulation and sensitivity check indicated that much of the variance in aggression could be predicted by composite irritation and that composite irritation was largely orthogonal to the experimental manipulation, composite irritation was added as a covariate. However, this did not increase the observed effect size. In the 2x2 ANOVA, effects of Violence, Difficulty, and their interaction were small: $t(193)$s = 1.40, 1.81, and -1.62; *r*s = .09 (-.04, .24), .13 (-.01, .26), and -.11 (-.25 .03), respectively. When the interaction term was dropped, main effects again shrank (Violence: $t(194) = 0.36$, $r = .03$ (-.11, .16); Difficulty: $t(194) = 0.93$, $r = .07$ (-.07, .20)).

Bayesian ANOVA. Models were compared using the BayesFactor package for R (Morey & Rouder, 2014). Because effects are expected to be small, I adjusted the scale of the effect size under the alternative hypothesis to ~Cauchy(.4). Models were generated to represent all possible combinations of main effects and/or interactions. Models including interactions were constrained to also include lower-order interactions and main effects. All models were compared to a null-hypothesis model including no effects. Bayes factors involving 2D:4D were similar regardless of whether the right or left hand was used; to be conservative, I report the Bayes factor closer to 1.

Of all the models, the null-hypothesis model was best supported by the data. Models of main effects of Violence, Difficulty, or 2D:4D were each outperformed by the null model (Bayes factors = 4.51, 3.87, and 5.64 in favor of the null, respectively). Models containing interactions were further outperformed by the null. The full model of

2 (Violence) x 2 (Difficulty) x 2D:4D was not preferred to the null (Bayes factor = 558). The 2 (Violence) x 2 (Difficulty) model was similarly outperformed by the null (Bayes factor = 8.69). Thus, the null model was supported over the hypothesized effect of each predictor.

When composite irritation was added as a predictor, Bayes factor strongly favored the composite-irritation-model to the null model, $B = 73{,}980$. This model was also preferred to models adding effects of violence ($B = 5.01$), difficulty ($B = 3.55$), additive effects of violence and difficulty ($B = 17.93$), or interactive effects of violence and difficulty ($B = 22.72$). This indicates that variance in coldpressor duration could be predicted by composite irritation but not by game condition.

Censored regression. To attempt to ameliorate the potential ceiling effect, a censored regression model was fit with the 'censReg' package for R (Henningsen, 2013). This fits a censored-regression Tobit model and attempts to model values that exceed the maximum of the scale.

Again, the 2 (Violence) x 2 (Difficulty) model was fit. As in the conventional ANOVA, a small and negative interaction was found ($t(219) = -2.26$, $r = -.15$ (-.27, -.02)) such that violent content increased aggressive behavior in the easy condition ($t(110) = 1.95$, $r = .18$ (-.01, .36)) but decreased aggressive behavior in the difficult condition ($t(111) = -1.32$, $r = -.12$ (-.30, .06)). Main effects of Violence and Difficulty were small, $t(219) = 1.84$, $r = .12$ (-.01, .25) and $t(219) = 2.41$, $r = .16$ (.03, .28), respectively.

As before, this interaction seems uninterpretable in light of the theoretical predictions. Removing this interaction from the model again dramatically reduced the size of the estimated main effects. Effects of Violence and Difficulty were very close to zero, $t(223) = 0.34$, $r = .02$ (-.11, .15) and $t(220) = 1.15$, $r = .07$ (-.05, .21), respectively.

Main effects of 2D:4D were again approximately zero, $t(151) = -0.19$, $r = -.02$ (-.17, .14) and $t(151) = 0.13$, $r = .01$ (-.15, .17) for left- and right-hand 2D:4D, respectively. Higher-order interactions with Violence or Difficulty were not supported (all $|t| < 1.5$).

Logistic regression. Another possibility is that participants completed the coldpressor assignment in one of two ways: either they followed instructions and randomly assigned the other participant to a value between 1 and 9, or they decided to aggress and assign the other participant the maximum value. To model this possibility, I treated the response variable as a dichotomous outcome. Participants assigning values 1-8 were treated as one category (nonaggressive response) and participants assigning value 9 were treated as the other (aggressive response). Logistic regression was performed to test whether the odds of aggressing were influenced by the experimental assignment.

Analysis began with a 2 (Violence) x 2 (Difficulty) general linear model with a logit link function. A small negative interaction was again observed, although it was not statistically significant ($t(219) = -1.60$, $r = -.11$ (-.23, .03). Here, the nature of this interaction was such that Violence decreased aggression when the game was easy, $t(103) = -0.63$, $r = -.06$ (-.25, .13) and decreased it to a lesser extent when the game was hard, $t(106) = -.284$, $r = -.03$ (-.22, .16). The main effect of Violence was quite small, $t(219) = 0.72$, $r = .05$ (-.08, .18), although participants were slightly more likely to aggress in the Difficult condition, $t(219) = 2.21$, $r = .15$ (.01, .27).

Again, because the negative interaction is difficult to interpret given the relevant theory, it was dropped from the model and main effects again estimated. Violence did not appear to influence aggression, $t(220) = -0.70$, $r = -.05$ (-.18, .08). Difficulty also had a minimal effect on aggression, $t(220) = 1.58$, $r = .11$ (-.03, .23). Application of composite irritation as a covariate to these models revealed an effect of composite irritation, $t(196) = 4.19$, $r = .29$ (.15, .40), but did not increase the estimated effects of violence, difficulty, or their interaction.

Main effects of 2D:4D on aggression were again negligible. Left 2D:4D did not predict aggression, $t(151) = -0.15$, $r = -.01$ (-.17, .15), nor did right 2D:4D, $t(150) = -0.05$, $r = -.05$ (-.21, .11). Application of composite irritation as a covariate did not influence the estimated effect. Higher-order interactions of 2D:4D with factors of Violence or Difficulty were not supported by the results (all $|t| < 1.53$).

Bayesian probit regression. I am currently working on a software implementation of Bayesian probit regression. This software would allow for Bayesian analysis of the categorical outcome. This would make it possible to compare the probability of the null vs. a reasonable alternative hypothesis, yielding Bayes factors as a summary of observed evidence. A further extension of this technique could model the outcome as a mixture of a binomial and a uniform distribution to inspect whether the game played influenced the probability of assigning a 9 or some other value, and additionally, whether the game influenced the mean of the assignments below 9.

Non-local Bayesian prior. In the Bayesian hypothesis tests provided above, we use a non-directional, non-specific alternative hypothesis scaled roughly to the magnitude of the expected effect. While this is a useful hypothesis to test, it would also be useful to compare the obtained results against a more specific alternative hypothesis representing the effect as estimated from previous meta-analysis, $\delta = .43$ (.35, .52) (Anderson et al., 2010).

The main effect of Violence in the traditional ANOVA, omitting the Violence x Difficulty interaction, was $d = 0.09$ (-0.17, 0.35). An online Bayes factor calculator (Dienes, 2008) was used to compare the evidence for H_0: $\delta = 0$ relative to H_1: $\delta = .43$ (.35, .52). The obtained Bayes factor substantially preferred the null, $B_{01} = 17.7$.

Discussion

Results indicate that when game stimuli are carefully controlled the effects of fifteen minutes of violent gameplay are likely to be small and not meaningfully different from zero. Similarly null effects were observed for game difficulty and for 2D:4D. Because observed effects were small and sampling precision was high, the present study provided considerable evidence against hypothesized effects of violence, difficulty, and 2D:4D.

The presented manipulation and sensitivity checks give me confidence that the null results are not due to failures of the methodology. First, participants indicated that the violent game was much more violent than the nonviolent game. Second, participants were generally irritated with their essay feedback. Lastly, the coldpressor measure of aggression was sensitive to participants' irritation with their partners. This sensitivity suggests that the null result is not due simply to the unusual distribution of the data or an overall invalidity of the coldpressor measure.

Effects of Violent Video Games

The current study indicates that, when game stimuli are tightly controlled, effects of violence in a brief laboratory experiment are minimal. Models without such effects are better supported by the data than are models with such effects. These results parallel our findings from a similar study with the same game stimuli but using different outcomes: noise-blasts in the Competitive Reaction-Time Task, ratings of aggressive affect, and measurements of aggressive-word accessibility (Engelhardt, Mazurek, Hilgard, Rouder, and Bartholow, in press).

The present research provides a closer experimental control than previous experiments. It has previously been argued that researchers have matched their stimuli on all reasonably possible confounds (Anderson et al., 2004). As outlined above, studies that conduct a pilot test and find no significant difference cannot demonstrate the truth of the null hypothesis of no true difference between stimuli. Similarly, studies using ANCOVA to "control for" confounds cannot be certain that all variance associated with the confounds have been removed. The tighter experimental controls of this research may have reduced the apparent effect size.

My other work draws similar conclusions from meta-analysis (Hilgard, in prep). In their meta-analysis, Anderson et al. (2010) argue significant effects of violent games on aggressive behaviors in laboratory experiments. Moreover, they argue that better-designed studies find larger effects than do studies on average. It seems that this criterion, among others, actually *increases* the degree of bias in publication and selection. While the naïve meta-analytic estimate of the effect on aggressive behavior is larger in the "best" studies than in all the studies on average, the "best" studies also have a more dramatically asymmetrical funnel plot, a sign of research bias.

To account for this research bias, I applied PET-PEESE meta-regression (Stanley & Doucouliagos, 2012), a statistical technique that examines the degree of funnel plot asymmetry and estimates a bias-corrected effect size. The concept is similar to that of the trim-and-fill procedure (Duvall & Tweedie, 2000), which Anderson et al. (2010) did apply; however, trim-and-fill is expected to perform poorly compared to PET-PEESE due to violations of the assumptions of trim-and-fill. In simulations, trim-and-fill tends to under-correct for bias when bias is present and over-correct for bias when bias is absent, even when its assumptions are met (Simonsohn, Simmons, & Nelson, 2014).

Application of PET-PEESE finds that the best-practices subset has greater research bias and a *smaller* estimated effect size than all studies on average (see Figure 3). This might suggest that better-designed studies find smaller effects but are subject to greater publication or analytic bias. Another possibility is that studies were selected as best- or not-best-practices *post hoc* on the basis of their obtained results. This may explain the apparent inconsistency in the application of some of the inclusion criterion (see Lakens, Hilgard, and Staaks, in press).

Considering that the bias-corrected meta-analytic effect size estimate is small and may yet involve confounds, it is plausible that the current study has accurately measured the true effect as being small and well-described by the null hypothesis.

This finding has implications for future laboratory research of violent media and aggressive behavior. If main effects of brief violent media manipulations are small, then laboratory paradigms may not be appropriate for developing elaborated and refined theories of violent media effects. A study hoping to find moderators or boundary conditions of the effect may need hundreds or even thousands of subjects to detect the anticipated interaction. Previous research detecting such interactions may involve an amount of *hypothesizing after results are known* (“HARKing”; Kerr, 1998) or post-hoc application of moderators (“moderator munging”). A skeptical reader might wonder whether, say, the interaction between game violence and rumination in predicting aggressive behavior truly is moderated by gender (Bushman & Gibson, 2010), or whether gender was added as a moderator when the predicted 2 (Violence) x 2 (Rumination) interaction could not be obtained.

It still seems likely that violent media has effects on its audience – just that such effects are difficult to detect in a single 15-30 minute laboratory gameplay session. By comparison, it seems rather more plausible that violent games can influence behavior over the course of hundreds of hours of gameplay over months and years of development. Thus, despite my skepticism, I would not endorse the sale of violent games to minors. Instead, I hope to make researchers aware that we may be deceiving ourselves as to the precision and predictive power of our hypotheses. It may be necessary to re-evaluate current measures, research, and theories, particularly with regard to brief experimental paradigms.

Effects of Difficult Video Games

The obtained results also appear inconsistent with the results of research indicating effects of competitive (Adachi & Willoughby, 2011b) or competence-thwarting (Przybylski et al., 2014) video games. Regarding effects of competitive games, sample sizes in the research presented by Adachi and Willoughby are small, and effects may have been misestimated. Furthermore, while games used in that research were thought to vary in their competitive content, they were not so tightly controlled as these, and so confounds may have increased the size of the obtained effect. Finally, some of the manipulations in that research contrasted competitive games against cooperative games, which may have larger effects than a comparison between a competitive and neutral game as in the present research.

Concerning the effects of competence-thwarting games, we must consider the potential differences between difficulty and competence-thwarting. In their research, Przybylski et al. (2014) measured players' comfort with the video games' controls, then used that comfort or discomfort to predict aggressive affect and behavior. In other experiments, they deliberately made the game controls awkward and unintuitive to use.

In the present research, it was expected that more difficult gameplay would, at least indirectly, lead to increased feelings of thwarted competence. Perhaps players would find themselves struggling with the controls more under the pressure, or they would find the in-game challenges unfair and frustrating. This may not have been the case. The game's controls were deliberately kept as simple as possible across all conditions, so perhaps the difficult-game condition represented an exciting and fair challenge rather than a competence-thwarting chore.

The present results also speak against the possibility I argued in my previous research on effects of violent and difficult video games (Engelhardt et al., 2015). In that research, we had observed an apparent effect of difficult gameplay such that players who were challenged by the game did more poorly on a subsequent modified Stroop task. We suggested that difficult games may exhaust mental resources, impairing later recruitment of cognitive control. We further suggested that such impairments of cognitive control may lead to greater aggressive behavior. The difficult games used in this experiment were designed to be more challenging than those in the previous experiment so that the effect might be easier to obtain. Instead, we do not find effects of game difficulty on aggressive behavior, suggesting either that difficult games do not deplete cognitive control or that depleted cognitive control does not lead to increased aggression (cf. Bushman, DeWall, Pond, & Hanus, 2014).

Digit Ratio

The present study finds strong evidence against presumed effects of 2D:4D. Theory suggests that 2D:4D should be negatively associated with aggression – that is, that participants with more masculine 2D:4D will be more aggressive. The generality of this prediction has been gradually shrinking over the past few years, with the most recent theory suggesting that 2D:4D only predict aggressive behavior among men in contexts involving provocation, as these contexts have aggression as a behavior that is accessible and available to participants (Millet, 2011; Millet & Dewitte, 2007; see Benderlioglu & Nelson, 2004; McIntyre et al., 2007). The present study features only male subjects, all provoked and given opportunity to aggress, but no such effect could be found. The present study supports other research indicating the invalidity of 2D:4D.

Replication across Laboratories

Research in this area has been somewhat divided, with certain researchers tending to find effects (e.g. Anderson and Bushman) and other researchers tending to not find effects (e.g. Ferguson). One recent meta-analysis has suggested heterogeneity in effect size according to research team (Greitemeyer & Mügge, 2014). Independent research by new research teams may help to reduce the dichotomization of research findings and bring a greater degree of consensus to research findings. It is worth noting that research in this laboratory has historically found effects of violent games (e.g. Engelhardt, Bartholow, Kerr, & Bushman, 2011; Engelhardt, Bartholow, & Saults, 2011; Sestir &

Bartholow, 2010), but not of such precisely-matched stimuli as used here (see also Engelhardt et al., in press).

Limitations

First, the distribution of coldpressor assignments was found to not resemble a normal distribution. Many participants, albeit a minority, assigned their partner the maximum coldpressor duration. Others seemed to randomly assign their partner's coldpressor duration. The obtained data roughly resemble a uniform distribution with a spike at 9. I attempted several ways to model this data: traditional ANOVA, treating the outcome as being censored from above, and logistic regression comparing the probability of a 9 against the probability of a 1-8. Results were comparable across modeling approaches. Data will be publicly archived for further modeling attempts. It is possible that the distribution of the data reflects a ceiling effect and that the effect size was diminished due to the restricted range of the measure, but again, the measure's sensitivity to participants' irritation may suggest otherwise.

It is possible that a mere 15 minutes of gameplay in the laboratory is not enough to elicit and test the effects of violent video games. This is not a unique weakness of this research, as most experimental studies involve approximately 15-30 minutes of gameplay. However, this would make it possible for the proposed study to yield null findings when the true effect in the real world after many hours is nonzero. Future longitudinal research may be needed to inspect the influence of game violence as an effect unique from game content or game genre. Other research might intend to inspect the influence of several hours of violent game play over several weeks.

The proposed study is also limited in that the research assistants were not blind to the participants' conditions. When bringing participants their cover stories and when recording their gameplay variables, the research assistants may have been able to detect whether participants are assigned to the violent or nonviolent game. However, previous research has not been blinded either (personal communication). While it is an enticing theoretical possibility that the divergent results between research labs is due to the beliefs of research assistants, this idea would have to be explored in a later research project with greater resources.

It is also possible that the violent game used in this study differs meaningfully from that used in other studies. For example, perhaps *Doom II* is too fantastic of a setting, and a more realistic and grounded game such as *Grand Theft Auto* would instead show larger effects. However, effects have been observed for fantasy games (Anderson & Ford, 1987; Konijn, Nije Bivank, & Bushman, 2007), as well as realistic ones. This explanation seems unlikely.

Finally, it is possible that the nonviolent *Chex Quest* game involves substantial amounts of violence. Current definitions and practices indicate that even E-rated games can contain substantial violence (Anderson et al., 2010; Thompson & Haninger, 2001), and that the effect of cartoon E-rated violence is as strong as that of explicit M-rated violence (Anderson, Gentile, & Buckley, 2007). These definitions and practices would seem to contradict the current theories of violent media that they are said to support; for example, exposure to more extreme violent content should be more desensitizing than mild violent content. In any case, it is possible that an effect was not found in the present study because even a relatively mild game such as *Chex Quest* has effects on aggression equal to those of *Brutal Doom*, however implausible this may seem. Future research may seek to compare the *Brutal Doom* game against a control game which involves no harm or conflict whatsoever, although this may risk confounding the effects of in-game conflict with those of violent content.

Summary

I find evidence against the prediction that brief exposure to violent games cause aggressive behavior. This evidence is corroborated by similar research with different measurements of aggressive outcomes (Engelhardt et al., in press). It seems that previous research on this topic either yielded results inflated by confounds (Adachi & Willoughby, 2011a; Hilgard et al., submitted) or by publication and selection bias (Hilgard, in prep). While longitudinal effects may truly exist, it is uncertain whether laboratory paradigms involving brief exposure can help to elucidate these mechanisms in light of the present null results.

2D:4D similarly predicts little in a laboratory experiment. Considered alongside other evidence of the invalidity of 2D:4D (Hönekopp & Watson, 2011; Voracek, submitted), it would seem that 2D:4D does not have much utility in understanding the causes and prevalence of aggression.

I note that, viewed in retrospect, the refinement of theories underlying the use of 2D:4D forms a distinct and familiar life-cycle. First, an extension of theory leads to the testing of an exciting and unconventional hypothesis. This hypothesis is tested in a small sample and an improbably large effect is obtained. Subsequent research manages to find effects, but they are often of dazzlingly complex interactions that could not have been predicted *a priori*. Meta-analysis finds small, potentially biased results, and researchers are encouraged instead to consider the specific and idiosyncratic contexts in which a result has been found. Eventually, the contexts grow too specific and too idiosyncratic, researchers recognize that the theory overfits the available data, and the theory collapses under its own weight. Finally, everyone goes out for a pint.

In the years ahead, I hope that there will be interest and resources for further study. Previous research findings may overstate the effects of violent games; it would be useful to know whether this is due to poor experimental control or due to bias in research practice. For future research, I hope to see more researchers using the current manipulation or manipulations like it. Recent years have demonstrated that obtained effects may vary dramatically across laboratories. Antagonistic collaboration could be especially helpful in creating informative results and soothing personal disputes. In the end, I hope that science communication to researchers and laypeople alike can be frank about what is and is not known about media effects.

References

Abenante, M. (2012). Brutal doom (Version 0.16) [Computer software]. http://http://www.moddb.com/mods/brutal-doom/.

Adachi, P. J. C., Allaire, J. C., Anderson, J., Annetta, L., Arnett, J. J., Arsenault, D., … Zerovnik, G. (2013). Scholars' open statement to the APA Task Force on Violent Media. Retrieved from http://www.christopherjferguson.com/APA%20Task%20Force%20Comment1.pdf

Adachi, P. J. C., & Willoughby, T. (2011a). The effect of violent video games on aggression: Is it more than just the violence? *Aggression and Violent Behavior, 16,* 55-62. doi: 10.1016/j.avb.2010.12.002

Adachi, P. J. C., & Willoughby, T. (2011b). The effect of video game competition and violence on aggressive behavior: Which characteristic has the greatest influence? *Psychology of Violence, 1,* 259-274. doi: 10.1037/a0024908

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

Anderson, C. A., & Ford, C. M. (1986). Affect of the game player: Short-term effects of highly and mildly aggressive video games. *Personality and Social Psychology Bulletin, 12,* 390-402. doi: 10.1177/0146167286124002

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

American Academy of Pediatrics, Council on Communications and Media. (2009). Media violence. *Pediatrics, 124,* 1495-1503. doi: 10.1542/peds.2009-2146

American Psychological Association, Task Force on Violent Media. (2005). Resolution on violence in video games and interactive media. Retrieved from https://www.apa.org/about/policy/interactive-media.pdf.

Archer, J. (2009). Does sexual selection explain human sex differences in aggression? *Behavioral and Brain Sciences, 32,* 249-266. DOI: http://dx.doi.org/10.1017/S0140525X09990951

Bandura, A., & McClelland, D. C. (1977). *Social learning theory.* New York City: General Learning Press.

Bandura, A., Ross, D., & Ross, S. A. (1961). Transmission of aggression through imitation of aggressive models. *The Journal of Abnormal and Social Psychology, 63,* 575-582. http://dx.doi.org/10.1037/h0045925

Benderlioglu, Z., & Nelson, R. J. (2004). Digit length ratios predict reactive aggression in women, but not in men. *Hormones and Behavior, 46(*5), 558-564. doi:10.1016/j.yhbeh.2004.06.004

Berkowitz, L. (1984). Some effects of thoughts on anti- and prosocial influences of media events: A cognitive-neoassociation analysis. *Psychological Bulletin, 95*(3), 410-427. doi: 10.1037/0033-2909.95.3.410

Buckley, K. E., & Anderson, C. A. (2006). A theoretical model of the effects and consequences of playing video games. Chapter in P. Vorderer & J. Bryant (Eds.), *Playing Video Games – Motives, Responses, and Consequences* (pp. 363-378). Mahwah, NJ: LEA.

Bushman, B. J., DeWall, C. N., Pond, R. S., Jr., & Hanus, M. D. (2014). Low glucose relates to greater aggression in married couples. *PANAS, 111*(17), 6254-6257. doi: 10.1073/pnas.1400619111

Bushman, B. J., Gollwitzer, M., & Cruz, C. (in press). There is broad consensus: Media researchers agree that violent media increase aggression in children, and pediatricians and parents concur. *Psychology of Popular Media Culture*. DOI: 10.1037/ppm0000046

Bushman, B. J., & Gibson, B. (2010) Violent video games cause an increase in aggression long after the game has been turned off. *Social Psychological and Personality Science, 2*(1), 29-32. doi: 10.1177/1948550610379506

Campbell, A. (2006). Sex differences in direct aggression: What are the psychological mediators? *Aggression and Violent Behavior*, *11*(3), 237-264. DOI: 10.1016/j.avb.2005.09.002

Carnagey, N. L., & Anderson, C. A. (2005). The effects of reward and punishment in violent video games on aggressive affect, cognition, and behavior. *Psychological Science, 16*(11), 882-889. doi: 10.1111/j.1467-9280.2005.01632.x

Digital Café. (1996). Chex Quest [Computer software]. http://www.chexquest.org/index.php?action=downloads;cat=1.

Duvall, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455-463. DOI: 10.1111/j.0006-341X.2000.00455.x

Elson, M., Breuer, J., Van Looy, J., Kneer, J., and Quandt, T. (2013). Comparing apples and oranges? Evidence for pace of action as a confound in research on digital games and aggression. *Psychology of Popular Media Culture*. doi: 10.1037/ppm0000010

Elson, M., Mohseni, M. R., Breuer, J., Scharkow, M., & Quandt, T. (2014). Press CRTT to measure aggressive behavior: The unstandardized use of the Competitive Reaction Time

Task in aggression research. *Psychological Assessment, 26*, 419-432. doi: 10.1037/a0035569

Engelhardt, C. R., Bartholow, B. D., Kerr, G. T., & Bushman, B. J. (2011) This is your brain on violent video games: Neural desensitization to violence predicts increased aggression following violent video game exposure. *Journal of Experimental Social Psychology, 47*(5), 1033-1036. doi:10.1016/j.jesp.2011.03.027

Engelhardt, C. R., Bartholow, B. D., & Saults, J. S. (2011). Violent and nonviolent video games differentially affect physical aggression for individuals high vs. low in dispositional anger. *Aggressive Behavior, 37*(6), 539-546. DOI: 10.1002/ab.20411

Engelhardt, C. R., Hilgard, J., & Bartholow, B. D. (2015). Acute exposure to difficulty (but not violent) video games dysregulates cognitive control. *Computers in Human Behavior, 45*, 85-92. doi:10.1016/j.chb.2014.11.089

Engelhardt, C. R., Mazurek, M. O., Hilgard, J., Rouder, J. N., & Bartholow, B. D. (in press). Effects of Violent Video Game Exposure on Aggressive Behavior, Aggressive Cognition, and Aggressive Affect among Adults with and without Autism Spectrum Disorder. *Psychological Science*.

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

Ferguson, C. J. (March, 2014). Violent video games don't exist. *The Huffington Post*. Retrieved from http://www.huffingtonpost.com/christopher-j-ferguson/violent-video-games-dont-_b_5051733.html

Finkel, E. J. (2013). The I^3 Model: Metatheory, theory, and evidence. In J. M. Olson & M. P. Zanna (Eds.), *Advances in Experimental Social Psychology, 49.* San Diego: Academic Press.

Gabbiadini, A., Riva, P., Andrighetto, L., Volpato, C., & Bushman, B. J. (2013) Interactive effect of moral disengagement and violent video games on self-control, cheating, and aggression. *Social Psychological and Personality Science, 5,* 451-458. doi: 10.1177/1948550613509286

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423,* 534-537. Retrieved from http://www.nature.com/nature/journal/v423/n6939/full/nature01647.html

Greitemeyer, T., & Mügge, D.O. (2014). Video games do affect social outcomes: A meta-analytic review of the effects of violent and prosocial video game play. *Personality and Social Psychology Bulletin*, published online before print. doi: 10.1177/0146167213520459

Greitemeyer, T., & Osswald, S. (2010) Effects of prosocial video games on prosocial behavior. *Journal of Personality and Social Psychology, 98*(2), 211-221. doi: 10.1037/a0016997

Henningsen, A. (2013). Censored Regression (Tobit) Models. R package version 0.5-20. http://cran.r-project.org/package=censReg

Hilgard, J. (in prep). There should not be broad consensus: Meta-regression reveals overestimated evidence of violent video game effects.

Hilgard, J., Engelhardt, C. R., & Bartholow, B. D. (In revision). Defining game violence and other game contents.

Hilgard, J., Engelhardt, C. R., Bartholow, B. D., & Rouder, J. N. (Submitted). A Bayesian reanalysis of studies in violent media research. https://osf.io/vgfyw/

Hönekopp, J., & Watson, S. (2011). Meta-analysis of the relationship between digit-ratio 2D:4D and aggression. *Personality and Individual Differences, 51*, 381-386. DOI: 10.1016/j.paid.2010.05.003

Huesmann, L. R. (1986). Psychological processes promoting the relation between exposure to media violence and aggressive behavior by the viewer. *Journal of Social Issues, 42*, 125-139. DOI: 10.1111/j.1540-4560.1986.tb00246.x

Huesmann, L. R. (1998). The role of social information processing and cognitive schema in the acquisition and maintenance of habitual aggressive behavior. In R. G. Geen & E. Donnerstein (Eds.), *Human aggression: Theories, research, and implications for social policy* (pp. 73-109). Waltham, Massachusetts: Academic Press.

Huesmann, L. R. (2010). Nailing the coffin shut on doubts that violent video games stimulate aggression: Comment on Anderson et al. (2010). *Psychological Bulletin, 136*(2), 179-181. doi: 10.1037/a0018567

iD Software (1994). Doom II [Computer software]. Rockville, Maryland: ZeniMax Media.

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

Judd, S. (2011). SLADE 3 (Version 3.0.2) [Computer software]. http://slade.mancubus.net/index.php?page=downloads/.

Kerr, N. L. (1998) HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*, *2*, 196-217. doi: 10.1207/s15327957pspr0203_4

Konijn, E. A., Nije Bijvank, M., & Bushman, B. J. (2007). I wish I were a warrior: The role of wishful identification in the effects of violent video games on aggression in adolescent boys. *Developmental Psychology, 43*, 1038-1044. DOI: 10.1037/0012-1649.43.4.1038

Lakens, D., Hilgard, J., & Staaks, J. (In press). On the reproducibility of meta-analyses: Six practical recommendations. *BioMed Central – Psychology*.

Lindsay, J. J., & Anderson, C. A. (2000) From antecedent conditions to violent actions: A general affective aggression model. *Personality and Social Psychology Bulletin, 26*(5), 533-547. doi: 10.1177/0146167200267002

Lutchmaya, S., Baron-Cohen, S., Raggatt, P., Knickmeyer, R., & Manning, J. T. (2004) 2nd to 4th digit ratios, fetal testosterone and estradiol. *Early Human Development, 77,* 23-28. doi:10.1016/j.earlhumdev.2003.12.002

Manning, J. T., Bundred, P. E., Newton, D. J., & Flanagan, B. F. (2003). The second to fourth digit ratio and variation in the androgen receptor gene. *Evolution and Human Behavior, 24,* 399-405. DOI: 10.1016/S1090-5138(03)00052-7

Manning, J. T., Scutt, D., Wilson, J., & Lewis-Jones, D. I. (1998). The ratio of 2nd to 4th digit length: a predictor of sperm numbers and concentrations of testosterone, luteinizing hormone and oestrogen. *Human Reproduction, 13,* 3000-3004. doi: 10.1093/humrep/13.11.3000

McIntyre, M. H., Barrett, E. S., McDermott, R., Johnson, D. D. P., Cowden, J., & Rosen, S. P. (2007). Finger length ratio (2D:4D) and sex differences in aggression during a simulated war game. *Personality and Individual Differences, 42*, 775-764. doi: 10.1016/j.paid.2006.08.009

Millet, K. (2011). An interactionist perspective on the relation between 2D:4D and behavior: An overview of (moderated) relationships between 2D:4D and economic decision making. *Personality and Individual Difference, 51,* 397-401. DOI: 10.1016/j.paid.2010.04.005

Millet, K., & Dewitte, S. (2007). Digit ratio (2D:4D) moderates the impact of an aggressive music video on aggression. *Personality and Individual Differences, 43,* 289-294. DOI: 10.1016/j.paid.2006.11.024

Millet, K., & Dewitte, S. (2009). The presence of aggression cues inverts the relation between digit ratio (2D:4D) and prosocial behavior in a dictator game. *British Journal of Psychology, 100,* 151-162. DOI:10.1348/000712608X324359

Moore, M. C., & Marler, C. A. (1987). Effects of testosterone manipulations on nonbreeding season territorial aggression in free-living male lizards, *Sceloporus jarrovi. General and Comparative Endocrinology, 65,* 225-232. DOI: 10.1016/0016-6480(87)90170-5

Morey, R. D., & Rouder, J. N. (2014). BayesFactor: Computation of Bayes factors for common designs. R package version 0.9.9. http://CRAN.R-project.org/package=BayesFactor

Phelps, V. R. (1952). Relative finger length as a sex-influenced trait. *American Journal of Human Genetics, 4,* 72-89.

Potts, R., Huston, A. C., & Wright, J. C. (1986). The effects of television form and violent content on boys' attention and social behavior. *Journal of Experimental Child Psychology, 41,* 1-17. DOI: 10.1016/0022-0965(86)90047-0

Progress and Freedom Foundation & Electronic Frontier Foundation (2009). Brief as amici curiae in support of respondents. Retrieved from https://www.eff.org/files/filenode/schwarzenegger_v/effpffamicus.pdf.

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

Rushton, B. (May, 2013) Backdooring it: Defense maneuvers around setback. *The Illinois Times.* Retrieved from http://illinoistimes.com/article-11440-backdooring-it.html

Sestir, M. A., & Bartholow, B. D. (2010). Violent and nonviolent video games produce opposing effects on aggressive and prosocial outcomes. *Journal of Experimental Social Psychology, 46*(6), 934-942. doi:10.1016/j.jesp.2010.06.005

Shanhan, J., & Morgan, M. (1999). *Television and its viewers: Cultivation theory and Research.* Cambridge, UK: Cambridge University Press.

Stanley, T. D., & Doucouliagos, H. (2012). *Meta-regression analysis in economics and business.* New York, NY: Routledge.

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2014, Dec 3). Trim-and-fill is full of it (bias) [Web log post]. Retrieved from http://datacolada.org/2014/12/03/30-trim-and-fill-is-full-of-it-bias-2/

Thompson, K. M., & Haninger, K. (2001). Violence in E-rated video games. *JAMA, 286*, 591-598. Retrieved from http://www.bvsde.paho.org/bvsacd/cd42/violence3.pdf

Voracek, M. No effects of anderogen receptor gene CAG and GGC repeat polymorphisms on digit ratio (2D:4D): Meta-analysis. Retrieved from http://arxiv.org/ftp/arxiv/papers/1310/1310.3465.pdf

Whitaker, J. L., & Bushman, B. J. (2011). "Remain calm. Be kind." Effects of relaxing video games on aggressive and prosocial behavior. *Social Psychological and Personality Science, 3,* 88-92. doi: 10.1177/1948550611409760

Wingfield, J. C., Ball, G. F., Dufty, A. M., Hegner, R. E., & Ramenofsky, M. (1987). Testosterone and aggression in birds. *American Scientist, 75,* 602-608. Retrieved from http://www.jstor.org/stable/27854889.

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition.* Stanford, CA: Stanford University Press.

vd Heiden, P. (2012) Doom builder 2 (Version 2.1.2.1553) [Computer software]. http://www.doombuilder.com.

Zillmann, D. Transfer of excitation in emotional behavior. In J. T. Cacioppo & R. E. Petty (Eds.), *Social psychophysiology: A sourcebook* (pp. 215-240). New York City: Guilford Press.

Table 1.

	Estimate	SE	*t*	*p*
(Intercept)	-0.05	0.41	-0.12	0.907
Violence	0.20	0.58	0.35	0.726
Difficulty	0.32	0.58	0.54	0.588
Violence x Difficulty	-0.84	0.82	-1.03	0.306

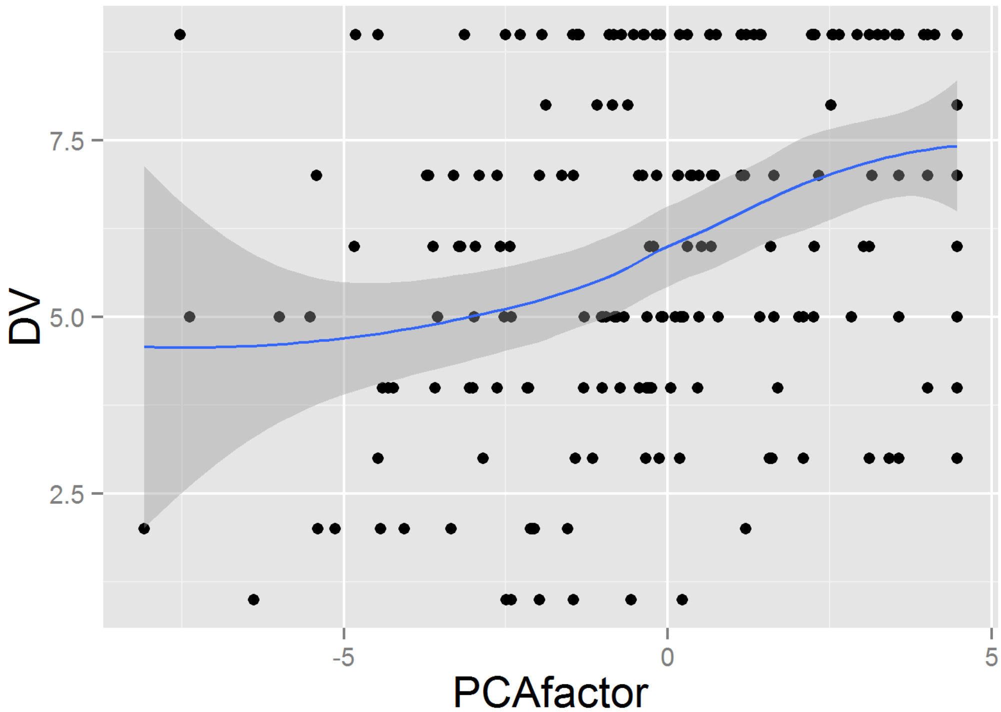
ANOVA output testing effects of game condition on composite irritation. Although it might be expected that players of a violent game might be more sensitive to irritation (e.g., a hostile expectancy bias), composite irritation is largely independent of game condition.

Table 2.

	Easy	Hard
Nonviolent	5.43 (2.54)	6.45 (2.58)
Violent	6.38 (2.23)	5.95 (2.46)

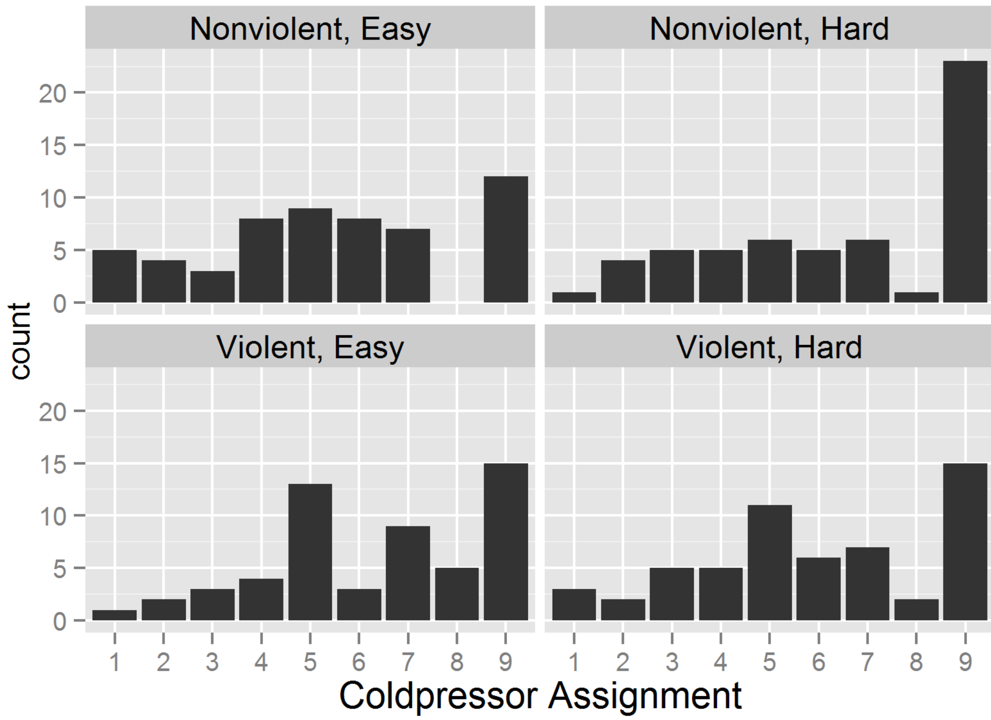
Mean coldpressor assignment per experimental condition. Coldpressor values ranged in integers from 1 (zero seconds) to 9 (80 seconds). Higher values are expected to represent greater aggression.

Figure 1. Scatterplot of coldpressor sensitivity to composite irritation.



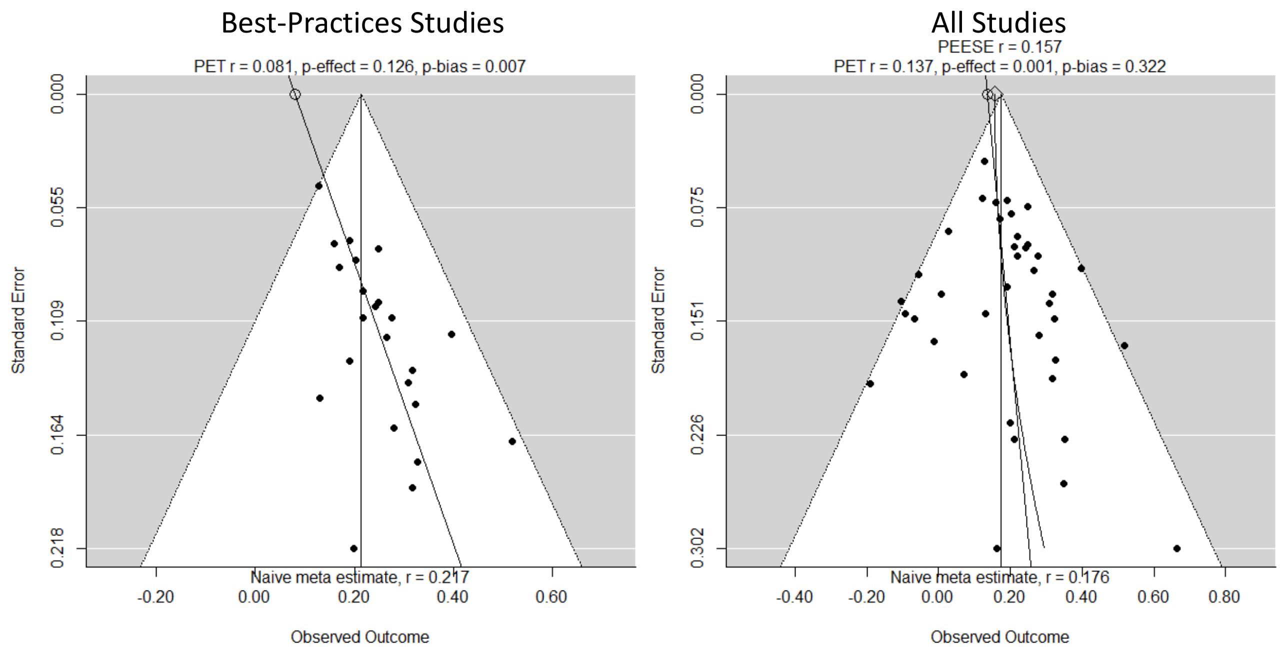
Scatterplot of participants' first principal component representing composite irritation with partner feedback. Participants more irritated with the feedback assigned greater coldpressor durations, indicating sensitivity and validity of the coldpressor measure of aggression. A locally-weighted regression curve (LOESS) with shaded standard error region is overlaid.

Figure 2. Histograms of coldpressor duration per condition.



Histograms of aggression in each cell of the 2 (Violence) x 2 (Difficulty) design. The obtained data are non-normal and suggest that analyses should include approaches for categorical and mixed-model data.

Figure 3.



PET-PEESE meta-regression of studies combined in Anderson et al. (2010) meta-analysis. Studies of effects of violent games on behavior in experimental paradigms are shown. On the left are studies selected as meeting "best-practices" criteria; on the right are all studies. While naïve meta-analysis concludes that the effect is larger among best-practices studies (*r* = .22) than among studies in general (*r* = .18), the funnel plot is more asymmetrical, suggesting that application of inclusion criteria increased selection bias. After adjusting for bias with PET-PEESE, it appears that effects in best-practices studies are very small (*r* = .08), and smaller than that in studies in general (*r* = .16). Thus, the results of the present study may not be as unusual as they initially seem.

VITA

I was born an infant nerd to a wiry, athletic nerd, Dr. James Hilgard, and a new-wave nerd, Mrs. Jennifer Hilgard. Two more nerds, Sophie and Tim, soon followed.

I came of age in a crucible of intellectual competition with my siblings. This competition extended even to my father, who had the rhetorical tactic of winning arguments by reminding me that he had completed the 21st grade. Being in 1st grade myself, I was not yet able to recognize argument from authority as a fallacy. Instead, I tried to behave, showed all my work on my math problems, and stayed indoors playing video games.

Today, I still enjoy questioning authority and exploring problems on my own. I have finally one-upped my father by loitering my way through the 22nd grade. However, my approach to learning today is the same as it was then: behave, and show your work. Behave, in that I try to stick closely to the data I have and model each competing hypothesis responsibly. Show my work, in that I post my data and **R** code to the Open Science Framework.

The past few years have been an exciting and terrifying time in psychological research. I would say without question that the most important manuscript of our decade is Bem's (2011) demonstration of ESP, without which the field may never have realized how skilled we had become at self-deception in service of significant test results. At the time, it seemed one's career depended solely on statistical significance. Today, I am co-author of a manuscript published at a prestigious journal (Engelhardt et al., in press). In this manuscript, there are no *p*-values, and the null hypothesis is favored over every alternative. I have a post-doc waiting for me and have not yet been ejected from research or discussion. Null results have now been published in prestigious journals such as *Psychological Science, Journal of Personality and Social Psychology,* and *Journal of Experimental Psychology: General.* How far we've come in just five short years! Psychology seems to be rapidly approaching an exciting new era in which research bias is diminished and researchers' careers do not depend on the good or ill fortune of the truth of the hypothesis.

I also have a life outside of Psychology, as I am a shiftless devil who refuses to work more than 40-50 hours a week. I enjoy fencing, weight lifting, and elegant European board games.

Games are an incredible thing. You drop one in front of three or four friends and watch them start losing their hair over the placement of a little wooden man or whooping and hollering over the acquisition of a tiny cardboard cathedral. You make a level for *Doom II* and get to see the look on your friends' faces when you surprise them with a nasty ambush. It is a tremendous pleasure to observe these curious models of human behavior.