Null effects of game violence, game difficulty, and 2D:4D digit ratio on aggressive behavior

Joseph Hilgard

Illinois State University

Christopher R. Engelhardt

CARFAX, Inc.

Jeffrey N. Rouder

University of California, Irvine

Ines Segert

Bruce D. Bartholow

University of Missouri

Address correspondence to Joseph Hilgard, DeGarmo Hall, 205 S. University St., Normal, IL, 61761. Email: [jhilgard@gmail.com](mailto:jhilgard@gmail.com)

Special thanks to the research assistants who helped collect, code, collate, and enter data. Taylor Green, Hyunji Suh, Conrad Neiderhauer, James Cole, Julian Segert, Landon Kirlin, Kathrine Helms, Christina Haser, Danielle Tobias, Melissa Noftsinger(?), Tyler Anderson-Sieg…

Violence is common in the media, and many are concerned about the effects such media may have on its audience. Indeed, psychological research reports that violent media may decrease prosocial behavior and increase aggressive behavior, and academic societies have made public statements on the harmful effects of violent media (American Academy of Pediatrics Council on Communications and Media, 2009; American Psychological Association Task Force on Violent Media, 2005).

Here, we study the effects of violent video games on aggressive behavior. Perhaps the strongest evidence for causal effects of violent video games come from laboratory experiments. In such experiments, researchers randomly assign some participants to play a commercially-available violent video game, say *Doom*, and others to play a commercially-available nonviolent video game, say *Myst* (Anderson & Dill, 2000). Following game play, there is some measure of aggressive thoughts, feelings, or behavior. Dozens of such studies have been performed, and meta-analyses of these experiments, accumulating results from thousands of participants, reveal greater levels of aggression following violent-video game play than following nonviolent-video game play (Anderson et al., 2010; Greitemeyer & Mügge, 2014).

Nonetheless, this evidence remains controversial for the following two reasons: First, it is often unclear whether the effects are caused by video games’ violent content in specific. An alternative explanation is that these effects may reflect confounded characteristics of violent video games such as competition or frustration rather than violent content per se (Adachi & Willoughby, 2011; Przybylski, Deci, Rigby, & Ryan, 2014). Second, and orthogonally, there is a debate about whether evidence for violent-video-game effects have been overstated, perhaps through a combination of low power, publication bias and questionable research practices (Ferguson & Kilburn, 2010; Hilgard, Engelhardt, & Rouder, 2017).

Our contribution is to address both of these possible challenges in establishing and interpreting violent-video-game effects. To address the critique that violent-video game effects may not reflect violent content, we implemented a higher level of stimulus control by customizing video games. One version of our games was based on *Doom*, and participants had to kill invading aliens. Another was a nonviolent version where participants had to save aliens that happened to be lost. Saving an alien required the participant to transport it back by aiming a remote controller at it. The gameplay in both games was exactly the same; they differed only in graphics, sounds, and cover story. To address the critique that previous studies were underpowered and may have involved post-hoc selection of outcome variables, we performed a large-scale experiment with 446 participants with preregistered sample sizes, manipulations, hypotheses, and outcome measures.

While running the experiment, we took the opportunity to test a related hypothesis about hormones, body morphology, and aggression. The male sex hormone testosterone is theorized to be one cause of aggression (see Carré, McCormick, & Hariri, 2011 for a review), and it is hypothesized that development of aggressive tendencies may be caused, in part, by prenatal testosterone exposure (see, e.g., Cohen-Bendahan, Buitelaar, van Goozen, Orlebeke, & Cohen-Kettenis, 2005). One supposed index of this prenatal exposure is the ratio of the lengths of the index and ring fingers, called the 2D:4D ratio. As such, this ratio is thought to be associated with aggressive behavior. Because the overall correlation between 2D:4D ratio and aggression is small (*r* = -.06 among males and no effect among females, Hönekopp & Watson, 2011), proponents of the 2D:4D ratio hypothesis of aggression have suggested the effects of 2D:4D may be moderated by context, only predicting aggressive behavior in aggressive situations (Millet, 2011). Hence, our large-scale experiment with well-controlled video games, a social provocation, and an opportunity to aggress provides an ideal test of the correlation between 2D:4D ratio and aggression.

**Violent Video Games**

Violent video games are hypothesized to cause increases in aggression through a number of causal pathways. These include the activation of aggressive thoughts, the operant and observational learning of aggressive scripts, increased processing of ambiguous cues as hostile, desensitization to suffering through repeated exposure to violence, increased arousal, and activation of hostile affect (Bushman & Anderson, 2002). Effect sizes have been reported as being consistent with typical effect sizes in social psychology (*r* = .21, Anderson et al., 2010; *r* = .19, Greitemeyer & Mügge, 2014) and practically meaningful based on their putative implications for public health.

**Difficult Video Games**

Researchers have attempted to test the specific effects of violent game content without confounding by other game features. Some have suggested that, despite these efforts, differences in violent content between games remain confounded by differences in competitiveness, pace of action, difficulty, or frustration. These confounds, rather than the violent content, may cause increases in aggression.

One study suggests that differences in aggression may be attributable to competitive, rather than violent, content (Adachi & Willoughby, 2011). However, the small sample size of this research yields little evidence against an effect of game violence, and another study reports differences in aggressive behavior between comparably competitive games (Anderson & Carnagey, 2009). Another series of studies reports that frustration with controls, but not game violence, may cause aggressive behavior (Przybylski et al., 2014). Finally, one exploratory analysis suggests that difficult, but not violent, gameplay may deplete cognitive control (Engelhardt, Hilgard, & Bartholow, 2015), which might be expected to lead to increased aggression (but see Hagger et al., 2016 for evidence that cognitive control does not deplete). The possibility that difficulty rather than violent content may stimulate subsequent aggression motivated us to manipulate video game violence and video game difficulty in a factorial design.

**Manipulating Game Content Without Confounds**

Most research manipulates violent content by assigning participants to play a violent or nonviolent game. 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 or fighting games, while nonviolent games are often racing, puzzle, or sports games. Therefore, while tested games do differ in their *violent content,* they are also different in their gameplay, presenting a possible confound*.*

Researchers have attempted several ways to account for these potential differences. First, one might conduct a pilot test, collecting ratings of some potential confounds and hoping not to observe any significant differences between the games. This approach is flawed in that small-sample pilot studies cannot provide substantial evidence for the null hypothesis, even if they yield nonsignificant *p*-values (Hilgard, Engelhardt, Bartholow, & Rouder, 2017). Another approach is to apply the potential confounds as covariates. This approach has two flaws. First, if the confound does cause aggression, and the confound is measured with error, residual variance will remain in the model. This residual variance will lead to an overestimated effect of violence alone. Second, covariates may not represent confounds, but rather meaningful consequences of violent content that mediate the relationship between violent content and aggressive outcomes. Applying these mediators as covariates would reduce the relationship between violent content and aggressive outcome and underestimate the effect.

Because pilot tests and ANCOVA are not effective ways of balancing game stimuli, we take a more direct approach by modifying the content of a single video game. Rather than comparing two separate games, 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 held constant (as suggested by Elson & Quandt, 2016 and demonstrated in Carnagey & Anderson, 2005; Elson et al., 2015; Przybylski et al., 2014). This approach allows manipulation of specific game features in much the same way that a researcher would manipulate features of a laboratory paradigm between conditions, permitting clearer inferences concerning the effects of the manipulated game feature.

**2D:4D Ratio**

Although violent-media research is concerned about the social causes of aggression, aggression also has biological causes. Because there are sex differences in aggression (see Campbell, 2006), it has been suggested that aggression is affected by the sex hormone testosterone. Some support for this testosterone effect 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 culture establishes sex differences in 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, with men having lower 2D:4D ratios than women (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).

Insofar as 2D:4D is a valid index of prenatal testosterone, and prenatal testosterone affects later aggressive behavior, we would expect a correlation between 2D:4D ratio and aggression. The research literature is conflicted in this regard. Most studies do not find main effects of 2D:4D ratio, but instead simple slopes in subgroups. For example, 2D:4D ratio was reported to interact with the effect of an aggressive music video on aggressive intent: Participants with more masculine ratios displayed 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 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 that these moderation models are overfitting the data, especially if they are attempted post-hoc when the anticipated main effects are not found.

Recent meta-analytic efforts call into question the validity of 2D:4D ratio as a measurement of prenatal testosterone action. A small initial study reported that 2D:4D ratio was associated with a gene variant that influences responsivity to androgens; greater responsivity implying greater effects of testosterone, in turn causing lower 2D:4D ratio (Manning, Bundred, Newton, & Flanagan, 2003). Subsequent research has failed to replicate this relationship, and a meta-analysis estimates the effect as *r* = .02 [-.02, .06] (Voracek, 2014). Thus, it is possible that 2D:4D is not a valid measurement of prenatal testosterone activity in typical populations. If this is the case, then 2D:4D ratio should not predict aggression even if prenatal testosterone levels are associated with subsequent aggressive tendencies.

**Superadditive Causes of Aggressive Behavior**

A number of models of aggressive behavior suggest that, as causes of aggression coincide, their effects might yield greater levels of aggression than their simple sum might suggest. For example, I3 Theory (Slotter & Finkel, 2011) categorizes causes of aggression as being instigating, impelling, or (dis)inhibiting. Similarly, the General Aggression Model suggests interactions between the person and the situation, such that a violent prime might be most influential on those already temperamentally disposed towards aggression. In both models, aggression-stimulating factors are thought to have superadditive effects in combination. In this study, we examine whether three purported causes of aggression (violent content, difficult content, and 2D:4D ratio) interact to predict aggressive behavior following video game play.

**Purpose**

The proposed study examines the effects of game violence, game difficulty, and 2D:4D ratio on aggressive behavior among college-aged males. These can be summarized as four hypotheses. H1: Violent video game content will increase aggressive behavior. H2: Video game difficulty will increase aggressive behavior. H3: More masculine 2D:4D ratios will be associated with more aggressive behavior. H4: These effects will yield superadditive interactions, such that the presence of multiple of these factors will produce more still more aggression.

**Method**

**Participants**

Participants were 446 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. The semester ended before the last four experimental sessions could be conducted. 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 removes gender as a potential source of variance. Participants were primarily Caucasian (79.7%), with some African-American (8.6%), Asian (4.6%), and Latino (3.3%), and 3.8% identified as another race. On average, participants were 19.0 (SD: 1.7) years old.

**Disclosures**

Hypotheses and sample size were preregistered at https://osf.io/cwenz/. All measures, materials, data, and analytic code are also available at that URL.

**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 (The GIMP Team, n.d.), 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. Five coders provided measurements in this fashion, with each scan coded by at least two coders. Inter-rater reliability was assessed using a one-way, mixed, consistency, average-measures intra-class correlation (Hallgren, 2012; McGraw & Wong, 1996) using the psych package for R (Revelle, 2017). The resulting ICCs were excellent (ICC3k = .94 for left 2D:4D, .88 for right 2D:4D), indicating high agreement across coders and minimal loss of power due to measurement error.

**Cold pressor task.** Participants had an opportunity to aggress against their partner by assigning the partner to immerse his fist in a bucket of painfully-cold water for an amount of time (Pedersen, Bushman, Vasquez, & Miller, 2008). 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 has the benefit of being quantified only in one way (e.g. 1-9 rating), eliminating the concerns about flexible quantification methods associated with the competitive reaction time measure of aggression (Elson, Mohseni, Breuer, Scharkow, & Quandt, 2014).

**Manipulation checks.** Participants completed a questionnaire assessing the efficacy of the various parts of the experimental manipulation. First, participants rated their partner’s feedback as pleasant or irritating (6 items). Then, participants rated the video game they played on a number of dimensions, including how violent, exciting, and challenging it was (18 items). All items were rated on a 1 (strongly disagree) to 7 (strongly agree) Likert scale. 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.

**Probe for suspicion.** Research assistants attempted an oral funneled debriefing. Following this oral debriefing, participants completed a questionnaire intended to imitate a funneled debriefing. This debriefing questionnaire started with broad questions about the study and then grew increasingly specific, asking whether anything seemed strange about the study, the aggression measure, or the other participant in the study.

**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, Difficult) x 2 (Violence: Nonviolent, Violent) design.

Across the four video games, all gameplay variables are held constant. Players had a rapid-fire tool and a slow-but-powerful tool (in the violent condition, these were a chaingun and a shotgun). 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. Levels were designed to be easy to navigate, reducing time spent exploring the map and maximizing the player’s time spent interacting with game characters. 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 was manipulated by changing the graphics, sounds, and story of the game while leaving the controls and enemy behavior constant. In the violent version, enemy graphics and sounds were borrowed from *Brutal Doom* (Abenante, 2012), a modified form of *Doom II* that makes the game more violent. In this game, defeated enemies exploded into fountains of gore, severed limbs, and scattering teeth. Participants in this condition were told that they must kill all the demons from hell. Players maintained their health and ammunition by picking up medkits, bullets, and shotgun shells. 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 were similarly replaced with “zorchers,” science-fiction tools that resemble remote controllers. Participants in this condition were told that the aliens are lost and confused and need to be sent home with the zorcher. Players maintained their health and ammunition by picking up fruits, vegetables, “zorch pellets,” and “zap tapes.”

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, using guns, claws, or fireballs in the violent game and throwing boogers in the nonviolent game. 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, enemies had their artificial intelligence changed so that they would 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-but-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 provoked by their partner in a procedure adapted from Bushman and Baumeister (1998). Participants were then given an envelope, a sheet of loose-leaf paper, and a printed essay prompt. They were informed that the first task was 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 do 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.

During the exchange, each participant received a fake, premade essay designed to oppose their beliefs. Participants who wrote a pro-life essay received a pro-choice essay, whereas 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. In the nonviolent condition, the story explained that the booger aliens are lost and confused, and when the player has “zorched” them all home, he sees a scene of the aliens playing together on their homeworld. By comparison, in the violent condition, the story explained that the aliens must all be slain, and 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.”[[1]](#footnote-1) To prepare the cold pressor task, a dozen ice cubes were added to the cold pressor pitcher 5 minutes before the end of the game session.

When the game session ended, the research assistant brought the cold pressor pitcher and a towel into the room and pressed a key to print the game variables, which the assistant then logged. The assistant then quit the game by pressing Alt+F4. The research assistant 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 cold pressor by placing his fist in it for five seconds. The participant was then asked if he would be okay with the cold pressor. No participants indicated unwillingness to participate in the cold pressor 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. This provoked the participant. 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 assign their partner to an amount of cold pressor exposure.

Finally, 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**

**Quality Control**

Of the 446 participants, 128 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 27 subjects were excluded because the research assistants indicated some failure of deception or of methodology. (For results when these subjects are included, see the supplement.) Three additional subjects’ data were excluded due to experimenter error (the wrong game file had been used). We note that our failure of deception rate of 28.7% is considerably higher than our anticipated 11% rate or of rates reported in previous work. After these exclusions, the effective sample size was 275.

**Manipulation Check**

**Game manipulation.** Participant ratings on the post-questionnaires were submitted to 2 (Violence) × 2 (Difficulty) ANOVA. The manipulation was highly effective: participants indicated that the violent game (*M* = 5.3 (1.6)) was much more violent than the nonviolent game (*M* = 2.2 (1.3); *d* = 2.1 [1.8, 2.4]).

**Provocation.** Mean evaluations of the participants’ interactions with the partner were also assessed. Participants generally indicated that they were irritated (M = 5, SD = 1.7), angered (M = 4.2, SD = 1.8), and annoyed (M = 4.9, SD = 1.8) by their partner. Furthermore, they were neither happy (M = 2.4, SD = 1.4) nor pleased (M = 2.2, SD = 1.4) with their partner and found the feedback unhelpful (M = 1.7, SD = 1.3).

To determine whether the cold pressor dependent variable was a sensitive measure of aggression, we tested whether these participants more provoked by the feedback gave higher cold pressor assignments. Parallel analysis suggested a two-factor solution for participants' ratings of their interaction with their partner. Factors were extracted using oblimin rotation. The first factor accounted for 52% of the variance and had the expected pattern of loadings: .77, .76, and .67 for irritation, anger, and annoyance, -.25, .02, and .02 for happiness, helpfulness, and pleasure. This provocation factor was then used as a linear predictor of cold pressor assignment. More irritated participants gave greater cold pressor assignments, *t*(249) = 5.73, *p* < .001, *r* = .33 [.22, .43], suggesting that the measure may indeed be influenced by participants’ intent to aggress. A scatterplot and loess regression line are provided in Figure 1.

A 2 (Violence) × 2 (Difficulty) ANOVA was conducted to determine whether the game played influenced participants’ ratings of the interaction. Effects were small and not statistically significant (violence, *t*(247) = -0.28, *p* = .777, *d* = -0.03 [-0.27, 0.2]; difficulty, *t*(247) = -0.17, *p* = .867, *d* = -0.02 [-0.26, 0.22]; Violence × Difficulty, *t*(247) = -0.86, *p* = .392, *d* = -0.1 [-0.34, 0.13]), suggesting that the game played had a minimal influence on participants’ provocation.

**Conventional General Linear Models**

General linear models were used to look for main effects and interactions of game difficulty, game violence, and 2D:4D ratio. These tests were preregistered. Two models were used to look for effects of left and right 2D:4D ratio separately. Factors were contrast-coded and 2D:4D ratios were standardized to preserve orthogonality of parameter estimates. Cell means and SDs are provided in Table 1.

Neither model found any significant effects. Neither left-hand 2D:4D (*t*(265) = -1.11, *p* = .266, *r* = -.07 [-.18, .05]) nor right-hand 2D:4D (*t*(266) = 0.52, *p* = .602, *r* = .03 [-.09, .15]) had a significant main effect on aggressive behavior. Additionally, effects of game violence (*t*(265) = -0.83, *p* = .407, *d* = -0.1 [-0.34, 0.14]), game difficulty (*t*(265) = 0.46, *p* = .645, *d* = 0.06 [-0.18, 0.29]), and their interaction (*t*(265) = -0.17, *p* = .867, *d* = -0.02 [-0.26, 0.22]) were small and not statistically significant. No higher-order interactions involving 2D:4D ratio of either hand were statistically significant. Full model output is summarized in Tables 2 and 3.

The earlier manipulation and sensitivity check indicated that much of the variance in aggression could be predicted by experienced provocation. Because this provocation was generally independent of the experimental condition, its inclusion as a covariate in analysis might increase statistical power. However, adding provocation as a covariate did not reveal significant effects. The effect of violence was *t*(246) = 0.78, *p* = .434, *d* = 0.09 [-0.14, 0.33], the effect of difficulty was *t*(246) = 1.08, *p* = .283, *d* = 0.13 [-0.11, 0.37], and their interaction was *t*(246) = -1, *p* = .318, *d* = -0.12 [-0.36, 0.12]. Effects of left-hand and right-hand 2D:4D remained nonsignificant (*t*(246) = -1.86, *p* = .065, *r* = -.12 [-.24, .01] and *t*(248) = -0.31, *p* = .755, *r* = -.02 [-.14, .11], respectively).

**Bayesian ANOVA**

Models were compared using the BayesFactor package for R (Morey & Rouder, 2014). The scale of the effect size under the alternative hypothesis was specified as *d* ~ 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 show the support in odds ratio for a hypothesis; BF­10 indicates support for the alternative hypothesis, while BF01 indicates support for the null hypothesis. A Bayes factor of 1 indicates that the data are equally likely under either hypothesis.

Of all the models, the null-hypothesis model was best supported by the data. Models of main effects of Violence, Difficulty, left-hand 2D:4D, or right-hand 2D:4D were each outperformed by the null model (Bayes factors = 3.61, 3.81, 4.4, and 6.53 in favor of the null, respectively). Higher-order interactions were not supported by the data, either. Evidence was ambiguous regarding a Violence × Difficulty interaction (BF = 1.41 favoring the null). Neither violence nor difficulty interacted with 2D:4D of the left hand (BF = 3.81, 4.44, respectively) or 2D:4D of the right hand (BF = 4.82, 4.23). The 3-way interaction was not supported (left-hand BF = 3.34, right-hand BF = 2.93).

Experienced provocation was added to the model as a predictor. An effect of provocation was strongly supported by the evidence (BF = 1.04×106). However, addition of this covariate did not improve the strength of evidence for main effects of violence (BF = 4.98), difficulty (BF = 3.65), or 2D:4D (BF = 1.26, left hand; BF = 6.13, right hand). Taken together, these results indicate that aggression could be predicted by experienced provocation but not by game condition.

**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, δ = .43 [.35, .52] (Anderson et al., 2010).

The main effect of Violence was *d* = 0.11 [-0.13, 0.35]. An online Bayes factor calculator (Dienes, 2008) was used to compare the evidence for H0: *δ* = 0 relative to H1: *δ* = .43 [.35, .52]. The obtained Bayes factor substantially preferred the null, B01 = 14.2.

Proponents have suggested that the Anderson et al. (2010) estimate may be an overestimate due to publication bias, but that after adjustment for publication bias the effect is still approximately *d* = .30 (Kepes, Bushman, & Anderson, 2017). The Bayes factor calculator was used to compare the evidence for H0: *δ* = 0 relative to H2: *δ* = .30 [.20, .40]. The obtained Bayes factor still preferred the null, but less so relative to this more modest estimate, B02 = 2.0.

**Supplementary methods**

Cold pressor assignments were found to be non-normally distributed. To address this non-normality, the data were tested in two additional models to attempt to deal with the spike at 9. Censored regression was used to attempt to model responses greater than 9, and logistic regression was used to model the probability of a 9 response vs. all other responses. These methods did not yield substantively different conclusions (i.e., no parameters were significant). See the supplement for details.

**Exploratory analyses**

A number of exploratory analyses were conducted. These examined whether aggression was predicted by participants' experience of difficulty during the game, participants' self-reported history of video games, and participants' in-game behaviors. Neither experienced difficulty nor history of game use predicted aggression. Participants who defeated more monsters and fired more bullets were slightly less aggressive (monsters defeated, *t*(272) = -2.51, *p* = .013, *r* = -.15 [-.26, -.03]; bullets fired, *t*(272) = -2.51, *p* = .013, *r* = -.15 [-.26, -.03]), but this finding should be regarded with caution given this test's exploratory nature and modest *p*-value.

**Discussion**

Results indicate that when game stimuli are carefully controlled, the effects of fifteen minutes of violent and/or difficult gameplay on aggressive behavior may be small and indistinguishable from zero. This suggests that the effects of brief violent video game play on laboratory measures of aggressive behavior may be smaller and less robust than the published research literature would indicate. Researchers may need to reevaluate whether violent game manipulations are useful for revealing the causes and mechanisms of aggression.

2D:4D digit ratio also failed to predict aggressive behavior among participants. The current results cast doubt on 2D:4D as an index of prenatal testosterone and a predictor of aggressive behavior (see also Hönekopp & Watson, 2011; Voracek, 2014). The sample size of the current research is considerably larger than many other studies reporting significant associations between 2D:4D ratio and aggression (e.g., Millet & Dewitte, 2007, 2009).

The presented manipulation and sensitivity checks suggest 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. These indicate that both the game manipulation and the essay provocation were effective. Third, the cold pressor measure of aggression was moderately sensitive (*r* = .33) 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 cold pressor measure (see also Pedersen et al., 2011, 2008).

**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 were not affected by brief violent game play (Engelhardt, Mazurek, Hilgard, Rouder, & Bartholow, 2015).

The present research provides a closer experimental control than do many 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, null results in small-sample pilot studies provide little evidence against confounds (Hilgard, Engelhardt, Bartholow, et al., 2017). 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.

These results are consistent with evidence from meta-analysis that suggests that violent video game effects have been overestimated through publication bias (Hilgard, Engelhardt, & Rouder, 2017). Meta-analysis has previously reported an effect of *d* = .43 [.25, .52] (Anderson et al., 2010); our data provide strong evidence against an effect of this size (BF­01 = 13.5). Proponents of violent-game effects have agreed that there may be publication bias, but that the publication bias may be modest, leaving a true effect of about *d* = 0.3 (Kepes et al., 2017). The current evidence is less opposed to such an adjusted estimate, but results still slightly prefer the null hypothesis by a factor of 2-to-1 over this more modest claim.

This finding has implications for future laboratory research of violent media and aggressive behavior. With regard to the study of violent media, brief violent media manipulations may have effects too small to reliably detect. If so, then laboratory paradigms may not be appropriate for validating measures of aggressive thoughts and developing theories of violent media effects. Researchers may need to develop stronger methods if they are to understand the long-term effects of violent media through short-term lab studies.

**Effects of Difficult Video Games**

The present results contradict our previous findings about possible effects of difficult gameplay on self-control (Engelhardt, Hilgard, et al., 2015). In that research, we reported that difficult gameplay exhausted self-control resources (“ego depletion”), such that players who were challenged by the game did more poorly on a modified Stroop task. If true, one might also expect such deficits in self-control might cause increases in aggression. Recent research challenges this “ego depletion” account of self-control resources (Hagger et al., 2016). Similarly, we did not find that difficult gameplay increased aggression.

The obtained results also appear inconsistent with the results of research indicating effects of competitive (Adachi & Willoughby, 2011) or competence-thwarting (Przybylski et al., 2014) video games. We note that sample sizes in research regarding the effects of competition are small, and effects may have been misestimated. Research regarding the effects of competence-thwarting games, on the other hand, was appreciably powered. Manipulation checks indicated that subjects found the difficult game to be more challenging and stressful. However, it is possible that the present study’s difficulty manipulation was difficult without being specifically competence-thwarting. Future research might seek to distinguish between the constructs of challenge and competence-thwarting and determine the conditions under which each leads to aggression.

**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 so 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. The most recent theory suggests that 2D:4D only predicts 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, Sciulli, & 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 as a predictor of aggressive behavior.

**Limitations**

First,the distribution of cold pressor assignments was found to not resemble a normal distribution. We attempted several models to address this non-normality. Results were comparable across modeling approaches, none of which indicated significant effects. 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. However, the measure’s sensitivity to participants’ provocation suggests otherwise. Additionally, the correlation we observe between experienced irritation and cold pressor (*r* = .33) is comparable to the effect of provocation on Taylor Aggression Paradigm scores (*d* = 0.52, or *r* = .25) observed by Chester (2017).

Second, it is possible that the nonviolent *Chex Quest* game involves sufficient violence to cause an increase in aggression, eliminating the difference between conditions. One study has claimed that the effect of cartoon E-rated violence is as strong as that of explicit M-rated violence (Anderson, Gentile, & Buckley, 2007). This seems unusual; compared to mild violent content, exposure to more extreme violent content should be more desensitizing, activate more aggressive thoughts, and stimulate more aggressive feelings. Still, it is possible that an effect was not found in the present study because the violence in *Chex Quest* has effects on aggression equal to that of *Brutal Doom*. Future research may test the dose-response curve of violent content and aggressive behavior.

Finally, a lot of data were discarded for failures of deception. Many participants indicated awareness of the research hypothesis. This may have been due, in part, to the redundant process of oral funneled debriefing and questionnaire funneled debriefing, which may have increased awareness of the hypothesis following collection of the primary outcome. This makes it difficult to know, on the basis of the questionnaire debriefing, when participants became aware of the hypothesis. Hypothesis-awareness mid-experiment would threaten the data’s validity, whereas hypothesis-awareness only following the oral debriefing would not be a problem. Still, we attempted to address this uncertainty by being conservative in our quality checks so as not to overstate the evidence for the null hypothesis. However, we recognize that there are inferential challenges associated with such a high exclusion rate. One might be concerned that still more participants were hypothesis-aware; this might reduce the observed effect size through reduction of internal validity or through reactance (Bender, Rothmund, & Gollwitzer, 2013). Researchers may find value in establishing standardized practices in deception and debriefing.

**Summary**

We found evidence that brief exposure to violent games does not cause aggressive behavior. This evidence is corroborated by similar research with different measurements of aggressive outcomes (Engelhardt, Mazurek, et al., 2015). Effect sizes reported in previous experiments on this topic may be either inflated by confounds (Adachi & Willoughby, 2011; Hilgard, Engelhardt, Bartholow, et al., 2017) or by publication and selection bias (Hilgard, Engelhardt, & Rouder, 2017). It is uncertain whether laboratory paradigms involving brief exposure to violent video games can reveal the causes of aggression.

2D:4D similarly does little to predict aggression in a laboratory experiment. Considered alongside other evidence of the invalidity of 2D:4D (Hönekopp & Watson, 2011; Voracek, 2014), one might question the validity of 2D:4D as an index of prenatal testosterone or whether prenatal testosterone predicts aggression.

Violent-game manipulations, on their own, may reveal little about the causes of aggression. We recommend that laboratory studies of aggression return to basic methodology. Progress may be made through validation of methods and measures, stimulation of aggression through more powerful and predictable manipulations, and establishment of best practices in maintaining deception.

References

Abenante, M. (2012). Brutal Doom. Retrieved from http://www.moddb.com/mods/brutal-doom

Adachi, P. J. C., & Willoughby, T. (2011). The effect of video game competition and violence on aggressive behavior: Which characteristic has the greatest influence? *Psychology of Violence*, *1*(4), 259–274. https://doi.org/10.1037/a0024908

American Academy of Pediatrics Council on Communications and Media. (2009). Media Violence. *Pediatrics*, *124*(5), 1495–1503. https://doi.org/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

Anderson, C. A., & Carnagey, N. L. (2009). Causal effects of violent sports video games on aggression: Is it competitiveness or violent content? *Journal of Experimental Social Psychology*, *45*(4), 731–739. https://doi.org/10.1016/J.JESP.2009.04.019

Anderson, C. A., Carnagey, N. L., Flanagan, M., Benjamin, A. J., Eubanks, J., & Valentine, J. C. (2004). Violent Video Games: Specific Effects of Violent Content on Aggressive Thoughts and Behavior. *Advances in Experimental Social Psychology*, *36*, 199–249. https://doi.org/10.1016/S0065-2601(04)36004-1

Anderson, C. A., & Dill, K. E. (2000). Video Games and Aggressive Thoughts, Feelings, and Behavior in the Laboratory and in Life. *Journal of Personality and Social Psychology*, *78*(4), 772–790. https://doi.org/10.1037//O022-3514.78.4.772

Anderson, C. A., Gentile, D. A., & Buckley, K. E. (2007). *Violent video game effects on children and adolescents : theory, research, and public policy*. Oxford University Press.

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

Archer, J. (2009). Does sexual selection explain human sex differences in aggression? *Behavioral and Brain Sciences*, *32*(3–4), 249. https://doi.org/10.1017/S0140525X09990951

Bender, J., Rothmund, T., & Gollwitzer, M. (2013). Biased Estimation of Violent Video Game Effects on Aggression: Contributing Factors and Boundary Conditions. *Societies*, *3*(4), 383–398. https://doi.org/10.3390/soc3040383

Benderlioglu, Z., Sciulli, P. W., & Nelson, R. J. (2004). Fluctuating asymmetry predicts human reactive aggression. *American Journal of Human Biology*, *16*(4), 458–469. https://doi.org/10.1002/ajhb.20047

Bushman, B. J., & Anderson, C. A. (2002). Violent Video Games and Hostile Expectations: A Test of the General Aggression Model. *Personality and Social Psychology Bulletin*, *28*(12), 1679–1686. https://doi.org/10.1177/014616702237649

Bushman, B. J., & Baumeister, R. F. (1998). Threatened egotism, narcissism, self-esteem, and direct and displaced aggression: Does self-love or self-hate lead to violence? *Journal of Personality and Social Psychology*, *75*(1), 219–229. https://doi.org/10.1037/0022-3514.75.1.219

Campbell, A. (2006). Sex differences in direct aggression: What are the psychological mediators? *Aggression and Violent Behavior*, *11*(3), 237–264. https://doi.org/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. https://doi.org/10.1111/j.1467-9280.2005.01632.x

Carré, J. M., McCormick, C. M., & Hariri, A. R. (2011). The social neuroendocrinology of human aggression. *Psychoneuroendocrinology*, *36*(7), 935–944. https://doi.org/10.1016/j.psyneuen.2011.02.001

Chester, D. (2017). A Preregistered Validation Study of the Taylor Aggression Paradigm. https://doi.org/10.17605/OSF.IO/8JGCA

Cohen-Bendahan, C. C. C., Buitelaar, J. K., van Goozen, S. H. M., Orlebeke, J. F., & Cohen-Kettenis, P. T. (2005). Is there an effect of prenatal testosterone on aggression and other behavioral traits? A study comparing same-sex and opposite-sex twin girls. *Hormones and Behavior*, *47*(2), 230–237. https://doi.org/10.1016/J.YHBEH.2004.10.006

Dienes, Z. (2008). *Understanding psychology as a science : an introduction to scientific and statistical inference*. Palgrave Macmillan.

Digital Café. (1996). Chex Quest. Retrieved from http://www.chexquest.org/index.php?action=downloads;cat=1

Elson, M., Breuer, J., Van Looy, J., Kneer, J., & Quandt, T. (2015). Comparing apples and oranges? Evidence for pace of action as a confound in research on digital games and aggression. *Psychology of Popular Media Culture*, *4*(2), 112–125. https://doi.org/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*(2), 419–432. https://doi.org/10.1037/a0035569

Elson, M., & Quandt, T. (2016). Digital Games in Laboratory Experiments: Controlling a Complex Stimulus Through Modding. *Psychology of Popular Media Culture*, *5*(1), 52–65. https://doi.org/10.1037/ppm0000033

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

Engelhardt, C. R., Mazurek, M. O., Hilgard, J., Rouder, J. N., & Bartholow, B. D. (2015). Effects of Violent-Video-Game Exposure on Aggressive Behavior, Aggressive-Thought Accessibility, and Aggressive Affect Among Adults With and Without Autism Spectrum Disorder. *Psychological Science*, *26*(8), 1187–1200. https://doi.org/10.1177/0956797615583038

Ferguson, C. J., & Kilburn, J. (2010). Much ado about nothing: The misestimation and overinterpretation of violent video game effects in Eastern and Western nations: Comment on Anderson et al. (2010). *Psychological Bulletin*, *136*(2), 174–178. https://doi.org/10.1037/a0018566

Greitemeyer, T., & Mügge, D. O. (2014). Video Games Do Affect Social Outcomes. *Personality and Social Psychology Bulletin*, *40*(5), 578–589. https://doi.org/10.1177/0146167213520459

Hagger, M. S., Chatzisarantis, N. L. D., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., … Zwienenberg, M. (2016). A Multilab Preregistered Replication of the Ego-Depletion Effect. *Perspectives on Psychological Science*, *11*(4), 546–573. https://doi.org/10.1177/1745691616652873

Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22833776

Hilgard, J., Engelhardt, C. R., Bartholow, B. D., & Rouder, J. N. (2017). How much evidence is p &gt; .05? Stimulus pre-testing and null primary outcomes in violent video games research. *Psychology of Popular Media Culture*, *6*(4), 361–380. https://doi.org/10.1037/ppm0000102

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

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

iD Software. (1994). Doom II. Rockville, MD: ZeniMax Media.

Judd, S. (2011). SLADE 3. Retrieved from http://slade.mancubus.net/index.php?page=downloads/

Kepes, S., Bushman, B. J., & Anderson, C. A. (2017). Violent video game effects remain a societal concern: Reply to Hilgard, Engelhardt, and Rouder (2017). *Psychological Bulletin*, *143*(7), 775–782. https://doi.org/10.1037/bul0000112

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*(1–2), 23–28. https://doi.org/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*(6), 399–405. https://doi.org/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*(11), 3000–3004. https://doi.org/10.1093/humrep/13.11.3000

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30–46. https://doi.org/10.1037/1082-989X.1.1.30

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*(4), 755–764. https://doi.org/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 Differences*, *51*(4), 397–401. Retrieved from http://www.sciencedirect.com/science/article/pii/S0191886910001996

Millet, K., & Dewitte, S. (2007). Digit ratio (2D:4D) moderates the impact of an aggressive music video on aggression. *Personality and Individual Differences*, *43*(2), 289–294. https://doi.org/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 behaviour in a dictator game. *British Journal of Psychology*, *100*(1), 151–162. https://doi.org/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*(2), 225–232. https://doi.org/10.1016/0016-6480(87)90170-5

Pedersen, W. C., Bushman, B. J., Vasquez, E. A., & Miller, N. (2008). Kicking the (Barking) Dog Effect: The Moderating Role of Target Attributes on Triggered Displaced Aggression. *Personality and Social Psychology Bulletin*, *34*(10), 1382–1395. https://doi.org/10.1177/0146167208321268

Pedersen, W. C., Denson, T. F., Goss, R. J., Vasquez, E. A., Kelley, N. J., & Miller, N. (2011). The impact of rumination on aggressive thoughts, feelings, arousal, and behaviour. *British Journal of Social Psychology*, *50*(2), 281–301. https://doi.org/10.1348/014466610X515696

Phelps, V. R. (1952). Relative index finger length as a sex-influenced trait in man. *American Journal of Human Genetics*, *4*(2), 72–89. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14943709

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

Revelle, W. (2017). psych: Procedures for Personality and Psychological Research. Evanston, Illinois: Northwestern University.

Slotter, E. B., & Finkel, E. J. (2011). I3 Theory: Instigating, impelling, and inhibiting factors in aggression. In P. R. Shaver & M. Mikulincer (Eds.), *Human aggression and violence: Causes, manifestations, and consequences*. American Psychological Association.

The GIMP Team. (n.d.). GNU Image Manipulation Program. Retrieved from www.gimp.org

vd Heiden, P. (2012). Doom Builder 2. Retrieved from http://www.doombuilder.com

Voracek, M. (2014). No effects of androgen receptor gene CAG and GGC repeat polymorphisms on digit ratio (2D:4D): a comprehensive meta-analysis and critical evaluation of research. *Evolution and Human Behavior*, *35*(5), 430–437. https://doi.org/10.1016/J.EVOLHUMBEHAV.2014.05.009

Wingfield, J., Ball, G., Dufty, A., Hegner, R., & Ramenofsky, M. (1987). Testosterone and aggression in birds. *American Scientist*, *75*(6), 602–608. Retrieved from http://www.jstor.org/stable/27854889

.

Table 1.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| term | estimate | std.error | statistic | p.value |
| (Intercept) | 0.001 | 0.058 | 0.012 | 0.990 |
| Violence1 | -0.016 | 0.058 | -0.284 | 0.777 |
| Difficulty1 | -0.010 | 0.058 | -0.167 | 0.867 |
| Violence1:Difficulty1 | -0.049 | 0.058 | -0.857 | 0.392 |

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

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| term | estimate | std.error | Statistic | p.value |
| (Intercept) | -137.170 | 128.550 | -1.067 | 0.287 |
| Difficulty1 | 59.240 | 128.550 | 0.461 | 0.645 |
| Violence1 | -106.792 | 128.550 | -0.831 | 0.407 |
| L2d4d\_std | -5.087 | 4.565 | -1.114 | 0.266 |
| Difficulty1:Violence1 | -21.497 | 128.550 | -0.167 | 0.867 |
| Difficulty1:L2d4d\_std | 2.099 | 4.565 | 0.460 | 0.646 |
| Violence1:L2d4d\_std | -3.797 | 4.565 | -0.832 | 0.406 |
| Difficulty1:Violence1:L2d4d\_std | -0.755 | 4.565 | -0.165 | 0.869 |

ANOVA model testing effects of difficulty, violence, and left-hand 2D:4D on cold pressor assignments to partner.

Table 3

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Term | estimate | std.error | statistic | p.value |
| (Intercept) | 70.173 | 122.756 | 0.572 | 0.568 |
| Difficulty1 | -75.798 | 122.756 | -0.617 | 0.537 |
| Violence1 | -32.107 | 122.756 | -0.262 | 0.794 |
| R2d4d\_std | 2.334 | 4.468 | 0.522 | 0.602 |
| Difficulty1:Violence1 | -63.336 | 122.756 | -0.516 | 0.606 |
| Difficulty1:R2d4d\_std | -2.764 | 4.468 | -0.619 | 0.537 |
| Violence1:R2d4d\_std | -1.174 | 4.468 | -0.263 | 0.793 |
| Difficulty1:Violence1:R2d4d\_std | -2.297 | 4.468 | -0.514 | 0.608 |

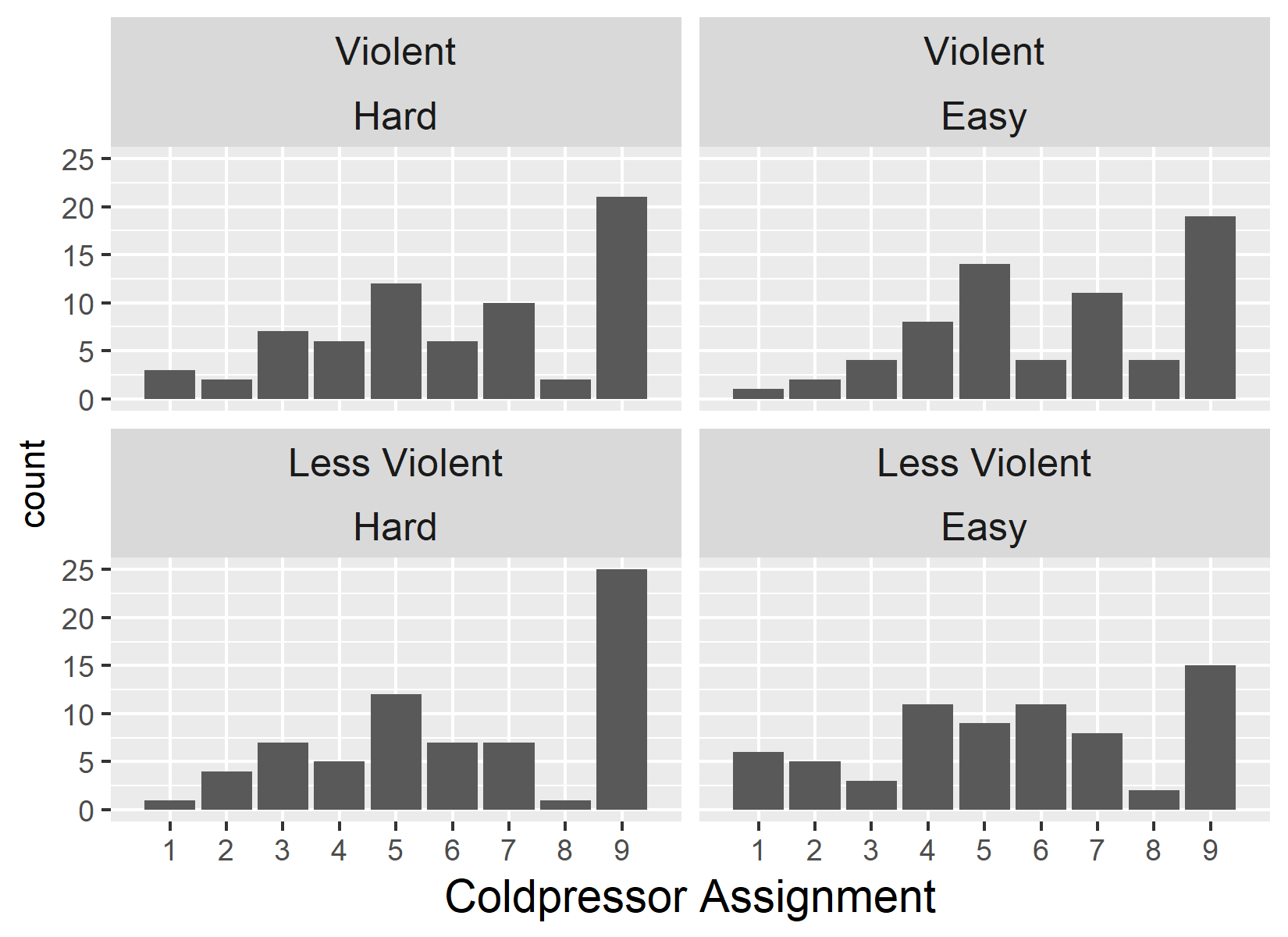
ANOVA model testing effects of difficulty, violence, and right-hand 2D:4D on cold pressor assignments to partner.

Figure 1. Scatterplot of cold pressor sensitivity to composite irritation.



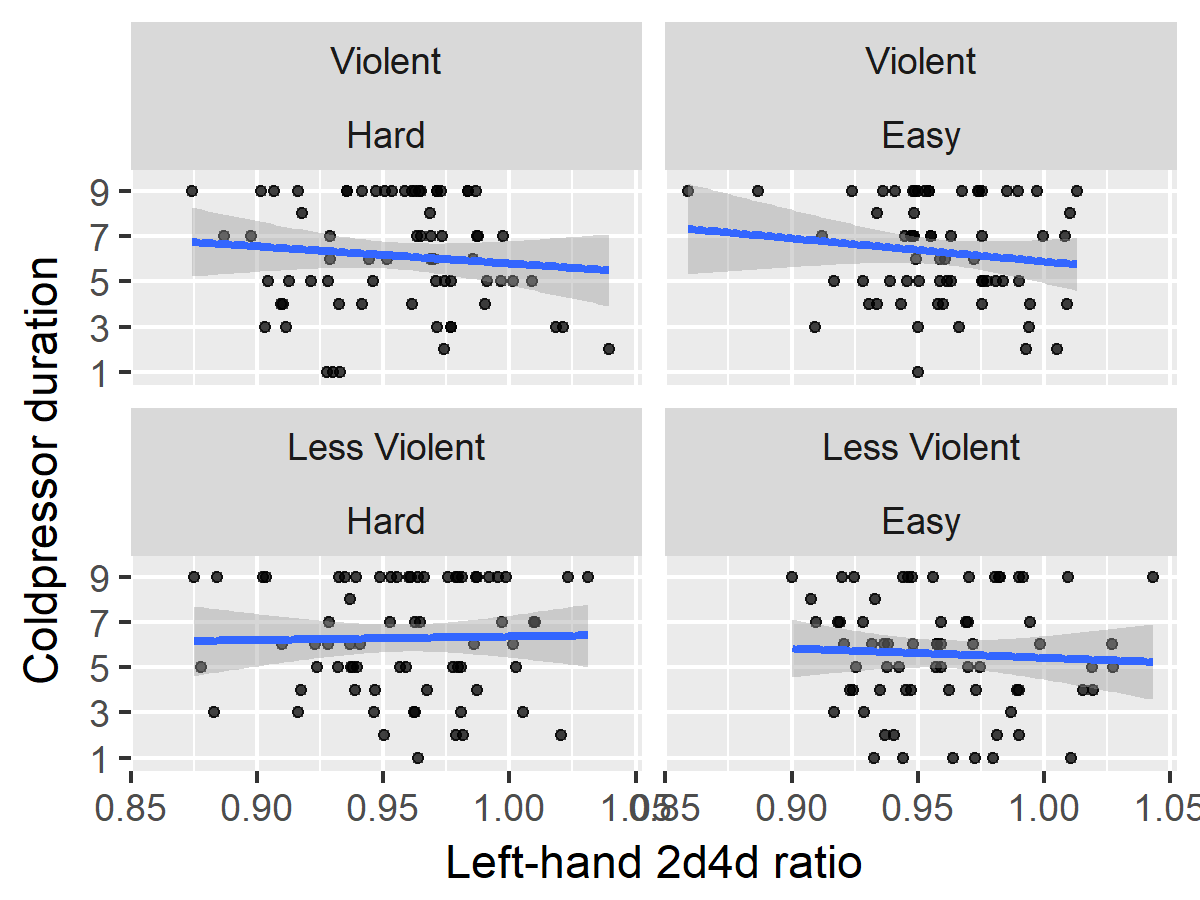
Scatterplot of participants’ experienced provocation and cold pressor assignment. Participants more irritated with the feedback assigned greater cold pressor durations (*r* = .33 [.22, .43]), indicating sensitivity and validity of the cold pressor measure of aggression. A locally-weighted regression curve (LOESS) with shaded standard error region is overlaid.

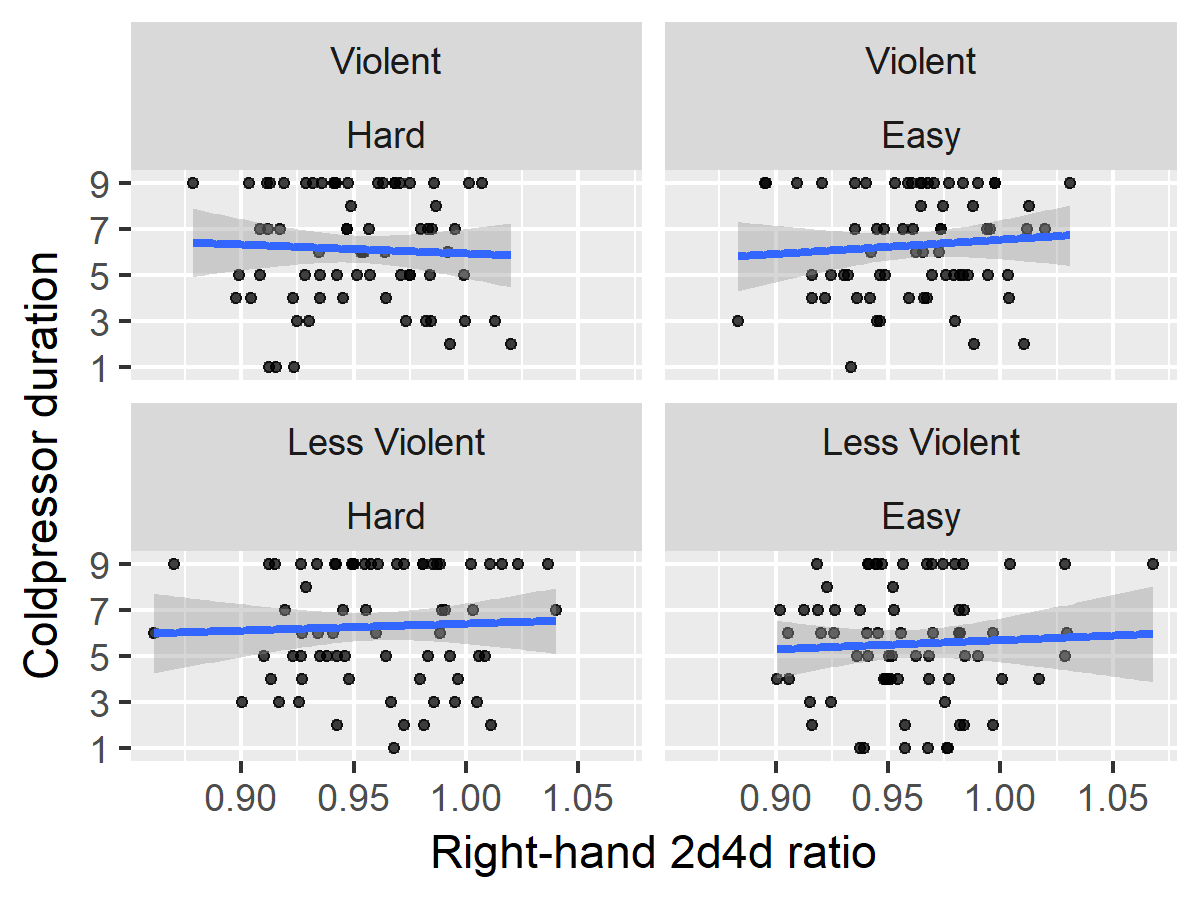
Figure 2. Histograms of cold pressor duration per condition.



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 data.

Figure 3. Null relationship between 2D:4D and aggression





Scatterplots illustrating the relationship between 2D:4D and aggression in each condition. Relationships are consistently near zero.

1. Originally, the comment read, “This is one of the worst essays I have ever read!” consistent with previous research. Participants generally found this to be suspicious and unbelievable, so we changed it to a more flippant and more credible insult. [↑](#footnote-ref-1)