societies

Article

Biased Estimation of Violent Video Game Effects on Aggression: Contributing Factors and Boundary Conditions

Jens Bender [1],*, Tobias Rothmund [1] and Mario Gollwitzer [2]

[1] University of Koblenz-Landau, Fortstraße 7, 76829 Landau, Germany;
E-Mail: rothmund@uni-landau.de

[2] Philipps University Marburg, Gutenbergstraße 18, 35032 Marburg, Germany;
E-Mail: mario.gollwitzer@uni-marburg.de

* Author to whom correspondence should be addressed; E-Mail: bender@uni-landau.de;
Tel.: +49-6341-280-31497.

Abstract: In order to improve the understanding of media violence effects, it is crucial to extend knowledge about factors that threaten the validity of such effects in empirical research. Research artifacts can be expected when participants are (a) aware of a scientist's hypothesis, (b) motivated to confirm or disconfirm the hypothesis, and (c) capable of manipulating their responses in line with their motivation. Based on social identity theory (SIT) and self-categorization theory (SCT), we assumed that identifying with the social group of video game players would provide a motivation to disconfirm the "violent video games increase aggression" hypothesis. We further assumed that the use of nontransparent aggression measures and cover stories would prevent research artifacts. Our results showed that highly identified (compared to lowly identified) players of video games reported less aggression on a transparent aggression measure but not on a nontransparent aggression measure. However, providing participants with a cover story did not prevent hypothesis awareness nor eliminate hypothesis-disconfirming response patterns. These results provide empirical support for the ideas that (a) motivational factors may contribute to a biased estimation of media violence effects, (b) cover stories may not always be effective, and (c) the use of nontransparent aggression measures can provide a valid methodological approach for avoiding biases in media effects research.

1. Research Artifacts in the Social Sciences

Social scientists who study human behavior are faced with a methodological challenge. Their research "objects" are thinking and feeling organisms who are "actively trying to understand the situation they find themselves in, using these explanations partly to help guide their subsequent behaviors" ([1], p. 861). For example, if participants in a media violence study are asked to complete an aggression measure after being exposed to a violent movie, they might correctly guess that the scientists' hypothesis is that violent movies increase aggression. This might alter their subsequent behavior in the experiment, thus causing a research artifact (*i.e.*, a systematic bias that threatens the internal and external validity of the study [1,2]). In other words, the empirical results might not reflect a genuine psychological response to the experimental stimulus (e.g., the violent movie) but rather a response to the scientists' research hypothesis (e.g., the expectation that violent movies increase aggression).

Although research artifacts have received a considerable amount of attention in various research areas [1,3–7], only a few attempts have been made to directly examine these effects in the area of media violence. In the present research, we investigated factors that contribute to research artifacts in studies on the effects of video game violence on aggression. Beyond identifying these artifacts, we also investigated the effectiveness of methodological tools that are commonly used to eliminate research artifacts. Our theoretical assumptions were derived from a framework model that describes the necessary conditions involved in the emergence of research artifacts [1].

2. Psychological Artifactology: An Overview

The pioneering work of Martin T. Orne [2] promoted the idea that participants in social experiments engage in hypothesis guessing, which can distort their subsequent behavior in the experiment in ways that threaten the internal validity of the study. For example, Horowitz and Rothschild [8] conducted an Asch-type conformity experiment [9] based on the hypothesis that "people make wrong judgments about the length of an object as a result of conformity." Participants who were aware of this hypothesis displayed less conformity than participants who were unaware of what the experiment was about.

Following Orne's classical work, the study of research artifacts became a growing field in psychology. After decades of theorizing and empirical research, there is now a huge body of evidence that supports the existence of research artifacts in the study of human behavior as well as their potential to limit the validity of studies [1,10,11].

In an extension of Orne's [2] work, three different factors that account for the emergence of research artifacts have been identified in recent years. These factors are outlined in a theoretical model that was developed to explain when and how research artifacts occur in social scientific research [1,12,13]. In order for a research artifact to occur, participants need to be (a) aware of the scientists' hypothesis, (b) motivated to confirm or disconfirm the hypothesis, and (c) capable of faking their responses in the experiment accordingly [1]. Theoretically, these three factors operate serially, with each posing a necessary but not sufficient condition for the emergence of a research artifact.

2.1. Hypothesis Awareness

The first precondition for research artifacts to occur in social experiments is participants' awareness of the hypothesis [1]. Participants might become aware of the hypothesis if the experiment contains cues that suggest a certain hypothesis and, thus, invite a certain kind of response from participants. These so-called demand characteristics [14] can be communicated unknowingly through the experimental procedure or the experimenter him- or herself [2]. For example, the hypothesis of a simple and straightforward experimental procedure is easy to discern (e.g., measuring attitudes toward smoking after being confronted with an anti-smoking campaign).

2.2. Motivation

Hypothesis awareness is not sufficient for producing a research artifact in an experimental situation. In order for participants to alter their behavior to either confirm or disconfirm a scientist's hypothesis, they need to be motivated to do so [1]. Weber and Cook [7] outlined different types of motivation that can lead participants to artificially confirm or disconfirm a given hypothesis. A motivation to confirm a hypothesis may result, for example, if participants expect to receive a favorable evaluation by the experimenter for their hypothesis-confirming behavior [15,16]. A motivation to disconfirm a hypothesis may result, for example, if participants show reactance to the experimenter's expectations to comply with the hypothesis [17,18].

2.3. Capability

Even if participants are aware of the hypothesis and motivated to confirm or disconfirm the hypothesis, a research artifact can occur only if participants are capable of altering their response pattern in line with their motivation [1]. The capability of participants to strategically alter their response pattern in an empirical study varies depending on the kind of measure that is used. For example, in experiments in which self-report measures are used, participants can easily fake their responses [19,20]. On the other hand, in experiments in which physiological measures are used, participants are less able to alter their responses [21].

Strohmetz's [1] theoretical framework provides researchers not only with guidelines for identifying research artifacts but also with possible approaches for attenuating and even eliminating research artifacts. The present research applied this theoretical framework to media violence research and more specifically to research on the effects of violent video games on aggression. The goal of the present research was (a) to investigate factors that contribute to the emergence of research artifacts in studies on violent video game effects on aggression and (b) to identify methodological approaches that may be able to eliminate the occurrence of research artifacts in this field of research.

3. Research Artifacts in Violent Video Game Research

Research artifacts can occur in almost all areas of psychological research [7]. However, strong arguments can be made for the idea that research on violent video games is especially susceptible to research artifacts. In the following, we will outline these arguments and discuss methodological approaches that have been used to eliminate research artifacts in this area of research. In line with

Strohmetz's [1] theoretical model, we distinguish between participants' (a) awareness of the scientist's hypothesis, (b) motivation to confirm or disconfirm the scientist's hypothesis, and (c) capability of faking outcome measures in studies on violent video games.

3.1. Hypothesis Awareness and Cover Stories

Cover stories are commonly used in violent video game research in order to prevent participants from being aware of the research hypothesis (e.g., [22–24]). For example, in a classical experiment by Anderson and Dill [22], participants were told that the study tested the hypothesis that video game play promotes the learning of motor skills and performance on cognitive tasks. However, there are different reasons to question the effectiveness of cover stories in violent video game experiments. In general, research has shown mixed results, with some studies providing evidence for the effectiveness of cover stories (e.g., [8,25]) and other studies providing evidence to the contrary (e.g., [26,27]). Therefore, cover stories should not be considered a guarantee that participants will be unaware of the research hypothesis [2] because their effectiveness may depend on the specific context. In the context of violent video game research, there are two reasons in particular to question the effectiveness of cover stories. First, whether or not violent video games increase aggression is a frequently debated issue in society and the mass media, especially in the aftermath of school shootings [28]. Therefore, it can be argued that many people are generally aware of the hypothesis that violent video games lead to aggression, thus increasing their awareness that such a hypothesis would exist in an experiment involving violent video games. Second, the experimental procedures used in violent video game experiments tend to provide strong cues that indicate such a hypothesis. A prototypical experiment requires participants to fill out an aggression measure immediately after being exposed to a violent (*vs.* nonviolent) video game. It can be argued that in such a straightforward experimental procedure, the hypothesis would be easy to discern despite the existence of a cover story.

Taken together, these arguments can be applied to question the effectiveness of cover stories in violent video game research and to call for empirical research on the matter. The present research systematically investigates the influence of a cover story in violent video game research by experimentally manipulating the presence *vs.* absence of a cover story.

3.2. Motivation and Social Identity

Participants' motivation to either confirm or disconfirm the research hypothesis may stem from different sources. In the present research, we focused on factors that provide participants who play video games on a regular basis with the motivation to disconfirm the hypothesis. We argue that the emergence of this motivation can be explained on the basis of social identity theory [29] and self-categorization theory [30].

Video games have become increasingly popular in recent years [31,32]. Video game players gather in various online communities (e.g., [33]) and participate in real-life events (e.g., trade fairs such as the E3 or Gamescom). As a reaction to the public debate, players of violent video games have even begun to demonstrate against stigmatization (e.g., [34]). Thus, there is reason to assume that video game players constitute a social group and that the level of identification with this group can vary across individuals [35]. Being highly identified with a social group affects how people act in situations that

are relevant to their group [30]. Accepting the "violent video games increase aggression" hypothesis would imply that players of video games are more likely to be socially deviant. This poses a threat to the social identity of highly identified video game players [36]. As a reaction to such a social identity threat, highly identified group members may engage in defensive responding, such as rationalizing the criticism to their group [37] or devaluing and attacking the source of the threat [38]. In the case of the "violent video games increase aggression" hypothesis, one form of defensive responding might consist in providing arguments against this particular hypothesis. Thus, highly identified video game players who take part in an experiment on violent video games might be motivated to manipulate their responses in order to provide evidence against the "violent games increase aggression" hypothesis.

3.3. Capability and the Transparency of Measures

People not only have to be motivated to disconfirm the hypothesis, they also have to be able to manipulate their responses on the relevant outcome measure. One way to reduce this ability would be to use nontransparent measures. Nontransparent measures are specifically designed to prevent people from discerning the measurement intention and thus from being able to fake responses on the measure. A whole line of research has been dedicated to the development of nontransparent measures such as physiological measures [21], projective tests [39], and reaction-time-based measures (e.g., [40,41]). Violent video game research has made use of nontransparent measures such as physiological measures (e.g., [42]), the implicit association test (e.g., [43]), the reading reaction time task (e.g., [22]), the lexical decision task (e.g., [44]), and the word completion task (e.g., [45]). However, several measures in violent video game research are transparent, such as self-report measures for normative aggressive beliefs (e.g., [46]), aggressive attitudes (e.g., [47]), state hostility (e.g., [22]), state anger (e.g., [48]), and revenge motivation (e.g., [49]). The present research contributes to the literature by systematically comparing responses on a transparent and a nontransparent aggression measure.

4. The Present Research

Despite plausible theoretical arguments for the existence of research artifacts in media violence research, no effort has yet been made to systematically investigate the existence of research artifacts in studies addressing media violence effects. In the present research, we investigated the hypothesis that the effects of video game violence on aggression are systematically biased due to research artifacts. Furthermore, we investigated whether the frequent use of cover stories and nontransparent measures in research on violent video game effects can add to the prevention of research artifacts. Specifically, we made three assumptions. First, we assumed that a strong identification with the social group of video game players would provide a motivation to disconfirm the "violent video games increase aggression" hypothesis. Second and in line with the common practice in violent video game research, we assumed that cover stories would prevent research artifacts because they prevent hypothesis awareness. Third, we assumed that nontransparent measures would prevent research artifacts because they eliminate the capability to fake responses on the relevant outcome variables. These three assumptions led to our central hypothesis.

Hypothesis: If participants in a study on violent video game effects are aware of the research hypothesis (compared to a condition with a cover story), highly identified (compared to lowly identified) players of video games will report less aggression on transparent measures but not on nontransparent measures.

In order to test this hypothesis, we conducted an experiment in which participants played a violent video game and completed subsequent aggression measures. Furthermore, we (a) manipulated whether participants received a cover story or not, (b) measured participants' identification with the group of video game players, and (c) used both a transparent and a nontransparent measure of aggression.

5. Method

Following the guidelines of ethical conduct in research [50], we will report all experimental conditions, all eliminated observations, all dependent variables collected in the study, and all eliminated items on the dependent variables.

5.1. Design and Participants

We conducted a laboratory experiment using a 2 (aggression measure: transparent *vs.* nontransparent) × 2 (cover story: yes *vs.* no) factorial design with repeated measures on the first factor. Identification with the group of video game players was assessed and treated as a continuous moderator variable (see below). A total of $N = 100$ male [51] students of a German university enrolled in various programs were recruited for a study on the evaluation of different media formats. Most participants studied mechanical engineering (32.0%) education (17.0%), civil engineering (9%), and physics (8.0%). Participants were randomly assigned to either the cover story condition ($n = 49$) or the condition without a cover story ($n = 51$). Nine participants had to be dropped from the sample due to missing values on the identification ($n = 7$) or the nontransparent aggression measure ($n = 2$). The final sample consisted of $N = 91$ participants [52–54] with ages ranging from 18 to 28 years ($M = 21.57$, $SD = 1.86$).

5.2. Identification with Video Game Players

Identification with video game players was measured 1 week prior to the laboratory experiment. The identification measure consisted of two items (Cronbach's $\alpha = 0.77$; "Being a video game player is an important reflection of who I am" and "In general, being a video game player is an important part of my self-image") with a 7-point response scale ranging from 0 (*strongly disagree*) to 6 (*strongly agree*). Both items were adapted from Luhtanen and Crocker [55] and translated into German. Importantly, identification with video game players did not differ between the experimental conditions (*i.e.*, cover story *vs.* no cover story), $t(89) = 0.15$, $p = 0.88$, $d = 0.04$.

5.3. Cover Story Manipulation

Upon arriving at the laboratory, participants were informed about the alleged research topic of the study: In the *cover story* condition, participants were told that the study was about the effects of video games on general processing speed. In order to make the research topic more salient, participants read a newspaper article about the alleged positive effects of video games on general processing speed.

In the *no cover story* condition, participants were fully informed that the study was about the effects of violent video games on aggression; participants also read a newspaper article about aggression-enhancing effects of violent video games.

5.4. Transparent and Nontransparent Measure of Aggression

Following the cover story manipulation, participants played a violent video game [56] for 15 min. The video game "Call of Duty" is a first-person shooter game set in World War II. Participants' task was to control an infantry soldier of the "Allied Forces" whose goal is to defeat the troops of the "Axis Powers." Participants were provided with a laptop, computer mouse, and headphones. In the next part of the experiment, aggressive cognitions were measured using both a nontransparent and a transparent measure with two versions of a word completion task [45]. The word completion task is a well-established and frequently used instrument for measuring aggressive cognitions in violent video game research [49]. Participants are presented with word fragments, and their task is to fill in the missing letters in order to form a word. Each word fragment can be completed to form either an aggressive or a neutral word (example from English version: "e x p l _ _ e" can be completed as *explode* or *explore*). An indicator of the availability of aggressive cognitions is obtained by dividing the number of aggressive word completions by the number of total word completions. Therefore, scores can vary between 0 and 1, with higher values indicating stronger availability of aggressive cognitions.

In the "nontransparent" version of the word completion task, participants were led to believe that the task was a measure of general processing speed. In order to increase credibility, the experimenter instructed participants to complete as many word fragments in 5 min as possible. The task also contained 25 additional word fragments that could only be completed to form neutral words (*i.e.*, filler items) to further conceal the aggression-related measurement intention. An item analysis was conducted for the 25 word fragments that could be completed to form an aggressive word, not including the filler items. Following test-construction recommendations [57], only items with an item difficulty index between 5 and 95 and a corrected item-total correlation of $r_{it} > 0.35$ were used to calculate a score for the nontransparent aggression measure. After the exclusion of 12 items that failed to meet these criteria, the nontransparent aggression measure consisted of 13 items that were used in all subsequent analyses.

In the "transparent" version, participants were (a) not informed about the task's measurement intention, (b) allowed as much time as they wanted to complete the task, and (c) given two response options (one aggressive, one non-aggressive) for each word fragment and asked to mark the alternative that came to their mind first when reading the word fragment. Explicitly presenting two response options made the aggression-related measurement intention more obvious. Again, an item analysis was conducted for the 25 word fragments that could be completed to form an aggressive word, excluding the filler items. After we eliminated 11 items that failed to meet the aforementioned criteria, the transparent aggression measure consisted of 14 items that were used in all subsequent analyses. The nontransparent measure was always administered prior to the transparent measure because completing the transparent aggression measure could have revealed the measurement intentions of both instruments.

5.5. Post-Experimental Questionnaire

Finally, participants completed a post-experimental questionnaire that contained manipulation checks for the cover story manipulation and the transparency of the two aggression measures. As a manipulation check for the cover story, we assessed hypothesis awareness with two items ("I believe the study was about the effect of video games on processing speed"; "I believe the study was about the effect of video games on aggression.") on 6-point response scales ranging from 0 (*totally disagree*) to 5 (*totally agree*). The transparency of the measurement intention of both the transparent and the nontransparent aggression measure was assessed by one item respectively ("In your opinion, what did the first word completion task intend to measure?"; "In your opinion, what did the second word completion task intend to measure?"). Open-ended questions were coded as 1 (*aware of measurement intention*) if participants named at least one aggression-related word in their answers (e.g., violence, aggression, war) and 0 (*unaware of measurement intention*) if participants named no aggression-related words in their answers.

6. Results

6.1. Descriptive Statistics

Descriptive statistics for the variables of interest are displayed in Table 1.

Table 1. Descriptive Statistics and Correlations among Variables.

Variables	M	SD	Scale	Correlations		
				1	2	3
1. Aggressive cognitions (nontransparent measure)	0.35	0.17	0–1	-	0.05	0.11
2. Aggressive cognitions (transparent measure)	0.50	0.16	0–1		-	−0.25*
3. Identification with video game players	0.48	0.78	0–6			-

* $p < 0.05$.

6.2. Manipulation Checks

6.2.1. Cover Story

In order to test whether the cover story prevented hypothesis awareness, a 2 (cover story: yes *vs.* no) × 2 (suspected research hypothesis: aggression *vs.* processing speed) ANOVA with repeated measures on the second factor was computed. Results revealed a significant main effect of suspected research hypothesis, $F(1, 89) = 78.63$, $p < 0.001$, $\eta^2_{partial} = 0.47$, indicating that in general, participants believed more strongly that the study was about video game effects on aggression ($M = 4.08$, $SD = 1.18$) than about video game effects on processing speed ($M = 2.02$, $SD = 1.52$). The main effect of cover story condition was not significant, $F(1, 89) = 0.02$, $p = .88$, $\eta^2_{partial} = 0.00$. Even more importantly, there was no interaction between cover story condition and suspected research hypothesis, $F(1, 89) = 1.65$, $p = 0.20$, $\eta^2_{partial} = 0.02$. Follow-up analyses revealed that participants in the cover story condition ($M = 3.90$, $SD = 1.17$) believed that the study was about aggression just as much as

participants in the condition without a cover story ($M = 4.22$, $SD = 1.18$), $t(89) = -1.30$, $p = 0.20$, $d = 0.27$. In conclusion, the provision of a cover story did not prevent hypothesis awareness.

6.2.2. Transparency of the Aggression Measure

As a manipulation check for the transparency of the aggression measures, we compared participants' awareness of the measurement intention between the transparent and nontransparent measures. As expected, more participants identified the transparent measure (33.0%) as a measure of aggression compared with the nontransparent measure (19.8%), $\chi^2(1, N = 91) = 4.65$, $p = 0.03$, $\varphi = 0.30$ [58].

6.3. Hypothesis Test

We expected that if participants were aware of the research hypothesis (compared to a condition with a cover story), highly identified (compared to lowly identified) players of video games would report less aggression on transparent but not on nontransparent aggression measures.

A moderated regression analysis was computed to test the hypothesis [59]. The difference score between the two aggression measures (transparent aggression measure minus nontransparent aggression measure; note that the scale range is equal on both measures) served as the dependent variable. Using a difference score as the dependent variable is equivalent to testing the hypothesis using a mixed-model ANOVA in which the transparent *vs.* nontransparent aggression measure would have been treated as a within-subjects factor. Positive values on the difference-score variable would indicate that participants reported more aggressive cognitions on the transparent compared to the nontransparent aggression measure. In a first step, identification with video game players was entered into the regression model. This main effect explained 6.7% of the variance in the difference-score variable. As expected, people with higher levels of identification with the group of video game players reported fewer aggressive cognitions on the transparent compared to the nontransparent measure, $B = -0.06$, $SE(B) = 0.02$, $p = 0.02$. Separate regression analyses for the transparent and nontransparent measure revealed that identification with video game players was negatively related to aggressive cognitions on the transparent measure, $B = -0.04$, $SE(B) = 0.02$, $p = 0.02$, but unrelated to aggressive cognitions on the nontransparent measure, $B = 0.02$, $SE(B) = 0.02$, $p = 0.30$.

In a second step, we entered the cover story condition (dummy coded: 0 = no cover story, 1 = cover story) and the Identification x Cover Story interaction into the regression model, explaining another 3.3% of the variance in the difference between the transparent and nontransparent aggression measure. Identification with video game players was standardized before the interaction term was calculated [60]. There was no main effect of cover story on the difference score variable, $B = 0.07$, $SE(B) = 0.05$, $p = 0.15$. More importantly and in contrast to our expectations, the interaction between identification and cover story on the difference-score variable was not significant, $B = -0.05$, $SE(B) = 0.05$, $p = 0.30$, indicating that the effect of identification on the difference between the transparent *vs.* the nontransparent measure did not differ between the cover story condition and the condition without a cover story. Notably, the statistical power of finding a medium-sized effect ($f^2 = 0.15$) with a given alpha of $\alpha = 0.05$ in the respective sample ($N = 91$) was 0.95.

7. Discussion

7.1. Summary of Findings and Theoretical Implications

The present study tested the hypothesis that if participants in a study on violent video game effects on aggression were aware of the research hypothesis (compared to a condition with a cover story), highly identified (compared to lowly identified) players of video games would report less aggression on a transparent measure but not on a nontransparent measure. The hypothesis combined three assumptions. First, a strong identification with the social group of video game players provides a motivation to disconfirm the "violent video games increase aggression" hypothesis. Second, nontransparent aggression measures leave participants blind to the measurement intention, thereby preventing them from faking responses in line with that motivation. Third, the use of a cover story prevents hypothesis awareness, thereby preventing research artifacts. We found support for the first and second assumptions, but not for the third assumption. The more participants identified with the group of video game players, the less aggressive cognitions they reported on the transparent (but not on the nontransparent) measure. This finding resonates with other research showing that social identification with the group of video game players provides a motivation to disconfirm the violent video games hypothesis [35].

It is important to bear in mind that identification with video game players is not the only motivational factor involved in violent video game research. Other motivational factors such as reactance [17] may also contribute to research artifacts. Some of these factors may even contribute to an overestimation of violent video game effects; for example, a participant's personal belief that violent video games increase aggression. In the present study, we focused on one motivational factor and confirmed its systematic influence on the outcome measure. The results highlight the importance of carefully considering all relevant motivational factors that might contribute to research artifacts when conducting media violence studies.

In addition to identifying motivational sources for artifacts, it is important to use methodological approaches to control artifacts. The present study tested the effectiveness of cover stories and nontransparent measures for preventing research artifacts. We found an effect of identification with video game players on aggressive cognitions only for the transparent but not for the nontransparent aggression measure. An important implication of this finding is that the use of nontransparent outcome measures can eliminate the systematic effects of motivational factors. Nontransparent measures leave participants blind to the measurement intention, thereby reducing the capability to fake responses on the measure. Therefore, the use of nontransparent measures can be a valid methodological approach that can be applied to eliminate research artifacts.

As a second methodological approach, we tested the effectiveness of cover stories in research on violent video games. Contrary to our expectations, results showed that the cover story we used did not prevent hypothesis awareness nor did it eliminate the systematic effect of identification on the difference between the transparent and nontransparent aggression measure. As indicated by the post-experimental questionnaire, leading participants to believe that the study was about video game effects on general processing speed was not sufficient for decreasing hypothesis awareness. However, post-experimental inquiry may not be a valid method for assessing hypothesis awareness [3,25,61]. Specifically,

participants may have reported hypothesis awareness after the experiment without actually being aware of the hypothesis during the experiment. However, the effect of identification with video game players on the difference score variable did not differ between the condition with and without a cover story. An important implication of this finding is that cover stories may not be as effective at preventing research artifacts as it is commonly assumed. Violent video game research has made frequent use of cover stories to prevent research artifacts (e.g., [22–24]). However, the present study is the first to systematically test the effectiveness of a cover story in an experiment on violent video game effects on aggression.

We would be going too far if we questioned cover stories on a more general level simply based on the ineffectiveness of the specific cover story that we used in our study. However, we invested great effort into leading participants in this condition to believe that the study was about video game effects on general processing speed by (a) explicitly indicating that this was the research topic, (b) asking participants to read an article about the effects of video games on general processing speed, (c) presenting the nontransparent word completion task as a measure of processing speed, and (d) instructing participants to respond as quickly as possible in order to improve credibility. At the very least, the present findings call for more caution in the use of cover stories and more methodological research on their effectiveness.

7.2. Limitations and Future Research

Future research can improve the understanding of research artifacts in violent video game research by addressing four limitations of the present study. First, the negative effect of identification on the transparent aggression measure can be interpreted in two ways. In line with our theoretical assumptions, the effect might be driven by highly identified video game players that provide hypothesis-disconfirming responses. However, the empirical results are also in line with the alternative explanation that lowly identified video game players provide hypothesis-confirming responses. While the first explanation implies an underestimation of violent video game effects, the second would imply an overestimation. Future research should include a condition with a non-violent video game as a reference point.

Second, we found generally low levels of identification with video game players in our sample ($M = 0.48$ on a scale from 0 to 6). Possibly, the effect of identification with video game players was attenuated due to the low mean and variance. Future research should include samples with a higher mean and variance of the identification scores, particularly non-student samples.

Third, we found a low correlation ($r = 0.05$) between the transparent and the nontransparent aggression measure. Possibly, the nontransparent measure reflects implicit processes, while the transparent measure reflects explicit processes [62]. Previous research has shown that the relationship between implicit and explicit measures tends to be very low [63]. Supporting the validity of the word completion task is its frequent and successful use as a measure of aggressive cognitions in violent video game research [49,64]. It would be interesting to investigate the low transparent-nontransparent consistency in future research. Notably, different word fragments were used for the transparent and nontransparent word completion task to avoid memory and consistency effects. Importantly, the two versions are not parallel tests. The transparent version presented participants with two alternatives (one

aggressive, one non-aggressive), whereas the nontransparent version usually allowed more alternatives for word completions (example form the English version: s _ _ y can be completed as slay, sony, stay, sway, or sexy). Therefore, items of the transparent version can be expected to have a lower difficulty index than items of the nontransparent version. The difference in item difficulties also accounts (at least partially) for the higher aggression mean of the transparent ($M = 0.50$) compared to the nontransparent measure ($M = 0.35$).

Forth, we cannot fully rule out alternative explanations for our assumption that group identification was the motivational source of the emergence of a research artifact. For example, highly identified players of video games might generally be more uncooperative [7,65] or prone to psychological reactance [17,66,67]. The present study provides evidence for biased response patterns among some participants. However, it would be interesting to disentangle the alternative explanations for the core motivation behind this effect by statistically controlling for possible confounds or experimentally manipulating group identification (e.g., [68]) in future research.

8. Conclusions

We agree with Hyman's notion that "all scientific inquiry is subject to error, and it is far better to be aware of this, to study the sources in an attempt to reduce it, and to estimate the magnitude of such errors in our findings, than to be ignorant of the errors concealed in the data" ([69], p. 4). In line with this reasoning, the present study was the first to systematically investigate and demonstrate research artifacts in media violence research. Three important insights can be gained from the present findings. First, we found evidence indicating that motivational factors have a systematic effect on relevant outcome measures in media violence studies, thus contributing to a biased estimation of media violence effects. Second, our findings suggest that cover stories may not always be effective in preventing research artifacts and that more methodological research on the effectiveness of cover stories is needed. Third, we found evidence indicating that nontransparent outcome measures can be a valid methodological approach for preventing research artifacts. This stresses the importance of (a) conducting media violence studies that carefully consider the relevant motivational factors and control them by (b) preventing hypothesis awareness (e.g., through a well-designed cover story), and/or (c) eliminating the capability to fake responses on relevant outcome measures (e.g., through the use of nontransparent measures).

Conflicts of Interest

The authors declare no conflict of interest.

References and Notes

1. Strohmetz, D.B. Research artifacts and the social psychology of psychological experiments. *Soc. Personal. Psychol. Compass* **2008**, *2*, 861–877.
2. Orne, M.T. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *Am. Psychol.* **1962**, *17*, 776–783.

3.   Nichols, A.L.; Maner, J.K. The good-subject effect: Investigating participant demand characteristics. *J. Gen. Psychol.* **2008**, *135*, 151–165.

4.   Rosenthal, R. *Experimenter Effects in Behavioral Research*; Appleton-Century-Crofts: East Norwalk, CT, USA, 1966.

5.   Rosenthal, R. Interpersonal expectancy effects: A 30-year perspective. *Curr. Dir. Psychol. Sci.* **1994**, *3*, 176–179.

6.   Rosenzweig, S. The experimental situation as a psychological problem. *Psychol. Rev.* **1933**, *40*, 337–354.

7.   Weber, S.J.; Cook, T.D. Subject effects in laboratory research: An examination of subject roles, demand characteristics, and valid inference. *Psychol. Bull.* **1972**, *77*, 273–295.

8.   Horowitz, I.A.; Rothschild, B.H. Conformity as a function of deception and role playing. *J. Pers. Soc. Psychol.* **1970**, *14*, 224–226.

9.   Asch, S.E. Effects of Group Pressure upon the Modification and Distortion of Judgments. In *Groups, Leadership, and Men*; Guetzkow, H., Ed.; Carnegie Press: Pittsburgh, PA, USA, 1951.

10. Roethlisberger, F.J.; Dickson, W.J.; Wright, H.A. *Management and the Worker: An Account of a Research Program Conducted by the Western Electric Company, Hawthorne Works, Chicago*; Harvard University Press: Cambridge, MA, USA, 1939.

11. Rosenthal, R.; Rosnow, R.L. *Artifact in Behavioral Research*; Academic Press: New York, NY, USA, 1969.

12. Rosnow, R.L.; Aiken, L.S. Mediation of artifacts in behavioral research. *J. Exp. Soc. Psychol.* **1973**, *9*, 181–201.

13. Strohmetz, D.B.; Rosnow, R.L. A Mediational Model of Research Artifacts. In *Probability in Theory-Building: Experimental and Nonexperimental Approaches to Scientific Research in Psychology*; Brzezinski, J., Ed.; Editions Rodopi: Amsterdam, The Netherlands, 1994; pp. 177–196.

14. Carlopio, J.; Adair, J.G.; Lindsay, R.C.L.; Spinner, B. Avoiding Artifact in the Search for Bias: The Importance of Assessing Subjects' Perceptions of the Experiment. *J. Pers. Soc. Psychol.* **1983**, *44*, 693–701.

15. Rosenberg, M.J. The conditions and consequences of evaluation apprehension. In *Artifact in Behavioral Research*; Rosenthal, R., Rosnow, R.L., Eds.; Academic Press: New York, NY, USA, 1969; pp. 279–349.

16. Rosnow, R.L.; Goodstadt, B.E.; Suls, J.M.; Gitter, A.G. More on the social psychology of the experiment: When compliance turns to self-defense. *J. Pers. Soc. Psychol.* **1973**, *27*, 337–343.

17. Brehm, J.W. *A Theory of Psychological Reactance*; Academic Press: New York, NY, USA, 1966.

18. Miron, A.M.; Brehm, J.W. Reactance theory: 40 years later. *Zeitschrift für Sozialpsychologie* **2006**, *37*, 9–18.

19. Viswesvaran, C.; Ones, D.S. Meta-Analyses of Fakability Estimates: Implications for Personality Measurement. *Educ. Psychol. Meas.* **1999**, *59*, 197–210.

20. Ziegler, M.; Schmidt-Atzert, L.; Bühner, M.; Krumm, S. Fakability of different measurement methods for achievement motivation: Questionnaire, semi-projective, and objective. *Psychol. Sci.* **2007**, *49*, 291–307.

21. Podlesny, J.; Raskin, D.R. Physiological Measures and the Detection of Deception. *Psychol. Bull.* **1977**, *84*, 782–799.

22. Anderson, C.A.; Dill, K.E. Video games and aggressive thoughts, feelings, and behavior in the laboratory and in life. *J. Pers. Soc. Psychol.* **2000**, *78*, 772–790.

23. Bushman, B.; Anderson, C.A. Comfortably Numb: Desensitizing Effects of Violent Media on Helping Others. *Psychol. Sci.* **2009**, *20*, 273–277.

24. Carnagey, N.L.; Anderson, C.A. The Effects of Reward and Punishment in Violent Video Games on Aggressive Affect, Cognition, and Behavior. *Psychol. Sci.* **2005**, *16*, 882–889.

25. Levy, L.H. Awareness, learning, and the beneficent subject as expert witness. *J. Pers. Soc. Psychol.* **1967**, *6*, 365–370.

26. Gallo, P.S.; Smith, S.; Mumford, S. Effects of deceiving subjects upon experimental results. *J. Soc. Psychol.* **1973**, *89*, 99–107.

27. Golding, S.L.; Lichtenstein, E. Confession of awareness and prior knowledge of deception as a function of interview set and approval motivation. *J. Pers. Soc. Psychol.* **1970**, *14*, 213–223.

28. Barlett, C.; Anderson, C. Violent Video Games and Public Policy. In *Wie wir Spielen, Was Wir Werden: Computerspiele in unserer Gesellschaft* (In German); Bevc, T., Zapf, H., Eds.; UVK Verlagsgesellschaft: Konstanz, Germany, 2009; pp. 227–240.

29. Tajfel, H. Social identity and intergroup behaviour. *Soc. Sci. Inform.* **1974**, *13*, 65–93.

30. Turner, J.; Hogg, M.; Oakes, P.; Reicher, S.; Wetherell, M. *Rediscovering the Social Group: A Self-Categorization Theory*; Blackwell: Oxford, UK, 1987.

31. Scharrer, E. Virtual Violence: Gender and Aggression in Video Game Advertisements. *Mass Comm. Soc.* **2004**, *7*, 393–412.

32. Williams, D.; Yee, N.; Caplan, S.E. Who plays, how much, and why? Debunking the stereotypical gamer profile. *J. Comput.-Mediat. Comm.* **2008**, *13*, 993–1018.

33. Gamespot. Available online: http://www.gamespot.com/ (accessed on 10 September 2013).

34. Thurm, F. Wir sind keine Amokläufer! Available online: http://www.zeit.de/online/2009/31/Killerspiele-Demo (accessed on 9 September 2009).

35. Nauroth, P.; Gollwitzer, M.; Bender, J.; Rothmund, T. Gamers against Science: The Case of the Violent Video Games Debate. *Eur. J. Soc. Psychol.* **2013**, submitted for publication.

36. Ellemers, N.; Spears, R.; Doosje, B. Self and social identity. *Annu. Rev. Psychol.* **2002**, *53*, 161–186.

37. Branscombe, N.R.; Ellemers, N.; Spears, R.; Doosje, B. The Context and Content of Social Identity Threat. In *Social Identity: Context, Commitment, Content*; Ellemers, N., Spears, R., Doosje, B., Eds.; Blackwell Science: Oxford, UK, 1999; pp. 35–58.

38. Branscombe, N.R.; Wann, D.L. Physiological arousal and reactions to outgroup members during competitions that implicate an important social identity. *Aggress. Behav.* **1992**, *18*, 85–93.

39. Lilienfeld, S.O.; Wood, J.M.; Garb, H.N. The scientific status of projective techniques. *Psychol. Sci. Publ. Interest* **2000**, *1*, 27–66.

40. Greenwald, A.G.; McGhee, D.E.; Schwartz, J.L. Measuring individual differences in implicit cognition: The implicit association test. *J. Pers. Soc. Psychol.* **1998**, *74*, 1464–1480.

41. Williams, J.M.; Mathews, A.; MacLeod, C. The emotional Stroop task and psychopathology. *Psychol. Bull.* **1996**, *120*, 3–24.

42. Carnagey, N.L.; Anderson, C.A.; Bushman, B.J. The effect of video game violence on physiological desensitization to real-life violence. *J. Exp. Soc. Psychol.* **2007**, *43*, 489–496.

43. Uhlmann, E.; Swanson, J. Exposure to violent video games increases automatic aggressiveness. *J. Adolesc.* **2004**, *27*, 41–52.

44. Glock, S.; Kneer, J. Game Over? The Impact of Knowledge about Violent Digital Games on the Activation of Aggression-Related Concepts. *J. Media Psychol.* **2009**, *21*, 151–160.

45. Anderson, C.A.; Carnagey, N.L.; Eubanks, J. Exposure to violent media: The effects of songs with violent lyrics on aggressive thoughts and feelings. *J. Pers. Soc. Psychol.* **2003**, *84*, 960–971.

46. Anderson, C.A.; Gentile, D.A.; Buckley, K.E. *Violent Video Game Effects on Children and Adolescents: Theory, Research, and Public Policy*; Oxford University Press: New York, NY, USA, 2007.

47. Gentile, D.A.; Lynch, P.J.; Linder, J.R.; Walsh, D.A. The effects of violent video game habits on adolescent hostility, aggressive behaviors, and school performance. *J. Adolesc.* **2004**, *27*, 5–22.

48. Fleming, M.J.; Rickwood, D.J. Effects of Violent Versus Nonviolent Video Games on Children's Arousal, Aggressive Mood, and Positive Mood. *J. Appl. Soc. Psychol.* **2001**, *31*, 2047–2071.

49. Anderson, C.A.; Carnagey, N.L.; Flanagan, M.; Benjamin, A.J.; Eubanks, J.; Valentine, J.C. Violent Video Games: Specific Effects of Violent Content on Aggressive Thoughts and Behavior. In *Advances in Experimental Social Psychology*; Zanna, M.P., Ed.; Elsevier Academic Press: San Diego, CA, USA, 2004; Volume 36, pp. 199–249.

50. Simmons, J.P.; Nelson, L.D.; Simonsohn, U. False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychol. Sci.* **2011**, *22*, 1359–1366.

51. We recruited exclusively male participants, because many female participants in our previously conducted experiments reported difficulties while handling the controls of a video game.

52. Little's MCAR test revealed that the missing values were missing completely at random, $\chi^2(14, N = 100) = 12.55$, $p = 0.56$. Based on guidelines for dealing with missing data, data were imputed using the expectation-maximization (EM) algorithm, and all analyses were run with and without imputation. Analyses with imputed data produced the same results as analyses on data with listwise deletion. Consequently, all analyses are reported for data with listwise deletion.

53. Tabachnick, B.G.; Fidell, L.S. *Using Multivariate Statistics*, 5th ed.; Allyn & Bacon: Boston, MA, USA, 2007.

54. Little, R.J.A. A Test of Missing Completely at Random for Multivariate Data With Missing Values. *J. Am. Stat. Assoc.* **1988**, *83*, 1198–1202.

55. Luhtanen, R.; Crocker, J. A collective self-esteem scale: Self-evaluation of one's social identity. *Pers. Soc. Psychol. Bull.* **1992**, *18*, 302–318.

56. *Call of Duty*, version green pepper; Computer software; Activision: Los Angeles, CA, USA, 2003.

57. Kelava, A.; Moosbrugger, H. Deskriptivstatistische Evaluation von Items (Itemanalyse) und Testwertverteilungen. In *Testtheorie und Fragebogenkonstruktion* (in German); Moosbrugger, H., Kelava, A., Eds.; Springer: Heidelberg, Germany, 2007; pp. 73–98.

58. It might be argued that in the "no cover story" condition, highly identified video game players would question the ostensible purpose of the nontransparent measure. If this was the case, then both the transparent and the nontransparent measure should be negatively related to identification in the "no cover story" condition. This, however, was not the case: Whereas identification was

negatively correlated with the transparent measure ($r = -0.22$, $p = 0.13$), it was not correlated with the nontransparent measure ($r = 0.05$, $p = 0.74$).

59. Cohen, J.; Cohen, P.; West, S.G.; Aiken, L.S. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3rd ed.; Lawrence Erlbaum Associates Publishers: Mahwah, NJ, USA, 2003.

60. Aiken, L.S.; West, S.G. *Multiple regression: Testing and Interpreting Interactions*; Sage Publications: Newbury Park, CA, USA, 1991.

61. Straits, B.C.; Wuebben, P.L.; Crowle, A.J. Confession of Prior Knowledge about Experimental Procedures as a Function of Evaluation Apprehension and Commitment. *Soc. Sci. Res.* **1982**, *11*, 227–244.

62. De Houwer, J. What Are Implicit Measures and Why Are We Using Them. In *The Handbook of Implicit Cognition and Addiction*; Wiers, R.W., Stacy, A.W., Eds.; Sage Publishers: Thousand Oaks, CA, 2006; pp. 11–28.

63. Fazio, R.H.; Olson, M.A. Implicit measures in social cognition research: Their meaning and use. *Annu. Rev. Psychol.* **2003**, *54*, 297–327.

64. Anderson, C.A.; Shibuya, A.; Ihori, N.; Swing, E.L.; Bushman, B.J.; Sakamoto, A.; Rothstein, H.R.; Saleem, M. Violent video game effects on aggression, empathy, and prosocial behavior in Eastern and Western countries: A meta-analytic review. *Psychol. Bull.* **2010**, *136*, 151–173.

65. Cook, T.D.; Bean, J.R.; Calder, B.J.; Frey, R.; Krovetz, M.L.; Reisman, S.R. Demand characteristics and three conceptions of the frequently deceived subject. *J. Pers. Soc. Psychol.* **1970**, *14*, 185–194.

66. Hong, S.M. Hong's Psychological Reactance Scale: A further factor analytic validation. *Psychol. Rep.* **1992**, *70*, 512–514.

67. Hong, S.M.; Faedda, S. Refinement of the Hong Psychological Reactance Scale. *Educ. Psychol. Meas.* **1996**, *56*, 173–182.

68. Leonardelli, G.J.; Brewer, M.B. Minority and majority discrimination: When and why. *J. Exp. Soc. Psychol.* **2001**, *37*, 468–485.

69. Hyman, H. *Interviewing in Social Research*; University of Chicago Press: Chicago, IL, USA, 1954.