Impact of violent video game realism on the self-concept of aggressiveness assessed with explicit and implicit measures

Joerg Zumbach a,⇑, Caroline Seitz a, Matthias Bluemke b

a University of Salzburg, Austria
b University of Heidelberg, Germany

Full Length Article

Article history:
Received 4 February 2015
Revised 7 July 2015
Accepted 11 July 2015

Keywords:
Violent computer games
Aggression
Immersive virtual environments

ABSTRACT

We compared the standard 2D representation of a recent violent computer game to its 3D representation realized by shutter-goggles in a lab experiment. Assuming that the higher degree of realism of media violence would impact stronger on players in a pretest–posttest design, we analyzed the influence of violent video game exposure on implicit and explicit measures of aggressiveness. According to an explicit questionnaire on aggressiveness, participants reported having become more peaceful, whereas an Implicit Association Test on aggressiveness (Agg-IAT) indicated that the association between self and aggressive behavior became stronger after violence exposure, confirming the unique utility of Agg-IATs in media research. The 3D visualization mode, however, did not further strengthen this association, and a media-model of increases in aggressiveness by participants’ flow experiences was not supported. When inspecting flow experiences, an interaction effect between gender and visualization mode was evident: Male participants were more likely to have flow experiences in the high-realism (3D) format, whereas female participants were more likely to experience flow in the standard (2D) mode. We discuss the findings in the context of automatic information processing in aggression, and we contend possible changes in automatic behavioral precursors due to media influence.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

This contribution investigates the psychological consequences of a potent factor in violent video and computer games (VVG): a player’s immersion into realistic game scenarios. The present experiment has been designed to test whether the impact of VVG on automatic aspects of players’ aggressiveness concept is a function of the game’s realism, potentially mediated by a player’s higher flow experiences in more realistic scenarios. The introduction will first briefly review short- and long-term effects of violent media exposure and theoretical explanations. Then potential mediators between game realism in VVG and subsequent changes of aggressiveness will be discussed. Drawing on current dual-process models of social cognition and behavior, it is contended that the observation of possible effects should encompass more objective measures of aggressive dispositions, such as implicit measures, as they capture crucial automatic processes that matter in the generation of aggressive behavior and may be shifted by VVG.

1.1. Violence exposure in video games

Computer game features help players get immersed into the games. Carefully calibrated challenges and rewards, well-balanced with elements of chance, engage a player in various tasks. Not denying the positive effects on players’ enjoyment, social ties, mental relaxation, visual acuity and other cognitive benefits (e.g., Adachi & Willoughby, 2013; Granic, Lobel, & Engels, 2014; Green & Bavelier, 2006, 2012), what has concerned many researchers is that violence exposure can stimulate aggressive cognition, affect, and behavior, while reducing empathy and prosocial behavior (e.g., Anderson & Bushman, 2002; Bartholow, 2005; Bushman & Anderson, 2001). In line with social learning theory (Bandura, 1977), gratification after observing aggressive models is held responsible for acquiring aggressive behavioral scripts (Crick & Dodge, 1994). Contemporary approaches underscore learning even in the absence of immediate reinforcement (e.g., mimicry of aggressive scripts; Bushman & Anderson, 2001; Bushman & Huesmann, 2006; Huesmann & Kirwil, 2007).

One prominent model for comprehensively explaining detrimental VVG effects in the short and long run is the General Aggression Model (GAM; Anderson & Bushman, 2002; Anderson & Dill, 2000; Bushman & Anderson, 2002; for a critical note see...
In the short run, situational factors—moderated by personality traits (Bushman, 1995)—can trigger aggressive scripts (short-time effects refer here to minutes up to hours; i.e., on the level of a state). The supposed mechanisms are unspecifiable arousal, angry emotions, and aggressive cognitions. A multitude of primary lab experiments as well as meta-analyses have demonstrated a significant causal effect of computer game violence on aggression (Anderson, 2004; Anderson and Bushman, 2001; Anderson et al., 2003). Long-term effects have been described for emotional desensitization to violence and chronic activation of aggressive behavioral scripts (Carnagey & Anderson, 2004; Cline, Croft, & Courier, 1973; long-time effects refer to time periods of months and years; i.e., they are rather on a trait level). Paralleling research on TV violence (Huesmann, Moise-Titus, Podolski, & Eron, 2003), long-term studies have shown that consensual or dominantly used forms of TV violence influence the threat assessment and aggression during adolescence (Anderson et al., 2008).

Despite these stable effects, the cause-effect debate is still going on, partly because meta-analytical conclusions diverge (Anderson et al., 2010; Ferguson & Kilburn, 2009). To the extent that preformed attitudes and gamer’s explicit self-views contrast with autonomous effects, which may even be hidden from introspection, a reconciliation seems difficult to achieve, especially if players feel as if they need to defend themselves against accusations of negative effects from common leisure activities (Kneer, Glock,beskes, & Bente, 2012; Przybylski, Deci, Rigby, & Ryan, 2014).

One way to partially reconcile the contradictory evidence then, apart from critically discussing any sophisticated meta-analytical strategies, is to highlight the limits of self-report measures of aggressive dispositions (Bender, Rothmund, & Gollwitzer, 2013) and to use objective implicit measures as a complement. The latter are based on objective indicators; being less blatant, they are more difficult to control than subjective measures. It has been firmly established that implicit measures of the aggressiveness self-concept sensitively reflect the influence of different computer games, and that implicit measures, under controlled conditions, reflect the impact of violent, but not nonviolent game conditions (Bluemke, Friedrich, & Zumbach, 2010; Bluemke & Zumbach, 2012; Uhlmann & Swanson, 2004). What is still unclear, though, is whether implicit measures reflect the influence of game realism as sometimes found with subjective measures, and whether the altered psychological state actually serves as a mediator in this process.

1.2. Involvement, identification, and realism in computer gaming

We use as a starting point the recent computer games that often attempt to maximize the realism of computer games. Here, virtual reality is intended to help players immerse themselves totally in a game, and thus to increase realism (Hoorn, Konijn, & van der Veer, 2003; Konijn & Hoorn, 2005). Modern Immersive Virtual Environment (IVE) technologies, which are typically head-mounted displays with 3D visualization and motion-based input devices, are used more frequently nowadays, partly with the approved goal to establish aggressive skills, for instance, in military training units (Persky & Blascovich, 2008; Tamborini & Skalski, 2006). This drives our overall research question: Do VVG effects become stronger, the more real the games feel, and can this be demonstrated independently of subjective self-reports?

Our first working hypotheses are here: (1) Directing one's game character through a virtual world connects the player’s self with the virtually executed behaviors to some extent in the short run, and if exposed to violent actions repeatedly, a player’s self-concept may be altered such that aggressive schemata are advanced as the dominant ones (Crick & Dodge, 1994). (2) This automatic association between self and aggressive schemata will be picked up by implicit measures (though it can be censored from an explicit view of the self).

As a preliminary answer, research has shown that the mere involvement of a computer game player with a game character (avatar) is sufficient for stronger media effects than when watching similar content passively (Carnagey & Anderson, 2004). Identifying with the virtual protagonist can blur the distinction between the virtual and the real world and exacerbate subsequent aggressive behavior (Konijn, Nije Biervank, & Bushman, 2007). Support for the relevance of a player’s involvement in realistic VVG comes from a lab experiment by Fischer, Kastenmüller, and Greitemeyer (2011). When participants created their own avatars so that these were optically reminiscent of the players themselves, the detrimental effect of violence exposure on aggression was markedly intensified in comparison to pre-selected avatars, and this effect was mediated by players’ involvement (cf. Hoffner & Buchanan, 2005).

In another study, Tamborini et al. (2004) compared VVG effects of a modern IVE with those of a classic desktop version of the game. Participants in the IVE condition reported that they experienced being present in the game (so-called “telepresence”) more strongly and felt more hostile affect afterward (cf. Eastin & Griffiths, 2006). Persky and Blascovich (2008) showed that, in comparison to less immersive desktop platforms, IVE platforms promoted particularly the experience of presence and aggressive thoughts, not necessarily of aggressive behavior; they also found higher levels of aggressive cognition among active players than among passive game observers. Krcmar, Farar, and McGlin (2011) showed that IVE participants reported higher levels of aggression to the extent that they strongly identified with their characters, suggesting that experienced realism mediated between platform and individual effects (cf. Ivory & Kalyanaraman, 2007). Rajae-Joordens (2008) observed long-term effects of using 2D and 3D devices for watching movies and playing a VVG. The 3D device significantly intensified not only aggressive, but all sorts of emotional experiences. It appears that blurring the distinction between virtual world and reality by increasing telepresence aggravates any emotional experiences, and possibly also unwarranted VVG effects.

The different processes—identification, involvement, and experienced realism—feed into a common signal: the experience of immersive flow, that is, the total absorption in a task, which is a specific state of mind and powerful explanatory concept. According to Csikszentmihalyi (1991), flow is an increased sense of psychological immersion, a holistic sensation that people feel when acting with total involvement. The primary antecedents of flow in computer-based tasks are congruency between skills and challenges and narrowly focused attention, complemented by secondary antecedents such as interaction and telepresence (Hoffman & Novak, 1996; Novak, Hoffman, & Yung, 2000). Telepresence refers to the perception of being present in virtual reality, once one is fully immersed. Equally, immersion has been defined as the state of flow when navigating a virtual environment (Koh & Kim, 2003). Hence, flow experiences and immersion are mutually related in the sense that flow is an affective outcome that follows from the cognitive phenomenon of interpreting spatial, acoustical, tactile, visceral, and social cues as if one were enveloped by a virtual world (Grinberg, Careaga, Mehl, & O’Connor, 2014). In a cyclic fashion, the more flow a gamer experiences, the stronger the sense of immersion into virtual reality. Though immersion and flow can be disentangled, they are conceptually and empirically related (Faiola, Newlon, Pfaff, & Smyloslova, 2013); many authors simply equate the two (see also Huang, Chi, Sung, & Farn, 2011). Also the scale we used (see Methods) refers to telepresence.
and immersive aspects of flow; hence for the present purpose we will use the terms “flow” and “immersion” interchangeably.¹

The occurrence of flow itself is fostered by the realism of the environment and the presence of the player. We propose that experiencing a game as realistic increases identification with a game character and involvement with the game. Sufficient playing skills provided, experienced sensory realism in a 3D environment will nourish psychological realism. Along this psychological chain, the overlap between virtual and real world might become so strong that a player’s real self becomes highly blended with game character executing aggressive scripts in the virtual world, thereby informing a player’s affect, cognition, and behavior (Klimmt, Hefner, & Vorderer, 2009).² Any disruptions of flow (e.g., technical annoyances, distancing oneself mentally from the avatar) are likely to weaken the connection between the player’s self and his game character. In sum, we come to our second set of working hypotheses that (3) playing VVG in 3D mode shapes a player’s aggressive dispositions and self-concept of aggressiveness more than in 2D mode, and (4) that immersive flow is a crucial mediator.

Though the aforementioned studies inspire adopting a causal chain model, the previously mentioned IVE studies either did not manipulate realism at all, counteracted the realism manipulation by having gamers play in an unrealistic invulnerable “god-like” mode, counteracted the notion of realism by playing an avatar from a third-person rather than first-person perspective, or did not vary exclusively 2D and 3D realism (invoking confounders). For instance, sometimes participants played at typical 2D desktops but were compared to players wearing an unusual—and restricting—3D helmet plus gloves that both were meant to insulate the users from environmental distractions. Alternatively, one group played an old-fashioned game while the other group played a modern game simply related by label (Doom 1 and 3), while at the same time significant differences existed between the quality of games, movements, player interactions, graphic resolutions, and gory details. According to the Media Comparison Paradigm (Bluemke & Zumbach, 2012; Bluemke et al., 2010), media comparisons are only valid within “zones of comparability”. In these zones either the semantics are identical (i.e., what specific media content participants encounter) while the syntactics vary (i.e., how participants interact with the media content), or the semantics vary while the syntactics are identical.

The goal of our research was to compare realism conditions while both content and interactivity remain identical (cf. Przybylski et al., 2014, for well-controlled interactive and structural aspects of media effects). Furthermore, a renewed inspection with the help of implicit measures seemed warranted as research on the mediating model has yielded inconsistent results. The mediation account mostly held for explicit self-report measures, yet when using rather objective indicators, the mediational model occasionally broke down (Jeong, Biocca, & Banse, 2015; Persky & Blascovich, 2008). This role of different measures of aggression and their role within research on media influences on aggression is another important and central issue here.

1.3. Measuring automatic and controlled processes in aggression

Explicit measures require participants to introspect, deliberate, and report accurately. Implicit measures utilize the performance in simple categorization tasks. Participants’ associations are inferred on the basis of their objective response latencies, which are taken to reflect automatic attitudes, stereotypes, self-evaluations, or trait-like aspects of the self-concept (Greenwald & Banaji, 1995). More specifically, these constructs are captured in an associative, non-propositional form a few hundred milliseconds after stimulus presentation (Gawronski, 2009). Implicit measures have benefits over explicit measures when predicting behavior (a) in clearly automatic behavioral domains, (b) in situations of low cognitive or motivational control, or (c) when participants are unable or reluctant to admit the attribute in question to themselves or others (Greenwald, Poehlman, Uhlmann, & Banaji, 2009). Embraced across nearly all psychological domains (e.g., Haines & Sumner, 2013; Roels et al., 2011), they have not had much impact on the field of aggression yet (cf. Bluemke & Teige-Mocigemba, 2015).

The Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) is one of the most successful implicit measures (for limitations see De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009). By selecting aggressive and non-aggressive behavior stimuli IATs can target at the spontaneous association between the self and aggressive behavior, and the resulting IAT scores can reflect automatic components of aggressive dispositions that are not fully represented by questionnaires (Banse & Fischer, 2002; Banse, Messer, & Fischer, 2015). With regard to the impact of VVG on implicit measures, Uhlmann and Swanson (2004) were the first to show that such an aggressiveness-IAT (Agg-IAT) correlated in a trait-like manner with long-term exposure to VVG. Furthermore, in their lab experiment Agg-IAT scores adequately reflected randomly allocated violent or non-violent computer game play (see Lemmer, Gollwitzer, & Banse, 2015, on the question of trait vs. situation-specific variance in the Agg-IAT). Bluemke et al. (2010) showed that the Agg-IAT was sensitive to both peaceful and violent game content even when controlling arousal confounds. In another study an IAT-variant differentiated between groups of computer game users (Bluemke & Zumbach, 2012). Never did explicit measures fully converge with implicit measures.

These findings indicate a dissociation rather than an overlap between explicit and implicit measures, which is in line with mounting evidence on the utility of implicit measures to explain parts of the variance in aggressive behavior via an automatic (as opposed to a controlled) pathway that can trigger behavioral scripts but is not captured by explicit measures (Richetin & Richardson, 2008). For instance, in studies on incremental validity, the IAT predicted aggression when trait-variables indicated, or situational manipulations created, low self-monitoring, self-control, or executive functioning (e.g., Bluemke & Friese, 2012; Denson, Caper, Oaten, Friese, & Schofield, 2011; Richetin, Richardson, & Mason, 2010; Schmidt, Zimmermann, Banse, & Imhoff, 2015).

Recent dual-process models of social cognition can explain both the dissociation between explicit and implicit measures, as well as the incremental validity of implicit measures in behavioral prediction. They conjecture that information processing can be reflective, effortful, slow, and flexible as opposed to automatic, effortless, fast, and inflexible. For instance, according to the reflective-impulsive model (RIM; Strack & Deutsch, 2003, 2004), human cognition promotes the activation and execution of behavioral scripts either via a rather reflective or a rather impulsive pathway, with a complex interplay between the two systems of information processing. Compatile with this notion, the GAM integrates an automatic, effortless, and fast route (called “impulsive action”) next to a reflective, but more resourceful and slower way to the activation of aggressive scripts (called “thoughtful action”). The mechanisms leading from perceptual input (cues) to aggressive behavior may involve associative priming of stereotypes, hostile interpretations, and behavioral scripts (Berkowitz, 1993; Crick & Dodge, 1994) as

¹ The label “immersion” has also been used for one of three factors underlying player motivations, summarizing empirically correlated aspects of discovery, role-playing, customization, and escapism (Yee, 2006). However, other authors clearly conceptualize immersion as a subscale of flow experiences, with immersion being the strongest contributor for learning success in e-learning games (Fu, Su, & Yu, 2009).

² Note that positively valenced scripts can be acquired too (Bluemke et al., 2010; Sestir & Bartholow, 2010). We will return to this in the Discussion.
well as reduced self-control due to angry emotions and unspecific physiological arousal or excitation transfer (Batinic & Appel, 2008; Bushman & Huesmann, 2006; Zillmann, 1978). Whereas the GAM does not specify how impulsive precursors are to be assessed, the RIM designates implicit measures for tapping—imperfectly though—into associative structures residing in the impulsive system.

Taking up the notion of dual-process models and applying it to the domain of aggression, one can argue that violent media are likely to affect not only reflective precursors of behavior but also automatic information processing, that is, accessible scripts and spontaneous associations (e.g., Bushman, 1998). These automatic processes are known to predispose some individuals to spontaneously aggress with higher likelihood than others—even if individuals are unaware of these determinants, or unwilling to report on them (Berkowitz, 2008; Richetin & Richardson, 2008; Todorov & Bargh, 2002). From the perspective of dual-process models, by using rather blatant and deliberate measures to demonstrate violence-exposure or realism effects, one misses out spontaneous cognitive processes that occur involuntarily. Relying exclusively on transparent questionnaires thus entails the risk of underestimating aggressive tendencies and not uncovering VVG effects (Bender et al., 2013; Mills & Kroner, 2006). Yet with the Agg-IAT there is a measure for exploring nuances of altered automatic associations after VVG exposure in 2D or 3D mode. Having executed aggressive behavioral scripts repeatedly during game play in virtual reality, a player's game character is associated with aggressive behavioral scripts repeatedly during game play in the domain of aggression, one can argue that violent media are likely to affect not only reflective precursors of behavior but also automatic information processing (Gawronski, 2009) and changes in the self-concept (Klimmt et al., 2010). As the impact of realistic, highly involving first-person shooters on Agg-IAT scores is to date unknown, we also compared the effect size on the Agg-IAT to previous outcomes based on static game scenarios with relatively low realism, as one might suspect stronger VVG effects after higher realistic game play.

Furthermore, in previous studies involving the Agg-IAT men had rather been underrepresented in the lab (Bluemke et al., 2010), but overrepresented in an online study (Bluemke & Zumbach, 2012), warranting a further check with a gender-balanced sample. We did not expect a severe influence of gender on the predicted data pattern, yet it has been noted that gendered self-identification with an avatar can foster aggression (Anderson & Murphy, 2003; Krcmar et al., 2011). Thus, if gender effects moderated the link between violence exposure and aggressiveness such that men generally showed higher Agg-IAT changes than women, we would expect this pattern to be even more pronounced in the highly realistic game environment.

2. Method

2.1. Participants and design

Sixty-nine participants from the University of [removed for review] (52% males; $M_{age} = 24.46$ yrs., $SD = 5.86$) participated in an experiment on computer visualization modes in exchange for course credit. Another four participants are not included in this sample due to incomplete data (technical failure, unmotivated responding in the IAT with more than 50% errors in at least one critical block, or complete drop-out). Due to listwise deletion for the occurrence of missing data, the reliability analyses based on Cronbach's Alpha may be based on fewer than 69 participants. As regards the experimental procedure, participants were randomly assigned to the 2D or 3D condition. The pretest-posttest design afforded implicit and explicit measures of aggressiveness, taken before and after the computer game. In addition, we assessed potentially mediating variables at posttest.

Participants reported owning a personal computer (100%), a Sony playstation (45%), a Microsoft Xbox (14%), or a Nintendo...
Wii (13%). On average they played 0.60 h a day ($SD = 1.33$) and 5.91 h a week ($SD = 9.19$). As regards gender differences, 22% of men vs. 6% of women reported owning an Xbox, and the likelihood to play computer games was double for men (36% vs. 18%), in line with Gentile, Lynch, Ruh Lindner, and Walsh (2004). Men played twice as many hrs/week than women, $M_s = 7.64$ ($SD = 9.83$) vs. 4.01 ($SD = 8.17$), in line with Lemmens and Bushman (2006).

Thirty-three participants stated that they do not play computer or video games at all (6 of them male).

### 2.2. Material and experimental procedure

After providing informed consent, participants reported socio-demographic characteristics. First, they worked on the implicit measure (Agg-IAT), then an explicit aggression questionnaire. This order was fixed to prevent priming of specific aggressive content and subsequent effects of aggression suppression (e.g., Glock & Kneer, 2009). Next, the 10-min treatment provided one experimental group with the realistic and highly graphic first person shooter “Call of Duty: Modern Warfare 2” (Act 1, Level 2; Infinity Ward/Activision Blizzard, 2011) on an IBM-type personal computer (PC) attached to a 120 Hz LCD flat screen (21” screen width). Another group used special shutter-goggles (nVIDIA GeForce 3D Vision) that allowed players to experience the same game at the same PC and LCD-screen in a realistic and flicker-free 3D manner.

Ten minutes of a mildly difficult level were presented to participants. The player switched into the role of a soldier within a team (other players were CPU-controlled) and the screen provided a first person perspective of the game environment and player’s extremities and weapons. By moving through the virtual environment the task was to shoot enemy soldiers while protecting oneself against enemy fire. Usually a level was completed after having destroyed all enemy forces and conquered specific targets. Apart from asking for the subjective flow experiences during the game, implicit and explicit measures of aggressiveness were taken a second time. Including a debriefing about the study background, the experiment took between 45 and 55 min. We used SPSS-20 for the statistical analyses.

#### 2.2.1. Implicit measure of aggressiveness

The Agg-IAT assessed the speed while sorting stimulus words in a double-barreled sorting-task (Uhlmann & Swanson, 2004). To keep the repeated-measurement procedures succinct, we used a standard five-block structure with a fixed order of the crucial blocks to reduce extraneous variance that might attenuate correlations and mediating paths. The following blocks required participants to correctly categorize stimuli with the help of two response keys: (1) attribute words (i.e., behavioral stimuli such as HIT or TALK) as either aggressive or peaceful (20 practice trials), (2) target words (such as ME or YOU) as related to self or other (20 practice trials), (3) attribute and target words when self + peaceful (as well as other + aggressive) shared one response-key (40 critical trials), (4) target stimuli with inverted response-key assignments (20 practice trials), and (5) target and attribute words when self + aggressive (other + peaceful) shared one response-key (40 critical trials). Participants had to correct their responses if they made a categorization error. The stimuli were behavioral words taken from studies that had successfully used German Agg-IATs (see Banse et al., 2015; Bluemke & Friese, 2012; Bluemke & Zumbach, 2012; Bluemke et al., 2010).

The first two trials in each critical block were discarded to compensate for orientating reactions and slow responses in the absence of practice trials in the critical blocks. Individual IAT effects were computed as so-called D-scores (Greenwald, Nosek, & Banaji, 2003), that is, as latency difference between the mean latencies in the critical blocks 3 and 5, standardized by an individual’s pooled standard deviation in the critical blocks. In line with recommendations for IAT scoring procedures for cognitively taxing treatments, no error penalties were applied to erroneous trials (see Bluemke & Zumbach, 2012; Schoel, Bluemke, Mueller, & Stahlberg, 2011). Higher scores indicate higher aggressiveness, and for most participants negative IAT scores (indicating a peaceful self-concept) resulted, yet given a fixed block order and the absence of an IAT calibration procedure of zero scores, the absolute magnitude of the IAT effects should not be overinterpreted. D-scores were reliable at both pretest and posttest (Spearman–Brown corrected reliability of IAT effects based on odd and even trials), $r_1 = r_2 = .75$ (90%-confidence intervals $CL_{90} = .62$–.83).

#### 2.2.2. Explicit measure of aggressiveness

A German version of the Aggression Questionnaire (BPAQ; Buss & Perry, 1992) with 29-items and a 5-point rating scale format (ranging from “I fully agree” to “I do not agree at all”) assessed facets of self-perceived aggressiveness. The items form subscales pertaining to physically aggressive behavior (e.g., “I have threatened people that I know well”), verbally aggressive behavior (e.g., “My friends say that I am belligerent”), anger (e.g., “I erupt due to nothing”), and hostility (e.g., “If people are kind to me I wonder what they want”). Two items did not enter the analyses to satisfy the intended four-dimensional structure (Von Collani & Werner, 2005). Analyses were based on the BPAQ means across all items, which were reliable, Cronbach’s $\alpha = .84$ ($CL_{90} = .79$–.89, $n_1 = 62$), and $\alpha = .88$ ($CL_{90} = .84$–.91, $n_2 = 60$) at pretest and posttest, respectively.

#### 2.2.3. Involvement, immersion, and flow experiences

Participants’ flow experiences during the game were assessed with a modified version of the Flow Short Scale adapted to game play in the present study (Engeser & Rheinberg, 2008; Rheinberg, Vollmeyer, & Engeser, 2003). The scale is based on the flow-concept as suggested by Csikszentmihalyi (1991) and has been externally validated in predicting learning performance within a statistics program or success in an Assessment Center subtask (cf. Engeser & Rheinberg, 2008; Rheinberg & Vollmeyer, 2003). Thirteen items with 7-point rating scales were averaged so that they formed the three relevant subscales absorption by activity (4 items, e.g., “I was totally absorbed in what I was doing during the game”; $\alpha = .78$, $CL_{90} = .70$–.85, $n = 68$), fluency of performance indicating the right skill level (6 items, e.g., “The right thoughts/movements occurred of their own accord during the game”; $\alpha = .86$, $CL_{90} = .82$–.90), and outcome importance encompassing potentially distracting thoughts (3 items, e.g., “I was worried about failing”; $\alpha = .85$, $CL_{90} = .79$–.90). Note that absorption reassembles the concept of presence or “being present”, that is, the factor accounting for realism effects on explicit reports of aggressive feelings and thoughts (Persky & Blascovich, 2008). The first two subscales were substantially correlated, $r = .62$ ($p < .001$), whereas outcome importance correlated with the other subscales only weakly, $rs = .27$–.33 ($ps < .05$). Hence, we aggregated absorption and fluency items to index the overall flow experience ($\alpha = .88$, $CL_{90} = .84$–.91, $n = 67$).

### 3. Results

We inspected pretest–posttest changes due to game play. The main analysis of explicit and implicit measures of aggressiveness used a 2 (time: pretest vs. posttest) × 2 (realism: 2D vs. 3D) × 2 (gender) ANOVA with repeated measurement on the first factor. Flow experiences were analyzed in the same manner, yet—in the absence of a baseline score—without the repeated measurement factor time.
3.1. Explicit measure

BPAQ scores were affected though not in the direction predicted by aggression theories and previous findings (cf. left panel of Fig. 1; Hypothesis 1). Instead, there was a significant decline of self-reported aggressiveness as a function of time, $F(1,65) = 9.57$, $p = .003$, explaining $\eta^2_p = .13$ of variance; gender and realism did not influence BPAQ scores, $Fs < 1$, also not interactively, $Fs < 2.37$, $ps > .13$, $\eta^2_p < .04$.

3.2. Implicit measure

As a preliminary test, mean Agg-IAT scores differed negatively from zero, such that participants were exactly 100 ms quicker to associate themselves with peaceful rather than with aggressive. Hence, the standardized $D$-score was significantly negative, $D = -0.29$, $F(1,65) = 21.84$, $p < .001$, that is, the standardized magnitude of the latency difference between the critical IAT blocks amounted to 0.29 standard deviations of individual sorting speed.

As regards the repeated measurement, $D$-scores increased significantly from pretest ($D = -0.37$) to posttest ($D = -0.22$), $F(1,65) = 5.50$, $p = .02$, $\eta^2_p = .08$, compatible with the predicted main effect that the self*aggressive association became stronger relative to self*peaceful after ten minutes of VVG play (Hypothesis 2). At the same time, there was no gender main effect, $F(1,65) = 1.26$, $p = .29$, $\eta^2_p = .02$, and the increases from pretest to posttest were equivalent for both genders (no interactions, $F < 1$; cf. right panel of Fig. 1). Contrary to our expectation though, we found no impact of the 2D vs. 3D mode, and the crucial realism*time interaction was absent, $F < 1$ (Hypothesis 3). Without evidence for a moderation effect, we additionally checked by regression analysis whether the effect of realism on Agg-IAT changes was moderated by BPAQ pretest scores, but this was not the case.3

3.3. Flow experiences

The absence of realism effects on Agg-IAT scores can be further illuminated by analyzing the psychological flow experiences during the game. Again, the hypothesized difference between the 2D and 3D mode did not emerge, $F(1,65) = 1.05$, $p = .31$, $\eta^2_p = .02$. This lack of realism effects was qualified by a strong and unexpected interaction of realism with gender, $F(1,65) = 12.76$, $p < .001$, $\eta^2_p = .16$: Men showed the anticipated higher flow in the 3D rather than in the 2D version; unforeseen, however, women experienced less flow in the highly realistic 3D presentation mode and more flow in the more common 2D presentation mode (cf. Fig. 2). The opposite trends not only canceled out the realism main effect, they also tended to carry over to a gender main effect, $F(1,65) = 3.32$, $p = .07$, $\eta^2_p = .05$, indicating that men overall experienced slightly more flow than women.

To develop a better understanding of this realism*gender interaction on flow experiences, we separately inspected the three subscales of the Flow Short Scale, absorption, fluency of performance, and outcome importance. In all three cases, realism had no reliable effects on any of the subscales. If anything, the strongest difference emerged on absorption scores, yet in the direction opposite from what one would expect, $Ms = 4.55$ and 4.12 for 2D and 3D mode, respectively, $F(1,65) = 1.76$, $p = .19$, $\eta^2_p = .03$. Instead, both absorption and fluency—the components aggregated to form flow scores—showed the unexpected realism*gender interaction, $Fs > 9.76$, $ps < .013$, $\eta^2_g > .13$, and it was particularly fluency which showed signs of the interaction carrying over to a gender main effect, $F(1,65) = 6.32$, $p = .01$, $\eta^2_g = .09$. It appears then that, on average, men handled the shooter better. No significant differences were found for concerns during game play, all $Fs < 1$, reflecting the uniqueness of this component as also evident in low correlations with the other flow dimensions.

3.4. Mediation analysis

As regards the hypothesized mediation of realism on the implicit measure by flow, no significant relationship between realism and Agg-IAT change scores (direct path) or between realism and flow (mediating path) existed, absolute $rs < .10$, $ps > .42$ (Hypothesis 4). This failure of two crucial steps in Baron and Kenny’s (1986) approach to mediation analysis rules out any potential causal relationships among the variables and brings the mediation analysis to a halt. In light of the unpredicted realism*gender interaction on flow itself, one would at least hope for a realism*gender interaction on Agg-IAT changes that could hint at mediation among men alone, yet no such interaction existed. In other words, neither manipulated realism (visualization mode) nor experienced realism (flow) was associated with how strongly participants coupled themselves spontaneously with aggressive behavior, and flow was unrelated to Agg-IAT changes, even when taking gender into account.
3.5. Correlational analyses

Given gender effects and the delicate nature between implicit and explicit measures, we additionally analyzed some theoretically derived and empirically demonstrated relationships with one-tailed testing of correlations as effect sizes. Men were more likely to play games, as indicated in daily and weekly hours spent on game play, rs = .18 and .20 (ps = .07 and .05). Although there was no gender difference for BPAQ scores, rs = .08 and .12 (ps = .25 and .16), at pretest and posttest, respectively, there was the typical trend for higher physical aggressiveness among men, r = .29 and .28 (ps = .01). Across genders BPAQ pretest and posttest scores were significantly and positively related to daily and weekly hours of gaming, rs = .22–.25 (ps = .02–.04). The causal nature of the latter relationships cannot be established here; of course they may be bidirectional (Gentile, Swing, Lim, & Khoo, 2012). The relationship between Agg-IAT and BPAQ was positive but weak at pretest, r = .17 (p = .08), and slightly attenuated at posttest, r = .13 (p = .15), replicating the known positive, but low implicit–explicit correlation (cf. Bluemke & Zumbach, 2012).

4. Discussion

4.1. Summary

Using a gender-balanced sample, we analyzed the violence exposure–aggressiveness link and its moderation by game realism. Irrespective of participant gender and game realism we observed a shift of participants’ self-concepts after ten minutes of playing a contemporary VVG. This effect was only established with the help of an implicit measure, the Agg-IAT, not with an explicit measure, the BPAQ, which indicated lower aggressiveness. We expected players in a highly realistic environment to be more vulnerable to changes in aggressiveness than players in a less immersive environment (cf. Barlett & Rodeheffer, 2009; Tamborini et al., 2004), and this effect was expected to be mediated by psychological flow (cf. Klimmt et al., 2010). In spite of reliable variables, the moderation-by-realism and the mediation-via-flow hypotheses were not supported.

Whereas our study suggests that violent content can exert a systematic effect on automatic associations, the exact mechanisms need to be illuminated. The Agg-IAT showed a stronger association between self and aggressive behavior after violent game play in comparison to pretest scores, replicating a previous finding from a violent game that was low in realism and hardly immersive (Bluemke et al., 2010; see also Uhlmann & Swanson, 2004). Our explanation goes beyond previous explanations of VVG effects that often centered on priming hostile affect (e.g., Barlett, Harris, & Baldassaro, 2007) or the accessibility of aggressive scripts (e.g., Bushman, 1998). While we do not deny these basic mechanisms, the media learning mechanism proposed here maintains that playing VVG did not merely invoke access to memory of scripts, but shaped the readiness with which players associated themselves with aggressive and peaceful behavioral options (cf. Klimmt et al., 2009, 2010). As the game environment offered a virtual reality in which one executes (virtual) violent actions, these actions became still related to the self at least temporarily.

4.2. Alternative explanations for Agg-IAT changes

The Agg-IAT change is compatible with the idea of changes of the self-concept (Lemmer et al., 2015). Still, one might suspect that in the absence of a control condition, rather than reflecting a genuine change of self-related associations, the finding could reflect an attenuation artefact that can occur when IATs are repeated in close temporal order (Greenwald et al., 2003). Note that such a decline has not been observed in a previous control condition without any treatment after a five-minute break (Bluemke et al., 2010). In that study the Agg-IAT changed hardly, in absolute terms only by \( \Delta D = 0.04 \) (in standardized terms: Cohen’s \( d = 0.10 \)), whereas the present shift was significant and about three times as strong, \( \Delta D = 0.14 \) (\( d = 0.25 \)).

A bit more difficult to rule out, though, is whether practicing the difficult but fast responses during the VVG boosted the sorting performance in the incompatible and difficult self + aggressive block of the subsequent IAT. If this happened, mere sorting performance of the IAT task would have been affected, not associative structures proper. While at a first glance quick reactions seem equally important for the IAT as for mastering the game (shooting enemies before the player gets hurt), a second look shows that actually a “hide and seek”-strategy is more important than fast responding to navigate the game successfully. Also, the motor behavior focuses on movements controlled jointly by keyboard and mouse, so that a training transfer from the game to the IAT task, which uses two different keys, appears unlikely. Still, we checked the error rates for performance gains and found no systematic reduction from the first to the second IAT. Yet to properly identify a performance artefact one would need an IAT other than on aggressiveness, or an implicit measure unaffected by skill confounds (cf. Teige-Mocigemba, Klauer, & Sherman, 2010).

Last but not least, it is difficult to discern whether the first-person shooter primed participants procedurally so that they took the Agg-IAT from within a virtual first-person perspective. In this case, changes in IAT scores would not have echoed the effect the game had on participants; instead, participants would have continued the study as if they were their own game character, and that role might have determined the test performance and transformed the meaning of IAT scores (cf. Klimmt et al., 2009, 2010). We can imagine only one empirical way out of this question, namely by using a personalized and a role-specific self-concept IAT that differentiate experimentally induced personal from extrapersonal associations (Olson & Fazio, 2004). Apart from not providing error feedback (as there would be no normatively incorrect responses on the personal level), the personalized IAT variant would employ “I am . . .” and “Others are . . .” as category labels instead of the underspecified category labels self and other. This variant should help participants refocus on their own perspective after they have ended the game. Another IAT variant would use “The fighter is . . .” and “The enemies are . . .” as category labels, and these would be compatible with a continued first-person shooter perspective. If it was merely the shooter role that determined the associations in the IAT with underspecified category labels, a role-specific IAT should mimic the outcomes we obtained, whereas the personalized IAT should not. Only if our outcomes reflect a genuine change of participants’ self-concept, the personalized IAT should replicate the pattern of the IAT with underspecified category labels.

When weighing the overall evidence, we deem a genuine temporal change of self-related associations after VVG most likely. Although the violent content colored the self-concept as less peaceful, the 2D and 3D mode did not differ as expected, and the increase in aggressiveness was not mediated by subjectively experienced flow, which raises the question why the highly malleable IAT, which incorporates both state and trait variance (Lemmer et al., 2015) could not distinguish the high from the low realism group.

4.3. Reflection on the sensitivity of Agg-IATs

The Agg-IAT—and also self-reported aggressiveness and flow experiences—failed to reflect the expected differences between the two realism conditions. One possibility is that our
manipulation, despite being effective in altering the flow experiences of men and women, was simply not strong enough to moderate the shift of automatic associations between self and aggressive. This absence of an effect due to 3D shutter-glasses may then underscore previously demonstrated effects for fully immersive head-mounted displays. Yet, the absence of the realism effect on the Agg-IAT stands not only in stark contrast to earlier evidence demonstrating such effect on other measures (Barlett et al., 2007; Krkmar et al., 2011; Persky & Blascovich, 2008; Tamborini et al., 2004), but also to the IAT’s previously demonstrated sensitivity to nuances in game content (Bluemke et al., 2010). In the previous experiment, playing a hardly immersive 2D desktop game affected player’s Agg-IAT scores. Controlling for arousal (heart rate, galvanic skin response), game context (a static forest screen), and motor responses (mouse movements, clicks), the player’s self-associations of aggressiveness changed—depending on the meaning of the virtual actions they took toward the target objects in the game, be it “watering sunflowers”, “eliminating triangles”, or “shooting enemy soldiers”. Five minutes of violent game play already resulted in an Agg-IAT change of $\Delta D = 0.33$ ($d = 0.73$) as compared to $\Delta D = 0.14$ ($d = 0.24$) after ten minutes of realistic game play in the present study. If anything, one would expect the modern shooter, not the simplistic version, to bolster the self-aggressive association.

Part of the explanation for this perplexing finding could be that the visualization modes did not differ psychologically, given that an identical game was presented that already demonstrated high-end realism in the 2D mode and the only difference between the conditions was the stereoscopic vision enabled by shutter-glasses. While such an explanation might be handy to explain the lack of realism effects on Agg-IAT scores, it neglects that the realism was psychologically effective for men (and for women in the opposite way) according to flow experiences. Furthermore, this explanation cannot elucidate why the Agg-IAT change was smaller for the present realistic game than for the previous unrealistic game type.

We suspect that the identical meaning of the symbolic actions in both modes (identical semantics according to the Media Comparison Paradigm) prevented players from construing their self-concepts differently, whereas in our previous study players associated themselves with clearly different behaviors that were either of positive, neutral, or negative valence (different semantics). Stressing the importance of the symbolic meaning can accommodate the fact that the current strengthening of self ≠ aggressive associations was relatively weak despite playing an immersive game. The amount of persecutions, or the degree of violence enacted, may not matter to a player’s Agg-IAT scores once aggressive scripts have been sufficiently activated. In fact, effects of realistic VVG on Agg-IATs may level off after some time, possibly due to individuals’ flow experiences, disconfirming the idea that immersion is a moderator of, let alone a prerequisite for, shaping the association between self and aggressive behavior. The self-concept changed independent of immersion in the game.

From hindsight it seems only natural that it was men who experienced higher flow in the 3D mode, which has been observed at least once before (Nowak, Krcmar, & Farrar, 2006): Male participants reported consuming computer games prior to the experiment more devotedly than females, a well-known trend (Gentile et al., 2004). If we assume that men were also the higher-skilled VVG players (e.g., weaponry, tactics, attention), they may have immersed themselves more readily in, or been absorbed more strongly by, the realistic and vivid 3D effects. At the same time, the unusually high degree of realism may have interfered with relatively untrained female players and hampered their gaming skills. Alternatively, men may have had more prior experiences with 3D games specifically, allowing for quicker adaptation within the game. The gender-effect may then be due to men showing well-learnt, dominant responses under increased physiological arousal in the high realism condition (Barlett & Rodeheffer, 2009; Zajonc, 1965). Another factor could be that some participants experienced flow in the graphic shooter, but perceived the actions of their game character as so self-discrepant that they underreported flow. This might be particularly relevant for women who use severe physical aggression less frequently than men.

It is beyond the scope of this paper to illuminate all the relevant processes underlying the unpredicted interaction effect. Yet the interaction underscores that whatever the outcome of a manipulation is, it can be opposite for differently experienced game players. Recently, Przybilski et al. (2014) showed that lack of competence in handling game controls can increase aggressive affect and cognition, whereas higher flow (due to fluency of performance) may counteract any increases of aggressiveness. If skill discrepancies coincide with gender, then findings are unlikely to replicate across genders, calling for stricter control in media comparison studies. Another caveat is that players’ skills need to be monitored, or varied, to warrant stringent conclusions about typical and atypical psychological flow did not mediate the increase in Agg-IAT scores, and on average the increase was comparable for the 2D and 3D mode. Hence, realism operationalized purely by visualization modes might best be regarded as an independent layer in the Media Comparison Paradigm, which speaks neither to interactivity (syntactics) nor content (semantics). Furthermore, the 2D mode presented a realistic scenario in which spatial perception is the product of moving through a simulated world (Müssler, 2008). Like the 3D stereoscopic vision, the 2D movement among sufficiently realistic objects may have fostered comparable spatial vision, realistic experiences, and involvement with the game. The absence of flow differences could then be due to a psychological ceiling effect.

However, men had higher flow in the 3D than in the 2D mode, and vice versa for women. Again, different flow experiences did not mediate the interindividual changes of Agg-IAT scores, not even among men which showed the expected flow pattern. Whatever Agg-IAT changes occurred from pretest to posttest, they were unrelated to individuals’ flow experiences, disconfirming the idea that immersion is a moderator of, let alone a prerequisite for, shaping the association between self and aggressive behavior. The self-concept changed independent of immersion in the game.

4.4. Implications for mediation-by-flow

Psychological flow did not mediate the increase in Agg-IAT scores, and on average the increase was comparable for the 2D and 3D mode. Hence, realism operationalized purely by visualization modes might best be regarded as an independent layer in the Media Comparison Paradigm, which speaks neither to interactivity (syntactics) nor content (semantics). Furthermore, the 2D mode presented a realistic scenario in which spatial perception is the product of moving through a simulated world (Müssler, 2008). Like the 3D stereoscopic vision, the 2D movement among sufficiently realistic objects may have fostered comparable spatial vision, realistic experiences, and involvement with the game. The absence of flow differences could then be due to a psychological ceiling effect.

However, men had higher flow in the 3D than in the 2D mode, and vice versa for women. Again, different flow experiences did not mediate the increase in Agg-IAT scores, not even among men which showed the expected flow pattern. Whatever Agg-IAT changes occurred from pretest to posttest, they were unrelated to individuals’ flow experiences, disconfirming the idea that immersion is a moderator of, let alone a prerequisite for, shaping the association between self and aggressive behavior. The self-concept changed independent of immersion in the game.
players. It would be premature to discard on the basis of our outcomes the previous evidence on the mediation hypothesis, because realism is likely to impact on other measures of hostile affect and cognition. If one assumes that the realism treatment worked, then the semantic stimulus sorting-procedure called Agg-IAT appears to be rather insensitive to our specific operationalization, because across both realism conditions aggressive scripts were acquired or activated to the same extent.

4.5. Implications for explicit measures

The low positive correlation between implicit and explicit measures that exists in many domains (Greenwald et al., 2009), and specifically in the domain of aggression (Bluemke et al., 2010), surfaced in our study too. Participants’ self-reports at posttest pointed toward lower aggressiveness. As the correlations of the BPAQ scores were not severely biased, one might hold a general cautious-ness shift among all participants responsible for the outcome. As Bender et al. (2013) have observed, transparent measures do not only underestimate effect sizes in VVG studies, but there is evidence for motivated responding due to concerns about admitting socially undesirable aspects and due to defending the social status of the ad-hoc ingroup of gamers.

Furthermore, the order of implicit and explicit measures may have contributed to the reversed effects. Whereas our fixed order prevented Agg-IAT scores to be biased after participants consciously reflected on their aggressiveness (e.g., suppression effects in Glock & Kneer, 2009), the reverse was still possible despite the IAT being a relatively unobtrusive measure. If participants noticed their decreasing performance during the self + aggressive task and become quite alert about what this means in conjunction with VVG, they may have answered subsequent questions more cautiously. Though we have never observed such a strong dissociation before, we have never before used such a realistic shooter either.

We consider it unlikely that our finding reflects a true cathartic effect (Bushman, 2002; Bushman, Baumeister, & Stack, 1999; Ferguson, Olson, Kutner, & Warner, 2014). Yet, BPAQ posttest scores may accurately reflect how participants perceived their aggressiveness after they had just encountered what violent behavioral options exist (many of which they would never consider to use in reality). Accordingly, a shift of the comparison standard could account for this outcome too. Or maybe participants thought they had experienced a cathartic effect, whereas they failed to realize that aggressive options became less self-discrepant at an automatic level (cf. Shiffrin & Schneider, 1977). Again, our findings highlight that explicit measures should be complemented by implicit measures, if only to rule out blatant distortions.

5. Conclusion

Not all the complex learning effects that computer games have are problematic. A balanced discussion will readily reveal that goal-directed games such as “Medal of Honor” (Electronic Arts, 2010) or “Call of Duty” (Infinity Ward/Activision Blizzard, 2011) stimulate at the cognitive level and lead to improved visual acuity (Achtman, Green, & Bavelier, 2008; Green & Bavelier, 2006). They can teach players to collaborate in virtual teams, overcome temporary frustrations, and commit themselves to tasks, as failure appears redeemable with increasing coordination, proficiency, and persistence (Gee, 2007). Solving the tasks is fun; experiencing gratification involves neural reward circuitries (e.g., dopamine release in the cortical network/ventral striatum; Koepf et al., 1998).

Still, the same circuits can reinforce aggressive scripts and have detrimental attitudinal effects on players (e.g., Anderson et al., 2010). It has repeatedly been documented that automatic components of the self-concept of aggressiveness can be affected too. That increasingly realistic games can pose an additional risk factor of violence-exposure effects needs to be taken seriously (Konijn et al., 2007). Implicit measures are a welcome add-on to explicit measures when monitoring positive and negative gaming effects on cognitive dispositions known to be related to aggressive behavior; they tap into spontaneous associations not covered by questionnaires.

Our experiment informs other researchers about a potential boundary condition of the Agg-IAT. The Agg-IAT, in general, is capable of predicting interindividual differences in behavior (Banse et al., 2015) and reflecting long-term influences of video-games, supporting its use as a trait-measure (Uhlmann & Swanson, 2004). However, like many other measurement devices, the Agg-IAT contains a substantial state component on top of an underlying trait (cf. Lemmer et al., 2015). When capitalizing on the state component, as in experimental situations, the Agg-IAT may not be able to distinguish different intensities of violence exposure, especially if highly graphic content is compared across conditions. Future research should inspect the mediating paths between violence-exposure and automatic components of the self-concept in the short and long-run.

References


