Who's afraid of Donkey Kong?

Testing the Stereotype Threat Effect in Video Gaming

Elisabeth Holl[1]

Gary L. Wagener[2]

André Melzer[1]

Author Note

[1]elisabeth.holl@uni.lu, andre.melzer@uni.lu, Media and Experimental Lab, Institute for Health and Behaviour, University of Luxembourg, Esch-sur-Alzette, Luxembourg

[2]lee.wagener@gmail.com, University of Bonn

Abstract

FULL PAPER: In two studies (Study 1: $N = 130$; Study 2: $N = 56$) participants played a video game (*Bejeweled 3*; *SkyChasers*) and were either confronted with a stereotype threat (ST) or not. ST is defined as the risk of confirming a negative stereotype about one's own group and has been investigated in various field, i.a. in gaming. In the first study participants were confronted with the stereotype that women would perform worse in video games than men. In the second study we worked with a reversed stereotype, namely that women would have now outpaced males in some genres of video games. Our results show that performance varies across gender and genre. Although we did not find the hypothesized interaction effect of gender and ST condition in performance, self-reported measures, such as perceived frustration, and moderating variables indicate performance differences both for women and men, but on different psychological dimensions.

Keywords: stereotype threat effect, video games, gender, casual gaming

For a long time, the medium of video games was regarded as dominated by men. This "male space", a world created mostly by and for men, is still ground for stereotypes (Fox & Tang, 2014). But over the last few years, the number of women in gaming has increased indicating that now almost half of the gaming community is female (Entertainment Software Association, 2018). However, and even though the stereotypical view of pale boys playing alone in their basement is no longer rampant, women might still refrain from calling themselves *gamers* (Kowert, Festl, & Quandt, 2014). In this confusion of both gender and gaming related prejudices, the present studies try to shed light on the underlying mechanisms of gaming and gender related stereotypes through the examination of the ST effect.

Stereotype Threat

Originally documented in topics not related to gaming (e.g., mathematics) ST is defined as "being at risk of confirming, as self-characteristic, a negative stereotype about one's own group" (Steele & Aronson, 1995 p. 797). For example, Spencer, Steele and Quinn (1999) showed that after telling participants that a math test reliably indicates gender differences and that women will generally perform worse than men in those tests, women's math performance was significantly worse than those of men. Moreover, women threatened with the gender stereotype also performed significantly worse than women in a control condition who were told that there are no gender differences in math test performance (Spencer et al., 1999). According to the integrated process model of Schmader, Johns, and Forbes (2008), members of the stereotyped group will suffer from a cognitive imbalance between their self-concept and the stereotypical expectations regarding their performance. While trying to cope with the conflict and disproving the stereotype (a) physiological stress directly impairs prefrontal processing, (b) awareness increases for their own performance and internal state, and (c) negative thoughts and emotions are suppressed. All these processes deplete the available amount of cognitive attention and

resources additionally to the actual performance task resulting in a poorer general performance (Schmader et al., 2008). The ST effect was often replicated in studies dealing with gender and performance in science or sports and contributes to the race gap in academics or even differences in cognitive abilities of elderly people (Spencer, Logel, & Davies, 2016). However, other studies were not able to show an ST effect, for example in different cultural settings like China (Shen, Ratan, Cai, & Leavitt, 2016; Tsui, Xu, & Venator, 2011). Furthermore, there has been criticism that statistical distortions and the publication bias are confounded in ST research (Flore & Wicherts, 2015; Stoet & Geary, 2012).

Stereotype Threat and Gaming

Bertozzi (2008) argues that females are confronted with negative stereotypes in gaming every day. Avatars are mostly male and overly masculine and when playing shooter games like *Counter Strike* (Valve Corporation, 2012) female gamers are likely to be confronted with exclusively male opponents. Female gamers may not just play but would feel forced to concurrently disprove a number of stereotypes (Bertozzi, 2008). Consistent with this, women reported more stress and perceived their own skills as lower when they thought they were playing against a man, as opposed to a female in an experimental setup. Being confronted with a male opponent requires females to engage in more executive resources to disprove the stereotype that women would perform worse in video games than men (Vermeulen, Núñez Castellar, & Van Looy, 2014). Kaye and Pennington (2016) compared gaming performance of online-gamers in a 2D side scroller called *Supertux*. Prior to playing, participants in an ST condition received information on research proofing males to be the more competent in gaming, and a related, negative stereotype. As expected, results showed that females performed worse than men even in a control condition. The authors attribute this to internalized negative stereotypes that were active without intentionally inducing an ST. Interestingly, however, when previously threatened

females were told that the negative stereotype would not apply for them because they were experienced and competent gamers, they performed as well as females in the control condition (Kaye & Pennington, 2016). Kaye, Pennington, and McCann (2018) conducted a study with participants who were threatened with the same gender-related stereotype as mentioned above. Additionally, participants either played a female or a male ape in a game called *Kiba and Kumba: High Jump*, a casual platformer. The authors could neither find an effect of the explicit ST condition, nor of the factor of playing a male animal avatar, which might have served as a subtler threat as it could connect to a negative self-perception.

Gaming Performance and Gender

Early studies addressing gender differences in gaming performance found that women performed significantly worse in video games than men (Brown, Hall, Holtzer, Brown, & Brown, 1997). In contrast, more recent studies like Shen and colleagues (2016) were unable to find gender performance differences: females advanced at least as quickly and performed as well as men in massively multiplayer online games. Despite their findings, the authors hypothesized a self-fulfilling-cycle in which women find video games less attractive, are therefore less committed to compete, have then fewer experiences and as result perform worse than men. This performance gap would then again lead to stereotyping, women might even internalize the negative assumption that "gaming is not for women" and the circle starts again from the beginning (Lynch, Tompkins, van Driel, & Fritz, 2016). Women are still rather perceived as "casual gamers" and males as "hardcore gamers" (Paaßen, Morgenroth, & Stratemeyer, 2017).

In summary, the ST effect related to gender and gaming performance was found in some studies, but only under certain circumstances. More specifically, how ST is induced and what stimulus material is used seem to be crucial to whether female participants show impaired gaming performance. Therefore, two experiments were conducted that each aim at highlighting a

different aspect of the ST effect in relation to gaming and gender. In the first experiment participants were confronted either with the stereotype that males outperform females in games or a neutral condition (i.e., no gender differences in performance).[1] We wanted to investigate whether women would indeed perform worse in a video game if the ST "women are less competent players than men" is activated compared to men and women of the control condition (H1: interaction of gender and ST). After considering the results of this first study, the stereotype was flipped in the second experiment, and participants were told that women have now outpaced men in certain gaming genres. To further investigate the contradictory results of Study 1, a reversed ST and a new game from a slightly different genre was chosen. We wanted to test whether men will perform worse in a video game if the ST "women are now more competent players than men in certain genres" is activated compared to women and men of a control condition (H2: interaction of gender and ST).

General Method

Using a 2 (participant gender) x 2 (ST) between-subjects design in both experiments, a gender-related ST was either activated (ST condition) or not (neutral condition) before participants played a video game. Additionally, participants provided details on demographics, gaming experience, gamer identification, gender identification and perception and other personality factors. ST was manipulated by means of a bogus internet report that either supported the respective stereotype or not. In-game score was used as performance measure and participants indicated how they perceived the gaming episode as well as the report.

Participants. Participants in both studies were recruited at XXXXX. In the first study it was possible to participate in a lottery (1x50€ and 4x25€). In the second study participants were

[1] Parts of the data based on a smaller sample have been presented at XXXXX.

rewarded with a 7€ voucher. Student participants were additionally rewarded with course credit. The first sample (N = 130) consisted of 50 male (38.5%) and 80 female (61.5%) participants, who were 18-49 years old (M = 22.65, SD = 5.22). Participants played video games for an average of M = 5.19 hours per week (SD = 8.88) and considered themselves to be gamers on a medium level (M = 2.06, SD = 1.10; 4-point Likert scale). The sample of the second study (N = 56) consisted of 25 male (45%) and 31 female (55%) participants, who were 19-57 years old (M = 26.40, SD = 6.88). Participants played video games for an average of M = 4.14 hours per week (SD = 6.37) and considered themselves to be gamers on a medium level (M = 2.28, SD = 1.26; 4-point Likert scale).

Materials. ST was induced using two different, fictitious web articles. In Study 1, the articles were titled "Women's skill still behind men's" (ST condition), and "There is no gender difference in skill" (neutral condition). In the ST condition text, it was explicitly stated that women would play significantly less, and that they would show significantly poorer performance than male gamers. In contrast, the neutral text stated that women would show performances, which were at least on par with male gamers. In Study 2, the stereotypical article was titled "Female gamers outpace male", whereas the neutral article displayed "Both gender on a par". The ST condition text stated that women would have taken the lead in genres like casual games and platformer both in prevalence and performance. The neutral text stated that women and men would play many genres equally often and equally well. All texts were carefully put together: real articles that covered similar topics were studied beforehand to create reports in a similar style and structure.

In Study 1, *Bejeweled 3* (PopCap Games, 2010) was chosen for playing to avoid confounds in performance resulting from gender differences in motivation, game preference, or skill. In this puzzle game rows of three same colored gems must be formed by swapping

neighboring gems. Pairs of three or more stones then disappear, the player is awarded with points, and new stones are added at the top. An equivalent game has been used in prior research, as it has simple controls, but is still perceived as challenging (Przybylski, Weinstein, Murayama, Lynch, & Ryan, 2012). In Study 2, the platformer *SkyChasers* (Lucky Kat Studios, 2015) was chosen to investigate the effect in a different genre. Participants had to steer a rocket ship through various 2D levels, while collecting coins and avoiding hazards. A bar indicated the battery of the ship, which unloaded during the flight, but could be restored through collecting coins and docking on stations. The game was chosen because it required very simple controls, and both the game environment and the avatar appeared gender neutral.

Measures. In both experiments, participants provided information on demographics, gaming habits, perception of the gaming episode and the perceived impact of the article on their playing. As main performance measures the in-game score (points, coins, etc.) was noted. In Study 1, participants additionally rated 19 items of the Competitiveness Index and its revised version (R-CI; Harris & Houston, 2010; Cronbach's $\alpha = .91$). Study 2 included 7 items of the Gaming Playing Skill Scale (GPSS; Bracken & Skalski, 2006; Cronbach's $\alpha = .92$), 14 items of the Social Identification Scale (Leach et al., 2008; Cronbach's $\alpha = .84$) with gender as the in-group and 9 items assessing the perceived credibility of the article (Nauroth, Gollwitzer, Bender, & Rothmund, 2014; Cronbach's $\alpha = .77$).

Procedure. After participants entered the lab and signed the consent form, they answered all questionnaires that were not directly related to the game or the article (e.g., demographics, gaming habits and skills, competitiveness, etc.). Participants were randomly assigned to either the condition (ST vs. neutral). Next, they read the article according to their respective condition. After indicating that they understood the game controls provided on an instruction sheet, they played for 15 minutes (Study 1) and 20 minutes (Study 2). Participants then answered the final

questions concerning the evaluation of the game and the article they had read. Finally, participants were debriefed and rewarded. The entire procedure took between 30 and 45 minutes.

Study 1

Results. We found significant gender differences for gaming experience (Welch's $F(1, 60.37) = 19.92$, $p < .001$, $d = .95$) with males playing more and having a stronger identification with being a gamer (Welch's $F(1, 82.8) = 43.1$, $p < .001$, $d = .24$). Furthermore, males enjoyed competition more ($F(1, 128) = 49.07$, $p < .001$, $d = 1.26$). For an overview of the means and standard deviations see Table 1. However, these significant gender differences did not interact with the ST manipulation.

Table 1

Means and standard deviations for selected outcome variables of the total sample and across gender.

Measure	Total		Male		Female	
	M	*SD*	*M*	*SD*	*M*	*SD*
Playing hours/week	5.19	8.88	9.90	11.48	2.25	4.91
Identification with being a gamer	2.06	1.10	2.80	1.08	1.61	0.83
Enjoyment of Competition	3.17	0.83	3.72	0.72	2.83	0.70

For the primary analysis, a two-way ANOVA on in-game score was calculated with gender and ST as between-subject factors. There was only a significant main effect for gender ($F(2, 126) = 4.99$, $p = .03$, $\eta_p^2 = .04$) showing that overall females had higher in-game scores than males. Neither the main effect for ST condition nor the interaction effect reached significance ($Fs \leq 2.73$, $p \geq .10$) (see Figure 1).

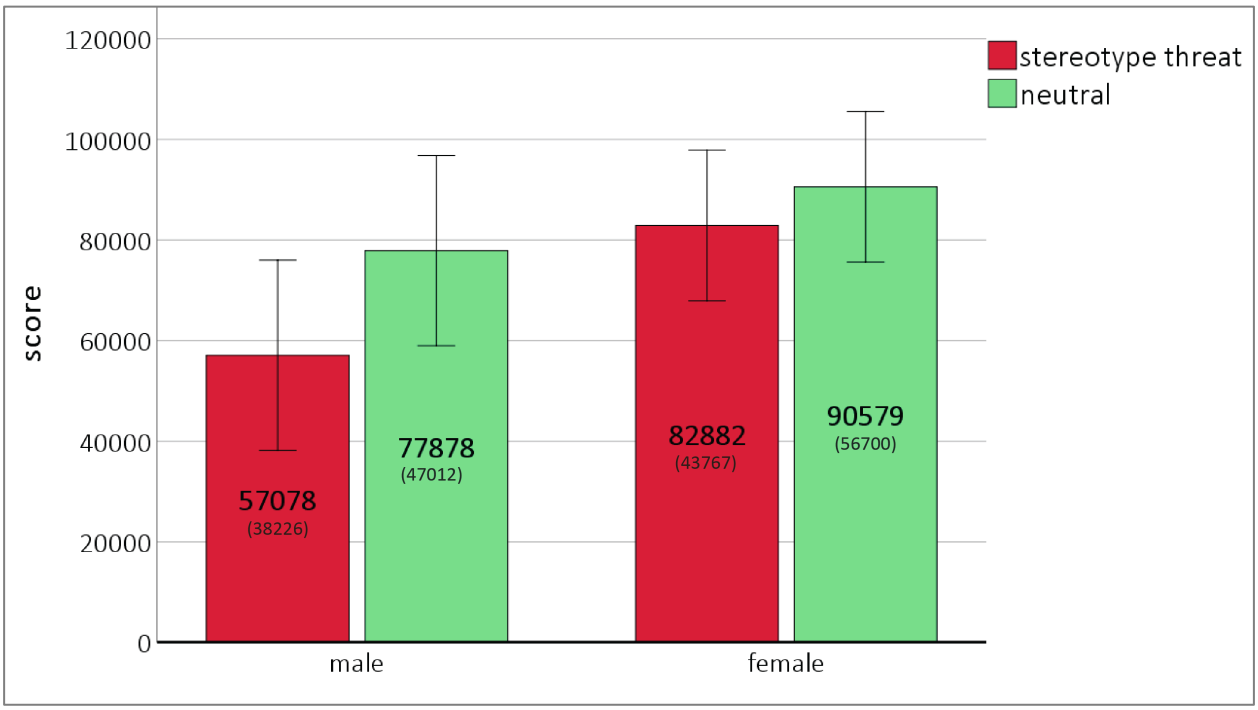

Figure 1. Mean scores (and standard deviations) across conditions.

Even with gaming experience as covariate, the main effect of gender was still significant ($F(1, 126) = 8.09$, $p = .005$, $\eta_p^2 = .06$). The main effect of ST was marginally significant ($F(1, 126) = 2.95$, $p = .088$, $\eta_p^2 = .02$), but the interaction was not ($p = .359$). With gamer identification as covariate, the main effect of gender was still significant ($F(1, 126) = 12.10$, $p = .001$, $\eta_p^2 = .09$) but neither the main effect of ST nor the interaction was significant ($ps \leq .136$). However, gamer identification was a significant predictor of in-game score ($F(1, 124) = 8.53$, $p = .004$, $\eta_p^2 = .06$).

Furthermore, there was a significant interaction effect between gender and ST condition on game knowledge prior to the study, $F(1, 126) = 5.72$, $p = .018$, $\eta_p^2 = .04$. An analysis of simple effects showed that there was a significant difference between conditions for males, $F(1, 126) = 6.96$, $p = .009$, $\eta_p^2 = .05$, with males in the neutral condition ($M = 2.80$; $SD = 1.30$) having significantly higher familiarity with the game *Bejeweled* prior to the study than males in the ST condition ($M = 2.00$; $SD = 1.00$). Additionally, there was a significant gender difference in the

neutral condition, $F(1, 126) = 4.05$, $p = .046$, $\eta_p^2 = .03$, with males ($M = 2.80$; $SD = 1.29$) having significantly higher familiarity with the game *Bejeweled* prior to the study than females ($M = 2.25$; $SD = .93$).

Looking at how participants perceived the article, there was a significant interaction effect for how frustrated they felt after reading ($F(1, 126) = 5.15$, $p = .025$, $\eta_p^2 = .04$; see Figure 2). Reading the ST article, females were more frustrated than males, and for the neutral article it was the other way around. However, it must be stated that the assumption of variance homogeneity was violated for this ANOVA (Levene's Test, $F = 8.76$, $p < .001$), therefore the results must be used cautiously. An analysis of simple effects showed that only for the ST condition, females ($M = 1.80$; $SD = .99$) felt significantly more frustrated after reading the article than males ($M = 1.36$; $SD = 70$), $F(1, 126) = 5.28$, $p = .023$, $\eta_p^2 = .04$. Furthermore, females felt significantly more frustrated after reading the article in the ST condition (M = 1.80; SD = .99) than in the neutral condition ($M = 1.23$; $SD = .48$), $F(1, 126) = 11.71$, $p = .001$, $\eta_p^2 = .09$.

There were also significant differences regarding the article perception (ST vs. neutral): The neutral article ($M = 3.36$, $SD = 2.91$) was rated more credible than the ST article ($M = 2.91$, $SD = 1.05$), (Welch's $F(1, 111.02) = 7.99$, $p = .006$ $d = -.5$), but the ST article ($M = 1.63$, $SD = .91$) invoked more frustration than the neutral article ($M = 1.29$, $SD = .58$), (Welch's $F(1, 108.42) = 6.39$, $p = .013$, $d = .44$).

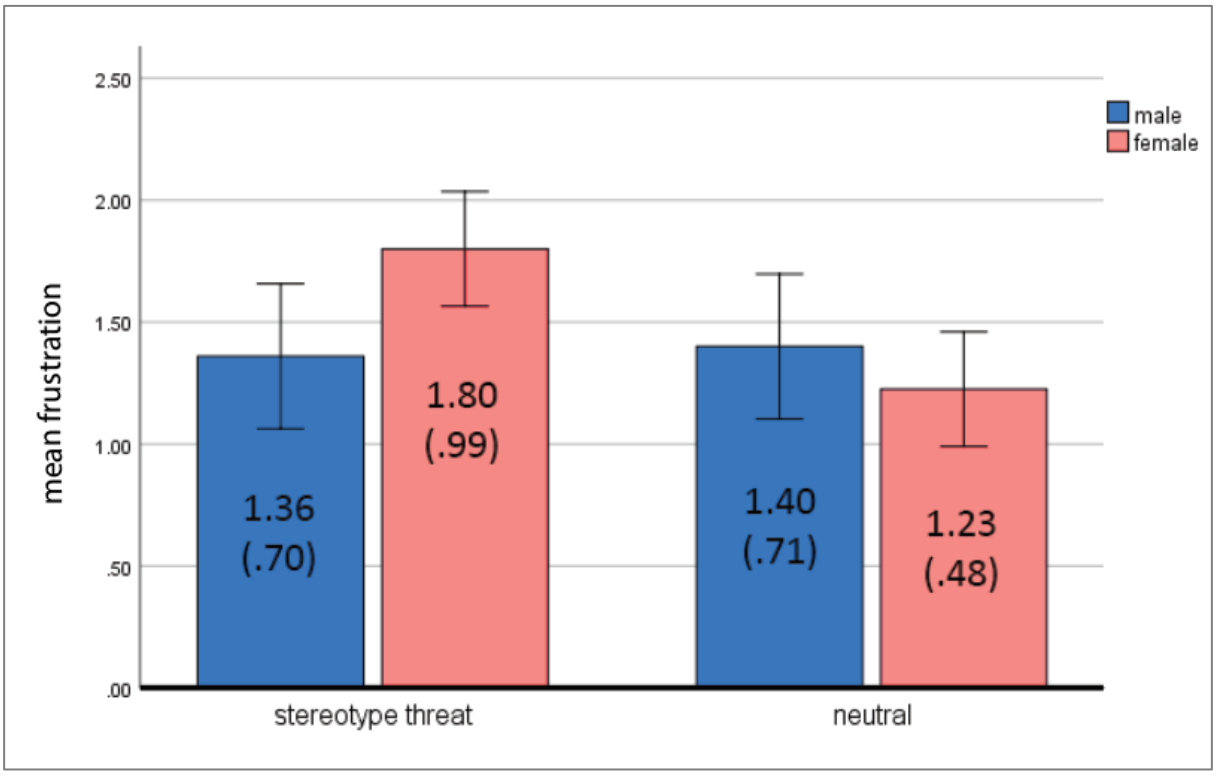

Figure 2. Interaction of gender and stereotype condition on article frustration.

Discussion. Unfortunately, the present findings provided no support for our hypothesis (H1). Neither the main effect for ST nor the interaction between gender and ST reached the level of significance. Only the main effect of gender on in-game score was significant, indicating that regardless of condition, females scored higher than males. The same result (i.e., females outperforming males in video games) was already shown by Shen et al. (2016). In our study, however, it is unclear why males underperformed, since they had significantly higher gaming experience and identified stronger with being a gamer. The additional analyses revealed an interesting picture. Although the expected interaction was not found in gaming performance, women—but not men—were cognitively affected in the predicted way. After reading the report in the ST condition, they felt more frustrated than after reading the neutral report. In contrast, and as expected, men's level of frustration was not affected by the report. This finding suggests a

possible dissociation between cognitive processing and actual behavior. Even though men and women share similar preferences for casual puzzle games and males even indicated having more knowledge about the game *Bejeweled* in the neutral condition, the subgenre of "match 3 games" is overly dominated by female players (Yee, 2017). We may speculate that generally females have higher expertise in playing puzzle games, helping female participants to compensate for their frustration. Expertise (e.g., domain knowledge) is known to compensate for adverse cognitive processing (Morrow, Leirer, Altiteri, & Fitzsimmons, 1994). As a consequence, and in order to avoid a skill-based advantage between male and female participants, we decided to replace the game in Study 2 (see Materials).

Study 2

Results. Although both men and women indicated that they spend the same amount of time on video games per week (Welch's $F(1, 27.80) = 6.04$, $p = .020$, $d = .72$), males identified stronger with being a gamer (Welch's $F(1, 45.42) = 8.14$, $p = .006$, $d = .79$) and scored higher on the GPSS ($F(1, 54) = 19.40$, $p < .001$, $d = 1.18$). In contrast, females showed higher identification with their gender ($F(1, 54) = 6.31$, $p = .015$, $d = -0.68$). However, these significant gender differences did not interact with the ST manipulation. For an overview of the means and standard deviations see Table 2.

Table 2

Means and standard deviations for selected outcome variables of the total sample and across gender.

Measure	Total		Male		Female	
	M	*SD*	*M*	*SD*	*M*	*SD*
Playing hours/week	4.14	6.37	6.56	8.55	2.19	2.68
Identification with being a gamer	2.29	1.26	2.80	1.32	1.87	1.06
GPSS	2.77	0.94	3.30	0.90	2.34	0.73
Gender Identification	3.33	0.64	3.11	0.60	3.52	0.63

Note: GPSS = Game Playing Skill Scale (Bracken & Skalski, 2006).

For the primary analysis, a two-way ANOVA on in-game score was calculated to examine the effects of gender, ST, and their interaction. Results revealed that there was only a significant main effect for gender ($F(1, 52) = 7.29$, $p = .009$, $\eta_p^2 = .12$) showing higher in-game scores for males. Neither the main effect for ST condition nor the interaction effect reached significance ($Fs \leq .25$, $p \geq .617$) (see Figure 5).

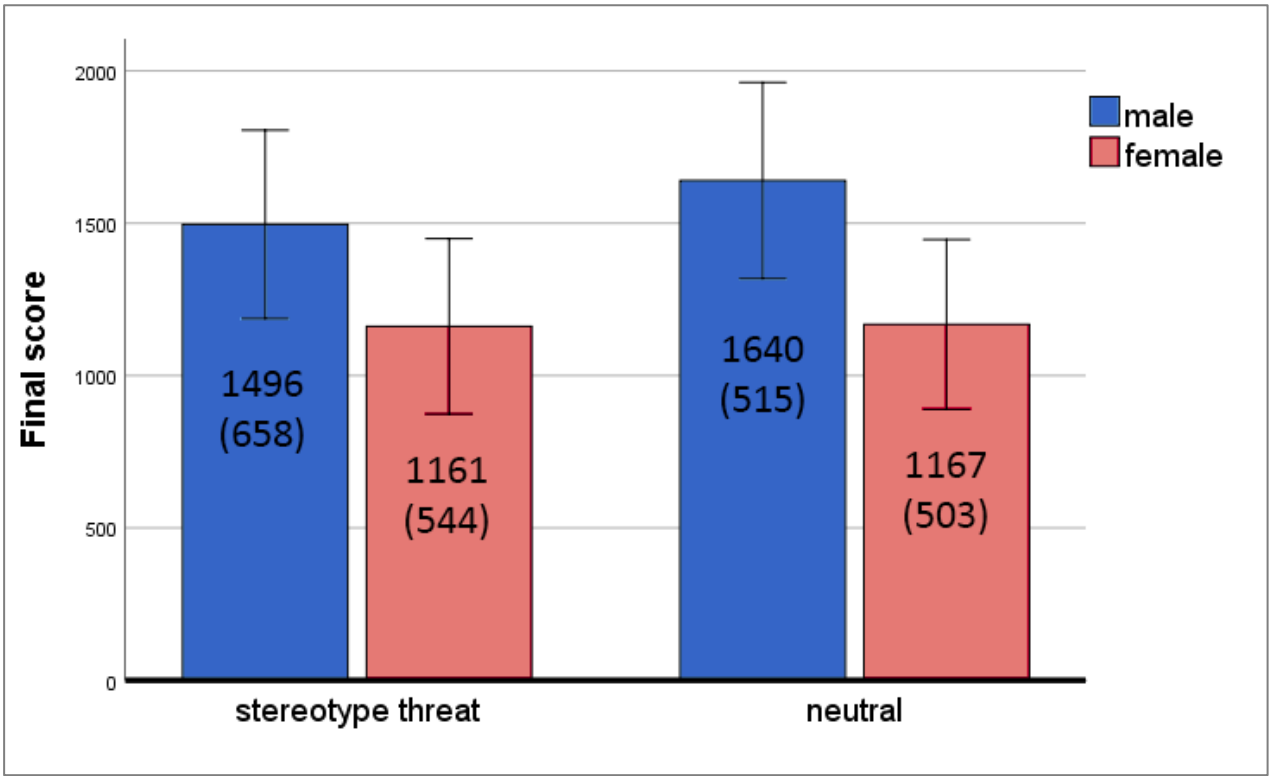

Figure 3. Mean scores (and standard deviations) across conditions.

With gamer identification as covariate, the main effect of gender on score was no longer significant ($F(1, 51) = 2.47$, $p = .122$, $\eta_p^2 = .05$), but gamer identification significantly predicted score ($F(1, 51) = 9.06$, $p = .004$, $\eta_p^2 = .15$). When including gaming experience as a covariate, the main effect of gender on score became marginally significant ($F(1, 51) = 3.26$, $p = .077$, $\eta_p^2 = .06$), and gaming experience significantly predicted score ($F(1, 51) = 6.22$, $p = .016$, $\eta_p^2 = .11$). Along all conditions no difference in article credibility was found ($Fs \leq 2.53$, $p \geq .118$) and

frustration both after reading the article ($Fs \leq .09$, $p \geq .768$) and after playing the game ($Fs \leq 1.15$, $p \geq .288$) did not vary significantly.

Looking at group differences for article perception, there was a significant interaction between gender and ST condition for the item „reading the article improves my motivation to apply myself, $F(1, 52) = 9.93$, $p = .030$, $\eta_p^2 = .09$. An analysis of simple effects showed that there was only a significant effect for females, $F(1, 52) = 5.88$, $p = .019$, $\eta_p^2 = .10$, with females in the control condition having significantly higher motivation improvement ($M = 3.50$; $SD = 1.32$) than females in the ST condition ($M = 2.27$; $SD = 1.58$). Furthermore, there was a significant difference between genders only in the neutral condition, $F(1, 52) = 7.70$, $p = .008$, $\eta_p^2 = .13$, with females ($M = 3.50$; $SD = 1.32$) having significantly higher motivation improvement than males ($M = 2.00$; $SD = 1.48$).

There was a significant interaction effect between gender an ST condition for the item "I felt confirmed after reading the article", ($F(1, 52) = 11.60$, $p = .001$, $\eta_p^2 = .18$). An analysis of simple effects showed that there was a significant effect for males, $F(1, 52) = 5.01$, $p = .029$, $\eta_p^2 = .09$, with males in the ST condition ($M = 2.69$; $SD = 1.11$) feeling more confirmation after reading the article than males in the neutral condition ($M = 1.67$; $SD = 1.23$). There was also a significant effect for females, $F(1, 52) = 6.79$, $p = .012$, $\eta_p^2 = .12$, with females in the neutral condition ($M = 2.94$; $SD = 1.29$) feeling significantly more confirmation after reading the article than females in the ST condition ($M = 1.87$; $SD = .92$). Additionally, there was only a significant difference between genders in the neutral condition $F(1, 52) = 8.46$, $p = .005$, $\eta_p^2 = .14$, with females ($M = 2.94$; $SD = 1.29$) feeling significantly higher confirmation after reading the article than males ($M = 1.67$; $SD = 1.23$).

Discussion. The results from Study 2 revealed a significant effect for gender on score, showing greater performance for males. However, if one takes the gaming experience and gamer

identification into account, this effect is no longer significant. Apparently, the observed gender difference in performance may be attributed, at least partly, to higher gaming experience and higher gamer identification of males. Furthermore, females in the control condition felt more confirmed and had higher motivation to apply themselves after reading the neutral article. Also, unlike females in Study 1, males were not affected in their level of frustration by the ST article and even felt more confirmed after reading the ST article that stated that females were more competent players. Even though males identified stronger with being a gamer, it seemes as though this gamer identification was not affected by the stereotype thread manipulation as hypothesized.

General Discussion

Two experimental studies tested the effect of ST on playing video games. Contrary to expectations, neither telling participants that women are worse players than men (Study 1), nor that women are better players than men (reverse ST, Study 2) had a detrimental effect on gaming performance. However, manipulation check in Study 1 (i.e., how frustrated participants felt after reading the ST report) indicated that presenting stereotype threatening information affected participants' cognition in the ST condition. But why did it fail to affect the behavioral measure of in-game performance?

Besides the ST effect two other effects could explain our results, the stereotype boost effect and, or the phenomenon of "choking under pressure". The stereotype boost occurs whenever one's own group is confronted with a positive stereotype, which could lead to higher performance for the nonstereotyped group (Shih, Ambady, Richeson, Fujita, & Gray, 2002). In our case, telling women that they are known to perform worse than men (ST condition) could have served as a positive stereotype for males. However, there was no such performance boost for male participants in the ST conditions of our studies. But inconsistencies in research on this

effect suggest that also the opposite could happen. When a group is confronted with a positive stereotype, they might perceive a heightened performance pressure, because they anticipate that others will have higher expectations on their performance compared to the stereotyped group (Beilock & Carr, 2005). Perhaps males were suffering from this "choking under pressure" effect. After reading the ST articles they felt performance pressure and were therefore eager to outperform women to prove the stereotype but failed.

Smith and Johnson (2006) argued that domain identification with the stereotype relevant group is an important moderator. They found that men, who were confronted with a gender related stereotype ("men are superior to women in mathematics"; Smith & Johnson, 2006, p. 54) and were low in math domain identification, scored lower in a math test that had been introduced as a well-proven indicator of male advantage compared to a condition were the stereotype was nullified. Although the addition of gamer identification as a covariate did not change our main results in Study 2, it significantly predicted the resulting score. Gaming performance seems to be rather a matter of domain specific interest and identification with the relevant topic. It might be promising to assess both of these factors for different gaming genres respectively, as this could explain, why we found different performance outcomes for males and females in Study 1 (match 3 game) and Study 2 (platformer).

Besides looking at the domain specific identification, in our case identification with gaming, considering the identification with one's own gender is important. Investigating identification with social subgroups (gamers, males, females) might be a promising route to disentangle previous mixed findings: Females' player motivation and play style may vary as a function of their gender identity rather than their biological sex (Poels, De Cock, & Malliet, 2012). More importantly, although women were aware of the negative stereotypes regarding their gender and gaming performance, they might not have integrated it into their own social

identity (Kaye et al., 2018). Apparently, knowing that the stereotype exists may affect behavior (i.e. gaming performance) only when it is closely related to social identity. Again, this would be in line with the speculation mentioned earlier that informing female participants about the alleged disadvantage may have cognitive but not behavioral consequences.

Although the present experiments provide important reference points for further research. However, a few limitations should be noted. The assumption of homogeneity of variance could not be met for some analyses, which might be due to high variation across the gender conditions. Even though we tried to control for these differences, it still limits the interpretability of the main analysis. As already mentioned above, the choice of the target game for Study 1 could have been problematic. Even though for casual puzzle games gender preferences are nearly equal, the subgenre of "match 3 games" is overly dominated by females (Yee, 2017). Therefore, a different game was chosen for Study 2. Again, we observed a gender effect in gaming performance: males significantly outperformed now females independent of ST condition. However, this effect was no longer significant when controlling for gaming experience and gamer identification. Apparently, more experienced players and those, who identify stronger as a gamer show greater gaming performance. For this subgenre (platformer) gamer identification is a greater predictor for performance than our independent variables.

In summary, our results corroborate that there are still gender differences in performance across genres. But the stereotype that males are more skillful players *per se* is not tenable. Although we were unable to find the hypothesized interaction effect of gender and ST condition in both studies, moderators and measures such as frustration or confirmation to apply oneself indicated that ST mechanisms were still triggered, at least cognitively. Thereby it was demonstrated that cognitive processing and behavior not always have a uniform effect (Morrow et al., 1994). Future research should therefore put effort in investigating the effects of ST in a

more balanced way, integrating different game genres, different ways of inducing ST, and especially alternating the targeted group of the stereotype itself (males vs. females vs. gamers).

References

Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and "choking under pressure" in math. *Psychological Science*, *16*(2), 101–105.

Bertozzi, E. (2008). "You play like a girl!": cross-gender competition and the uneven playing field. *Convergence: The International Journal of Research into New Media Technologies*, *14*(4), 473–487. https://doi.org/10.1177/1354856508094667

Bracken, C. C., & Skalski, P. (2006). Presence and video games: The impact of image quality and skill level.

Brown, R. M., Hall, L. R., Holtzer, R., Brown, S. L., & Brown, N. L. (1997). Gender and video game performance. *Sex Roles*, *36*(11–12), 793–812. https://doi.org/10.1023/A:1025631307585

Entertainment Software Association. (2018). *Essential facts about the computer and video game industry*. Retrieved from http://www.theesa.com/wp-content/uploads/2018/05/EF2018_FINAL.pdf

Flore, P. C., & Wicherts, J. M. (2015). Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis. *Journal of School Psychology*, *53*(1), 25–44. https://doi.org/10.1016/j.jsp.2014.10.002

Fox, J., & Tang, W. Y. (2014). Sexism in online video games: The role of conformity to masculine norms and social dominance orientation. *Computers in Human Behavior*, *33*, 314–320. https://doi.org/10.1016/j.chb.2013.07.014

Harris, P. B., & Houston, J. M. (2010). A reliability analysis of the revised competitiveness index. *Psychological Reports*, *106*(3), 870–874. https://doi.org/10.2466/pr0.106.3.870-874

Kaye, L. K., & Pennington, C. R. (2016). “Girls can’t play”: The effects of stereotype threat on females’ gaming performance. *Computers in Human Behavior*, *59*, 202–209. https://doi.org/10.1016/j.chb.2016.02.020

Kaye, L. K., Pennington, C. R., & McCann, J. J. (2018). Do casual gaming environments evoke stereotype threat? Examining the effects of explicit priming and avatar gender. *Computers in Human Behavior*, *78*, 142–150. https://doi.org/10.1016/j.chb.2017.09.031

Kowert, R., Festl, R., & Quandt, T. (2014). Unpopular, overweight, and socially inept: reconsidering the stereotype of online gamers. *Cyberpsychology, Behavior, and Social Networking*, *17*(3), 141–146. https://doi.org/10.1089/cyber.2013.0118

Leach, C. W., van Zomeren, M., Zebel, S., Vliek, M. L. W., Pennekamp, S. F., Doosje, B., … Spears, R. (2008). Group-level self-definition and self-investment: A hierarchical (multicomponent) model of in-group identification. *Journal of Personality and Social Psychology*, *95*(1), 144–165. https://doi.org/10.1037/0022-3514.95.1.144

Lucky Kat Studios. (2015). *SkyChasers*. Den Haag: Lucky Kat Studios.

Lynch, T., Tompkins, J. E., van Driel, I. I., & Fritz, N. (2016). Sexy, strong, and secondary: a content analysis of female characters in video games across 31 years: female game characters across 31 years. *Journal of Communication*, *66*(4), 564–584. https://doi.org/10.1111/jcom.12237

Morrow, D., Leirer, V., Altiteri, P., & Fitzsimmons, C. (1994). When expertise reduces age differences in performance. *Psychology and Aging*, *9*(1), 134.

Nauroth, P., Gollwitzer, M., Bender, J., & Rothmund, T. (2014). Gamers against science: The case of the violent video games debate. *European Journal of Social Psychology*, *44*(2), 104–116. https://doi.org/10.1002/ejsp.1998

Paaßen, B., Morgenroth, T., & Stratemeyer, M. (2017). What is a true gamer? The male gamer stereotype and the marginalization of women in video game culture. *Sex Roles*, *76*(7/8), 421–435. https://doi.org/10.1007/s11199-016-0678-y

Poels, K., De Cock, N., & Malliet, S. (2012). The Female Player Does Not Exist: Gender Identity Relates to Differences in Player Motivations and Play Styles. *Cyberpsychology, Behavior, and Social Networking*, *15*(11), 634–638. https://doi.org/10.1089/cyber.2012.0164

PopCap Games. (2010). *Bejeweled 3*. Seattle, WA.

Przybylski, A. K., Weinstein, N., Murayama, K., Lynch, M. F., & Ryan, R. M. (2012). The ideal self at play: the appeal of video games that let you be all you can be. *Psychological Science*, *23*(1), 69–76. https://doi.org/10.1177/0956797611418676

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, *115*(2), 336–356. https://doi.org/10.1037/0033-295X.115.2.336

Shen, C., Ratan, R., Cai, Y. D., & Leavitt, A. (2016). Do men advance faster than women? Debunking the gender performance gap in two massively multiplayer online games. *Journal of Computer-Mediated Communication*, *21*(4), 312–329. https://doi.org/10.1111/jcc4.12159

Shih, M., Ambady, N., Richeson, J. A., Fujita, K., & Gray, H. M. (2002). Stereotype performance boosts: The impact of self-relevance and the manner of stereotype activation. *Journal of Personality and Social Psychology*, *83*(3), 638–647. https://doi.org/10.1037//0022-3514.83.3.638

Smith, J. L., & Johnson, C. S. (2006). A stereotype boost or choking under pressure? Positive gender stereotypes and men who are low in domain identification. *Basic and Applied Social Psychology*, *28*(1), 51–63. https://doi.org/10.1207/s15324834basp2801_5

Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology*, *67*(1), 415–437. https://doi.org/10.1146/annurev-psych-073115-103235

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*(1), 4–28. https://doi.org/10.1006/jesp.1998.1373

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797.

Stoet, G., & Geary, D. C. (2012). Can stereotype threat explain the gender gap in mathematics performance and achievement? *Review of General Psychology*, *16*(1), 93–102. https://doi.org/10.1037/a0026617

Tsui, M., Xu, X. Y., & Venator, E. (2011). Gender, stereotype threat and mathematics test scores. In *Journal of Social Sciences*. Citeseer.

Valve Corporation. (2012). *Counterstrike: Global Offensive*. Bellevue, WA.

Vermeulen, L., Núñez Castellar, E., & Van Looy, J. (2014). Challenging the other: exploring the role of opponent gender in digital game competition for female players. *Cyberpsychology, Behavior And Social Networking*, *17*(5), 303–309. https://doi.org/10.1089/cyber.2013.0331

Yee, N. (2017, January 19). Beyond 50/50: breaking down the percentage of female gamers by genre. Retrieved October 8, 2018, from https://quanticfoundry.com/2017/01/19/female-gamers-by-genre/