# Do Gender Stereotypes Affect Gaming Performance? Testing the Stereotype Threat Effect in Video Games

**Article** *in* Game Studies · August 2024

CITATIONS
0

READS
20

**3 authors:**

Elisabeth Holl
University of Wuerzburg
**39** PUBLICATIONS **83** CITATIONS
SEE PROFILE

Gary Lee Wagener
**12** PUBLICATIONS **19** CITATIONS
SEE PROFILE

André Melzer
University of Luxembourg
**98** PUBLICATIONS **914** CITATIONS
SEE PROFILE

Game Studies
the international journal of computer game research
volume 24 issue 2
July 2024
ISSN:1604-7982



# Do Gender Stereotypes Affect Gaming Performance? Testing the Stereotype Threat Effect in Video Games

*by* Elisabeth Holl, Gary Lee Wagener, André Melzer

## Abstract

Despite a growing body of statistics highlighting comparable gender prevalence and performance in gaming, the cliché persists that women are gaming amateurs or only "casual" players. Does the lower performance of women who are observed in certain video game contexts result from a stereotype threat effect (Steele & Aronson, 1995)? To test this, we investigated video game performance in two studies ($N_1$ = 130; $N_2$ = 139). In Study 1, participants were confronted with the stereotype that women would perform worse in video games than men. In Study 2, we explored a reversed stereotype, namely that women would outpace men in some video game genres. We found no evidence of an effect of stereotype threat on gaming performance in either study. However, performance varied across gender and game genre: female participants generally outperformed males in Study 1, but male participants outperformed females in Study 2. Although we found no stereotype threat effect, perceived frustration indicated expected gender differences in Study 1: after reading the stereotypical article, female participants felt significantly more frustrated than males. This suggests a subtler stereotype threat, affecting cognitive and motivational rather than behavioral outcomes. The present findings add to the current literature, which assumes that the stereotype threat effect is not universal but occurs only under certain conditions.

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

**Elisabeth Holl**

Elisabeth Holl is a post-doctoral researcher at the Institute Human-Computer-Media at the University of Wuerzburg in Germany. She is a member of the Department of Media and Business Communication currently investigating effects of entertainment media, such as film and gaming. Other research interests are immersion and psychophysiological measuring. She received her PhD in 2022 from the University of Luxembourg with a thesis focusing on moral decision-making in video games.

Contact information: elisabeth.holl at uni-wuerzburg.de

**Gary Lee Wagener**

Gary Lee Wagener is a doctoral candidate at the Department of Behavioral and Cognitive Sciences of the University of Luxembourg. He is a member of the Institute of Health and Behavior. For his doctoral thesis he is currently investigating the relation of different gaming scenarios on gamer stress levels and mood using a multimethod approach of self-report and psychophysiological parameters. Furthermore, he is interested in investigating so-called "dark" personality traits and aggression in gaming.

Contact information: gary.wagener at uni.lu

**André Melzer**

André Melzer is an

## Introduction

For a long time, video games were considered a male domain. Even though recent polls show that almost half of gamers are female (Entertainment Software Association, 2024), the gaming landscape is still a breeding ground for gender stereotypes (Fox & Tang, 2014). For example, women are still typecast as "casual gamers," whereas males are considered "hardcore gamers" (Paaßen et al., 2017). Although gamers themselves might deny the validity of such stereotypical attributions, non-gamers appear to endorse them (Kowert et al., 2012). Furthermore, upholding gender stereotypes might lead to marginalization or discrimination against female players. Stereotyping like this can lead to negative psychological outcomes, including feelings of not fitting in or feeling isolated. As findings indicate that gaming culture and its related stereotypes are changing (Paaßen et al., 2017), the two studies analyzed here try to shed light on underlying mechanisms of gender-related stereotypes in gaming (males are better players [Study 1] and females are better casual players [Study 2]) by examining the stereotype threat (ST) effect.

### Stereotype Threat Effect

Per definition, ST is "being at risk of confirming… a negative

Assistant Professor in Psychology at the University of Luxembourg. He is affiliated with the Department of Behavioral and Cognitive Sciences and a member of the Institute of Health and Behavior. His research focuses on interactive media use and its effects across the lifespan, but also gender portrayals in modern media and their effect on adolescents and young adults. He also published research on social psychological mechanisms, and implicit versus explicit memory.

Contact information: andre.melzer at uni.lu

stereotype about one's own group" (Steele & Aronson, 1995, p. 797). In a study on math performance, Spencer et al. (1999) told female participants that they would generally perform worse compared to men, which indeed led to poorer math performance of females compared to males. In contrast, women who were told that no gender differences in math exists performed better than threatened females (Spencer et al., 1999). According to the integrated process model of Schmader et al. (2008), members of a stereotyped group suffer from a cognitive imbalance between their self-concept and stereotypical expectations. While trying to cope with the conflict and disprove the stereotype, such members experience (a) physiological stress that impairs prefrontal processing; (b) an increased awareness of their own performance and internal state; (c) attempt to repress negative thoughts and emotions. These processes deplete the available amount of cognitive attention and resources in addition to the actual task, resulting in poorer general performance. In addition to reduced performance, threatened individuals may suffer from long-term effects such as self-doubt, chronic stress or avoidance behavior (Davies et al., 2005; Schmader et al., 2008). ST has been shown to explain gender performance gaps in athletics (Hively & El-Alayli, 2014), science (Good et al., 2010), leadership (self-perceived and rated performance; Hoyt & Blascovich, 2007) and gaming (Kaye & Pennington, 2016). However, there has also been criticism on statistical distortions and publication bias in ST research that might lead to an underestimation of null findings (Flore & Wicherts, 2015; Stafford, 2018; Stoet & Geary, 2012).

**Gender and Gaming Performance**

Several gaming studies have not only shown a gender performance gap in favor of male players, but also that females expected themselves to perform poorer than men (e.g., Brown et al., 1997; Hopp & Fisher, 2017). In contrast, Shen et al. (2016) found no gender performance gap and Eden et al. (2010) found no relation between skill and perceived masculinity of online players. Moreover, female *League of Legends* (Riot Games, 2009) players leveled up at the same rate, despite being less confident in their skills than male players were (Ratan et al., 2015). In addition, Ratan et al. (2020) found that performance gaps dissipated when controlling for play time. To explain performance gaps, Shen et al. (2016) hypothesized a self-fulfilling cycle in which women find video games less attractive, and therefore become less committed to compete, have fewer experiences and as a result perform worse than men. This performance gap would again lead to stereotyping, probably causing women to further internalize the negative assumption that "gaming is not for women." Consequently, the cycle starts again (Lynch et al., 2016).

**Stereotype Threat and Gaming Performance**

Regardless of actual performance, Bertozzi (2008) argues that female players are frequently confronted with negative stereotypes. Most avatars look overly masculine, which already has a subtle threatening effect (Bertozzi, 2008; Kaye & Pennington, 2016). When, for example, playing shooters, females are likely to be confronted with exclusively male opponents, feeling forced to concurrently disprove stereotypes on aggressiveness or competitiveness. Accordingly, female players reported more stress and perceived their own skills as inferior when led to believe they were playing against a man as opposed to a woman (Bertozzi, 2008; Vermeulen et al., 2014).

However, studies that specifically address an ST effect in gaming find mixed results. For example, female performance in a first-person shooter decreased when confronted with a negative stereotype (Fordham et al., 2020). Furthermore, Kaye and Pennington (2016) compared gaming performance in a 2D side scroller across four conditions (threatened females, attenuated threatened females, female controls and male controls). Prior to playing, women in the ST condition received research information stating that males are more competent in gaming. In contrast, females in the attenuated threat condition were told that the negative stereotype would not apply to them because they were experienced and competent gamers. As

expected, threatened females performed worse than non-threatened male controls. However, females confronted with an attenuated threat performed as well as both males and females in control conditions (Kaye & Pennington, 2016). In contrast, Kaye et al. (2018) found no ST effect for either an explicit (stereotypical text) or implicit threat ([fe]male avatar). Similarly, Pennington et al. (2018) found no ST effect for different types of threats (that is, personal threat vs. group threat). In summary, the ST effect related to gender and gaming performance was found only in some studies and only under certain circumstances or for certain genres (such as shooters; Fordham et al., 2020).

It should be stated that the above-mentioned studies looked at a wide range of gaming genres, such as tactical shooters, online-role playing games or side scrollers. However, both studies in this current article focus on casual games for several reasons: 1) the high and increasing prevalence of causal or non-core gaming in the overall population (Chess & Paul, 2019); 2) an equal prevalence of male and female players (Kaye et al., 2018); and 3) straightforward controls (Przybylski et al., 2012) that allow to factor out additional demand through physically controlling the game.

**Other Factors Influencing Gaming Performance**

One explanation for the hitherto inconsistent results may be a failure to consider other factors as additional or even stronger predictors of gaming performance. Proneness to stereotypes is of course influenced by one's own identification with the stereotyped group (Schmader, 2002) like, gender or gaming identification. Additionally, pre-existing gender attitudes and beliefs may compete with momentarily presented ST information. Furthermore, the above-mentioned study from Vermeulen et al. (2014) concluded that self-perceived skill and gaming experience are influenced not only by gender dynamics, but also by trait competitiveness: low-competitive females appear more prone to use gender as a predictor for gaming performance. This illustrates that self-efficacy is critical for being susceptible to gender stereotypes (Bussey & Bandura, 1999). In the context of ST and leadership performance, higher levels of self-efficacy attenuated players' stress and promoted their coping (Burnette et al., 2010; Hoyt & Blascovich, 2007). Regarding gaming, Klimmt and Hartmann (2005) suggest self-efficacy as one crucial motivational factor, possibly determining the successful coping with gaming challenges. However, self-efficacy may also backfire on gaming performance. Vancouver et al. (2002) showed a negative relationship between self-efficacy and performance in an analytical game due to overconfidence errors. Lastly, evidence suggests that achievement motivation correlates with performance, which has already been investigated in educational ST studies (Thoman et al., 2013). However, this direct motivational measure has not yet been tested in terms of gaming performance.

**Current Studies**

To resolve prior mixed findings on (casual) gaming performance and ST and to contribute to a research field that currently suffers from possible publication bias and replication problems, we conducted two experiments, each highlighting different aspects of the ST effect. In the first experiment (Study 1), participants were confronted with either the stereotype that males would outperform females in games (ST condition), or with the claim that there would be no gender difference in performance (neutral condition). The purpose was to investigate whether women's gaming performance is worse when the ST -- "women are less successful players than men" -- is activated.

Our first hypothesis (hereafter, H1) is: *Female gamers threatened prior to playing will show poorer performance compared to non-threatened female participants and male participants in general.*

To the best of our knowledge, no study of game performance to date has integrated the individual factors mentioned above to assess their respective predictive quality for performance. Thus, we asked the following research question (hereafter, RQ1): *Do other factors besides*

*gender and ST confrontation (such as gamer or gender identification, gender attitudes, achievement motivation, competitiveness and self-efficacy) predict video game performance?*

To answer H1 and RQ1, we conducted Study 1. Data analyses and interpretation did not support a ST effect in Study 1, where females outperformed males regardless of condition (see below). We then tested a reversed ST (that is threatening males) in Study 2. In general, the ST effect rarely tests for negative male stereotypes (Koenig & Eagly, 2005). To our knowledge, no other gaming study has attempted to threaten the self-perception of male gamers, thus Study 2 offers a novel contribution by testing the ST, "women are now more competent players than men at casual games," on male players. Therefore, we offer a second research question (hereafter, RQ2): *Do male participants threatened prior to playing show poorer performance compared to non-threatened male participants and female participants in general (interaction effect of gender and ST on gaming performance)*?

## Study 1

### Participants

A priori power analysis indicated that a sample of 128 participants was required for our ANOVA calculations and 72 for our linear regression analyses. Participants stemmed from a convenience sample recruited on campus of the University of Luxembourg and were tested at the Media and Experimental Laboratory (MExLab). From overall $N_1$ = 130, 80 participants identified themselves as female (61.5%) and 50 as male (38.5%). Participants were mostly students (80.0%), but also employees (15.4%), pupils (2.3%), or other (2.3%) and between 18 and 49 years old ($M$ = 22.65, $SD$ = 5.22). On average, participants played for $M$ = 5.19 hours/week ($SD$ = 8.88) and regarded themselves as gamer on a medium level ($M$ = 2.06, $SD$ = 1.10; 4-point scale).

### Materials

Prior to playing, all participants read fictitious web articles that referred to fabricated scientific evidence and graphical gender prevalence. We deliberately chose this implicit and covert option to limit the risk of participants guessing the purpose of our study. We based this strategy on Hoyt and Blascovich (2007), who used pre-constructed material with statistics to induce a gender-specific ST on leadership beliefs. We believed a more overt approach (such as directly pointing the stereotype to the threatened group; Kaye & Pennington, 2016) would be problematic if participants were familiar with the ST concept. Articles were titled "Women's skill still behind men's" (ST condition) and "There is no gender difference in skill" (neutral condition). The ST condition text 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 no performance gap exists between female and male gamers anymore. We put these texts together carefully, studying real web and magazine articles that covered similar topics beforehand to create fictitious reports in similar style and structure.

We chose *Bejeweled 3* (PopCap Games, 2010) as the target game. *Bejeweled 3* is a puzzle game in which the player must form rows of three identical gems by swapping neighboring gemstones. Players are rewarded with points for doing so, then pairs disappear from the screen as new stones are added to the top. An equivalent game has been used in prior research, as it requires simple controls (point and click) but is still perceived as challenging (Przybylski et al., 2012). Furthermore, the game is free of a (possible gendered) avatar that could lead to confounding threatening effects (cf. Kaye et al., 2018) and has a rather gender-neutral gaming environment (i.e., use of diverse colors). Screenshots of stimulus games and fabricated articles can be found in the Supplementary Material.

### Measures

The main measure of performance was in-game score (points scored in *Bejeweled 3*). However, participants also provided information on demographics and gaming related items (e.g., gaming experience, i.e., gaming hours per week). Participants additionally rated several standardized self-report scales that are listed in Table 1.

| **Measure** | **Scale/Items** | **Source** | **Internal consistency** |
|---|---|---|---|
| *Study 1*: | | | |
| **Gender identification** | Positive/ Negative Sex Role Inventory (PN-SRI); 24 items in two subscales: feminity and masculinity | Berger & Krahé (2013) | α = .76 (for both subscales) |
| **Gender attitudes** | Gender attitudes concerning both cognitive and physical capabilities; 6 items | Behm-Morawitz & Mastro (2009) | α = .68 |
| **Competitiveness** | Revised Competitiveness Index (R-CI); 14 items in two subscales: contentiousness and enjoyment of competition | Harris & Houston (2010) | α = .91 (contentiousness), α = .76 (enjoyment of competition) |
| **Fear of failure/ Hope of success** | German Achievement Motivation Scale (AMS); 10 items in two subscales: fear of failure and hope of success | Dahme et al., (1993) Engeser (2005) | α = .82 (fear of failure), α = .75 (hope of success) |
| **Self-efficacy** | German Items on general self-efficacy; 10 items | Schwarzer & Jerusalem (1999) | α = .85 |
| *Study 2:* | | | |
| **Gaming skill** | Game Playing Skill Scale (GPSS); 7 items | Bracken & Skalski (2006) | α = .90 |

| | | | |
|---|---|---|---|
| **Gender identification** | Social Identification Scale (SIS); 14 items with gender as the in-group | Leach et al. (2008) | α = .82 |
| **Article credibility** | Evaluation of a research area; 10 items | Nauroth et al. (2014) | α = .80 |

Table 1. Standardized measures for Study 1 and 2. Note: PN-SRI values are summarized in one score with high values indicating a strong (fe)male identification and low values a rather androgynous gender role identification (Rydberg Sterner et al., 2018). Furthermore, the R CI subscales were averaged into one competitiveness score.

Lastly, participants provided single item responses on how they believed the article shaped their experience, along with their perception of the gaming episode (for example, "The game frustrated me").

**Procedure**

After entering the lab and giving informed consent, participants answered computerized questionnaires (for example, demographics, gaming habits and skills, personality traits, etc.). Participants were randomly assigned to conditions (ST vs. neutral) and read the respective fabricated article for five minutes. The instructor then handed out written instructions on how to control the game and participants played the game for 15 minutes. Next, participants evaluated the game and the falsified article using a computer-based questionnaire. Finally, participants were debriefed and rewarded (i.e., course credit and participation in a raffle with 1x50€ and 4x25€). Each participation took 30 to 45 minutes.

**Results**

To test our hypothesis (the ST effect on female participants), we calculated a two-way ANOVA on in-game score (points scored in *Bejeweled 3*) with gender and ST as between-subject factors. A significant main effect for gender, $F(2, 126) = 4.99$, $p = .03$, $\eta_p^2 = .04$, indicated that overall females ($M = 86730.00$; $SD = 50475.72$) had higher scores than males ($M = 67478.00$; $SD = 43688.05$). Contrary to our hypothesis, neither the main effect for ST condition nor the interaction effect of gender and ST condition reached significance ($Fs \leq 2.73$, $p \geq .10$, see Figure 1).

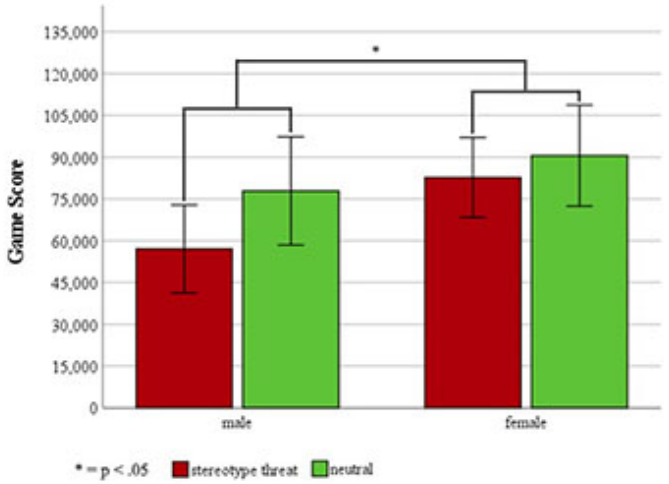


Figure 1. Mean performance measures (game scores) or male and female participants in Study 1. Note: Red (dark grey) columns

represent the ST condition, green (light grey) columns represent the neutral condition. Error bars indicate 95% CI.

To explore relevant characteristics of the gender subgroups, we carried out additional comparative analyses, finding significant gender differences for gaming experience. Overall, the male participants played more often, Welch's [1] $F(1, 60.37) = 19.92$, $p < .001$, $d = .95$ and had greater identification as gamers, Welch's $F(1, 82.8) = 43.1$, $p < .001$, $d = .24$. Furthermore, males showed greater enjoyment for competition, $F(1, 128) = 49.07$, $p < .001$, $d = 1.26$. For an overview of means and standard deviations see Table 2. However, none of these significant gender differences interacted with our ST manipulation.

| **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 |

Table 2. Means and standard deviations for selected outcome variables of the total sample and across gender for Study 1.

Looking closer at participants' perception of the articles, we found a significant interaction effect for frustration after reading, $F(1, 126) = 5.15$, $p = .025$, $\eta_p^2 = .04$ (see Figure 2). Reading the ST article, females ($M = 1.80$, $SD = 0.99$) felt significantly more frustrated than males ($M = 1.36$, $SD = 0.70$). However, for the neutral article this pattern was reversed (females: $M = 1.23$, $SD = 0.48$; males: $M = 1.40$, $SD = 0.71$). Nevertheless, it should be stated that the assumption of variance homogeneity was violated (Levene's Test, $F = 8.76$, $p < .001$) and therefore results must be interpreted with caution. An analysis of simple effects showed that 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 ($M = 1.23$; $SD = .48$), $F(1, 126) = 11.71$, $p = .001$, $\eta_p^2 = .09$.

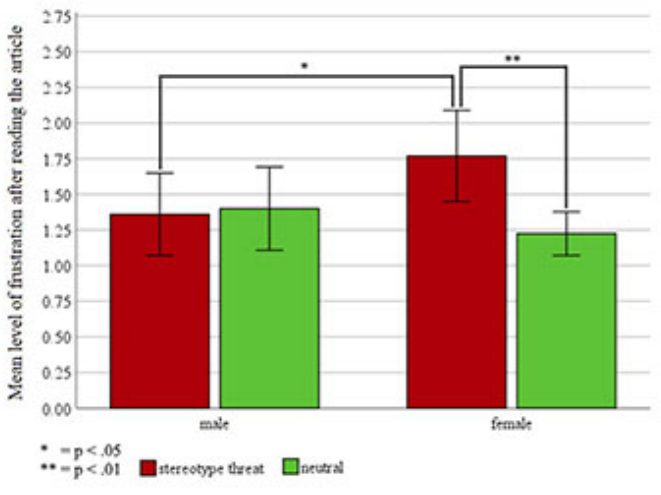

Figure 2. Mean level of frustration after reading the fabricated article in Study 1. Note: Red (dark grey) columns represent the ST condition, green (light grey) columns represent the neutral condition. Error bars indicate 95% CI.

We conducted a multiple linear regression analysis (with forced entry) with gaming performance as dependent variable to test our first research question (i.e., statistically testing the predictive value of further factors, such as identification, attitudes and achievement traits). Unfortunately, two participants had to be excluded from the following analyses due to missing data, leaving a sample of $N$ = 128. The model explained a small but significant amount of variance, $F(9, 118) = 2.00$, $p = .045$, $R^2 = .13$, $R^2_{Adj} = .07$. Table 3 shows that only gender and gamer identification significantly contributed to the model. Being female and having a higher identification with being a gamer both predicted better performance when playing *Bejeweled*.

| | ***B*** | ***SE B*** | **β** | ***p*** |
|---|---|---|---|---|
| **Constant** | - .597 | .188 | | **.002** |
| **Gender[a]** | .757 | .235 | .371 | **.002** |
| **Gamer identification** | .297 | .104 | .297 | **.005** |
| **ST condition[b]** | .267 | .176 | .134 | .13 |
| **Fear of failure** | -.114 | .113 | -.114 | .31 |
| **Self-efficacy** | -.083 | .111 | -.083 | .46 |
| **Competitiveness** | -.080 | .105 | -.080 | .45 |
| **Gender identification** | -.070 | .093 | -.070 | .45 |
| **Hope of success** | .036 | .108 | .036 | .74 |
| **Gender attitudes** | .023 | .092 | .023 | .81 |

Table 3. Regression table for factors influencing gaming performance (Study 1). Note: N = 128, Dependent variable: Score. a 0 = male, 1 = female; b 0 = Stereotype threat, 1 = neutral.

## Discussion

Study 1 provided no support for our hypothesis. Only the main effect of gender on in-game score was significant, indicating that females scored higher than males regardless of condition. This pattern remained stable even after integrating several covariates (for example, gaming experience). This is especially interesting, as our comparative analysis of gender groups revealed significantly higher gaming experience, greater identification with being a gamer and more prior experience with the game *Bejeweled* for male participants. Furthermore, regression analyses revealed that only gender and gamer identification were meaningful predictors of performance.

However, none of our additional player characteristics (for example, competitiveness, achievement motivation) predicted performance. Moreover, it is worth noting that the two significant predictors (gender and gamer identification) turned out to be rather opposed in our case: although higher identification with being a gamer predicted better performance (that is, a higher score), which in our sample was predominantly the case for male participants, being female turned out to be an even better predictor. Despite many beneficial predictors (for example, higher gaming identification) for male players, they were still clearly outplayed by females. Although our hypothesis was not confirmed, women, and not men, were cognitively affected in the initially predicted way: after reading the ST article, female participants felt more frustrated than after reading the neutral report. In contrast, men's level of frustration was not affected by the report. This finding suggests a possible dissociation between cognitive processing and actual behavior. Furthermore, except for gamer identification, other influencing factors did not play a role regarding gaming performance. Thus, most of these factors were replaced or dropped in Study 2.

## Study 2

### Participants

A priori sample size calculation (i.e., at least 128 for ANOVA calculations) and recruitment was identical to Study 1. Participants again stemmed from a convenience sample recruited on campus of the University of Luxembourg and were tested at the Media and Experimental Laboratory (MExLab). In Study 2 ($N_2$ = 139), 87 participants identified themselves as female (62.6%) and 49 as male (35.3%) [2]. Participants were mostly students (80.6%), but also employees (11.5%), pupils (3.6%), or other (4.3%) and between 16 and 57 years old ($M$ = 23.77, $SD$ = 6.02). On average, participants played for $M$ = 3.88 hours/week ($SD$ = 6.56) and considered themselves to be gamers on a medium level ($M$ = 2.08, $SD$ = 1.29; 4-point scale).

## Materials

The article in the ST condition read "Female gamers outpace male," whereas the neutral version was titled "Both gender on a par." The ST condition text stated that women would have taken the lead in genres like casual games and platformers both in prevalence and performance. The neutral text stated that women and men would play many genres equally often and equally well.

The casual platformer *SkyChasers* (Lucky Kat Studios, 2015) was chosen to investigate the effect in a slightly different but still casual game genre (i.e., so-called platformers). Participants had to steer a rocket ship-like paper box through 2D levels using two bumpers of an *Xbox* controller, while collecting coins and avoiding hazards. The ship's battery unloaded while flying but could be restored through collecting coins and docking at stations. The game won the category of best casual game at the 2016 Dutch Game Awards. We chose this game as it requires simple controls and both the game environment and avatar appearance are gender neutral.

### Results

The main variable of interest was again in-game score (coins collected in *SkyChasers*). In addition to demographical information, we included several standardized scales (see Table 1). The overall procedure was identical to Study 1. However, participants played 20 minutes and were rewarded with a 7€ voucher and optional course credit.

### Results

To test our second exploratory RQ (that is, ST effect for male participants), we calculated a two-way ANOVA on in-game score (collected coins in *SkyChasers*). There was one significant main effect for gender, $F(1, 132) = 15.82$, $p \leq .001$, $\eta_p^2 = .11$, showing higher in-game scores for males ($M$ = 1583.88, $SD$ = 584.64) compared to

female participants ($M$ = 1216.13, $SD$ = 477.25). Neither the main effect for ST condition nor the interaction effect reached significance, $Fs \leq 1.14$, $p \geq .29$ (see Figure 3). However, the main effect of gender disappeared when controlling for gaming experience and gamer identification, $F(1, 130) = 3.23$, $p = .075$, $\eta_p^2 = .03$. The covariate "gamer identification" was significantly related to gaming performance, $F(1, 130) = 8.65$, $p = .004$, $\eta_p^2 = .06$.

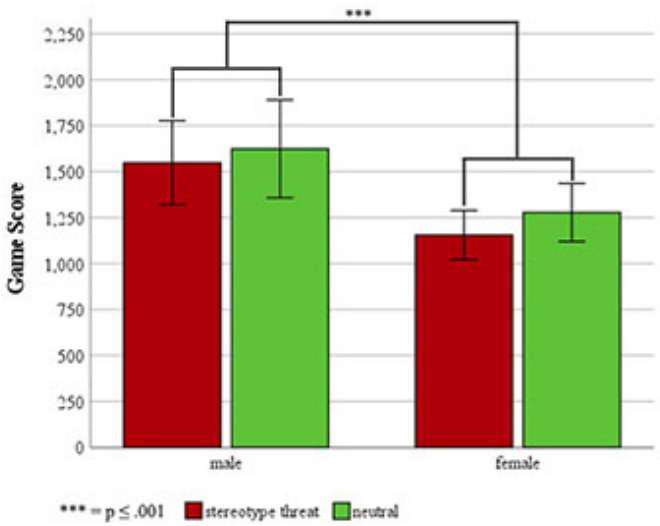


Figure 3. Mean performance measure in Study 2 (game scores). Note: Red (dark grey) columns represent the ST condition, green (light grey) columns represent the neutral condition. Error bars indicate 95% CI.

Additional comparative analyses [3] were carried out on conditional subgroups. Like in Study 1, male participants spent significantly more time on video games per week than female participants (Welch's $F[1, 62.79] = 15.46$, $p < .001$, $\eta_p^2 = .14$). Compared to female participants, they identified also more strongly as gamers (Welch's $F[1, 82.65] = 23.61$, $p < .001$, $\eta_p^2 = .17$) and had greater self-reported skill ($F[1, 134] = 42.60$, $p < .001$, $\eta_p^2 = .24$). In contrast, females showed comparably greater identification with their gender, $F(1, 134) = 32.27$, $p < .001$, $\eta_p^2 = .19$. However, none of these gender differences interacted with our ST manipulation. For an overview of the means and standard deviations, see Table 4. Comparing ST and neutral condition article credibility was rated equally, $F(1, 134) = 0.46$, $p = .50$. Also, frustration after reading the article ($Fs \leq .62$, $p \geq .433$) and after playing the game ($Fs \leq 1.40$, $p \geq .240$) did not vary significantly across the four conditions.

| **Measure** | **Total** | | **Male** | | **Female** | |
|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| **Playing hours/week** | 3.88 | 6.61 | 7.14 | 8.46 | 2.05 | 4.38 |
| **Identification with being a gamer** | 2.07 | 1.27 | 2.76 | 1.33 | 1.68 | 1.06 |
| **GPSS** | 2.70 | 0.93 | 3.31 | 0.85 | 2.36 | 0.80 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Gender Identification** | 3.40 | 0.63 | 3.03 | 0.56 | 3.60 | 0.57 |

Table 4. Means and standard deviations for selected outcome variables of the total sample (N = 136) and across gender for Study 2. Note: GPSS = Game Playing Skill Scale (Bracken & Skalski, 2006). Values can slightly differ from details in the Participants section, as non-binary participants are not considered here.

To test the influence of other variables in addition to the conditions, a (forced entry) multiple regression was conducted with performance (game score) as the dependent variable and several predictors (see Table 5). Again, non-binary data was transformed using *z*-standardization. All three gaming related measures (gamer identification, gaming hours per week and GPSS score) showed high and significant intercorrelations ($r \geq .58$, $p < .001$). Therefore, we decided to reduce the number of predictors and included only the GPSS score because it had the highest correlation coefficient with our outcome variable ($r = .47$, $p < .001$) and to avoid multicollinearity. The resulting model significantly explained slightly more than 25 percent of the variance, $F(5, 130) = 8.52$, $p < .001$, $R^2 = .25$. However, only the GPSS score contributed significantly to the regression model, as can be seen in Table 5.

| | ***B*** | ***SE B*** | **β** | ***p*** |
|---|---|---|---|---|
| **Constant** | .146 | .401 | | .72 |
| **GPSS** | .411 | .088 | .411 | **< .001** |
| **Article credibility** | .100 | .077 | .100 | .19 |
| **Gender[a]** | -.229 | .198 | -.110 | .25 |
| **Gender identification** | -.040 | .085 | -.040 | .64 |
| **ST condition[b]** | .155 | .153 | .078 | .31 |

Table 5. Regression table for factors influencing gaming performance (Study 2). Note: N = 136, Dependent variable: Gaming performance (score). GPSS = Game Playing Skill Scale (Bracken & Skalski, 2006). a 1 = male, 2 = female; b ST condition: 1 = Stereotype threat, 2 = neutral.

**Discussion**

The results from Study 2 revealed a significant gender effect on score, showing better performance for male participants. However, taking gaming experience and identification as a gamer into account, the effect disappears. More importantly, however, we found no confirmation for RQ2 as the interaction of gender and ST was non-significant. Furthermore, level of frustration was constant across all conditions. Regression analyses on other factors predicting gaming performance revealed that only greater self-reported gaming skills were associated with a greater score. Additionally, conditions failed to predict game score significantly.

## General Discussion

The two experimental studies tested the effect of gender ST on gaming performance. Contrary to our expectations, neither participants confronted with the stereotype that men are better players than women (Study 1), nor participants that read a text stating that women are better players than men (Study 2) were affected in their gaming performance. Thus, our findings do not support a ST effect of gender in gaming. However, Study 1 indicated that stereotype threatening information indeed affected female participants' cognition (in the ST condition) but failed to affect subsequent behavioral measures of in-game performance. We encourage future studies to highlight the interrelations of ST, frustration and performance in, for example, mediation models.

Other related effects, such as "choking under pressure," have explained inconsistencies in ST research (Rosenthal & Crisp, 2007). This explanation could also apply to our results: when a group is confronted with a positive stereotype, participants may suffer from increased pressure, as they anticipate that others have higher expectations toward their performance compared to that of the stereotyped group (Beilock & Carr, 2005). It is possible that the male participants in Study 1 suffered from this "choking under pressure" effect: after reading the ST articles, they may have felt pressured to outperform women, but failed to prove the (positive) stereotype. Our results could furthermore be explained by individual reception of the two contrasting effects, which could be based on varying sensitivity for gender stereotypes, self-concepts, or personality characteristics. We cannot rule out that our null findings are caused by different effects cancelling each other out. Thus, future studies should include post-hoc items that explicitly address the perceived pressure in participants.

Another explanation could be that there is simply no ST effect (Stoet & Geary, 2012; Tsui et al., 2011). Although gender performance gaps and gender related stereotypes are still prevalent and might influence cognitive factors, such as frustrations, they do not have the potential to alter behavioral outcomes (anymore). In recent years, the gender share in gaming has equalized (Entertainment Software Association, 2021; 2024) and the number of non-stereotypical, diverse personas increases, potentially eliminating ST effects (Elizaga & Markman, 2008; Taylor et al., 2011). Therefore, effects of gender stereotypes may have disappeared, especially among younger generations. Thus, future research should also emphasize inter-generational differences in gaming stereotypes.

In general, it may be promising to decipher earlier mixed findings by examining to what extent study participants identify with social subgroups (gamers, males, females, etc.). According to Poels et al. (2012) females' motivation and play style may vary as a function of their gender identity rather than their biological sex. Similarly, Rosenthal and Crisp (2007) found that math performance was only affected negatively by stereotypes when group membership (here: male gender and math student) was important to participants' self-perception. More importantly, while women might be aware of negative stereotypes regarding their gender and gaming performance, they may not have integrated these stereotypes into their own social identity (Kaye et al., 2018). Apparently, knowing that a stereotype exists will only affect behavior (that is, gaming performance) when it is closely related to social identity. Again, this would be in line with the speculation (based on the results of Study 1) that informing female participants about the alleged disadvantage has cognitive but not behavioral consequences. It also emphasizes that interindividual differences may be responsible for more variance in performance than gender itself.

Although both studies found a main effect of gender on performance (i.e. significantly better performance for one of the two genders), our additional analyses in both studies indicated that gender alone is a poor predictor. Smith and Johnson (2006) previously argued that domain identification with the stereotype-relevant group (for example, gamers) 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 compared to a condition in which the stereotype was nullified. Although the addition of gamer identification as a covariate did not change the main results of Study 1, it significantly predicted the resulting score. In Study 2, self-reported gaming skills served as the best predictor for performance. Gaming performance therefore seems to be rather a matter of domain-specific interest and identification with the relevant topic. It may be promising to assess both factors for different game genres respectively, as this may explain why we found different performance results for male and female participants in Study 1 (match 3 game) and Study 2 (platformer).

While gender preferences for casual games in general are nearly equal, the subgenre of match 3 games (that is, *Bejeweled 3* in Study 1) is overly dominated by females (Yee, 2017). This probably contributed to how female participants outperformed males in Study 1. After changing the target game in Study 2 to a platformer, we observed a main gender effect on game performance -- now males generally outperformed females independent of ST. However, this effect was no longer significant when controlling for gaming experience and gamer identification: more experienced players and those who identify themselves more strongly as gamers showed greater game performance. In summary, our findings lend further credence to the claim that generalization of ST effects must be met with caution, even across relatively similar stimulus. Therefore, future studies should control for gaming experience and identification as a gamer, as well as familiarity with different input devices (mouse and keyboard, gamepad, etc.), to rule out performance gaps due to motor abilities.

Although the present experiments provide important reference points for further research, some limitations need to be addressed. Homogeneity of variance assumptions were not met for some analyses, which might be due to the great variability across gender conditions. Even though we tried to control for these differences, it still limits the interpretability of our main analyses. Additionally, there was no true control condition in either experiment due to limited recruitment capacities. This limits the generalizability of the present results, as it remains unclear whether both conditions could have produced an effect. However, it should be noted that the term "control group" is often used misleadingly in ST research. Prior experiments that confronted participants with performance stereotypes often include "control" conditions presenting material that proposes equality in performance (Hively & El-Alayli, 2014). Therefore, we encourage researchers to assemble a true control group with a neutral text free of stereotyped material. The use of a convenience sample predominantly consisting of students unfortunately fits into the WEIRD concept (Apicella et al., 2020) and thus also limits the generalizability of our findings. Finally, we would encourage future studies to broaden the research scope by including more culturally diverse demographics, other (e.g., non-binary) genders, wider age ranges, or longitudinal designs to account for a possibly changed culture in terms of gendered stereotypes in gaming.

In summary, the present results corroborate the view that there are still gender differences in gaming performance across genres. However, the stereotype that males are *necessarily* more skillful game players *per se* is not tenable. Despite gender differences in performance, simply "threatening" participants with gender stereotypes was unsuccessful. Independent of gender, participants appeared to be "immune" to our induced stereotypical gender disadvantages, at least at the behavioral level. While we did not observe the hypothesized interaction effect of gender and ST condition in either study, we did observe that measures such as frustration indicated that ST mechanisms were still triggered, at least cognitively. This highlights how cognitive processing and behavior do not always have a uniform effect (Morrow et al., 1994). Regression analyses on a wide range of personality traits and gamer characteristics revealed that variables closely related to gaming (that is, Study 1: greater gamer identification; Study 2: greater self-reported gaming skill) are

associated with better performance. Thus, circumstantial factors (e.g., gaming experience) can explain gender differences, instead of gender itself. Future research should investigate the effects of ST in a more balanced way, integrating different game genres, different ways of inducing ST, having true control conditions and alternating the targeted group of the stereotype itself (males vs. females vs. gamers). Despite the absence of ST effects in our studies, we maintain that research on gender-related stereotypes has both theoretical and practical implications. However, as the world is becoming more complex and diverse it is more difficult to find stereotypes that may cause strong gender-related ST effects. To close remaining gender gaps in gaming, the industry should avoid alienating women (or other gender identities) with stereotypical (and possibly sexist) portrayals. Instead, the positive portrayal of females in games may lead to an even greater increase of women's interest in gaming as a hobby or even as a profession and to more favorable and equitable attitudes toward female gamers (Lynch et al., 2016; Melzer, 2019). As stereotypes evoked in the context of gaming can have carry-over effects on other stereotyped fields (such as STEM; Fordham et al., 2020) and the stereotypical portrayal of female characters is unrelated to the success of a game (Lynch et al., 2016), the gaming industry should continue to implement diverse environments and characters.

## Endnotes

[1] Welch's test is used to account for cases where the heterogeneity of variances was too extreme.

[2] A total of $n = 3$ participants identified as non-binary. Due to the small group size, they were not included in the main analyses, but in the descriptive measures (age, gaming habits).

[3] Homogeneity of variances was violated for several gender comparisons; thus, Welch's test is reported.

## References


Apicella, C., Norenzayan, A., & Henrich, J. (2020). Beyond WEIRD: A review of the last decade and a look ahead to the global laboratory of the future. *Evolution and Human Behavior*, *41*(5), 319-329. https://doi.org/10.1016/j.evolhumbehav.2020.07.015

Behm-Morawitz, E., & Mastro, D. (2009). The effects of the sexualization of female video game characters on gender stereotyping and female self-concept. *Sex Roles*, *61*(11-12), 808-823. https://doi.org/10.1007/s11199-009-9683-8

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.

Berger, A., & Krahé, B. (2013). Negative attributes are gendered too: Conceptualizing and measuring positive and negative facets of sex-role identity: Positive and negative facets of sex-role identity. *European Journal of Social Psychology*, *43*(6), 516-531. https://doi.org/10.1002/ejsp.1970

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.* In Proceedings of the 9th Annual International Workshop on Presence (pp. 28-29). Cleveland OH: Cleveland State University.

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

Burnette, J. L., Pollack, J. M., & Hoyt, C. L. (2010). Individual differences in implicit theories of leadership ability and self-efficacy: Predicting responses to stereotype threat. *Journal of Leadership Studies*, *3*(4), 46-56. https://doi.org/10.1002/jls.20138

Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review*, *106*(4), 676-713.

Chess, S., & Paul, C. A. (2019). The End of Casual: Long Live Casual. *Games and Culture*, 14(2), 107-118. https://doi.org/10.1177/1555412018786652

Dahme, G., Jungnickel, D., & Rathje, H. (1993). Güteeigenschaften der Achievement Motives Scale (AMS) von Gjesme und Nygard (1970) in der deutschen Übersetzung von Göttert und Kuh-Vergleich der Kennwerte norwegischer und deutscher Stichproben (english title: Quality characteristics of the Achievement Motives Scale (AMS) by Gjesme and Nygard (1970) in the German translation by Göttert and Kuhl-Comparison of the characteristic values of Norwegian and German samples). *Diagnostica*, *39*(3), 257-270.

Davies, P. G., Spencer, S. J., & Steele, C. M. (2005). Clearing the air: Identity safety moderates the effects of stereotype threat on women's leadership aspirations. *Journal of Personality and Social Psychology*, *88*(2), 276-287. https://doi.org/10.1037/0022-3514.88.2.276

Eden, A., Maloney, E., & Bowman, N. D. (2010). Gender Attribution in Online Video Games. *Journal of Media Psychology*, *22*(3), 114-124. https://doi.org/10.1027/1864-1105/a000016

Elizaga, R. A., & Markman, K. D. (2008). Peers and performance: How in-group and out-group comparisons moderate stereotype threat effects. *Current Psychology*, *27*(4), 290-300. https://doi.org/10.1007/s12144-008-9041-y

Engeser, S. (2005). *Messung des expliziten Leistungsmotivs: Kurzform der Achievement Motives Scale* (english title: Assessment of explicit achievement motives: Short version of the Achievement Motive Scale). Accessed 12 July 2024 from https://www.uni-trier.de/fileadmin/fb1/prof/PSY/PGA/bilder/Engeser__2005__Kurzform_der_AMS.pdf.

Entertainment Software Association. (2024). *Essential facts about the computer and video game industry*. https://www.theesa.com/wp-content/uploads/2024/05/Essential-Facts-2024-FINAL.pdf

Entertainment Software Association. (2021). *Essential facts about the computer and video game industry*. https://www.theesa.com/resource/2021-essential-facts-about-the-video-game-industry/

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

Fordham, J., Ratan, R., Huang, K.-T., & Silva, K. (2020). Stereotype threat in a video game context and its influence on perceptions of science, technology, engineering, and mathematics (STEM): Avatar-induced active self-concept as a possible mitigator. *American Behavioral Scientist*, *64*(7), 900-926. https://doi.org/10.1177/0002764220919148

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

Good, J. J., Woodzicka, J. A., & Wingfield, L. C. (2010). The effects of

gender stereotypic and counter-stereotypic textbook images on science performance. *The Journal of Social Psychology*, *150*(2), 132-147. https://doi.org/10.1080/00224540903366552

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

Hively, K., & El-Alayli, A. (2014). "You throw like a girl:" The effect of stereotype threat on women's athletic performance and gender stereotypes. *Psychology of Sport and Exercise*, *15*(1), 48-55. https://doi.org/10.1016/j.psychsport.2013.09.001

Hopp, T., & Fisher, J. (2017). Examination of the relationship between gender, performance, and enjoyment of a first-person shooter game. *Simulation & Gaming*, *48*(3), 338-362. https://doi.org/10.1177/1046878117693397

Hoyt, C. L., & Blascovich, J. (2007). Leadership efficacy and women leaders' responses to stereotype activation. *Group Processes & Intergroup Relations*, *10*(4), 595-616. https://doi.org/10.1177/1368430207084718

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

Klimmt, C., & Hartmann, T. (2005). Effectance, self-efficacy, and the motivation to play video games. In P. Vorderer & J. Bryant (Eds.), *Playing video games: Motives, responses, and consequences* (pp. 133-145). Lawrence Erlbaum Associates.

Koenig, A. M., & Eagly, A. H. (2005). Stereotype threat in men on a test of social sensitivity. *Sex Roles*, *52*(7-8), 489-496. https://doi.org/10.1007/s11199-005-3714-x

Kowert, R., Griffiths, M. D., & Oldmeadow, J. A. (2012). Geek or chic? Emerging stereotypes of online gamers. *Bulletin of Science, Technology & Society*, *32*(6), 471-479. https://doi.org/10.1177/0270467612469078

Leach, C. W., van Zomeren, M., Zebel, S., Vliek, M. L. W., Pennekamp, S. F., Doosje, B., Ouwerkerk, J. W., & 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* [Microsoft Windows]. Digital game published by 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. *Journal of Communication*, *66*(4), 564-584. https://doi.org/10.1111/jcom.12237

Melzer, A. (2019). Of princesses, paladins, and player motivations: Gender stereotypes and gendered perceptions in video games. In J. Breuer, D. Pietschmann, B. Liebold, & B. P. Lange (Eds.), *Evolutionary psychology and digital games: Digital Hunter-Gatherers* (S. 205-220). Routledge.

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

Pennington, C. R., Kaye, L. K., & McCann, J. J. (2018). *Applying the multi-threat framework of stereotype threat in the context of digital gaming*. PloS one, 13(2), e0192137. https://doi.org/10.1371/journal.pone.0192137

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* [Online game]. Digital game directed by Jason Kapalka and published by PopCap Games. https://en.gameslol.net/bejeweled-3-68.html

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

Ratan, R., Shen, C., & Williams, D. (2020). Men do not rule the world of tanks: Negating the gender-performance gap in a spatial-action game by controlling for time played. *American Behavioral Scientist*, *64*(7), 1031-1043. https://doi.org/10.1177/0002764220919147

Ratan, R., Taylor, N., Hogan, J., Kennedy, T., & Williams, D. (2015). Stand by your man: An examination of gender disparity in League of Legends. *Games and Culture*, *10*(5), 438-462. https://doi.org/10.1177/1555412014567228

Riot Games. (2009). *League of Legends* [Microsoft Windows]. Digital game directed by Tom Cadwell and published by Riot Games.

Rosenthal, H. E. S., & Crisp, R. J. (2007). Choking under pressure: When an additional positive stereotype affects performance for domain identified male mathematics students. *European Journal of Psychology of Education*, *22*(3), 317-326. https://doi.org/10.1007/BF03173429

Rydberg Sterner, T., Gudmundsson, P., Seidu, N., Bäckman, K., Skoog, I., & Falk, H. (2018). A psychometric evaluation of a swedish version of the Positive-Negative Sex-Role Inventory (PN-SRI). *Societies*, *8*(1), 13. https://doi.org/10.3390/soc8010013

Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology*, *38*(2), 194-201. https://doi.org/10.1006/jesp.2001.1500

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

Schwarzer, R., & Jerusalem, M. (Eds.). (1999). *Skalen zur Erfassung von Lehrer- und Schülermerkmalen: Dokumentation der psychometrischen Verfahren im Rahmen der wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen* (english title: *Scales for asessing teacher and pupil characteristics: Documentation of psychometric procedures as part of the scientific monitoring of the pilot project Self-Effective Schools*). Freie Universität Berlin.

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

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., 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

Stafford, T. (2018). Female chess players outperform expectations when playing men. *Psychological Science*, *29*(3), 429-436. https://doi.org/10.1177/0956797617736887

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. https://doi.org/10.1037/0022-3514.69.5.797

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

Taylor, C. A., Lord, C. G., McIntyre, R. B., & Paulson, R. M. (2011). The Hillary Clinton effect: When the same role model inspires or fails to inspire improved performance under stereotype threat. *Group Processes & Intergroup Relations*, *14*(4), 447-459. https://doi.org/10.1177/1368430210382680

Thoman, D. B., Smith, J. L., Brown, E. R., Chase, J., & Lee, J. Y. K. (2013). Beyond performance: A motivational experiences model of stereotype threat. *Educational Psychology Review*, *25*(2), 211-243. https://doi.org/10.1007/s10648-013-9219-1

Tsui, M., Xu, X. Y., & Venator, E. (2011). Gender, stereotype threat and mathematics test scores. *Journal of Social Sciences*, *7*(4), 538-549.

Vancouver, J. B., Thompson, C. M., Tischner, E. C., & Putka, D. J. (2002). Two studies examining the negative effect of self-efficacy on performance. *Journal of Applied Psychology*, *87*(3), 506-516. https://doi.org/10.1037/0021-9010.87.3.506

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. *Quantic Foundry*. https://quanticfoundry.com/2017/01/19/female-gamers-by-genre/