Received: 25 March 2020 | Revised: 4 February 2021 | Accepted: 4 February 2021

DOI: 10.1002/bdm.2239

RESEARCH ARTICLE

WILEY

Gender differences in decision-making: The effects of gender stereotype threat moderated by sensitivity to punishment and fear of negative evaluation

Laura Villanueva-Moya | **Francisca Expósito**

Mind, Brain, and Behavior Research Center (CIMCYC), Department of Social Psychology, University of Granada, Granada, Spain

Correspondence
Laura Villanueva-Moya, Mind, Brain, and Behavior Research Center (CIMCYC), Department of Social Psychology, University of Granada, Granada, Spain.
Email: lauravm@ugr.es

Funding information
Spanish Ministry of Economy, Industry and Competitiveness

Abstract

Research has demonstrated gender differences in the decision-making process, showing that women make more disadvantageous risk decisions than men. However, these differences have not been examined in terms of psychosocial or socio-structural variables, such as the gender stereotype threat. We conducted an experimental study (N_s = 105) to test the well-established stereotype threat effect on decision-making through the Iowa Gambling Task and the possible moderation of this effect by sensitivity to punishment and fear of negative evaluation. The results revealed that women under a stereotype threat condition make more disadvantageous risk decisions than men in the same conditions or women in the nonstereotype threat condition. Moreover, women greatly fearing negative evaluation seemed to make more disadvantageous risk decisions compared with other groups. These findings highlight the relevance of psychosocial variables that legitimize gender inequality, such as the stereotype threat and fear of negative evaluation, in women's decision-making process.

KEYWORDS
decision-making, fear of negative evaluation, gender differences, Iowa Gambling Task, stereotype threat

1 | INTRODUCTION

Every day, people are involved in a multitude of social interactions through which they are exposed to possible positive or negative evaluations that can affect their cognitive processes. Indeed, it has been found that situations involving social evaluative or stereotype threats undermine performance on cognitive tasks (Baumeister et al., 2002; Schmader et al., 2008). These threats differ in nature: a *stereotype threat* occurs when a person is aware of a negative stereotype about their social group and is concerned about confirming it, which undermines their performance on stereotype tasks (e.g., women in math, Black people on standardized tests, or White males in athletics; Spencer et al., 1999; Steele & Aronson, 1995; Stone, 2002, respectively). According to Schmader et al. (2008), a *social evaluative threat* occurs when a person believes they will perform poorly on a task due to expectations that others could have for their performance (e.g., a 5-min speech in front of an evaluative audience: Brown et al., 2012). In this sense, the manipulations of stereotype threats are designed to remember the negative stereotype about one's group (e.g., Spencer et al., 1999), whereas those in social evaluative are designed to create a sense of public criticism through a social presence or a negative social evaluation (Dickerson & Kemeny, 2004). Another significant difference is that a stereotype threat increases one's motivation to do something well to try to disconfirm the negative stereotype, while this motivation is not clearly related to social evaluative threats (Schmader et al., 2008). Therefore, the threats are different from each other. On this basis, this research aims to investigate how belonging to a stigmatized group (stereotype threat) can affect individual performance. Specifically, we analyze how the gender stereotype threat can affect women's cognitive processes, such as decision-making.

The stereotype threat experienced by women in the science domain is triggered by negative stereotypes about their mathematical ability (Steele, 1997). Socially, women are not considered competent in stereotypical masculine domains (e.g., mathematics), and consequently, feeling threatened leads to a decrease in their performance (e.g., Spencer et al., 1999). Research has demonstrated that women constantly experience this sensation of threat, which could extend over other domains, such as decision-making (Carr & Steele, 2010). Decision-making is one of the most important components of executive function (Bechara, Damasio, & Damasio, 2000); therefore, it is important to understand the possible factors that affect it. The stereotype threat could explain gender differences in decision-making; women (vs. men) require more effort and time to make decisions (e.g., Evans & Hampson, 2015; van den Bos et al., 2013), increasing gender inequality situations. For instance, it has been shown that the stereotype threat decreases women's risk-aversion behavior (Carr & Steele, 2010), which could lead to women's decreased participation in stereotypically masculine domains (e.g., leadership positions), maintaining their underrepresentation and therefore legitimizing gender inequality. That is why we propose broadening empirical evidence through the Iowa Gambling Task (IGT), shown to be able to predict real-life decision deficits (Bechara et al., 1994). The IGT was created to evaluate a person's ability to make decisions in non-clinical and clinical samples (e.g., Bechara et al., 1994; Weller et al., 2010). The IGT score has been associated with people's tendency to take risks (e.g., Weller et al., 2010); a higher IGT score is associated with more advantageous risk decisions. The IGT is a computer task in which participants must learn about gains and losses progressively using four decks of cards (A, B, C, and D). Two are "bad choices" (A and B), which provide immediate, high gains but great long-term losses, and two are "good choices" (C and D), which provide immediate, low gains but small long-term losses. Subjects are not aware of the decks' characteristics until they choose an option; then, the computer gives them feedback on the consequences of their choice. Specifically, this research could add to knowledge about how the stereotype threat affects decisions in a more realistic way, given that IGT findings have been shown to be generalizable outside the laboratory. In fact, it was designed to assess real-life decision making in individuals with lesions of the ventromedial prefrontal cortex, showing that they made disadvantageous risk decisions in this task (Bechara et al., 1994; Bechara, Damasio, & Damasio, 2000; Bechara, Tranel, & Damasio, 2000). In addition, the IGT has been associated with disadvantageous risk decisions in individuals with attention deficit hyperactivity disorder (e.g., Dekkers et al., 2018), sexual risk behaviors (e.g., Hardy et al., 2006), alcohol abuse (e.g., Kovács et al., 2017), or even in individuals with excessive social network sites use (Meshi et al., 2019). Further, to our knowledge, no research has been done on stereotype threats related to IGT; thus, this work would contribute to knowledge in this field.

1.1 | Stereotype threat

Each group in society is associated with positive or negative stereotypes depending on their characteristics. According to the *stereotype content model* (Fiske et al., 2002), stereotypes are captured by two dimensions—competence and warmth—through which differences between groups are established, legitimizing the social status of each one. Regarding gender, men have been characterized by high competence and low warmth, whereas women have been characterized by a lack of competence and high warmth. Fiske et al. (2002) indicated that competence differentiates between groups to a greater extent, showing public signs such as performance on tasks. Indeed, it has been established that in tasks requiring high competence, such as mathematics, women do not perform as well as men (e.g., Spencer et al., 1999).

In this respect, it seems that referring to gender differences in math tasks triggers a threatening environment in which women may fear being judged due to their gender (for a review, see Spencer et al., 2016). Indeed, Spencer et al. (1999) were the main authors in the stereotype threat field who demonstrated that referring to gender differences in math tasks triggered a stereotype threat in women (vs. men). Researchers studying the gender stereotype threat have suggested that for women, the activation of negative stereotype in math undermines their performance because they feel threatened and are afraid to confirm the stereotype, in contrast to women who do not feel threatened or men (e.g., Spencer et al., 1999). This also seems to occur with other stereotypes, such as the racial stereotype threat (i.e., Black people have lower intellectual ability than White people; Steele, 1997). There is empirical research showing that the activation of negative stereotypes related to intellectual ability undermines the performance of Black people on standardized academic tests compared with White participants and others who are not threatened (e.g., Ho & Sidanius, 2010; Steele & Aronson, 1995). To try to explain how the stereotype threat undermines performance in stigmatized individuals, Schmader et al. (2008) developed an *integrated process model of the stereotype threat*. According to their model, the stereotype threat activates three processes: the physiological stress response, such as increased activation of the sympathetic nervous system (Murphy et al., 2007); monitoring the self-relevance of performance, such as increased motivation to do well (Schmader et al., 2008); and suppression of thoughts or emotions, such as feeling self-doubt and worry or having a scattered mind (Beilock et al., 2007; Mrazek et al., 2011; Steele & Aronson, 1995). This model explains that these mechanisms, triggered by the stereotype threat, can interrupt the working memory of the stigmatized individual (e.g., women in math) that is essential for achieving difficult cognitive tasks (e.g., decision-making).

Thus, the stereotype threat should be considered, as it causes cognitive, emotional, and behavioral consequences among women. The activation of negative stereotypes in math (i.e., women are bad at math) among women causes consequences in stereotypically masculine domains (e.g., mathematics), and these effects spill over into other domains, such as verbal ability (Beilock et al., 2007), perceptual learning (Rydell & Boucher, 2017), basic executive functions (Rydell et al., 2014), and decision-making (Carr & Steele, 2010), giving rise to what is known as *stereotype threat spillover* (Inzlicht & Kang, 2010; Rydell & Boucher, 2017). On this basis, we propose broadening the

research on the stereotype threat and decision-making, because is one of the most important components of executive functions (Bechara, Damasio, & Damasio, 2000).

1.2 | Decision-making

Decision-making is an individual process that can determine a person's life course. When people make decisions, they assess the advantages and disadvantages of options and the costs and benefits associated with them (Bechara, 2005). Literature on decision-making has found gender differences: in general, women carry out integrated information processing, using all available information in an environment, even when this information may lead them to make disadvantageous decisions. Men process information selectively, using specific information that benefits their decisions (Byrne & Worthy, 2016). That is, men make decisions based on a goal-oriented performance, being able to ignore environmental information (e.g., consequences or risks of a decision) to achieve their goal, whereas women take all environmental information into account, focusing their attention on all details (e.g., benefits and consequences of a decision). Therefore, men do not mind making decisions that involve great risks as long as they achieve their goal, whereas women consider the risks of these decisions which seem to influence their final decision. This gender difference can be observed in risk decisions: men make riskier decisions than women do through explicit (Domain-Specific Risk-Taking scale: Lozano et al., 2017) and implicit measures (Cambridge Gambling Task: Deakin et al., 2004; IGT: van den Bos et al., 2013). Although many implicit risk-taking tasks have been used to test gender differences (e.g., Balloon Analogue Risk Task: e.g., Lighthall et al., 2012; Columbia Card Task: Buelow & Cayton, 2020; Game of Dice Task: e.g., Zhang et al., 2017), only the Cambridge Gambling Task and IGT found significant differences.

Based on this range of tasks, we decided to use an implicit measure (IGT or Cambridge Gambling Task) instead of an explicit measure (Domain-Specific Risk-Taking) because it has been shown that implicit measures better predict intuitive decision-making than explicit measures (Richetin et al., 2007). Participants cannot control their answers in implicit measures as much as in explicit measures, removing any possible social desirability bias. Specifically, we used the IGT instead of the Cambridge Gambling Task because we considered it more appropriate to measure risk-taking for several reasons. A test–retest reliability of the decision-making tasks indicated that the IGT is the only task that assesses the decision-making process in its entirety (ambiguity and risk; Buelow & Barnhart, 2018). Furthermore, Buelow and Blaine's (2015) factor analysis found that the IGT is not comparable with other risk-taking tasks; in the IGT, participants must learn through exploration to differentiate advantageous from disadvantageous options as a function of risks and benefits (e.g., van den Bos et al., 2013). Therefore, the main difference between IGT and other tasks is the learning process in decision-making: participants make decisions based on their experiences (i.e., past outcomes), so it is considered to more closely resemble real-life decisions (Bechara et al., 1994). In this sense, Bishara et al. (2009) found a change in decision consistency over time in the IGT response, which reflects how people change their decisions based on their experience, such as in real-life situations. Lastly, research has demonstrated greater gender differences in IGT results than in other measures (for a review, see van den Bos et al., 2013). The IGT may be best suited for detecting gender differences because in contrast to the Cambridge Gambling Task, where participants must select among different boxes without any information; in the IGT, participants receive information about gains and losses associated with each decision. Therefore, participants can make their decisions based on the information provided to them, where gender differences can be reflected. It has been shown that men process information selectively to make decisions, whereas women take all environmental information into account (Byrne & Worthy, 2016; Meyers-Levy, 1989), being more sensitive to losses than wins (Garrido-Chaves et al., 2020).

Gender differences in the IGT have suggested that women make more disadvantageous risk decisions than men (e.g., Singh et al., 2020; van den Bos et al., 2013). The explanations for these differences have been limited to stress (e.g., Santos-Ruiz et al., 2012), anxiety (e.g., Zhang et al., 2017), or information processing (e.g., Byrne & Worthy, 2016); there is not enough evidence regarding psychosocial or socio-structural variables, such as the stereotype threat. Research has shown that the stereotype threat could explain gender differences in decision-making (Carr & Steele, 2010), but to our knowledge, no study has tried to explain these differences using the IGT for decision-making assessment. Carr and Steele (2010) investigated the effect of the stereotype threat on risk-aversion behavior—measuring it through loss aversion and risk aversion (Kahneman & Tversky, 1979), finding that women under the stereotype threat condition were more risk- and loss-averse than men in the same condition and women in the nonstereotype condition. These differences could due to the stereotype threat triggering several processes that interrupt working memory (e.g., stress, hypervigilance, or worry), which is essential for achieving cognitive tasks (Schmader et al., 2008). Indeed, women under stereotype threat conditions showed lower working memory than men in the same conditions and women under nonstereotype conditions (e.g., Schmader & Johns, 2003). These findings could be associated with decision-making, given that several studies on the IGT have shown that decision-making is significantly associated with working memory (for a review, see Woodrow et al., 2019). Participants with low working memory capacity seem to make disadvantageous risk decisions than those with high capacity (Bagneux et al., 2013; Duarte et al., 2012). Considering the stereotype threat and the IGT, it could be said that the stereotype threat would decrease women's working memory by causing constant worry about their performance under threat situations, consequently causing them to make more disadvantageous risk decisions during the IGT than others. Based on this reasoning, we expected women under the stereotype threat condition to score lower on the IGT (make more disadvantageous risk decisions) than men in the same condition or women and men in the nonstereotype condition.

It is noteworthy that the effect of the stereotype threat on decision-making could be moderated by sensitivity to punishment or fear of negative evaluation (FNE). Women have higher context sensitivity than men (Miller & Ubeda, 2012); that is, they seem be more sensitive to signals (e.g., threats, punishments, and negative social evaluations) and consequently could modify their behavior in accordance to what is expected of them (i.e., being warm and incompetent and therefore not making risky decisions). Could these variables help us explain how the stereotype threat affects women's decision-making process, as they assess how people fear failing or being evaluated and punished?

1.3 | Sensitivity to punishment and fear of negative evaluation

Sensitivity to punishment measures the functioning of the behavioral inhibition system—the cognitive processes triggered by the threat of failure or punishment (Torrubia et al., 2001). It assesses an individual's ability to detect cues that indicate failure or punishment. Gender differences have shown that women exhibit more sensitivity to punishment than men (Castellà & Pérez, 2004; Torrubia et al., 2001). Therefore, it seems that women have higher context sensitivity than men do, being more sensitive to contextual signals (e.g., stereotype threat). Indeed, it was found that sensitivity to punishment, that is, a hypervigilant personality style, seems to affect people (e.g., anxiety or depression; Katz et al., 2020) through an attentional bias toward threatening information (Hundt et al., 2007). In this sense, women who score higher on sensitivity to punishment could decrease their performance due to stereotype threat to a greater extent, given that they would have a greater sensitivity and would look for information (e.g., cues that indicate potential failure) to confirm or disprove negative stereotypes about their group (Schmader et al., 2008). Regarding decision-making, it has been found that women make decisions on the basis of integrating and understanding all information in an environment (Byrne & Worthy, 2016), and therefore, their decisions are determined by context to a great extent (Miller & Ubeda, 2012). This cognitive pattern can be observed in the IGT (van den Bos et al., 2013), which has demonstrated that men's abilities to make decisions are based more on goal-oriented performance and the ability to ignore environmental information (e.g., losses). In contrast, women focus their attention on details (e.g., losses). As a result, men tend to attempt to maximize long-term gains, and women focus on maximizing short-term gains without perceiving the significant losses such decisions might incur. If it is taking into account the sensibility to punishment and the IGT; it has found that people with high sensitivity to punishment score lower on the IGT (make disadvantageous risk decisions; Davis et al., 2007). Therefore, women's performance in terms of maximizing short-term gains (making disadvantageous risk decisions) could be explained by their greater sensitivity to punishment. Specifically, under stereotype threat conditions this sensitivity could increase women's perception of this threat because they could consider losses in the IGT as a sign of stereotype confirmation and try to avoid them, a priori selecting the option that presents fewer losses. It should be noted that we have considered punishment as the number of losses that women receive as feedback after their decision. Therefore, we expected sensitivity to punishment to moderate the effect of the stereotype threat in the IGT; that is, that women in the stereotype threat condition with high sensitivity to punishment would score lower on the IGT than others.

Beyond the aforementioned, the *fear of negative evaluation* (FNE) could also be a variable moderator. FNE refers to the sensation that a person experiences when they fear they could be evaluated negatively by others (Gallego et al., 2007). It has been found that women score higher on FNE than men (e.g., Biolcati, 2017). This sensation can be said to resemble the concern that women under threat conditions feel about their performance in terms of avoiding confirmation of negative stereotypes. Maresh et al. (2017) found that participants with high levels of FNE in threat conditions showed greater difficulty to perform tasks effectively than in nonthreat conditions. Although we did not find studies that tested the moderator effect of FNE together with gender stereotype threat, this is one of the studies indicating that FNE could explain the individual performance under threatening conditions. In this sense, it is expected that the performance of participants with high levels of FNE—usually women—under stereotype threat conditions is affected to a greater extent. Regarding the relationship between FNE and decision-making, Maner et al. (2007) found a relationship between FNE and risk-taking (measured through the Balloon Analog Risk Task; Lejuez et al., 2003), indicating that FNE is linked to a tendency to make less risky decisions. Concerning the IGT, we could not find any studies have assessed both constructs, and neither under the stereotype threat. Considering the empirical evidence, we expected FNE to moderate the effect of the stereotype threat in the IGT; that is, that women in the stereotype threat condition with high FNE would score lower on the IGT than others.

Based on this theoretical conceptualization, this research aims to analyze the effect of the gender stereotype threat in decision-making—measured through IGT—and the possible moderation of this effect by sensitivity to punishment and FNE. Specifically, we expected that women in the stereotype threat condition would score lower on the IGT (make more disadvantageous risk decisions) than: men in the stereotype threat condition (Hypothesis 1), women in the nonstereotype threat condition (Hypothesis 2), and men in the nonstereotype threat condition (Hypothesis 3). Furthermore, we expected sensitivity to punishment and FNE to moderate the effect of the stereotype threat in IGT: in other words, that women in the stereotype threat condition with high sensitivity to punishment would score lower (make more disadvantageous risk decisions) on the IGT than others (Hypothesis 4) and that women in the stereotype threat condition with high FNE would score lower on the IGT than others (Hypothesis 5).

1.4 | Preliminary study

Given the importance that the manipulation work as we intended, we decided to carry out a preliminary study to test the effect of Carr and

Steele's manipulation (Carr & Steele, 2010), because their study was one of the first to find that the stereotype threat decreased women's risk-aversion behavior in decision-making. Participants were told that they would complete a task of mathematical, logical, and rational reasoning (*stereotype threat condition*) or that they would complete a puzzle-solving task (*nonstereotype threat condition*). Following Carr and Steele's procedure, participants also had to indicate their gender before beginning the task, whereas in the nonstereotype threat condition, they indicated it after completing the questionnaire. After this manipulation participants completed the corresponding measures (IGT, sensitivity to punishment and fear of negative valuation). Results demonstrated that, with our sample, Carr and Steele's manipulation (Carr & Steele, 2010) was not effective in eliciting a real threat in women (see Supporting Information). This manipulation could not work in this study due to women being unaware or not worrying about negative stereotypes associated with their performances in math (Steele, 1997). Perhaps it is necessary to refer to gender differences in math tasks to trigger the stereotype threat—making the threat explicit to remind women of the negative stereotype associated with their group. It would generate a threatening environment in which women may fear being judged in relation to their gender (Spencer et al., 2016). Referring to gender differences in math tasks to develop a stereotype threat condition was the manipulation used by Spencer et al. (1999), who performed the main research on the stereotype threat and women's math performance. Therefore, based on this preliminary study, we decided to use the classic manipulation developed by Spencer et al. (1999) to carry out the present study.

Main study as follows.

2 | METHOD

2.1 | Participants

A convenience sample of 110 undergraduate psychology students (62 women and 48 men; ages 18–37, M_{age} = 21.49, *SD* = 2.74) from the University of Granada took part in the study voluntarily and without financial compensation. Familiarity with the task was an exclusion criterion. Five participants failed to pass an attention check; therefore, they were removed, leaving a sample of 105 undergraduate psychology students (58 women and 47 men; ages 18–37, M_{age} = 21.51, *SD* = 2.77). Using G*Power (Faul et al., 2007), we conducted a sensitivity power analysis (1 − β = 80%; α = .05; *N* = 105), which revealed that that the design could detect a sufficiently acceptable[1] effect size of *f* = .28 using a two-way ANOVA with four groups and one degree of freedom.

2.2 | Procedure and design

This was an experimental 2 (gender: male vs. female) × 2 (stereotype threat: stereotype threat vs. nonstereotype threat) between-group design with IGT total score as the dependent variable and sensitivity to punishment and FNE as moderators. Participants were randomly assigned to experimental conditions (stereotype threat condition: 21 men and 31 women; nonstereotype threat condition: 26 men and 27 women).

To carry out the study participants were seated in separate cubicles equipped with computers. Before starting the experiment, we informed them about the aim of the study (evaluating how people make decisions) and about the anonymity and confidentiality of their responses, asking them to sign the informed consent form if they agreed after reading it ("After being informed of the above, I agree to participate in the study"). Once they accepted it, the experimental manipulation was presented to them. Participants then received the instructions for performing the computer task. When they completed the task, we administered the questionnaire with the rest of the measures. They performed the experiment without the female experimenter present. The approximate duration of the experiment was 20 min. We conducted this study following the recommendations of the Human Research Ethics Committee of the University of Granada and protected the collected data under the 1964 Declaration of Helsinki.

2.3 | Measures

2.3.1 | Stereotype manipulation (stereotype threat vs. nonstereotype threat)

In alignment with Spencer et al. (1999), the main researchers of the stereotype threat and women's math performance as well as other studies (e.g., Beilock et al., 2007; Rydell et al., 2014), the negative stereotype of women in the domain of mathematics was manipulated through a brief text. In the stereotype threat condition (*threatening to women*), we told participants that there were gender differences on a mathematics test:

> There are empirical studies that show that women perform worse than men in tasks of mathematical, logical, and rational reasoning. This research aims to investigate the reasons for these differences and the factors involved in the development of these tasks.

In the nonstereotype threat condition (*nonthreatening to women*), we did not refer to gender differences:

> There are empirical studies that show that some people perform worse than others in tasks of mathematical, logical, and rational reasoning. This research aims to investigate the reasons for these differences and the factors involved in the development of these tasks.

Furthermore, Danaher and Crandall (2008) demonstrated that inquiring about gender at the end of a test (versus at the beginning) decreases gender differences in the results. Thus, in the stereotype

threat condition, we told participants to indicate their gender before beginning the task, whereas in the nonstereotype threat condition, they indicated it after completing the questionnaire (Carr & Steele, 2009, 2010).

Next, participants completed the following measures:

2.3.2 | Iowa Gambling Task

This measure implicitly simulates the decision-making process in everyday life through the evaluation of gains and losses under conditions of uncertainty or risk (Bechara et al., 1994; Bechara, Tranel, & Damasio, 2000). Participants must choose between four decks of cards: two of them are "bad choices" (A and B), which provide immediate, high gains but great long-term losses, and the other are "good choices" (C and D), providing immediate, low gains but small long-term losses. After the participant chooses a deck, the program gives feedback (gains or losses), based on which participants must decide on the best options. The end goal is to try to win as much money as possible. The dependent variable was measured by the overall IGT score after 100 trials, which, according to the authors, is the difference between the number of advantageous and disadvantageous choices. A higher IGT score indicates advantageous risk decisions, whereas a low IGT score indicates disadvantageous risk decisions.

2.3.3 | Control question

In the following measures, an extra control question was included to identify subjects not paying attention (e.g., "If you are reading this question, answer with 3"; Lozano et al., 2017).

2.3.4 | Fear of Negative Evaluation (FNE)

The Brief Fear of Negative Evaluation scale (Leary, 1983), adapted into Spanish by Gallego et al. (2007), was used. Twelve items assess the degree to which people experience fear of being evaluated negatively by others through 5-point Likert-type statements (e.g., "I am frequently afraid of other people noticing my shortcomings"; 1 = *not at all characteristic of me*, 5 = *extremely characteristic of me*). Average scores were calculated: a higher score means a greater tendency to afraid in situations in which one can be evaluated by others. This scale revealed adequate psychometric properties in the original measure (α = .90–.91) and in its validation for the Spanish population (α = .90). In this sample, the Cronbach's alpha coefficient was .89.

2.3.5 | Sensitivity to punishment

The subscale of the Punishment Sensitivity and Reward Sensitivity Questionnaire developed by Torrubia et al. (2001) was used. It consists of 24 dichotomous items (*Yes*/*No*) measuring individual differences in the functioning of behavioral inhibition systems (cognitive processes triggered by the threat of failure or punishment; e.g., "Generally, do you pay more attention to threats than to pleasant events?"). The total score was obtained through the sum of affirmative answers. The original measure revealed acceptable reliability and validity properties (α = .82–.83). With this data, the internal consistency was .82.

2.4 | Data analysis

We tested the effects of the stereotype threat condition (stereotype threat vs. nonstereotype threat) on the IGT with a two-way ANOVA between groups (see Figure 1). Then, to test the possible moderation of these effects by sensitivity to punishment or FNE, a moderation analysis was performed with the PROCESS (Model 3; Hayes, 2017) macro for SPSS (version 3.4.1) with 10,000 bias-corrected bootstrap samples and 95% confidence intervals. To calculate the magnitude of the interaction (≥.02/.15/.35 indicate small/medium/large effects; Cohen, 1988), we determined the standardized effect size f^2 on the basis of the change in R^2 (Δf^2). Post hoc tests were calculated by simple slope tests (see Figure 2; Dawson, 2014). We performed the analyses with IBM SPSS Statistics version 22.0.

3 | RESULTS

3.1 | Effects of the stereotype threat on decision-making

We conducted a two-way ANOVA to test if women in the stereotype threat condition would score lower on the IGT than: men in the stereotype threat condition (Hypothesis 1), women in the nonstereotype threat condition (Hypothesis 2), and men in the nonstereotype threat condition (Hypothesis 3). The experimental condition (0 = stereotype threat; 1 = nonstereotype threat) and gender (0 = male; 1 = female) were introduced as independent variables and the IGT total score as dependent variable. The results revealed the predicted gender-by-condition interaction, $F(1, 101) = 19.39$, $p < .001$, $\eta_p^2 = .16$ (see Figure 1).[2] That is, women assigned to the stereotype threat condition ($M = 11.03$, $SD = 26.74$) scored lower on the IGT than men in the stereotype threat condition, $t(50) = 2.39$, $p = .02$, 95 CI [2.39, 27.93], $d = .65$, ($M = 26.19$, $SD = 19.08$), and women assigned to the nonstereotype threat condition, $t(55) = -4.32$, $p < .001$, 95 CI [−39.27, −14.37], $d = 1.13$, ($M = 37.85$, $SD = 20.46$); no gender differences were found in men assigned to the nonstereotype threat condition ($p = .99$, 95 CI [−14.66, 14.59]). These findings supported Hypotheses 1 and 2, but not 3.[3]

3.2 | Sensitivity to punishment

To analyze whether women in the stereotype threat condition with high sensitivity to punishment would score lower on the IGT than

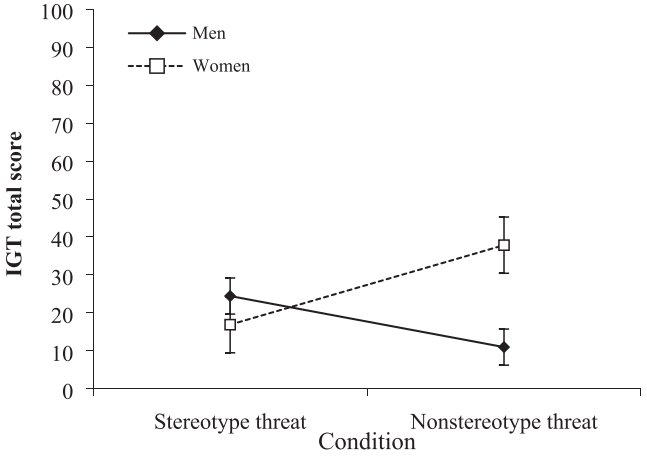

FIGURE 1 FiEffect of gender-by-condition interaction over the Iowa Gambling Task (IGT total score: the difference between the number of advantageous and disadvantageous choices). Error bars show ±standard error of the mean

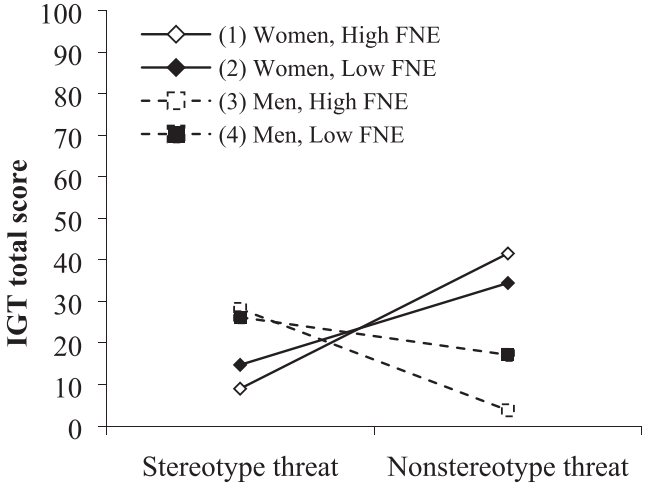

FIGURE 2 FiInteraction of stereotype threat, gender, and fear of negative evaluation (FNE) on the Iowa gambling task (IGT) total score. High: +1 *SD*; low: −1 *SD*; IGT total score: the difference between the number of advantageous and disadvantageous choices

others (Hypothesis 4), we performed a moderation analysis. We took the condition (0 = stereotype threat; 1 = nonstereotype threat) as the predictor variable, sensitivity to punishment and gender (0 = male; 1 = female) as moderators, and IGT total score as the criterion variable. The results showed that the three-way interaction between Gender × Condition × Sensitivity to Punishment was not significant (b = 1.02, SE = 2.09, t = 0.49, p = .625, 95 CI [−3.12, 5.17], Δf^2 = .002).

3.3 | Fear of negative evaluation

Lastly, we tested whether women in the stereotype threat condition with high FNE would score lower on the IGT than others (Hypothesis 5). We followed the same steps as in the previous moderation analysis, considering FNE as the moderator. The results revealed the significant three-way interaction predicted by Gender × Condition × FNE (b = 26.60, SE = 12.87, t = 2.07, p = .04, 95 CI [1.06, 52.14], Δf^2 = .03).

To ease the interpretation of the three-way interaction effect, we illustrated the results in Figure 2 and examined the plausible differences between groups using simple slope tests. According to Dawson (Dawson, 2014; Dawson & Richter, 2006), this analysis checks whether the difference between slopes is statistically significant. As shown in Figure 2, in the stereotype threat condition, women high in FNE (+1 *SD*) seemed to score lower on the IGT than others. In contrast, in the stereotype threat condition, women low in FNE (−1 *SD*) seemed to score higher on the IGT than others. This observation can be supported by simple slope tests, as there was a significant difference between the slopes; that is, women high in FNE (1) scored lower on the IGT than women low in FNE (2), t = 2.04, p = .04; men high in FNE (3), t = 2.24, p = .03; and men low in FNE (4), t = 2.164, p = .03.[4]

4 | DISCUSSION

This research aims to investigate the effect of the gender stereotype threat in decision-making—measured through IGT—and the possible moderation of this effect by sensitivity to punishment and FNE. The results observed showed that women under a stereotype threat condition scored lower on the IGT than men in the same condition, and women in the nonstereotype threat condition. Additionally, women with higher FNE scored lower on the IGT than others. These findings highlight the importance of the stereotype threat and FNE (psychosocial variables) in decision-making, explaining gender differences in this process. Research using the IGT has demonstrated that gender differences in decision-making exist; women make more disadvantageous risk decisions than men (e.g., Singh et al., 2020; van den Bos et al., 2013). These differences have been attributed to stress (e.g., Santos-Ruiz et al., 2012), anxiety (e.g., Zhang et al., 2017), or information processing (e.g., Byrne & Worthy, 2016). Nevertheless, previous research has not taken psychosocial or socio-structural variables into account regarding the stereotype threat to try to explain these gender differences. Empirical evidence has shown the effect of the gender stereotype threat on decision-making (Carr & Steele, 2010) as one of the possible factors of the persistence of gender inequality. Nevertheless, so far, this effect has not been evaluated with the IGT, one of the most common tasks used to measure decision-making.

Regarding the effect of the stereotype threat on decision-making, findings showed that women under the stereotype threat condition scored lower on the IGT; that is they made more disadvantageous risk decisions than men in the same condition and women in the nonstereotype threat condition. These findings are in line with empirical evidence (Carr & Steele, 2010), showing that women under a stereotype threat condition were more averse to risk than men in the same condition and women in a nonstereotype condition. This could

indicate that the gender stereotype threat affects women's decision-making process. Women are aware of negative stereotypes about their social group and are concerned about confirming negative stereotypes about their group when they perform a task (Steele & Aronson, 1995). As a result, under stereotype threat situations, women report monitoring their performance (for a review, see Spencer et al., 2016), worrying more about it (Beilock et al., 2007), feeling self-doubt (Steele & Aronson, 1995), or having a scattered mind (Mrazek et al., 2011). All of these cognitive activities interrupt working memory (Schmader et al., 2008), which is essential for good performance on difficult tasks (Régner et al., 2010). This could apply to decision-making (Woodrow et al., 2019), leading to women making more disadvantageous risk decisions.

Concerning the effect of the stereotype threat on decision-making, we have considered sensitivity to punishment and FNE as possible moderators to help explain gender differences. Firstly, we expected that individuals with high levels of sensitivity to punishment—usually women—would be more affected by threatening information (Hundt et al., 2007), such as stereotype threat, which might have interfered in their IGT performance because they would look for information (e.g., cues that indicate potential failure) to confirm or disprove negative stereotypes about their group (Schmader et al., 2008). Nevertheless, we did not find significant effects of sensitivity to punishment as a moderator of decision-making. Our results may be due to the measure of sensitivity used: Because it was an explicit measure (e.g., "In tasks that you are not prepared for, do you attach great importance to the possibility of failure? or Generally, do you pay more attention to threats than to pleasant events?"), participants could have answered as a function of social desirability. Perhaps an implicit measure would be necessary to detect the level of sensibility to punishment for women under stereotype threat conditions. Regarding the FNE, the results showed that women high in FNE scored lower on the IGT than others; that is, they made more disadvantageous risk decisions when they were fearful of being evaluated. These results are in line with and expand on those of previous studies that indicated that people high in FNE showed greater difficulty to perform tasks effectively (Maresh et al., 2017) and a tendency to make less risky decisions (Maner et al., 2007).

Our findings could be framed in the *integrated process model of the stereotype threat* (Schmader et al., 2008). The stereotype threat seems to trigger a fear of being negatively evaluated by others (FNE) in women under the stereotype threat condition. This sensation could be similar to the concern that threatened women feel about their performance in terms of avoiding confirming negative stereotypes (e.g., Beilock et al., 2007). Consequently, threatened women may try to suppress their thoughts or emotions, a process that, according to Schmader et al. (2008), could interrupt working memory, which is necessary to achieve cognitive tasks such as decision-making. Indeed, it has been shown that decision-making is significantly associated with working memory (for a review, see Woodrow et al., 2019). Considering the stereotype threat and IGT, it could be said that the stereotype threat would affect women's working memory negatively through constant exposure to the FNE under threat situations, and consequently, they would make more disadvantageous risk decisions than others. This study contributes to Schmader et al.'s model of how the stereotype threat could influence stigmatized individuals (i.e., women). To support these findings, it would be interesting to take working memory into account as a control variable. Régner et al. (2010) found that the decrease in women's performance triggered by the stereotype threat could be related to working memory, supporting Schmader et al. (2008). Specifically, they found that stereotype threat patterns emerged among women with low levels of working memory, who underperformed to a greater extent; we discuss this further in the limitations section.

Regarding the nonstereotype threat condition, it should be noted that although Carr and Steele (2010) found no significant differences between women under a stereotype threat condition and men assigned to a nonstereotype threat condition, they obtained different results in terms of tendencies: Women were more averse to risk than men under the nonstereotype threat condition (see Footnote 3). We discuss plausible explanations below. On one hand, the IGT was described as a mathematical task; thus, it ought not to have produced a perception of threat in men, given that mathematical domains are perceived as masculine and languages as feminine (Chaffee et al., 2019). The nonstereotype threat condition was designed to not threaten women, but it could have worked as a threat to men's gender identity. Even though we did not refer to gender differences in this condition, we mentioned that "some people have worse performance than others in mathematical, logical, and rational reasoning tasks." According to a stereotype content model (Fiske et al., 2002) men should perform better than women because math requires high competence. Men could have felt pressured to answer correctly or feared obtaining a worse score than women, according to what is expected of them. Therefore, men could have perceived this condition as a threat to their gender identity, consequently undermining their performance, as occurs with women in stereotype threat situations (Schmader et al., 2008). On the other hand, when women make decisions, they consider all signals of the environment (e.g., Byrne & Worthy, 2016; Miller & Ubeda, 2012). In this condition, apparently, there were not any cues or information that threatened women, which could have increased their ability to make decisions. Therefore, the condition could have worked as a safe environment for women, showing that they have sufficient skills to make decisions. Both the plausible threat to men's gender identity and benefits to self-confidence for women could explain the larger gender differences in this condition. The gender differences in the threat condition could have been smaller because although the stereotype threat undermines women's performance, this effect could be decreased due to lower gender inequality in society.

Last, we would like to indicate that FNE seems to have various roles across genders and conditions. The stereotype threat could have triggered FNE in women, interfering in their capacity to achieve a cognitive task. In a similar way, perhaps men in the nonstereotype threat condition felt threatened and had a high FNE, which could have undermined their performance, as occurs with women in stereotype threat situations (Schmader et al., 2008). By

contrast, when there is not a threat, FNE could have worked as a motivation to do well.

4.1 | Limitations, future research, and implications

This research has important limitations that will be addressed in subsequent studies. First, the participants were psychology students, which could undoubtedly produce biased results due to their having knowledge of the processes involved in tasks of this type. To improve the generalizability of the research results, studies based on the general population should be carried out. Second, it is necessary to include an implicit measure of sensitivity to punishment to avoid possible biases. Third, to support that the stereotype threat undermines IGT performance by affecting working memory, we recommend that future researchers consider working memory as a control or moderating variable. Lastly, although incentives could improve or worsen performance (or neither; Camerer & Hogarth, 1999), Fryer et al. (2008) found that women in stereotype threat conditions without financial incentives performed better than women in other conditions. Future research could analyze how real incentives might affect women's performance in IGT under stereotype threat conditions, contributing to this field. Furthermore, we would like to note that our results shed light on how stereotype threat manipulation can affect women, but this result cannot be generalized, as it was not designed to create a threat perception in men. Lastly, although the interaction term Gender × Condition × FNE was statistically significant, caution is warranted in the interpretation of this finding because the effect size was small (Cohen, 1988). In this sense, and given that it is a preliminary finding, it would be convenient for future studies could replicate it, strengthening the preliminary findings with a greater significance and effect size.

Despite the fact that literature on decision-making shows gender differences explained by stress or anxiety, the influence of psychosocial variables on this process has scarcely been studied. The results highlight the importance of the stereotype threat and FNE in decision-making as being able to explain gender differences in this process. Overall, this research raises important social psychology insights, highlighting the importance of women's identities and how they affect the individual process of decision-making. Research has shown that the stereotype threat affects women's goals or objectives, undermining their confidence in achieving their professional goals (von Hippel et al., 2011). If the FNE is added to these consequences, women could have more difficulty reaching their objectives, given that they would focus their attention on signals that indicate failure. The effects of both variables spill over into the decision-making domain and could affect any decision that involves some risk. Indeed, this may be extrapolated to decisions that women might consider risky in everyday life, like leaving or staying with their partner, choosing one career over another, or choosing to accept a promotion, and so on.

It is necessary to establish an effective intervention to reduce the concerns caused by the stereotype threat, and this finding can be generalized to other domains (Shapiro et al., 2013). Organizations' implementation of family-work conciliation policies does not improve the family–work conflict, perhaps because these policies do not address the threat of stereotype (Miller, 2019). It is essential to decrease the occurrence of stereotype threats through safe environments in which women do not feel threatened (Spencer et al., 2016) and, therefore, do not feel overloaded or pressured when making decisions. Holding onto a stereotype threat may handicap women by limiting their capacities to make decisions, legitimizing gender inequality in society. In sum, the findings of this research highlight the importance of considering the psychosocial variables of an environment and how their effects can spill over into women's decisions. These effects could be decreased if they are considered in the educational field in terms of creating safe environments. Rosenthal (1994) showed that teachers' beliefs that boys are better than girls at math undermine girls' learning, as these teachers generate threatening situations. An egalitarian belief is necessary to avoid this type of situation. Teachers may show students role models who have overcome these stereotypes, such as Maryam Mirzakhani, the first woman to be awarded a Fields Medal in mathematics. Moreover, they can teach students about the stereotype threat, which has been shown to decrease its effects (Johns et al., 2005). Likewise, it has been found that several educational methods, such as collaboration, can decrease the stereotype threat. Students who complete tests collaboratively show better performance than those who complete the same tests individually (Pociask & Rajaram, 2014). In this sense, Pociask and Rajaram (2014) found that the effects of completing the collaborative test spilled over to a second individual test for girls. Thus, addressing the stereotype threat in education can help keep it from affecting girls' behavior for the rest of their lives.

ACKNOWLEDGEMENTS

This research was supported by the Spanish Ministry of Economy, Industry and Competitiveness for the project "New Forms of Gender Violence: Risk and Protector Factors for Psychological Well-Being" (PSI-2017-84703-R [MINECO/AEI/FEDER/UE]).

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest, given that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ENDNOTES

[1] To determine that this effect size was sufficiently acceptable we referred to previous literature. Carr and Steele (2010) tested the effect of the gender stereotype threat on decision-making, and obtained effect sizes of $f = .10$ and $f = .12$. Moreover, a review of social psychology studies (Richard et al., 2003) found an average effect size of .21 in general, and an average of .12 for gender differences. Lastly, Cohen (1969) defined a value of $f = .25$ as a medium effect.

[2] The power was calculated using G*Power, obtaining a value of 99% and an effect size of $f = .44$.

[3] Under the nonstereotype threat condition, women scored higher on the IGT than men: $t(46) = -3.98$, $p < .001$, 95 CI [−40.44, −13.26], $d = 1.09$, ($M_{women} = 37.85$, $SD = 20.46$; $M_{men} = 11.00$, $SD = 27.96$).

[4] Simple slope tests showed significant differences between slopes (2) and (4), t = 2.47, p = .02, but not between (2) and (3), t = 0.77, p = .44, or (3) and (4), t = 1.94, p = .06.

REFERENCES

Bagneux, V., Thomassin, N., Gonthier, C., & Roulin, J. L. (2013). Working memory in the processing of the Iowa Gambling Task: An individual differences approach. *PLoS One*, *8*(11), e81498. https://doi.org/10.1371/journal.pone.0081498

Baumeister, R. F., Twenge, J. M., & Nuss, C. K. (2002). Effects of social exclusion on cognitive processes: Anticipated aloneness reduces intelligent thought. *Journal of Personality and Social Psychology*, *83*(4), 817–827. https://doi.org/10.1037/0022-3514.83.4.817

Bechara, A. (2005). Decision making, impulse control and loss of willpower to resist drugs: A neurocognitive perspective. *Nature Neuroscience*, *8*(11), 1458–1463. https://doi.org/10.1038/nn1584

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1–3), 7–15. https://doi.org/10.1016/0010-0277(94)90018-3

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*(3), 295–307. https://doi.org/10.1093/cercor/10.3.295

Bechara, A., Tranel, D., & Damasio, H. (2000). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain*, *123*(11), 2189–2202. https://doi.org/10.1093/brain/123.11.2189

Beilock, S. L., Rydell, R. J., & McConnell, A. R. (2007). Stereotype threat and working memory: Mechanisms, alleviation, and spillover. *Journal of Experimental Psychology: General*, *136*(2), 256–276. https://doi.org/10.1037/0096-3445.136.2.256

Biolcati, R. (2017). The role of self-esteem and fear of negative evaluation in compulsive buying. *Frontiers in Psychiatry*, *8*, 1–8. https://doi.org/10.3389/fpsyt.2017.00074

Bishara, A. J., Pleskac, T. J., Fridberg, D. J., Yechiam, E., Lucas, J., Busemeyer, J. R., Finn, P. R., & Stout, J. C. (2009). Similar processes despite divergent behavior in two commonly used measures of risky decision making. *Journal of Behavioral Decision Making*, *22*(4), 435–454. https://doi.org/10.1002/bdm.641

Brown, K. W., Weinstein, N., & Creswell, J. D. (2012). Trait mindfulness modulates neuroendocrine and affective responses to social evaluative threat. *Psychoneuroendocrinology*, *37*(12), 2037–2041. https://doi.org/10.1016/j.psyneuen.2012.04.003

Buelow, M. T., & Barnhart, W. R. (2018). Test–retest reliability of common behavioral decision making tasks. *Archives of Clinical Neuropsychology*, *33*(1), 125–129. https://doi.org/10.1093/arclin/acx038

Buelow, M. T., & Blaine, A. L. (2015). The assessment of risky decision making: A factor analysis of performance on the Iowa Gambling Task, Balloon Analogue Risk Task, and Columbia Card Task. *Psychological Assessment*, *27*(3), 777–785. https://doi.org/10.1037/a0038622

Buelow, M. T., & Cayton, C. (2020). Relationships between the big five personality characteristics and performance on behavioral decision making tasks. *Personality and Individual Differences*, *160*, 109931. https://doi.org/10.1016/j.paid.2020.109931

Byrne, K. A., & Worthy, D. A. (2016). Toward a mechanistic account of gender differences in reward-based decision-making. *Journal of Neuroscience, Psychology, and Economics*, *9*(3–4), 157–168. https://doi.org/10.1037/npe0000059

Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, *19*(1–3), 7–42. https://doi.org/10.1023/A:1007850605129

Carr, P. B., & Steele, C. M. (2009). Stereotype threat and inflexible perseverance in problem solving. *Journal of Experimental Social Psychology*, *45*(4), 853–859. https://doi.org/10.1016/j.jesp.2009.03.003

Carr, P. B., & Steele, C. M. (2010). Stereotype threat affects financial decision making. *Psychological Science*, *21*(10), 1411–1416. https://doi.org/10.1177/0956797610384146

Castellà, J., & Pérez, J. (2004). Sensitivity to punishment and sensitivity to reward and traffic violations. *Accident Analysis & Prevention*, *36*(6), 947–952. https://doi.org/10.1016/j.aap.2003.10.003

Chaffee, K. E., Lou, N. M., Noels, K. A., & Katz, J. W. (2019). Why don't "real men" learn languages? Masculinity threat and gender ideology suppress men's language learning motivation. *Group Processes & Intergroup Relations*, *23*(00), 1–18. https://doi.org/10.1177/1368430219835025

Cohen, J. (1969). *Statistical power analysis for the behavioural sciences*. New York: Academic Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Danaher, K., & Crandall, C. S. (2008). Stereotype threat in applied settings re-examined. *Journal of Applied Social Psychology*, *38*(6), 1639–1655. https://doi.org/10.1111/j.1559-1816.2008.00362.x

Davis, C., Patte, K., Tweed, S., & Curtis, C. (2007). Personality traits associated with decision-making deficits. *Personality and Individual Differences*, *42*(2), 279–290. https://doi.org/10.1016/j.paid.2006.07.006

Dawson, J. F. (2014). Moderation in management research: What, why, when, and how. *Journal of Business and Psychology*, *29*(1), 1–19. https://doi.org/10.1007/s10869-013-9308-7

Dawson, J. F., & Richter, A. W. (2006). Probing three-way interactions in moderated multiple regression: Development and application of a slope difference test. *Journal of Applied Psychology*, *91*(4), 917–926. https://doi.org/10.1037/0021-9010.91.4.917

Deakin, J., Aitken, M., Robbins, T., & Sahakian, B. J. (2004). Risk taking during decision-making in normal volunteers changes with age. *Journal of the International Neuropsychological Society*, *10*(4), 590–598. https://doi.org/10.1017/S1355617704104104

Dekkers, T. J., Agelink van Rentergem, J. A., Huizenga, H. M., Raber, H., Shoham, R., Popma, A., & Pollak, Y. (2018). Decision-making deficits in ADHD are not related to risk seeking but to suboptimal decision-making: Meta-analytical and novel experimental evidence. *Journal of Attention Disorders*, *25*(4), 486–501. https://doi.org/10.1177/1087054718815572

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, *130*(3), 355–391. https://doi.org/10.1037/0033-2909.130.3.355

Duarte, N. A., Woods, S. P., Rooney, A., Atkinson, J. H., & Grant, I. (2012). Working memory deficits affect risky decision-making in methamphetamine users with attention-deficit/hyperactivity disorder. *Journal of Psychiatric Research*, *46*(4), 492–499. https://doi.org/10.1016/j.jpsychires.2012.01.006

Evans, K. L., & Hampson, E. (2015). Sex differences on prefrontally-dependent cognitive tasks. *Brain and Cognition*, *93*, 42–53. https://doi.org/10.1016/j.bandc.2014.11.006

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. https://doi.org/10.1037/pspa0000163

Fryer, R. G. Jr., Levitt, S. D., & List, J. A. (2008). Exploring the impact of financial incentives on stereotype threat: Evidence from a pilot study. *American Economic Review*, *98*(2), 370–375. https://doi.org/10.1257/aer.98.2.370

Gallego, M. J., Botella, C., Quero, S., Baños, R. M., & García-Palacios, A. (2007). Psychometric Proprieties of the Brief Version of the Fear of Negative Evaluation Scale (BFNE) in a clinical population. *Revista de Psicopatología Y Psicología Clínica*, *12*(3), 163–176. https://doi.org/10.5944/rppc.vol.12.num.3.2007.4042

Garrido-Chaves, R., Perez-Alarcón, M., Perez, V., Hidalgo, V., Pulopulos, M. M., & Salvador, A. (2020). FRN and P3 during the Iowa Gambling Task: The importance of gender. *Psychophysiology*, *58*(3), e13734. https://doi.org/10.1111/psyp.13734

Hardy, D. J., Hinkin, C. H., Levine, A. J., Castellon, S. A., & Lam, M. N. (2006). Risky decision making assessed with the gambling task in adults with HIV. *Neuropsychology*, *20*(3), 355–360. https://doi.org/10.1037/0894-4105.20.3.355

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford Publications.

Ho, A. K., & Sidanius, J. (2010). Preserving positive identities: Public and private regard for one's ingroup and susceptibility to stereotype threat. *Group Processes & Intergroup Relations*, *13*(1), 55–67. https://doi.org/10.1177/1368430209340910

Hundt, N. E., Nelson-Gray, R. O., Kimbrel, N. A., Mitchell, J. T., & Kwapil, T. R. (2007). The interaction of reinforcement sensitivity and life events in the prediction of anhedonic depression and mixed anxiety-depression symptoms. *Personality and Individual Differences*, *43*(5), 1001–1012. https://doi.org/10.1016/j.paid.2007.02.021

Inzlicht, M., & Kang, S. K. (2010). Stereotype threat spillover: How coping with threats to social identity affects aggression, eating, decision making, and attention. *Journal of Personality and Social Psychology*, *99*(3), 467–481. https://doi.org/10.1037/a0018951

Johns, M., Schmader, T., & Martens, A. (2005). Knowing is half the battle: Teaching stereotype threat as a means of improving women's math performance. *Psychological Science*, *16*(3), 175–179. https://doi.org/10.1111/j.0956-7976.2005.00799.x

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, *47*(2), 263–292. https://doi.org/10.1142/9789814417358_0006

Katz, B. A., Matanky, K., Aviram, G., & Yovel, I. (2020). Reinforcement sensitivity, depression and anxiety: A meta-analysis and meta-analytic structural equation model. *Clinical Psychology Review*, *77*, 101842. https://doi.org/10.1016/j.cpr.2020.101842

Kovács, I., Richman, M. J., Janka, Z., Maraz, A., & Andó, B. (2017). Decision making measured by the Iowa Gambling Task in alcohol use disorder and gambling disorder: A systematic review and meta-analysis. *Drug and Alcohol Dependence*, *181*, 152–161. https://doi.org/10.1016/j.drugalcdep.2017.09.023

Leary, M. R. (1983). A brief version of the Fear of Negative Evaluation Scale. *Personality and Social Psychology Bulletin*, *9*(3), 371–375. https://doi.org/10.1177/0146167283093007

Lejuez, C. W., Aklin, W. M., Zvolensky, M. J., & Pedulla, C. M. (2003). Evaluation of the Balloon Analogue Risk Task (BART) as a predictor of adolescent real-world risk-taking behaviours. *Journal of Adolescence*, *26*(4), 475–479. https://doi.org/10.1016/S0140-1971(03)00036-8

Lighthall, N. R., Sakaki, M., Vasunilashorn, S., Nga, L., Somayajula, S., Chen, E. Y., Samii, N., & Mather, M. (2012). Gender differences in reward-related decision processing under stress. *Social Cognitive and Affective Neuroscience*, *7*(4), 476–484. https://doi.org/10.1093/scan/nsr026

Lozano, L. M., Megías, A., Catena, A., Perales, J. C., Baltruschat, S., & Cándido, A. (2017). Spanish validation of the domain-specific risk-taking (DOSPERT-30) scale. *Psicothema*, *29*(1), 111–118. https://doi.org/10.7334/psicothema2016.132

Maner, J. K., Richey, J. A., Cromer, K., Mallott, M., Lejuez, C. W., Joiner, T. E., & Schmidt, N. B. (2007). Dispositional anxiety and risk-avoidant decision-making. *Personality and Individual Differences*, *42*(4), 665–675. https://doi.org/10.1016/j.paid.2006.08.016

Maresh, E. L., Teachman, B. A., & Coan, J. A. (2017). Are you watching me? Interacting effects of fear of negative evaluation and social context on cognitive performance. *Journal of Experimental Psychopathology*, *8*(3), 303–331. https://doi.org/10.5127/jep.059516

Meshi, D., Elizarova, A., Bender, A., & Verdejo-Garcia, A. (2019). Excessive social media users demonstrate impaired decision making in the Iowa Gambling Task. *Journal of Behavioral Addictions*, *8*(1), 169–173. https://doi.org/10.1556/2006.7.2018.138

Meyers-Levy, J. (1989). Gender differences in information processing: A selectivity interpretation. In P. Cafferata & A. M. Tybout (Eds.), *Cognitive and affective responses to advertising*. Lexington, MA: Lexington Books.

Miller, A. L. (2019). Stereotype threat as a psychological feature of work–life conflict. *Group Processes & Intergroup Relations*, *22*(2), 302–320. https://doi.org/10.1177/1368430217735578

Miller, L., & Ubeda, P. (2012). Are women more sensitive to the decision-making context? *Journal of Economic Behavior & Organization*, *83*(1), 98–104. https://doi.org/10.1016/j.jebo.2011.06.014

Mrazek, M. D., Chin, J. M., Schmader, T., Hartson, K. A., Smallwood, J., & Schooler, J. W. (2011). Threatened to distraction: Mind-wandering as a consequence of stereotype threat. *Journal of Experimental Social Psychology*, *47*(6), 1243–1248. https://doi.org/10.1016/j.jesp.2011.05.011

Murphy, M., Steele, C., & Gross, J. (2007). Signaling threat: How situational cues affect women in math, science, and engineering. *Psychological Science*, *18*(10), 879–885. https://doi.org/10.1111/j.1467-9280.2007.01995.x

Pociask, S., & Rajaram, S. (2014). The effects of collaborative practice on statistical problem solving: Benefits and boundaries. *Journal of Applied Research in Memory and Cognition*, *3*(4), 252–260. https://doi.org/10.1016/j.jarmac.2014.06.005

Régner, I., Smeding, A., Gimmig, D., Thinus-Blanc, C., Monteil, J. M., & Huguet, P. (2010). Individual differences in working memory moderate stereotype-threat effects. *Psychological Science*, *21*(11), 1646–1648. https://doi.org/10.1177/0956797610386619

Richard, F. D., Bond, C. F. Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, *7*(4), 331–363. https://doi.org/10.1037/1089-2680.7.4.331

Richetin, J., Perugini, M., Adjali, I., & Hurling, R. (2007). The moderator role of intuitive versus deliberative decision making for the predictive validity of implicit and explicit measures. *European Journal of Personality: Published for the European Association of Personality Psychology*, *21*(4), 529–546. https://doi.org/10.1002/per.625

Rosenthal, R. (1994). Interpersonal expectancy effects: A 30-year perspective. *Current Directions in Psychological Science*, *3*(6), 176–179. https://doi.org/10.1111/1467-8721.ep10770698

Rydell, R. J., & Boucher, K. L. (2017). Stereotype threat and learning. *Advances in Experimental Social Psychology*, *56*, 81–129. https://doi.org/10.1016/bs.aesp.2017.02.002

Rydell, R. J., van Loo, K. J., & Boucher, K. L. (2014). Stereotype threat and executive functions: Which functions mediate different threat-related outcomes? *Personality and Social Psychology Bulletin*, *40*(3), 377–390. https://doi.org/10.1177/0146167213513475

Santos-Ruiz, A., García-Ríos, M. C., Fernández-Sánchez, J. C., Pérez-García, M., Muñoz-García, M. A., & Peralta-Ramírez, M. I. (2012). Can decision-making skills affect responses to psychological stress in healthy women? *Psychoneuroendocrinology*, *37*(12), 1912–1921. https://doi.org/10.1016/j.psyneuen.2012.04.002

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, *85*(3), 440–452. https://doi.org/10.1037/0022-3514.85.3.440

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

Shapiro, J. R., Williams, A. M., & Hambarchyan, M. (2013). Are all interventions created equal? A multi-threat approach to tailoring stereotype threat interventions. *Journal of Personality and Social Psychology*, *104*(2), 277–288. https://doi.org/10.1037/a0030461

Singh, V., Schiebener, J., Müller, S. M., Liebherr, M., Brand, M., & Buelow, M. T. (2020). Country and sex differences in decision making under uncertainty and risk. *Frontiers in Psychology*, *11*, 486. https://doi.org/10.3389/fpsyg.2020.00486

Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology*, *67*, 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. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*(6), 613–629. https://doi.org/10.1037/0003-066X.52.6.613

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797–811. https://doi.org/10.1037/0022-3514.69.5.797

Stone, J. (2002). Battling doubt by avoiding practice: The effects of stereotype threat on self-handicapping in White athletes. *Personality and Social Psychology Bulletin*, *28*(12), 1667–1678. https://doi.org/10.1177/014616702237648

Torrubia, R., Ávila, C., Moltó, J., & Caseras, X. (2001). The Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ) as a measure of Gray's anxiety and impulsivity dimensions. *Personality and Individual Differences*, *31*(6), 837–862. https://doi.org/10.1016/S0191-8869(00)00183-5

van den Bos, R., Homberg, J., & de Visser, L. (2013). A critical review of sex differences in decision-making tasks: Focus on the Iowa Gambling Task. *Behavioural Brain Research*, *238*(1), 95–108. https://doi.org/10.1016/j.bbr.2012.10.002

von Hippel, C., Issa, M., Ma, R., & Stokes, A. (2011). Stereotype threat: Antecedents and consequences for working women. *European Journal of Social Psychology*, *41*(2), 151–161. https://doi.org/10.1002/ejsp.749

Weller, J. A., Levin, I. P., & Bechara, A. (2010). Do individual differences in Iowa Gambling Task performance predict adaptive decision making for risky gains and losses? *Journal of Clinical and Experimental Neuropsychology*, *32*(2), 141–150. https://doi.org/10.1080/13803390902881926

Woodrow, A., Sparks, S., Bobrovskaia, V., Paterson, C., Murphy, P., & Hutton, P. (2019). Decision-making ability in psychosis: A systematic review and meta-analysis of the magnitude, specificity and correlates of impaired performance on the Iowa and Cambridge Gambling Tasks. *Psychological Medicine*, *49*(1), 32–48. https://doi.org/10.1017/S0033291718002660

Zhang, F., Xiao, L., & Gu, R. (2017). Does gender matter in the relationship between anxiety and decision-making? *Frontiers in Psychology*, *8*, 1–8. https://doi.org/10.3389/fpsyg.2017.02231

AUTHOR BIOGRAPHIES

Laura Villanueva-Moya is a PhD candidate at the University of Granada (Spain). Her research interests include gender differences in decision-making.

Francisca Expósito is a Full Professor of the Faculty of Psychology at the University of Granada (Spain). She specializes in gender discrimination, sexism, and intimate partner violence.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Villanueva-Moya L, Expósito F. Gender differences in decision-making: The effects of gender stereotype threat moderated by sensitivity to punishment and fear of negative evaluation. *J Behav Dec Making*. 2021;1–12. https://doi.org/10.1002/bdm.2239