AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

2026, Vol. 155, No. 3, 819–838
https://doi.org/10.1037/xge0001886

Explaining Gender Differences in Negotiation: A Close Replication of Amanatullah and Morris (2010)

Jens Mazei[1], Julia B. Bear[2], Rebecca Schaumberg[3], and Joachim Hüffmeier[1, 4]

[1] Department of Psychology, TU Dortmund University
[2] College of Business, Stony Brook University
[3] ESMT Berlin
[4] Department of Sport Science and Physical Education, University of Agder

Amanatullah and Morris (2010) advanced and tested central propositions from the field of gender differences in negotiation. They observed that women more readily anticipated backlash and requested lower salaries than men, yet only when they negotiated for themselves and not when they advocated for others (i.e., interaction effects). These insights are key building blocks of current theory explaining why and when women and men differ in salary negotiations. However, the research by Amanatullah and Morris had low statistical power and never received a close replication. Moreover, other conceptually related research has revealed divergent results. Thus, we conducted a close replication (total $N = 517$) of the seminal research by Amanatullah and Morris. We did not observe a Gender × Advocacy interaction on anticipated backlash and salary requests. We only observed a main effect of gender on salary requests, which was mediated by anticipated backlash. Moreover, consistent with the original study, women (as compared to men) rated their negotiation style as less competitive, but only if they negotiated for themselves (and not when they advocated for others), and there were no effects regarding negotiators' chosen verbal statements. We discuss the relevance of these novel insights for theory and research on gender differences in negotiation, as well as its implications for women's pay and workplace success.

Public Significance Statement

We ran a replication study that closely mirrored an influential past study, which found that women (as compared to men) ask for lower salaries because they more readily think that other people would evaluate them negatively. The past study also suggested that these gender differences occur only when women and men negotiate for themselves, but not when they negotiate on behalf of a friend. In our replication study, we did not observe that gender differences depended on whether they negotiated for themselves or for a friend. We only observed that women (as compared to men) more readily anticipated negative evaluations and, as a result, asked for lower salaries.

Keywords: gender, sex, negotiation, bargaining, gender gap

Supplemental materials: https://doi.org/10.1037/xge0001886.supp

Do women still display less assertiveness in negotiations than men, leading to lower negotiated outcomes, such as lower salaries? Amanatullah and Morris (2010) found that women, as compared to men, more readily anticipated incurring *backlash* (i.e., negative reactions by others due to stereotype or gender role violations; Rudman & Fairchild, 2004), which led them to request lower salaries. However, this effect was observed only when women negotiated for themselves, not when they advocated for others, which suggests that women do not lack the basic skills to negotiate effectively. Thus, Amanatullah and Morris substantially advanced

This article was published Online First January 15, 2026.

Musawenkosi Donia Saurombe served as action editor.

Jens Mazei https://orcid.org/0000-0003-3579-6857

Earlier versions of this research were presented at the conference of the International Association for Conflict Management as well as the conference of the section on Work and Organizational Psychology of the German Psychological Society. This research was supported by the German Research Foundation (Grant MA 9683/2-1 awarded to Jens Mazei and Joachim Hüffmeier). Moreover, this research was supported by the Wharton Behavioral Lab of the Wharton School of the University of Pennsylvania.

Jens Mazei played a lead role in conceptualization, formal analysis, methodology, writing–original draft, and writing–review and editing. Julia B. Bear played an equal role in conceptualization, methodology, writing–original draft, and writing–review and editing. Rebecca Schaumberg played a lead role in investigation and an equal role in writing–original draft and writing–review and editing. Joachim Hüffmeier played an equal role in conceptualization, methodology, writing–original draft, and writing–review and editing.

Correspondence concerning this article should be addressed to Jens Mazei, Department of Psychology, TU Dortmund University, Emil-Figge-Straße 50, 44227 Dortmund, Germany. Email: jens.mazei@tu-dortmund.de

research on gender differences in negotiation by illuminating how role incongruity (Eagly & Karau, 2002) and backlash (Rudman, 1998) influence women's behavior in negotiations. Given the importance of their findings, it is no surprise that the research by Amanatullah and Morris has attracted considerable attention: Their article has been cited 817 times (Google Scholar: September 24, 2025). Altogether, the research by Amanatullah and Morris plays a central role in the understanding of gender differences in negotiation. Moreover, their findings have implications for research in many domains, as "bargaining is a cardinal illustration of social interaction" (Rubin & Brown, 1975, p. 3) and an activity in the workplace that determines many central outcomes (e.g., one's pay; Thompson, 2009).

Yet, the more central theoretical propositions for a certain field are, the more compelling the underlying evidence should be. In this respect, the study by Amanatullah and Morris (2010) also had clear limitations: With a sample size of $N = 59$ for a 2×2 design, the original study was severely underpowered, given that gender effects in negotiations are typically small or medium sized (e.g., Kugler et al., 2018; Walters et al., 1998; for statistical details, see below). Furthermore, despite its central role for the field and the insufficient statistical power, to the best of our knowledge, there has been no close[1] replication of the research by Amanatullah and Morris. Finally, other conceptually related research has yielded diverging results: For instance, Artz et al. (2018) and Bowles et al. (2007) did not observe gender differences in feared or anticipated backlash. Hence, the actual evidence for a key mechanism that might explain gender differences in negotiation is rather thin. However, it could also be that the mixed results were due to differences in the studies' methodologies and dependent variables, rather than because the underlying propositions are incorrect (for an overview, see below). In summary, it remains unclear whether gender differences in negotiation could be explained definitively by anticipated backlash, which is problematic given the central role of this finding for the field. Thus, we provide a close replication (Brandt et al., 2014; Hüffmeier et al., 2016) of the research by Amanatullah and Morris and thereby contribute needed empirical evidence regarding the theoretical propositions that Amanatullah and Morris advanced.

Notably, in our close replication ($N = 517$), we did not replicate the critical interaction effects reported by Amanatullah and Morris (2010): Gender did not interact with advocacy to predict anticipated backlash and salary requests. Collapsing across advocacy role, however, we found that women generally anticipated backlash more readily than men. Under certain circumstances, women also made lower salary requests than men. Moreover, replicating the original study, women, as compared to men, rated their negotiation style as less competitive, but only if they negotiated for themselves (and not when they advocated for others). Consistent with the original study, we also did not observe effects regarding negotiators' chosen verbal statements.

Theoretical Underpinnings: Role Congruity and Backlash

A common thread that runs through negotiation research and the role of gender is that there is an "overwhelming assortment of contradictory findings" (Rubin & Brown, 1975, p. 172) and that "findings have been highly inconsistent" (Small et al., 2007, p. 601; e.g., Kugler et al., 2018; Stuhlmacher & Walters, 1999). Thus, strong theory is needed to delineate the conditions under which gender differences are expected to emerge or not to emerge and also why a specific pattern is expected (for overviews, see Bowles et al., 2022, 2025; Eagly & Wood, 2013). To this end, role congruity theory (Eagly & Karau, 2002) and its social role foundations (Eagly & Steffen, 1984) have gained prominence as an integrative account of gender differences in negotiation (e.g., Kugler et al., 2018; Nohe et al., 2022).

According to socially shared and normative gender roles that play a key role in this account (Eagly, 1987), women are *communal* (e.g., nurturing or polite), and men are *agentic* (e.g., aggressive or assertive; for a recent meta-analysis, see Eagly et al., 2020; see also Rudman et al., 2012). Gender roles influence people's behavior because they capture what a society regards as "desirable" behavior (for a recent cross-cultural investigation, see Bosson et al., 2022). As such, not adhering to one's gender role can be penalized by others (Amanatullah & Tinsley, 2013; Bowles et al., 2007). Such penalties, or negative interpersonal reactions, are denoted as *backlash* (Rudman, 1998). Yet agentic behaviors facilitate the attainment of individual gains in negotiations (Kray et al., 2001; Kulik & Olekalns, 2012). Thus, women's communal gender role is often incongruent with agentic behavior that can lead to economic negotiation success (Mazei et al., 2015; Stuhlmacher & Linnabery, 2013). Finally, however, it is worth noting that not all assertive behaviors that can be interpreted as a deviation from women's gender role result in backlash: For instance, a meta-analysis by Williams and Tiedens (2016) showed that only explicit forms of dominance (but not implicit forms) result in backlash.

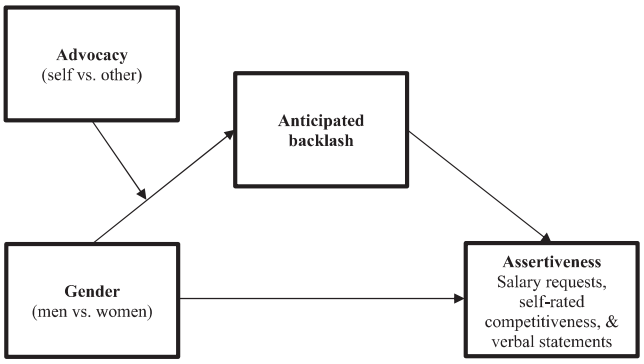
The Seminal Research by Amanatullah and Morris (2010)

Amanatullah and Morris (2010) made a distinctive contribution to the literature by advancing and testing central propositions based on role congruity and backlash (Eagly & Karau, 2002; Rudman & Fairchild, 2004). The authors proposed that women recognize the risk of incurring backlash and, as a result, negotiate less assertively. Negotiators can act assertively in multiple ways, for instance, by making particularly high salary requests (e.g., Hüffmeier et al., 2014) or by using verbal statements that are commonly perceived as assertive (e.g., Toosi et al., 2019). A strength of Amanatullah and Morris's study is that they considered multiple facets of negotiator assertiveness (see also Figure 1).

Moreover, Amanatullah and Morris (2010) examined whether gender differences vanish when women negotiate on behalf of someone else (i.e., other-advocacy). Negotiating assertively as an advocate for others represents a nurturing act, which matches women's communal gender role (e.g., Amanatullah & Tinsley, 2013; see also Bear & Babcock, 2017; Bowles et al., 2005). Thus,

[1] Depending on the degree of similarity between a replication and an original study, a replication can be classified on a continuum ranging from "exact replication" to "very far replication" (LeBel et al., 2018, p. 394). Given that a replication cannot be a fully "exact" copy of an original study (e.g., Nelson et al., 2018), we use the term "close replication" (Brandt et al., 2014; also known as direct replication; Schmidt, 2009). Irrespective of the terminology, our replication adhered in its methodology as much as possible to the methodology of Amanatullah and Morris (2010; see Brandt et al., 2014; Hüffmeier et al., 2016; LeBel et al., 2019).

Figure 1
Conceptual Model



Note. Gender represents the independent variable, anticipated backlash represents the mediator, and advocacy represents the moderator. Assertiveness represents the dependent variable and is captured in different ways: Negotiators can act assertively by requesting a high salary (Hüffmeier et al., 2014) and by utilizing statements that are perceived as "assertive" (Amanatullah & Morris, 2010). In turn, the more assertive a negotiator has been, the more competitive they should self-rate their behavior afterward.

women would not be expected to anticipate backlash when advocating for others. Altogether, Amanatullah and Morris (p. 258) "sought to extend prior theory and research primarily through the proposal of anticipated backlash as the mechanism leading women to hedge their assertive behavior in negotiations and more fully developing the rationale for why advocacy moderates this relationship." In this way, the authors highlighted women's strategic impression management as the cause of gender differences in negotiation, as opposed to alternative explanations, such as a lack of basic skill or interest among women to negotiate (Amanatullah & Morris, 2010).

Amanatullah and Morris (2010) reported a laboratory experiment with $N = 59$ participants ($n = 31$ men and $n = 28$ women). Participants first indicated plans and anticipations regarding a salary negotiation that they believed they would conduct via a computer with another participant. During the negotiation, which was actually computer-simulated, participants negotiated their own (self-advocacy condition) or a friend's (other-advocacy condition) salary by making requests and exchanging verbal statements with their alleged counterpart in up to five rounds of negotiation. Afterward, participants indicated how competitively they had behaved during the negotiation.

A first key result in Amanatullah and Morris's (2010) study was that women did not differ from men in their reservation points, target points, and their anticipated opening offers. These findings are important because if women were simply less interested in achieving economic success, for instance, then they would be expected to enter negotiations with a lower target salary. Moreover, as women and men did not differ in the indicated target points—the salaries that they aimed for—Amanatullah and Morris's study suggests that a moderating effect of advocacy is unlikely to especially "motivate" women because of a potentially heightened interdependent self-construal (see below; Bowles et al., 2005).

However, as predicted, self-advocating women more readily anticipated backlash than self-advocating men as well as compared to both other-advocating women and other-advocating men. These findings are key because they suggest a mechanism that might underlie gender differences in negotiation. Furthermore, and importantly, self-advocating women requested lower salaries than the three comparison groups,[2] and this interaction was, in fact, mediated by anticipated backlash. This is arguably the most important finding in Amanatullah and Morris's (2010) original study, because they were able to trace gender differences in assertive negotiation behavior (see also Walters et al., 1998) back to the mechanism of anticipated backlash. In addition to these intriguing results, self-advocating women reported that they had behaved less competitively, which is in line with the proposition that women are aware of the risk of incurring backlash and their corresponding behavioral choices (Amanatullah & Morris, 2010). Finally, there were no effects regarding the exchanged verbal statements. Please recall, however, that not all behaviors that constitute a role deviation for women result in backlash (e.g., Williams & Tiedens, 2016). As such, it could also be expected that gender differences emerge only for certain kinds of assertive acts in a negotiation.

Despite its strengths, however, the study by Amanatullah and Morris (2010) is also limited. The sample size was small, which leads to issues with statistical power. Meta-analyses have revealed that gender effects in negotiation are typically small in magnitude but can be larger in situations characterized by role incongruity (Kugler et al., 2018; Mazei et al., 2015), such as when women advocate for themselves. For explanatory purposes, let us assume that approximately half of the participants in the original study by Amanatullah and Morris were in the self-advocacy condition (i.e., $n = 14$ women and $n = 16$ men). To examine the statistical power of the original study to detect the reported main, moderator, and mediator effects, we conducted power-determination analyses (Giner-Sorolla et al., 2024). Calculations using G*Power (Faul et al., 2007) revealed that the achieved statistical power for a t test (two-tailed) to detect an effect size of $d = 0.50$ with the noted group sample sizes was $1 - \beta = .26$ (for a one-tailed test: $1 - \beta = .38$). Similarly, the achieved statistical power for an analysis of variance ("fixed effects, special, main effects and interactions" in G*Power) to detect an *interaction* of $f = 0.25$ with $N = 59$ was $1 - \beta = .47$. Obviously, the achieved statistical power to detect smaller effects was even lower (e.g., $1 - \beta = .12$ for an interaction of $f = 0.10$). All of these estimates are below the conventionally desired level for statistical power of $1 - \beta = .80$ and also the more ideal level of $1 - \beta = .95$ (Giner-Sorolla et al., 2024).

The statistical power to detect the reported *mediation* was also small: Conducting a Monte Carlo power analysis using mc_power_med (Schoemann et al., 2017) for a mediation in which all paths are assumed to have a correlation of $r = .30$ with $N = 59$ revealed a statistical power of $1 - \beta = .27$. When assuming a correlation of zero only between the independent variable and the dependent variable, as might be expected in the case of a "full" mediation, the statistical power was $1 - \beta = .43$.[3] These analyses reveal that the original study was underpowered.

[2] Self-advocating women differed significantly from self-advocating men, other-advocating women, and other-advocating men when compared with t tests. Yet, the interaction effect in a regression analysis regarding salary requests in the first round of negotiation had a p value of "<.10" (Amanatullah & Morris, 2010, p. 261).

[3] For these analyses, the standard deviation for each variable was set to 1 (as would be the case with standardized variables), the number of replications was set to 5,000, the "Monte Carlo Draws per Rep" were set to 20,000, the random seed was set to 1,234, and the confidence interval (CI) was 95%.

Given the relevance but also the limitations of the original study, the insights obtained by Amanatullah and Morris (2010) should be carefully evaluated and potentially extended or updated. Surprisingly, however, the study by Amanatullah and Morris has (to the best of our knowledge) never been subject to a close replication. Moreover, findings from other related research call into question whether anticipated backlash is in fact the key mechanism driving gender differences in negotiation.

Findings From Conceptually Related Research

As shown in Table 1, several studies have examined conceptually related questions, but all of these studies used markedly different methodology (and even different dependent variables). Thus, it is currently not possible to draw strong conclusions as to the validity of the findings by Amanatullah and Morris (2010) based on other extant research. Dreber et al. (2022) observed, for instance, that women believed to incur more social costs than men when requesting a high salary, but these beliefs did not translate into gender differences in salary requests. Similarly, Babcock et al. (2006) examined whether a gender difference in *the propensity to initiate negotiations* was mediated by apprehension (e.g., "I feel anxious when I have to ask for something I want"; p. 247). Women reported greater apprehension than men, but this factor was unrelated to the propensity to initiate negotiations, thus ruling it out as a mediator. Furthermore, Bowles et al. (2007) did not observe gender differences in *anticipated backlash* (e.g., the extent to which people expected another person to be unwilling to employ or work together with them) as a response to asking for a high salary and a bonus. Likewise, Artz et al. (2018) did not find that women, as compared to men, were more likely to report greater concern about "negative effects on my relationship with my manager/employer" (p. 636) as the reason that they had *not attempted to get a raise or a promotion*. These findings raise doubts about the explanatory value of anticipated backlash for gender differences in salary negotiations.

Yet, results reported by Toosi et al. (2019) are more in line with those of the original study. Building on the paradigm as developed by Amanatullah and Morris (2010), Toosi et al. realized a 2 (women vs. men) × 2 (White American vs. Asian American) design in an online study. As such, this study also provided important insights into the relevance of people's intersectionality (e.g., Desai & Gunia, 2023; Leigh & Desai, 2023; see also our Supplemental Material and the Discussion section). Among White Americans, women requested lower salaries than men, which was mediated by one item (but not the two items as in the original study) tapping into anticipated backlash (among Asian Americans, there were no gender differences). This study did not manipulate advocacy, however, thus limiting the scope of its examination of gender and anticipated backlash. Similarly, Ma et al. (2024) found that women anticipated backlash more readily than men, which helped to explain gender differences in requests. Yet again, this study did not manipulate advocacy, and the pattern involving anticipated backlash seemingly replicated in only one of two similar supplementary studies (Ma et al., 2024).

In still other primary (Bear & Babcock, 2017; Bowles et al., 2005; Kouchaki & Kray, 2018) and meta-analytic work (Mazei et al., 2015; Nohe et al., 2022), researchers did observe a moderating effect for advocacy. Yet, the question of why advocacy acts as a moderator has not been fully clarified. Specifically, findings by Bowles et al. (2005) were more consistent with an explanation based on an interdependent self-construal among women: Bowles et al. found that women were not strikingly less successful than men when they negotiated for themselves (as would be expected from an account based on role incongruity and backlash), yet, at least descriptively, their findings appear to suggest that other-advocating women outperformed men, which may point to the "motivating" force of an activated interdependent self-construal (Bowles et al., 2005). Thus, the specific pattern of the interaction differed between the studies by Bowles et al. (Figure 2, p. 959) and Amanatullah and Morris (2010, Figure 2, p. 262).

Finally, findings by Bear and Babcock (2017) more closely resembled the interaction pattern observed by Amanatullah and Morris (2010). However, Bear and Babcock, like Bowles et al. (2005), examined economic outcomes and performance aspirations but not the assertiveness of salary requests. Kouchaki and Kray (2018) generally focused on unethical behaviors, and they observed interactions between gender and advocacy that also resembled the findings by Amanatullah and Morris, but they did not observe a significant interaction when analyzing the perceived appropriateness of competitive behaviors (e.g., making extreme requests). We are not aware of any additional study that has examined anticipated backlash as the mediator of the moderating effect for advocacy (the meta-analyses by Mazei et al., 2015; Nohe et al., 2022, did not examine mediators), and as noted above, the related evidence regarding anticipated or feared backlash is rather mixed.

It is important to stress again, however, that all of the described related studies differed from the research by Amanatullah and Morris (2010) in relevant ways, including the measures used to capture the mediator, the experimental design, and even the dependent variable of interest (see Table 1). For instance, whereas Amanatullah and Morris examined "anticipated backlash" (see also their specific items provided in their Method section), Babcock et al. (2006) measured "apprehension." The latter appears to capture more of an emotional reaction, which may or may not be different from the anticipated reactions studied by Amanatullah and Morris. As another example, Artz et al. (2018) focused on negative effects on the relationship to a counterpart, which again is not fully consistent with Amanatullah and Morris's construct. Therefore, it remains unclear whether the divergent findings result from the notable differences between studies or whether they indicate that the underlying theoretical propositions need to be revised. Clearly, a replication that closely adheres to the original research is needed (e.g., LeBel et al., 2019).

The Need for a Close Replication: The Current Research

Taken together, a close replication is needed for multiple reasons (see Giner-Sorolla et al., 2018): First, the research by Amanatullah and Morris (2010) has clear theoretical relevance, as it examined key propositions from the currently dominant account of gender differences in negotiation. In fact, as of today, "the leading explanation for women's lack of assertiveness in negotiating for higher compensation is social backlash" (Bowles et al., 2019, p. 1646). Second, the actual, extant evidence for these propositions is rather thin, as related research often has yielded diverging results (but note the differences between studies; Table 1). Third, the original study was underpowered. Fourth, the original study has had a considerable impact as it is frequently cited in research on negotiations and other

Table 1
Overview of Relevant Past Studies

Study	Study type	Mediator	Advocacy studied	Type of key DV	*N*/*K*
Amanatullah and Morris (2010)	Primary study (journal publication)	Anticipated backlash	Yes	• Assertive behavior during negotiation • Self-rated competitiveness	59
Artz et al. (2018)	Primary study (journal publication)	Concern about negative effects on the relationship with the manager or employer	No	• Initiation • Economic outcomes	4,582
Babcock et al. (2006)	Primary study (book chapter)	Apprehension (along with recognition of opportunities and entitlement)	No	• Initiation	227
Bear and Babcock (2017)	Primary study (journal publication)	Perception of fit	Yes (imagining being an advocate)	• Economic outcomes • Performance aspirations	112 and 996
Bowles et al. (2005)	Primary study (journal publication)	Not studied	Yes	• Economic outcomes • Target points	170 and 238
Bowles et al. (2007)	Primary study (journal publication)	Anticipated backlash (along with nervousness and gender identity)	No	• Initiation	341
Dreber et al. (2022)	Primary study (journal publication)	Perceived social costs (along with several others)	No	• Assertive behavior during negotiation • Initiation • Economic outcomes	792
Kouchaki and Kray (2018)	Primary study (journal publication)	Anticipatory guilt (along with the perceived social expectation to deceive and moral disengagement)	Yes	• Unethical behavior • Appropriateness of competitive behavior	159–395
Ma et al. (2024)	Primary study (journal publication)	Anticipated backlash (along with relational orientation)	No	• Assertive behavior during negotiation • Impasses • Economic outcomes	395, 595, and 391
Mazei et al. (2015)	Meta-analysis (journal publication)	Not studied	Yes	• Economic outcomes	123 ES
Nohe et al. (2022)	Meta-analysis (journal publication)	Not studied	Yes	• Unethical behavior	116 ES
Toosi et al. (2019)	Primary study (journal publication)	Anticipated backlash	No	• Assertive behavior during negotiation	145 and 980

Note. *N* refers to sample sizes for primary studies, whereas *K* refers to sample sizes for meta-analyses. DV = dependent variable; ES = effect size.

topics (e.g., Brescoll, 2011; London et al., 2012). Fifth and finally, the original study has practical relevance: Gender differences in negotiation contribute to prevailing gender inequalities in the workplace and beyond, such as the gender pay gap (e.g., Blau & Kahn, 2017; Kennedy & Kray, 2015). Therefore, if anticipated backlash is, in fact, the main driving force of gender differences in negotiation, it would be advisable to devise interventions that address women's anticipation of backlash, as well as the underlying tendency of people to respond to women negotiators with backlash (e.g., Amanatullah & Tinsley, 2013). However, if other processes are, in fact, actually more important, alternative interventions might be needed.

Given these reasons, our goal was to provide a registered replication report containing a close replication of the seminal research by Amanatullah and Morris (2010), as "close replications are an important part of cumulative science" (Brandt et al., 2014, p. 217). Accordingly, we aimed to test the same hypotheses put forth by Amanatullah and Morris (see, e.g., p. 260). These hypotheses concern the interactive effects of gender and advocacy as well as anticipated backlash as the underlying mechanism. The interactions are hypothesized to follow a specific pattern, such that self-advocating women are hypothesized to differ from self-advocating men, other-advocating women, and other-advocating men (Amanatullah & Morris, 2010). Following Amanatullah and Morris, we hypothesized:

Hypothesis 1: Self-advocating women more readily anticipate backlash than self-advocating men, other-advocating women, and other-advocating men.

Hypothesis 2: Self-advocating women request a lower salary than self-advocating men, other-advocating women, and other-advocating men.

Hypothesis 3: The interaction between gender and advocacy on salary requests (see Hypothesis 2) is mediated by anticipated backlash.

Hypothesis 4: Self-advocating women rate their negotiation style as less competitive than self-advocating men, other-advocating women, and other-advocating men.

Hypothesis 5: The verbal behavior of self-advocating women is characterized by fewer assertive statements than the verbal behavior of self-advocating men, other-advocating women, and other-advocating men.

Method

Transparency and Openness

To maximize the informative value of our research, we followed Open Science Practices (e.g., Nelson et al., 2018; Nosek et al., 2015) and guidelines for conducting and evaluating close replications (e.g., Brandt et al., 2014; LeBel et al., 2019). We submitted a first version of this article, prior to the data collection, to the *Journal of Experimental Psychology: General* as a proposal for a registered replication report. This version of the article gave a detailed overview of our hypotheses and the method of our close replication (including planned statistical analyses, sample size calculations, and exclusion criteria). This proposal—and the close replication described therein—was then revised further based on issues that were raised during the Stage 1 editorial review process for registered reports and issues that we encountered while preparing and beginning to run the close replication. We then uploaded the revised article to the Open Science Framework. It can be retrieved at https://osf.io/6bfvq/?view_only=5d54007b1dca40b2b21a4eaf1413e7a1 (see the file titled "20230327_RRR_GenderNegotiation.pdf").

Our study materials, data, and syntax are also available as per the above Open Science Framework link (Mazei et al., 2025). Following Simmons et al. (2012; see also Simmons et al., 2011), we fully disclose not only how we determined our sample size but also the realized sample size (including all exclusions), all measures, and all manipulations. We also adhere to the remaining transparency and openness guidelines that concern citation as well as design and analysis transparency.

Independence of Researchers

We hereby confirm that we are independent researchers (Brandt et al., 2014; LeBel et al., 2019). Thus, there is no overlap in authorship with the authors of the original study (see also Hüffmeier et al., 2016). However, we contacted the authors of the original study to inform them about this project to ensure transparency and fairness. Furthermore, to ensure the rigor of our replication, we asked the authors of the original study to share with us the original study materials and to answer our questions about the original methodology (adapted from Brandt et al., 2014). Our questions are also available as per the previously provided link to the Open Science Framework. Answers to our questions are indicated in this article as personal communication with E. T. Amanatullah (E. T. Amanatullah also provided original study materials that helped to answer our questions). We greatly appreciate the original authors' help and openness that enabled us to conduct this close replication.

Overview of General Method and Procedure

We closely followed the protocol of Amanatullah and Morris (2010), with a few necessary exceptions (for an overview, see Supplemental Table S1). Participants engaged in a computer-mediated salary negotiation with a (supposed) participant from another university and were randomly assigned to either negotiate for themselves (self-advocacy) or on behalf of a friend (other-advocacy). Before negotiating, participants answered questions about their plans and expectations, including anticipated backlash. They then negotiated for up to five rounds, exchanging offers and audio-recorded messages. After the negotiation, participants completed questionnaires about their experience (e.g., how competitively they had behaved), followed by comprehension checks and additional measures. As shown in Figure 1, participant gender was the main predictor, advocacy role (self vs. other) was the moderator, and anticipated backlash was the mediator of initial salary requests, which was the primary dependent variable.

Power Analyses and Planned Sample Size

We conducted power analyses to determine the sample size needed for our close replication. To do this, we needed estimates of effect sizes. One option would have been to use effect sizes from Amanatullah and Morris (2010). Reconstructing the cell sizes from

Amanatullah and Morris ($N = 59$), we estimated effect sizes of $d = 0.87$ for the difference between self-advocating women and self-advocating men in anticipated backlash and $d = 1.05$ for the difference between self-advocating women and self-advocating men in salary requests (for these effect size calculations, the formulas provided by Borenstein, 2009, were used).[4] These are large effects, which may reflect the strength of the original paradigm (see Giner-Sorolla et al., 2024). However, with the small cell sizes, the original study was underpowered (for details, see above).

We thus took a different approach to determine the sample size for our close replication. Meta-analyses of gender and negotiation show that differences are typically small to medium in size (Kugler et al., 2018; Mazei et al., 2015; Walters et al., 1998). Based on this evidence, we adopted $d = 0.50$ as a more realistic effect size estimate. Detecting this effect in a two-tailed comparison between two independent groups requires $N = 128$ participants ($\alpha = .05$, $1 - \beta = .80$; Faul et al., 2007). Because our close replication tested an interaction, we followed recommendations to quadruple this number (Giner-Sorolla, 2018; Simonsohn, 2014), yielding a target of $N = 512$ (256 women, 256 men). To allow for technical exclusions, we arranged with the laboratory where we ran our replication to recruit 520 participants.

Because we also studied mediation, we also needed to ensure that our planned sample size was sufficient to detect an indirect effect. To this end, we ran Monte Carlo power analyses using mc_power_med (Schoemann et al., 2017). With $N = 512$, power would exceed .90 to detect an indirect effect when the *a*-path was set to $r = .15$ (i.e., half of an assumed medium-sized effect, given that Amanatullah & Morris, 2010, examined a mediated moderation—a key predictor in their model was the interaction between gender and advocacy; see pp. 261–262) and the *b*-path was set to $r = .30$ (a medium effect), across plausible correlations between the independent and dependent variables.[5] Thus, our sample size met recommended thresholds for replication studies (Simonsohn, 2015) by being over 8.5 times larger than the original study and providing sufficient power.

Recruitment

We designed our recruitment to mirror the original study as much as possible, while making necessary adjustments (cf. Brandt et al., 2014; Hüffmeier et al., 2016). The original study was conducted at Columbia Business School at Columbia University, located in New York City (E. T. Amanatullah, personal communication, June 18, 2020). To maintain comparability, we also conducted our study at a business school of an Ivy League institution, located in a metropolis on the East Coast of the United States. Specifically, we ran it at the Wharton Behavioral Lab at the University of Pennsylvania, which is located in Philadelphia.

Recruitment methods necessarily differed somewhat, given the time between the two studies and the standard practices at the respective study sites. In the original study, participants were recruited using flyers describing the study as "interesting psychological research" that takes 25–30 min to complete for a remuneration of $5, which were disseminated on campus and in nearby places (E. T. Amanatullah, personal communication, June 18, 2020). This approach allowed both students and nonstudents to participate (E. T. Amanatullah, personal communication, June 18, 2020). In our close replication at Wharton, participants were drawn from the Behavioral Lab's volunteer pool, which is advertised through flyers, classroom visits, and campus outreach. Participants received $12 (later $15, as rates increased at the Wharton Behavioral Lab) to reflect local compensation norms and the potentially longer duration of the sessions.

Finally, session logistics required a slight modification of session size. In the original study, only one or two participants participated at a time (E. T. Amanatullah, personal communication, June 18, 2020). Following this same standard would have made it impossible to reach the large sample size required for adequate power, so our replication had multiple participants complete a session simultaneously. To preserve the original study's conditions regarding privacy (E. T. Amanatullah, personal communication, June 18, 2020), participants were prevented from seeing other participants' computer screen while the study was being run and provided headsets so they would not overhear audio from their alleged negotiation counterpart.

Sample and Exclusions

At the Wharton Behavioral Lab, participants sign up for individual time slots within larger blocks of scheduled sessions. As a result, there is some uncertainty about how many participants will attend a given session. To achieve the large sample size required for our replication, we therefore ran multiple sessions, which ultimately yielded 562 cases in our raw data set. We applied our exclusion criteria to this sample.

Fifteen participants encountered technical difficulties, so that they had to be excluded (reducing the sample size to 547). Twenty-seven additional cases had to be excluded because paper–pencil questionnaires could not be matched with the corresponding computer survey data, leaving 520 participants (this issue was not anticipated).[6] Moreover, as the hypotheses concerned self-identified women and men, we excluded three additional participants who did not indicate "female" or "male" as their gender. Thus, after these exclusions, our total sample size was $N = 517$.

Sample Characteristics

Of the $N = 517$ participants, $n = 366$ were women and $n = 151$ were men. A total of $n = 261$ were in the self-advocacy condition, and $n = 256$ were in the other-advocacy condition (see also Table 2, which shows the number of women and men in both advocacy conditions for each key analysis). The average age was $M = 22.01$ years ($SD = 6.20$; range: 18–65). The participants had a range of backgrounds (for detailed information, see our supplementary information and analyses on participants' ethnicity/race in the

[4] We had to make assumptions about cell sizes in Amanatullah and Morris (2010) to calculate these effect sizes. Based on their total sample ($N = 59$), we assumed roughly half of women and men were in the self-advocacy condition (i.e., 14 women, 16 men) and the remainder in the other-advocacy condition (i.e., 14 women and 15 men).

[5] Full simulation settings: *mc_power_med* (Schoemann et al., 2017) was run with 5,000 replications, 20,000 draws per replication, $SD = 1$ for each variable, and random seed = 1,234. The analysis assumed an *a*-path of $r = .15$ and a *b*-path of $r = .30$, with the correlation between the independent and dependent variable being set at $r = .00$, $r = .15$, or $r = .30$.

[6] Participants were provided with a participant code (as was true for the original study). When they began to work on the computer, they were asked to enter their assigned code, which was necessary to ensure that the study could be run with many participants taking part simultaneously.

Table 2
Descriptive Statistics for Hypotheses 1, 2, and 4

Measure	Self-advocating women	Self-advocating men	Other-advocating women	Other-advocating men
	M (*SD*; *n*)	*M* (*SD*; *n*)	*M* (*SD*; *n*)	*M* (*SD*; *n*)
Anticipated backlash				
All participants ($N = 517$)	56,501.36 (7,814.56; 184)	58,881.39 (7,729.47; 77)	54,703.30 (7,775.50; 182)	57,510.14 (11,381.95; 74)
Attentive participants ($N = 461$)	55,831.82 (6,414.97; 165)	58,343.20 (7,506.95; 71)	54,805.56 (6,647.57; 162)	55,884.92 (8,635.05; 63)
Salary requests				
All participants ($N = 515$)	59,891.30 (70,164.40; 184)	58,849.35 (7,370.65; 77)	53,863.89 (7,963.75; 180)	58,702.70 (12,617.37; 74)
Attentive participants ($N = 459$)	54,536.36 (7,722.82; 165)	58,061.97 (6,785.58; 71)	53,378.13 (7,287.83; 160)	56,571.43 (7,283.91; 63)
Self-rated competitiveness				
All participants ($N = 517$)	4.28 (1.57; 184)	5.00 (1.40; 77)	4.65 (1.71; 182)	4.90 (1.36; 74)
Attentive participants ($N = 461$)	4.27 (1.60; 165)	5.07 (1.37; 71)	4.65 (1.70; 162)	4.85 (1.29; 63)

Note. Salary requests are from the first round of negotiation.

Supplemental Material). On average, participants had a mean negotiation experience, rated on a scale that ranged from 1 (*not at all*) to 7 (*very*) of $M = 3.19$ ($SD = 1.55$; $N = 516$ in this analysis).

Gauging Participant Attention

Going beyond Amanatullah and Morris (2010), we planned a priori to apply additional exclusion criteria. We specified that we would exclude participants if their response about for whom they primarily negotiated did not match their assigned experimental condition, or if they did not correctly identity "Brian" as their negotiation counterpart (for a full description of our two comprehension checks, see below). We also planned a priori to exclude participants whose first-round salary request was outside of the stated salary range of $37,000–$64,000 by more than $10,000, as these requests could also indicate a lack of attention or comprehension.

As planned with our registered report, we ran our primary analyses including only "attentive participants," meaning those who did not fail a comprehension check and made first-round salary requests that did not fall outside of the prespecified range (i.e., $10,000 or more above $64,000 or below $37,000). In total, 43 participants failed a comprehension check, 14 made requests outside of the prespecified range, and one participant fulfilled both criteria, resulting in a sample of $N = 461$ of "attentive participants." Because these exclusions go beyond Amanatullah and Morris (2010), we also planned a priori to perform the analyses while not excluding inattentive participants (Simmons et al., 2011). This procedure ensured that we provided analyses with the same exclusion criteria that were realized in the original study.

Suspicion Check

We also planned a priori to follow Amanatullah and Morris (2010) in flagging participants who expressed suspicions about the study design and in reporting analyses both including and excluding them (the original study noted that three participants expressed suspicions and that excluding them did not alter the results; p. 259). However, we subsequently realized that the suspicion check had limited value, because participants in the original study (and, therefore, also in our study) were first fully debriefed and only then asked about their suspicions (which we explain in more detail below). For this reason, we report the results with the suspicion exclusions only in the Supplemental Material.

Design

In line with Amanatullah and Morris (2010), our replication employed a 2 (gender of participant: men vs. women) × 2 (advocacy: self vs. other) design. Participants in the self-advocacy condition negotiated their own salary. Participants in the other-advocacy condition negotiated on behalf of a "friend," without specifying the friend's gender (as was also done by Bowles et al., 2005). Doing so was necessary to remain faithful to the received original materials (E. T. Amanatullah, personal communication, June 18, 2020). Notably, Amanatullah and Morris (p. 259) also reported that "the gender of the person on whose behalf other-advocates were negotiating" was found to be irrelevant. We note that at the end of our close replication study, we asked participants whether they had imagined the friend to have a specific gender. This exploratory measure allowed us to assess participants' perceptions, and the results are reported in the Supplemental Material.

Procedure and Experimental Manipulation

The procedure and experimental manipulation in our replication closely adhered to the original study by Amanatullah and Morris (2010; see pp. 259–260 for their description of their method). The authors provided us with their original study materials and answered our questions to clarify additional procedural details (E.T. Amanatullah, personal communication, June 18, 2020). Because the original study was programmed with outdated software, we reprogrammed the study in Qualtrics. Yet, unless otherwise noted, we preserved all the original details, including the survey design, a photograph, and the voice messages of the alleged negotiation counterpart.

Following the original study (Amanatullah & Morris, 2010), participants were told that they would conduct a negotiation with another participant from another university (i.e., Princeton) via a computer. In reality, the negotiation was computer-simulated, with no actual counterpart (Amanatullah & Morris, 2010). Depending on experimental condition, participants either negotiated their own salary—that is, they played the role of a recruit (self-advocacy)—or

they negotiated the salary of a friend—that is, they played the role of an agent (other-advocacy; Amanatullah & Morris, 2010).

Salary Information and Alternative Offer

To contextualize the negotiation, we gave participants the same salary information as the original study (with values updated for inflation). Specifically, participants were told, "Around 70% of new consultants earn somewhere between $39,000 and $60,000 with a low around $37,000 and a high near $64,000" (all of these numbers were adapted to address inflation). Hence, the values of $37,000 and $64,000 defined the broadest limits of the salary range. Participants also learned that new hires earn about $48,000 on average annually. Furthermore, participants were told that they had an alternative job offer from "another consulting firm that has offered [them] a job and that [they] have every reason to assume will pay a salary commensurate with the industry standard" (the exact amount of the alternative offer was not specified; still, given the provided information, participants likely expected the alternative offer to be near the mentioned average amount, as similarly described by Amanatullah & Morris, 2010).

The Simulated Counterpart

Participants learned their counterpart in the negotiation was a Princeton University student named Brian. They were told that their counterpart would play the role of the hiring manager. Participants saw a picture of their counterpart that was supposedly taken during his lab session. We used the same picture as was used in the original study. The picture showed a White, college-aged man. In the original study, participants were also photographed, and their picture was presented in the study next to the picture of their alleged counterpart. In our replication, however, this was not possible: We could not take pictures of participants due to confidentiality and feasibility concerns with having several participants complete the study simultaneously. To mimic the original protocol, we included web cameras on all the computers and the computer simulated taking a picture of them before the negotiation.

Preparation and Role Materials

Participants received their role instructions on the computer and as a hard copy to reference during the negotiation (Amanatullah & Morris, 2010). The hard copy, which we referred to as "Packet 1," also included the prenegotiation measures, including the measure of anticipated backlash, which we describe below.

The Salary Negotiation

The negotiation simulated a multiple-round salary bargaining interaction between the participant and their simulated counterpart ("Brian"). Each round followed the same structure. Participants could accept their counterpart's offer or decline it and make a salary request themselves. Participants also exchanged verbal statements, which they could select from a list, record via a microphone, and send to their counterpart (Amanatullah & Morris, 2010). They also listened to verbal statements from their alleged counterpart, which were again identical to those used in the original study. In each of the potentially five rounds of negotiation, participants could either accept or decline their counterpart's offer. If they accepted, the negotiation ended. If they declined, they had to propose a counteroffer, and the negotiation continued to the next round.

The counterpart's offers were always the same and fixed, so they did not depend on the participant's response. Across the five rounds, the counterpart offered $33,500; $36,750; $37,500; $38,250; and $39,000. (We note that these numbers were adapted from the original numbers for inflation.) If a participant made a request that was higher than the next offer by their counterpart, the counterpart rejected it and made the next offer. By contrast, the negotiation ended when (a) the participant made a request that was lower than (or equal to) the next offer by their counterpart, (b) the participant accepted an offer, or (c) the five rounds were over without an achieved agreement (see the Supplemental Material for exploratory analyses regarding impasse rates; e.g., Ma et al., 2024).

Postnegotiation

Following the negotiation, participants completed Amanatullah and Morris's (2010) original postnegotiation questionnaire, which we referred to as "Packet 2." Then, participants were debriefed and asked about any suspicions (Amanatullah & Morris, 2010). We again used the original debriefing and suspicion check, with the two exceptions that we added to the debriefing document the information that no actual photo was taken and that we included our own contact information. Finally, participants completed another questionnaire ("Packet 3"). This questionnaire, which included comprehension checks and additional measures (described below), was added by us and not part of the original study. Including this questionnaire did not alter the procedure of the study, as it was added only *after* the last measure that was part of the original study was collected.

Further Details About the Administration of the Study

In the original study, the experimenter manually set up the version of the computer-simulated negotiation that corresponded to each participant's condition (i.e., self-advocacy vs. other-advocacy) and distributed the matching hard-copy materials (E. T. Amanatullah, personal communication, June 18, 2020). As a result, the experimenter was present in the laboratory for the negotiation and necessarily aware of each participant's condition (E. T. Amanatullah, personal communication, June 18, 2020).

Participants in our close replication were instead randomly assigned to conditions by designated participant codes.[7] When participants entered their code in the computer, they were directed to the appropriate version. A research assistant was present in the laboratory to administer study materials but was likely unaware of each participant's experimental condition (as multiple participants were taking the study at a time).

We further ensured administrative consistency with the original study by giving detailed written instructions to the research assistants and embedded clear directions in the study materials to direct

[7] The code either included the letter "R" (for recruits) or an "A" (for agents). Likewise, the code included either an "M" (for men) or an "W" (for women). At the outset of the study, participants were handed out "packets" on which their participant code was written. When participants entered their code, the computer automatically recognized whether it included an "R" or an "A," so that the stimuli presented on the computer matched the paper–pencil "packets" that were handed out to participants.

participants through each stage of the study (e.g., when to start or turn in a questionnaire).

Finally, we created a new consent form because no original consent form was available. The new consent form included the same procedural information as the original study, including the deceptive statement that they would negotiate with another participant from Princeton.

Measures

We used the exact same measures as the original study. As noted above, we included an additional questionnaire after the conclusion of the study that included comprehension checks to ensure data quality. The measures are described in the order in which participants saw them.

Prenegotiation Measures

Prenegotiation Aspirations. Following Amanatullah and Morris (2010), participants answered questions about their reservation point, target point, and anticipated opening offer. These questions were included in a paper–pencil questionnaire attached to the hard copy of their role information ("Packet 1"). Participants indicated in dollars (a) their reservation point: "What is the lowest salary you are willing to accept [for your friend] in this negotiation? In other words, what is the least you would agree [for your friend] to earn at Alpha before choosing [advising your friend] to take the alternative job offer at Lambda?" (Amanatullah & Morris, 2010, p. 259); (b) their target point: "What is the highest salary that you will strive to get [for your friend] in this negotiation? In other words, what is the ideal salary you want Alpha to agree to [pay your friend]?" (Amanatullah & Morris, 2010, p. 259); and (c) their anticipated opening offer: "If you are given the opportunity to make the first offer in this negotiation, how much will you suggest Alpha pay for your [as your friend's] salary?" (Amanatullah & Morris, 2010, p. 260). The versions of the questions for the other-advocacy condition are in brackets.

Anticipated Backlash. In the same questionnaire with the prenegotiation aspiration questions (i.e., "Packet 1"), participants answered the following two questions concerning anticipated backlash: "How much do you think you can reasonably ask for [your friend] without the hiring manager perceiving you [your friend] to be a pushy person?" and "How much do you think you can reasonably ask for [your friend] without causing the hiring manager to punish you [your friend] for being too demanding?" (Amanatullah & Morris, 2010, p. 260). The versions of the questions for the other-advocacy condition are again in brackets.

Participants indicated their answers in dollars. Thus, "a lower score indicates greater anticipation of backlash" (Amanatullah & Morris, 2010, p. 260). In line with the original study, participants' responses to the two questions on anticipated backlash were averaged to build a scale. The correlation between the two variables was $r = .70$ ($\alpha = .82$; $N = 517$). For further details on our measurement of anticipated backlash, see the Supplemental Material.

Measures Collected During the Negotiation

Salary Requests. Participants' requested salary is a first key operationalization of negotiator assertiveness (see Figure 1). At each round of negotiating, participants stated a salary request (or counteroffer in response to the counterpart's offer), recorded in dollars. Following Amanatullah and Morris (2010), we focused on participants' first-round requests as the central outcome, because, according to the original authors, these served as the "primary measure of concessionary behavior" (p. 260).

Two participants accepted the counterpart's very first offer. As a result, they were unable to make a salary request because the negotiation ended (see above). This left a sample size of $N = 515$ for analyses regarding salary requests. Because participants could potentially make a request in each of the five rounds, we also conducted supplementary analyses of later requests (see Supplemental Material).

Verbal Behavior, Including Assertiveness. The second measure of negotiator assertiveness was participants' verbal behavior during the negotiation (see Amanatullah & Morris, 2010, p. 260). In each round of the negotiation, participants could send two verbal statements to their counterpart—once in response to the counterpart's offer and once when justifying their own request. They always chose from a set of five presented options (e.g., "Your offer is insulting. It is way too low for me" [assertive]; "I deserve to earn far more than that offer" [entitled]; "That offer does not sufficiently compensate my superior qualifications" [qualification-based]; "Thank you for that offer, but I was hoping to earn more" [cooperative]; "I am disappointed with that offer" [dejected/hopeful]). The presentation order of these options was varied across rounds. For an overview of all statements, see Amanatullah and Morris (2010, pp. 266–267).

Two student assistants examined all verbal statements and noted whether a statement was assertive, entitled, qualification-based, cooperative, dejected/hopeful, missing (if a participant accepted the offer by the "counterpart," they could not send further verbal statements, except for one indicating their accepting of the offer), or simply unclear or another kind of statement (i.e., one that was different from the presented options; please note that the results remained the same no matter whether or not these unclear or differing selections were included in our hypothesis tests). The interrater reliability was $\kappa \geq .99$ for the responses and the own requests in all five rounds. The average κ was .99. On a final note, as two participants accepted the very first offer (see above), the total sample size even for the statements in the first round of negotiation was again $N = 515$.

Postnegotiation Measures

After completing the negotiation, participants filled out a second paper–pencil questionnaire ("Packet 2"). This questionnaire included an item on self-rated competitiveness (our third dependent variable) as well as demographic information. Notably, the questionnaire also contained several additional measures used in the original study, including multiple self-ratings of negotiation style (E. T. Amanatullah, personal communication, June 18, 2020). Consistent with Amanatullah and Morris (2010), we focus only on participants' competitiveness ratings. However, to adhere closely to the original protocol, we administered the full original questionnaire.

Self-Rated Competitiveness. Following Amanatullah and Morris (2010), we measured self-assessment of competitiveness during the negotiation as another indicator of negotiator assertiveness. This measure captures participants' own perceptions of their negotiation style, that is, "the extent to which they characterized their negotiation style as competitive" (Amanatullah & Morris, 2010, p. 260). The used

scale ranged from 1 (*not at all*) to 7 (*a great deal*). Three values were missing; we replaced them with the sample mean.

Demographics, Including Gender of Participants. At the end of the original postnegotiation questionnaire ("Packet 2"), participants reported demographic information, including age, gender, and race/ethnicity. Age was provided as a free-response item. For gender (the term "sex" was used in the original questionnaire), in addition to the original categories of "male" and "female," we added the option "prefer to self-describe." For race/ethnicity (the term "race" was used in the original questionnaire), participants could select from "Caucasian American," "African American," "Asian American," "Hispanic American," "Native American," "prefer not to answer," or "other—please specify." Additional details on the measurement of participant demographics are available in the Supplemental Material.

Debriefing and Suspicion Check

Following the original study (E. T. Amanatullah, personal communication, June 18, 2020), participants then received debriefing information. At the bottom of the page including the debriefing information, participants were probed about their suspicions. Specifically, participants were asked in writing, "Were you suspicious of the study design used in this experiment?" The response options were "yes" and "no." A subsequent prompt ("If yes, please describe") asked participants to elaborate on their potential suspicions (E. T. Amanatullah, personal communication, June 18, 2020).

In our study, $n = 265$ of the participants said they were suspicious, which is a higher rate than in the original study.[8] A likely explanation is that a key goal of our research was to provide a *close* replication, thereby adhering as much as possible to the original study protocol and utilizing original study materials. In turn, original study materials might have made today's participants suspicious because they might have appeared as outdated (although it is possible today to let participants interact online in real time, as was simulated in our and the original study; e.g., Molnar, 2019). Alternatively, today's participants could think that the interaction was simulated with an artificial intelligence—a technology that was less relevant when the original study was conducted.

In our hypothesis tests, as planned a priori, we always ran the analyses with and without participants indicating suspicions (analyses excluding suspicious participants are reported in the Supplemental Material). To foreshadow the results, descriptive statistics and effect sizes for several hypothesis tests suggest that the findings were not drastically different. Still, effects were typically nonsignificant when excluding participants who indicated suspicions, which might be due to the lowered statistical power in these analyses.[9]

Please also note that the original suspicion check that we also used had noteworthy limitations: Participants were asked about potential suspicions only *after* they had been debriefed. Therefore, participants might have felt sure about their suspicions only in hindsight. Similarly, they might have indicated suspicions to avoid appearing naïve. Both of these issues shed doubt on the usefulness of the suspicion check.

Additional Questionnaire, Including Novel Comprehension Checks

We attached to the debrief document another questionnaire that was not part of the original study. The debrief document including the suspicion check and our additional questionnaire formed our "Packet 3." The additional questionnaire began with two comprehension checks. First, participants indicated for whom they primarily negotiated (cf. Kouchaki & Kray, 2018). The response options were "for myself—the negotiation concerned my own salary" and "for someone else—the negotiation concerned my friend's salary." Second, they indicated the name of their counterpart. The three response options were "Gary," "Brian," and "Michael."

These novel comprehension checks allowed us to check data quality (see above and see also our Supplemental Material for further details). Afterward, participants were asked to respond to the following question, if they negotiated on behalf of a friend: "Did you think of your friend as having a particular gender?" The response options were "I thought of a *male* friend (i.e., a man)," "I thought of a *female* friend (i.e., a woman)," "I thought of a *nonbinary* friend," and "I did not think of my friend having a particular gender."

Then, participants indicated whether they had ever heard of, or read about, the article by Amanatullah and Morris (2010; the reference was provided to participants), as well as whether they had ever heard of, or read about, the main finding from the original study. Finally, participants also indicated their nationality, how experienced in negotiation they considered themselves, and they responded to an additional measure of fear of backlash (adapted from Jarvis & Kray, 2022; Rudman & Fairchild, 2004) as well as a measure of independent and interdependent self-construal (Singelis, 1994).

Context

We are researchers who are interested in the topics of gender and negotiation. For many years, we have conducted research to better understand what causes gender differences in negotiation processes and outcomes, as well as how to mitigate these differences. Moreover, we strongly believe in the importance of conducting replications, especially for key findings that exert extensive influence on both theory and practice. For these reasons, we became aware of the need for a close replication of the original study by Amanatullah and Morris (2010). There is no question that this study has advanced theoretical propositions that play a central role in current research on gender differences in negotiation. Indeed, we often cite this study in our own work. However, upon closer examination, the actual evidence for these propositions is rather thin: The original study had low statistical power, and related research has yielded divergent results. Moreover, the original study has never been replicated despite its centrality to gender and negotiation research. Thus, we aimed to conduct a compelling close replication that closely adheres to current open science practices. We

[8] Five additional participants indicated "no," yet they wrote a comment that expressed a suspicion. Additionally excluding these further participants did not lead to diverging results in any of our hypothesis tests. Moreover, eight participants had missing values regarding the suspicion check—as these participants did not explicitly indicate suspicions, there were judged as nonsuspicious.

[9] Although the sample size was lower when excluding participants who indicated suspicions ($N = 252$), our replication was still substantially larger than the original study ($N = 59$). To be exact, it was 4.27 times larger. This is an important point because extant scholarship recommends that replications be 2.50 times larger than original studies to allow for meaningful comparisons (Simonsohn, 2015).

believe that our close replication makes an important contribution to research on gender and negotiation.

Results

Table 2 provides descriptive statistics. Table 3 provides regression analyses. In what follows, we always provide the results when including all participants (as was done by Amanatullah & Morris, 2010). Moreover, going beyond Amanatullah and Morris (2010), we also provide the results when including only attentive participants (i.e., excluding participants who failed at a comprehension check or made a request in the first round of negotiation that was substantially below or above the provided salary range). The analyses were planned ex ante to ensure data quality. Further analyses are reported in our Supplemental Material.

Hypothesis 1: Anticipated Backlash

Hypothesis 1 predicted that self-advocating women would more readily anticipate backlash than self-advocating men, other-advocating women, and other-advocating men. Put differently, self-advocating women were hypothesized to differ from "the other three experimental groups" (Amanatullah & Morris, 2010, p. 260). We applied multiple complementary tests of the hypothesis. First, we conducted planned contrasts to directly evaluate our focal prediction of whether self-advocating women differed from all other groups combined (contrast code: −3, 1, 1, 1). Second, we estimated linear regression models to examine Gender × Advocacy interactions and to parallel the analytic approach used by Amanatullah and Morris (2010). Third, we ran pairwise comparisons (i.e., *t* tests) to clarify which groups differed from one another, again to mirror the original study's reporting. Following our plans made prior to collecting data, all of these analyses were conducted with bootstrapping (5,000 samples, bias-corrected and accelerated), as we observed many outliers and violations of statistical assumptions (e.g., Field, 2013). Finally, we ran additional exploratory analyses.

In the analysis including all participants ($N = 517$), the contrast was not significant, $t(513) = 0.67$, $p = .527$. The means across the four conditions also did not match the predicted pattern (Table 2). The same findings emerged when including only attentive participants ($N = 461$), $t(457) = 0.73$, $p = .461$. Put differently, self-advocating women did not differ from the other three experimental conditions (see Table 2).

Following Amanatullah and Morris (2010), we next ran regression models to test whether gender and advocacy interacted to predict anticipated backlash. In the regressions, gender was coded as 0 = *men* and 1 = *women*, and advocacy was coded as 0 = *self-advocacy* and 1 = *other-advocacy* (Amanatullah & Morris, 2010). This coding reflects a so-called simple effects parameterization (Hayes, 2018). This means that the effects for gender and advocacy do not represent *main effects* (as in an analysis of variance) but *simple effects* (i.e., the effect of a predictor at the level coded as zero on the other predictor; Hayes, 2018).

In the analysis with all participants and also in the analysis including only attentive participants, there was no significant interaction (see Table 3). However, there was a significant effect of gender (in both the full and attentive samples), such that women anticipated more backlash than men in the self-advocacy condition.

Next, again following Amanatullah and Morris (2010), we performed *t* tests to compare condition differences. There was a significant difference between self-advocating women and self-advocating men when including all participants, $t(259) = -2.25$, $p = .028$, $d = -0.31$, and also when including only attentive participants, $t(234) = -2.62$, $p = .014$, $d = -0.37$, such that women more readily anticipated backlash than men. Moreover, when including all participants, there was a significant yet *reversed* difference between self-advocating women and other-advocating women, such that other-advocating women anticipated backlash more readily, $t(364) = 2.21$, $p = .031$, $d = 0.23$. There was no such effect in the analysis including only attentive participants ($p = .159$). Finally, there were no other significant differences in the analyses including all participants (both ps $\geq .384$) as well as in the analyses including only attentive participants (both ps $\geq .088$).

Exploratory Analyses

As we observed many, potentially influential outliers, we also ran nonparametric Kruskal–Wallis tests (e.g., Field, 2013), always with Bonferroni-corrected post hoc comparisons. These tests should be viewed as exploratory, as we did not plan to conduct them ex ante. In the analysis including all participants and also in the analysis including only attentive participants, there was a significant difference between self-advocating women and self-advocating men ($p = .009$ and $p = .014$, respectively). Moreover, again in the analysis including all participants and in the analysis with only attentive participants, there was a significant difference between

Table 3
Regression Analyses for Hypotheses 1, 2, and 4

Predictor	Anticipated backlash		Salary request		Self-rated competitiveness	
	b (*SE*)	*p*	*b* (*SE*)	*p*	*b* (*SE*)	*p*
All participants						
Gender	−2,380.03 (1,061.32)	.027	1,041.95 (5,283.32)	.756	−0.73 (0.20)	<.001
Advocacy	−1,371.26 (1,596.79)	.395	−146.65 (1,698.87)	.946	−0.10 (0.23)	.664
Gender × Advocacy	−426.80 (1,802.69)	.822	−5,880.77 (5,513.75)	.424	0.48 (0.29)	.099
Attentive participants						
Gender	−2,511.38 (1,020.84)	.014	−3,525.61 (997.38)	<.001	−0.80 (0.20)	<.001
Advocacy	−2,458.28 (1,433.95)	.083	−1,490.54 (1,209.96)	.219	−0.22 (0.23)	.344
Gender × Advocacy	1,432.02 (1,600.52)	.375	332.31 (1,463.42)	.823	0.60 (0.30)	.044

Note. Salary requests are from the first round of negotiation. *SE* = standard error.

other-advocating women and self-advocating men (both $ps \leq .001$). There were no other differences in these analyses.

In light of the entire set of results, we ran additional exploratory t tests comparing women and men while collapsing across the advocacy conditions. There was a significant gender difference in the analysis including all participants, $t(515) = 3.20$, $p = .004$, $d = 0.31$, and also in the analysis including only attentive participants, $t(459) = 2.58$, $p = .024$, $d = 0.27$.

Summary

We did not replicate a key finding from Amanatullah and Morris (2010), namely, that gender interacts with advocacy to predict anticipated backlash. As such, Hypothesis 1 was not supported. Instead, we observed only a main effect of gender, with women more readily anticipating backlash than men across advocacy conditions.

Hypothesis 2: Salary Requests

Hypothesis 2 predicted that self-advocating women would make lower salary requests than the other three groups. As in the original study, we focused on participants' first-round requests (for additional analyses, see the Supplemental Material). To test the hypothesis, we again ran contrast analyses, regression analyses, and t tests (see above). All of these tests were again conducted with bootstrapping. We also note that there were again many outliers. For instance, one self-advocating woman requested a salary of $1,000,000, which was a particularly influential case.[10] This makes our analyses including only attentive participants especially relevant (see also the exploratory Kruskal–Wallis tests below).

In the analysis including all participants ($N = 515$), the contrast was not significant, $t(511) = -0.68$, $p = .619$. Likewise, the contrast was nonsignificant in the analysis including only attentive participant ($N = 459$), $t(455) = 1.98$, $p = .055$. The means were also not entirely in line with the prediction, as other-advocating women made *lower* salary requests than self-advocating women.

Turning to the regressions, there were no significant effects in the analysis including all participants—that is, gender did not interact with advocacy to predict salary requests (see Table 3). In the analysis including only attentive participants, there was again no significant interaction, yet only a significant simple effect for gender (Table 3).

Regarding the t tests, in the analyses including all participants, there were no significant differences (all $ps \geq .421$). In the analyses including only attentive participants, there was a significant difference between self-advocating women and self-advocating men, $t(234) = -3.33$, $p < .001$, $d = -0.47$, but no other differences were significant (all $ps \geq .069$).

Exploratory Analyses

We again ran Kruskal–Wallis tests. In the analysis including all participants, self-advocating women requested significantly lower salaries than self-advocating men ($p = .001$). Moreover, other-advocating women requested significantly lower salaries than both self-advocating men ($p < .001$) and other-advocating men ($p = .003$; no other differences were significant). In the analysis including only attentive participants, self-advocating women again requested significantly lower salaries than self-advocating men ($p = .003$). Moreover, other-advocating women requested lower salaries than both self-advocating men ($p < .001$) and other-advocating men ($p = .015$; no other differences were significant).

Once again in light of the entire set of results, we also ran exploratory t tests comparing women and men while collapsing across the advocacy conditions. There was no significant gender difference in the analysis including all participants, $t(513) = 0.45$, $p = .561$, $d = 0.04$. However, women (as compared to men) requested significantly lower salaries in the analysis including only attentive participants, $t(457) = 4.48$, $p < .001$, $d = 0.46$.

Summary

We did not replicate the second key finding from Amanatullah and Morris (2010): Gender did not interact with advocacy to predict salary requests. Thus, Hypothesis 2 was not supported. However, across some analyses, women made lower salary requests than men, particularly among attentive participants.

Hypothesis 3: Mediation

Hypothesis 3 predicted that anticipated backlash would mediate the interaction between gender and advocacy on salary requests. In keeping with current methodological standards, we used PROCESS (Hayes, 2018) Model 7 to examine conditional indirect effects, with gender as the independent variable, anticipated backlash as the mediator, advocacy as the moderator (of the a-path in a mediation model), and salary requests as the dependent variable. First, there was no significant difference between the conditional indirect effects (i.e., the effect of gender on salary requests through anticipated backlash as compared between the two advocacy–conditions): index of moderated mediation = −391.21, 95% CI [−7400.15, 4945.27], for the analysis including all participants; index of moderated mediation = 995.64, 95% CI [−940.92, 2762.44], for the analysis including only attentive participants.

Second, there was a significant indirect effect of gender on salary requests through anticipated backlash *within the self-advocacy condition*, both in the analysis including all participants, indirect effect = −3,391.42, 95% CI [−9015.54, −375.26], as well as in the analysis including only attentive participants, indirect effect = −1,558.31, 95% CI [−2940.53, −267.88]. There was no such indirect effect *within the other-advocacy condition* in any analysis.

Third, in both the analysis including all participants and also in the analysis including only attentive participants, regressions including gender and anticipated backlash as predictors of the salary requests (across conditions, as this regression examined the b-path in a mediation), anticipated backlash was a significant positive predictor (all $ps < .001$, and all 95% CIs did not include zero). This finding concerning the significant relationship between anticipated backlash and salary requests is relevant because it replicates the respective finding by Amanatullah and Morris (2010) and also because some past research did not observe such a relationship (Babcock et al., 2006; Dreber et al., 2022).

[10] When excluding participants who failed at a comprehension check and only this one case, we again observed a significant effect for gender, but no significant interaction (see below).

We note that we had planned to perform the same mediation analyses following the regression approach of Baron and Kenny (1986) to align with Amanatullah and Morris (2010). However, because there was no interaction between gender and advocacy, both with anticipated backlash and with salary requests as the dependent variable, the necessary preconditions for this analysis were not met. Thus, we also did not run an accompanying Sobel test.

Summary

We did not replicate the mediation observed by Amanatullah and Morris (2010), as there was no interaction between gender and advocacy on anticipated backlash or salary requests to begin with. Thus, Hypothesis 3 was not supported. Still, anticipated backlash predicted salary requests, and anticipated backlash mediated the gender difference within the self-advocacy condition (but not in the other-advocacy condition).

Hypothesis 4: Self-Rated Competitiveness

Hypothesis 4 predicted that self-advocating women would rate their negotiation style as less competitive than self-advocating men, other-advocating women, and other-advocating men. To examine Hypothesis 4, we again conducted contrast analyses, regression analyses, and *t* tests (all again with bootstrapping). We did not run exploratory Kruskal–Wallis tests, as outliers posed less of a problem here (we observed outliers only in the analyses excluding participants who indicated suspicions, which are reported in the Supplemental Material).

In the analysis including all participants, the contrast was significant, $t(357.26$; unequal variances$) = 3.99$, $p < .001$. The means were also, at least partly, in line with the predicted pattern (the mean among other-advocating women was not fully on par with the means among men; see Table 2, and also see the *t* tests reported below). The same findings emerged in the analysis including only attentive participants, $t(311.77$; unequal variances$) = 3.85$, $p < .001$.

Turning to the regressions (Table 3), we observed a simple effect for gender in the analysis including all participants as well as in the analysis including only attentive participants. There was no significant effect for advocacy in any analysis. Finally, and importantly, there was a significant Gender × Advocacy interaction in the analysis including only attentive participants, but not in the analysis including all participants (see Table 3).

The *t* tests also provided noteworthy results. In the analysis including all participants, self-advocating women significantly differed from self-advocating men, $t(259) = -3.50$, $p < .001$, $d = -0.48$; other-advocating women, $t(364) = -2.19$, $p = .032$, $d = -0.23$; and other-advocating men, $t(256) = -3.00$, $p = .002$, $d = -0.41$ (there was no difference between self-advocating men and other-advocating men). Likewise, in the analysis including only attentive participants, self-advocating women differed from self-advocating men, $t(153.83$; unequal variances$) = -3.92$, $p < .001$, $d = -0.52$; other-advocating women, $t(325) = -2.09$, $p = .037$, $d = -0.23$; and other-advocating men, $t(226) = -2.58$, $p = .005$, $d = -0.38$ (there was again no difference between self-advocating men and other-advocating men).

Summary

We replicated the effect from Amanatullah and Morris (2010) that self-advocating women rated their negotiation style as less competitive than the three comparison groups. Thus, Hypothesis 4 was supported.

Hypothesis 5: Verbal Behavior, Including Assertive Statements

Hypothesis 5 predicted that the verbal behavior of self-advocating women would be characterized by fewer assertive statements than the verbal behavior of self-advocating men, other-advocating women, and other-advocating men. Amanatullah and Morris (2010) reported that their analyses of the verbal statements did not yield any effects, making analyses regarding Hypothesis 5 less central for our replication.

Participants' verbal behavior was measured categorically. Thus, we conducted log-linear and χ^2 analyses testing the associations between gender, advocacy, and the selected verbal behavior, including assertiveness. As with salary requests as the dependent variable, we focused on the verbal statements from the first round of negotiation (see also our Supplementary Analyses). Table 4 shows frequencies by condition for the whole sample ($N = 515$). It is not fully clear how the verbal behavior was analyzed in the original study (E. T. Amanatullah, personal communication, July 20, 2020).

A log-linear analysis utilizes a "backward elimination" procedure, such that it begins with more complex models—examining the relevance of the highest order effect (here, the interaction between gender, advocacy, and the selection of a particular type of verbal statement)—and tests whether dropping it would reduce the model fit. If not, the dropped effect is interpreted as irrelevant (for an overview, see Field, 2013). In the log-linear analysis including all participants on the responses in Round 1, there was no significant highest order three-way effect (i.e., Gender × Advocacy × Statement), $\chi^2(5) = 5.85$, $p = .321$. The same was observed in the analysis including only attentive participants, $\chi^2(5) = 4.88$, $p = .431$.

Next, we ran χ^2 analyses testing for associations between gender and the selected responses in Round 1, separately for the two advocacy conditions. In the analysis including all participants, there was no such association within the self-advocacy condition, $\chi^2(5) = 7.07$, $p = .216$, nor within the other-advocacy condition, $\chi^2(5) = 4.94$, $p = .424$. The same was observed in the analysis including only attentive participants: $\chi^2(5) = 7.86$, $p = .164$, for the self-advocacy condition; $\chi^2(5) = 3.05$, $p = .692$, for the other-advocacy condition.

We next sought to examine Hypothesis 5 using the statements accompanying participants' first request. However, only two of the total 515 participants—both self-advocating men—selected the statement capturing assertiveness, which was the critical type of statement in Hypothesis 5 (see again Table 4). As such, Hypothesis

Table 4
Frequencies of Chosen Verbal Statements for Hypothesis 5

Statement type	Self-advocating women (*n*)	Self-advocating men (*n*)	Other-advocating women (*n*)	Other-advocating men (*n*)
Responses in Round 1				
Assertive	4	6	9	6
Entitled	24	6	23	15
Qualification-based	80	35	105	34
Cooperative	54	19	26	10
Dejected/hopeful	11	7	11	5
Other/unclear	11	4	6	4
Statements accompanying the own first request				
Assertive	0	2	0	0
Entitled	19	15	43	24
Qualification-based	7	4	8	10
Cooperative	113	37	102	28
Dejected/hopeful	38	17	22	9
Other/unclear	7	2	5	3

5 could not be meaningfully tested using the statements accompanying participants' first request.

Summary

Amanatullah and Morris's (2010) null finding regarding participants' selected verbal statements was replicated. That is, Hypothesis 5 was not supported.

Evaluation of Findings Vis-à-Vis the Original Study

The results of a replication can be evaluated in multiple ways (e.g., Nelson et al., 2018; Open Science Collaboration, 2015). To start, it is important to clarify that we focused on the following dependent variables when evaluating the results of our replication: anticipated backlash, salary requests in the first round of negotiation, and self-rated negotiation style. These are the dependent variables that were addressed in the section on hypothesis testing. People's verbal behavior is also part of the section on hypothesis testing (see Hypothesis 5). However, given that Amanatullah and Morris (2010) did not observe any effects regarding the exchanged verbal statements, this measure was less relevant for our close replication (likewise, dependent variables and analyses addressed in the Supplemental Material and, of course, the suspicion and comprehension checks were not included when evaluating the extent to which the original results replicated).

As a framework to evaluate the findings of our replication, we adopted the recommended outcome categories by LeBel et al. (2019). The effect sizes from our close replication stem from the analyses including only attentive participants (i.e., excluding participants who failed at a comprehension check or who made a request outside of the prespecified range, as these exclusion criteria increase reliability; participants who mentioned suspicions were not excluded, because they were kept in the analyses provided by Amanatullah & Morris, 2010, as well). For the calculations, we used the calculator provided by Wilson (2023). Table 5 gives an overview of the findings.

Regarding *anticipated backlash*, in line with the original study, we observed a *signal* (i.e., an effect that differed from zero; LeBel et al., 2019) in the comparison between self-advocating women and

Table 5
Evaluation of Key Replication Results

Comparison	ES: Replication	95% CI: Replication	ES: Original study	Evaluation
Anticipated backlash				
Self-advocating women versus self-advocating men	$d = -0.37$	[−0.65, −0.09]	$d = -0.87$	Signal (inconsistent, smaller)
Self-advocating women versus other-advocating women	$d = 0.16$	[−0.06, 0.37]	$d = -1.13$	No signal (inconsistent)
Self-advocating women versus other-advocating men	$d = -0.01$	[−0.30, 0.28]	$d = -0.97$	No signal (inconsistent)
Salary requests				
Self-advocating women versus self-advocating men	$d = -0.47$	[−0.75, −0.19]	$d = -1.05$	Signal (inconsistent, smaller)
Self-advocating women versus other-advocating women	$d = 0.15$	[−0.06, 0.37]	$d = -1.64$	No signal (inconsistent)
Self-advocating women versus other-advocating men	$d = -0.27$	[−0.56, 0.02]	$d = -1.96$	No signal (inconsistent)
Self-rated competitiveness				
Self-advocating women versus self-advocating men	$d = -0.52$	[−0.80, −0.24]	$d = -0.76$	Signal (consistent)
Self-advocating women versus other-advocating women	$d = -0.23$	[−0.45, −0.01]	$d = -0.88$	Signal (inconsistent, smaller)
Self-advocating women versus other-advocating men	$d = -0.38$	[−0.67, −0.09]	$d = -0.77$	Signal (inconsistent, smaller)

Note. Salary requests are from the first round. Evaluation categories stem from LeBel et al. (2019). A negative sign in this table indicates that self-advocating women had *lower* values (signs in the main text may be different, depending on the type of analysis that was run). ES = effect size; CI = confidence interval.

self-advocating men (see also our results presented above). This effect was *inconsistent*, such that it was smaller than in the original study (i.e., the 95% CI around the effect size estimated in our replication did not include the estimate from the original study; see Table 5; see also LeBel et al., 2019). Still, given that our replication suggests that the effect in principle exists, support for the underlying proposition was provided. Regarding the other two comparisons (i.e., self-advocating women as compared to both other-advocating women and other-advocating men), however, we observed *no signal* (inconsistent; i.e., the 95% CI around the effect size estimated in our replication included zero, and it did not include the estimate from the original study; LeBel et al., 2019), unlike the original study. As mentioned earlier, we simply did not observe the hypothesized interaction effect.

Regarding *salary requests*, the same findings emerged: There was a *signal* (inconsistent, smaller) for the comparison between self-advocating women and self-advocating men (see above), yet *no signal* (inconsistent) for the other comparisons.

Regarding *self-rated competitiveness*, there was a *signal* (consistent) for the comparison between self-advocating women and self-advocating men (i.e., the 95% CI around the effect size estimated in our replication did not include zero, and it did include the estimate from the original study; LeBel et al., 2019). Moreover, for both remaining comparisons, we again observed a *signal* (inconsistent, smaller). These results replicate the findings from the original study. As a final observation, please also note that the effects observed in our replication study are comparable in size to those estimated in recent meta-analyses (e.g., Kugler et al., 2018; Nohe et al., 2022).

Discussion

Amanatullah and Morris (2010) provided a particularly relevant study on the topic of gender differences in negotiation. Their most central findings were that women (as compared to men) negotiated less assertively because they anticipated backlash more readily, yet only when they negotiated for themselves (vs. for someone else). Given its central role for theory and research (e.g., Bowles et al., 2022; Eagly & Karau, 2002), as well as practice (e.g., England et al., 2020), and in light of the mixed findings from past studies (e.g., Artz et al., 2018; Babcock et al., 2006), we closely replicated Amanatullah and Morris's study.

Summary of Key Findings

Consistent with the original study, women anticipated more backlash than men and requested lower salaries. Anticipated backlash also mediated the effect of gender on salary requests. However, our findings did not replicate the moderation by advocacy. As a result, the proposed full model (Figure 1) and key hypotheses—such that advocacy moderates the effect of gender on anticipated backlash and salary requests—were not supported. Finally, we also examined two additional outcomes included in the original study: Regarding self-rated competitiveness, our findings largely replicated those of the original study, such that self-advocating women perceived themselves as less competitive than the other groups. Finally, also replicating the original study, we found no significant interaction regarding gender, advocacy, and verbal assertiveness.

Theoretical Implications

The seminal research by Amanatullah and Morris (2010) is relevant because it advanced and tested key theoretical propositions (e.g., Eagly & Karau, 2002; Rudman, 1998). A first central proposition was that women more readily anticipate backlash than men. Role congruity theory and the backlash literature suggest that other people can react negatively when women violate expectations associated with their communal gender role (e.g., Eagly & Karau, 2002; Williams & Tiedens, 2016). As a consequence, women anticipate this reaction (see also Rudman & Fairchild, 2004) and tend to negotiate less assertively (Amanatullah & Morris, 2010; Walters et al., 1998). Importantly, we also observed that women more readily anticipated backlash than men—a tendency that mediated the gender difference in salary requests. As such, this first proposition was replicated in our study, which is relevant because findings from past related research regarding this central proposition are mixed (e.g., Dreber et al., 2022; Toosi et al., 2019).

Similarly, Amanatullah and Morris (2010) derived from theory the additional proposition that women are *aware* of how they negotiate, which should lead them to rate their negotiation style as less competitive (this idea was supported in their study). In our close replication, we observed the same findings (see Table 5). As such, the theoretical proposition of women being aware of their negotiation style was replicated in our study.

Still, yet another central proposition was clearly not supported in our close replication. In line with role congruity theory (Eagly & Karau, 2002), Amanatullah and Morris (2010) reasoned that women would anticipate backlash less readily when they advocate for others, because being assertive for someone else's benefit could be interpreted as a communal, role-congruent act (see also Amanatullah & Tinsley, 2013). Amanatullah and Morris's original results supported the theoretical proposition of advocacy moderating gender differences in negotiation. But as we already pointed out, the original study had, in all likelihood, insufficient statistical power to reliably observe such a moderator effect. Still, prior to our replication, the theoretical proposition of advocacy being a moderator itself was less in question than the theoretical proposition of anticipated backlash being a mediator (see our lengthy discussion in the introduction). This reasoning is because other extant primary studies (e.g., Kouchaki & Kray, 2018) and meta-analyses (Nohe et al., 2022) also observed a moderating effect for advocacy. However, in our close replication, we did not observe a moderating effect for advocacy when gender differences in anticipated backlash or salary requests were concerned (please recall that self-rated competitiveness was an exception here). As such, Hypotheses 1 and 2 regarding the interactive effects of gender and advocacy were not supported in our replication. This raises the theoretically relevant question of why gender differences did not depend on the advocacy role assumed by the negotiators.

One reason could be that being an advocate triggers two distinct psychological processes, with diverging consequences. On the one hand, as suggested by past work and Amanatullah and Morris (2010), being an advocate for others might reduce the risk of incurring backlash (e.g., Amanatullah & Tinsley, 2013), as the congruity between women's gender role and the role of a negotiator is increased (e.g., Stuhlmacher & Linnabery, 2013). This reduced risk of incurring backlash could then be reflected in women's

anticipated feelings and behaviors prior to negotiating salary (Amanatullah & Morris, 2010).

On the other hand, being an advocate for others—or a friend, in particular—could also make *salient and relevant* women's communal gender role. Recent reviews highlighted that gender effects become more likely when gender is a relevant and salient characteristic in a social situation (Bowles et al., 2022, 2025). In turn, disadvantaging processes that have their root in women's gender role, such as fear of backlash, would be expected to remain influential. Similarly, women's communal gender role involves being sensitive to other people's needs (e.g., Eagly et al., 2020; Rudman et al., 2012). In a negotiation, following dual concern theory (Rubin et al., 1994; see also De Dreu et al., 2000), this could mean that women are also being sensitive to their counterpart's needs (i.e., they have a high "other-concern"), as the counterpart is the immediate and only other person that is present in the interaction. Acting in line with their communal gender role, women may thus accommodate to their counterpart's aims by making lower requests.

These possibilities might have become all the more relevant because a meta-analysis suggests that, in the United States, women are *increasingly* stereotyped as communal over time (Eagly et al., 2020; see also their Figure 4, which suggests a high level of consensus among relatively recent respondents to a nationally representative poll). Hence, in comparison to the time when the original study by Amanatullah and Morris (2010) was conducted, processes disadvantaging women in negotiation might have intensified. Taken all together, the two outlined diverging processes might cancel each other out, with the result of advocacy not moderating gender differences. Of course, this idea needs further testing.

Another possibility is that at least some participants found the idea of negotiating for a friend unrealistic. A negotiator assuming the role of an advocate presupposes that someone else is willing to be advocated for—a decision that involves vulnerability and, as such, necessitates high trust (Mayer et al., 1995). Hence, in real life, people might be hesitant to be advocated for (Lietz et al., 2023), especially when a relevant workplace outcome, such as one's salary, is at stake. A supplementary analysis, reported in the Supplemental Material, gives a hint regarding the role played by experienced realism (we thank an anonymous reviewer for sharing this observation): Among participants who were picturing their friend as having a specific gender, there appeared to emerge a "conventional" gender difference in anticipated backlash, but this tendency was not observed among participants who did not think of their friend as having a specific gender. Being unable to clearly picture the supposed friend might reflect a lack of perceived realism and resulting immersion. In turn, if being an other-advocate was seen as unrealistic, it would be unsurprising that assuming this role did not unfold a strong influence, thereby reducing the likelihood of observing a moderator effect.

Practical Implications

Our research revealed gender differences in anticipated backlash, salary requests, and self-rated competitiveness. These findings have practical relevance because many employees negotiate their salary at work (e.g., Kulik & Olekalns, 2012). As such, our findings help to explain why women, up to this day, earn less than men (for overviews, see Blau & Kahn, 2017; England et al., 2020). The insight that women still ask for lower salaries than men is also relevant because gender effects are dynamic and context-bound (Bowles et al., 2025; Kray et al., 2024). In this respect, it is also worth mentioning that our research was conducted at an elite business school. This population will soon enter the workforce and likely assume influential positions. Thus, our results suggest that practitioners such as organizational diversity officers or policymakers need to take gender differences in salary negotiations seriously and take action if they are to be mitigated in the near future. Still, women's tendency to make less assertive requests can also be an advantage, as this behavior can mitigate the risk of reaching an impasse (Ma et al., 2024). Hence, if reaching an agreement is a primary goal in a negotiation, women may eventually be better off (Ma et al., 2024).

In line with Amanatullah and Morris's (2010) findings, we also observed that gender differences in salary requests were driven by women's heightened anticipation of backlash. This finding has practical relevance, too, because it provides information on how to mitigate gender differences in negotiation. We stress that women themselves are not to be "fixed" (e.g., Recalde & Vesterlund, 2023)—simply addressing their fears of backlash is unlikely to work. Backlash reactions represent a clear disadvantage for women at work, so that it is reasonable for them to take these reactions into account when considering how to negotiate. Therefore, rather than aiming to "fix women" (Recalde & Vesterlund, 2023), a key remedy for gender differences in negotiation would be to make sure that evaluators do not react with backlash (Amanatullah & Tinsley, 2013; Bowles et al., 2007). For instance, organizations may standardize procedures for determining pay. Organizational members would then be provided with guidance on what is negotiable and how a topic can be negotiated (e.g., Bowles et al., 2022), thereby reducing the likelihood with which people consider individual demographic information (e.g., gender) when negotiating salary.

Limitations and Future Research

Our key goal was to *closely* replicate the original study by Amanatullah and Morris (2010). As such, our study not only has the same strengths as the original study but also the same limitations. For instance, the alleged counterpart in the simulated negotiation was a White man. This is an important design feature because past research has revealed that gender differences are more likely with a man (as compared to a woman) as a counterpart (Bowles et al., 2007). Another relevant issue is that anticipated backlash was measured right before the negotiation (see the Method section). Hence, the risk of incurring backlash is made salient to participants, which may again increase the likelihood of observing gender differences (Bowles et al., 2022). Moreover, it is worth highlighting that the original paradigm was designed to examine women's and men's negotiation *behavior* (see Figure 1). Behavioral assertiveness certainly is a relevant criterion. Using a paradigm that allows for an examination of outcomes (e.g., actually achieved pay) would still be important, however, because this criterion strongly matters in real life. Finally, the measure of participants' self-rated negotiation style likely also had limitations. This measure included broad terms, such as competitiveness, without an accompanying explanation of how, exactly, negotiators can be competitive. This could have created ambiguity among participants as to what they are supposed to report specifically. As a result, the measure might also capture

participants' general motives or how they would like to have behaved. Hence, future research examining the validity of this measure would be in order.

Conducting a close replication—just like conducting any kind of research—also generally comes with limitations. It is no surprise that our study participants voiced the suspicion that they were interacting with an artificial intelligence bot as opposed to with a real person via a computer (although, ironically, today's technology actually enables such interactions; e.g., Molnar, 2019). Therefore, future research would be helpful that generally also adheres to the original method (including the same negotiation task, questionnaires, etc.), yet participants have a *real* counterpart with whom they interact (this counterpart could also be a confederate).

As a close replication, we also strove to run our study at a site comparable to that of the original study (i.e., a business school at an Ivy League university in the United States). However, with increasing knowledge on how gender differences are not only context-bound (e.g., Bowles et al., 2022) but also culture-bound (e.g., Shan et al., 2019), a key direction for future research would be to run replications in other cultural contexts. Doing so would prevent the literature from being filled with findings based on participants with selective cultural backgrounds.

Moreover, it is worth stressing that by focusing on a single study that we sought to closely replicate, our research does not provide insights on the replicability of other important findings from the field. For instance, as recent research suggests that women (as compared to men) do not have a lower propensity to initiate negotiations anymore (Kray et al., 2024), it again becomes important to closely replicate the original studies that suggested such a gender difference in the first place. As with our case, methods differ across studies, thereby raising the possibility that a variation in the methods (partly) explains the variation in findings.

Conclusion

There is no doubt that the study by Amanatullah and Morris (2010) plays an important role in the field of gender differences in negotiation. However, from today's perspective—with more knowledge about the role of statistical power and also about the outcomes of subsequent studies—we felt compelled to closely replicate their study. Although women more readily anticipated backlash and requested lower salaries than men, we did not observe gender to interact with advocacy to predict these criteria. Thus, our results underscore the need to continually reexamine central findings from the field. In this respect, we fully agree with Kray et al. (2024, pp. 27–28) who recently noted that "topics regarding gender in the workplace likely require more frequent updates, given the speed of change and progress that women have experienced in modern organizational life." We hope that our contributions encourage other researchers to closely replicate further central findings, not least due to the practical importance of understanding why, exactly, women are still being paid less than men.

Constraints on Generality

As our research represents a close replication, we conducted our study in a laboratory of a business school of an Ivy League institution that is located in a metropolis (i.e., Philadelphia) on the East Coast of the United States. Also consistent with the original study by Amanatullah and Morris (2010), our participants had diverse backgrounds. Still, our findings may not generalize to other populations, such as nonstudents or people from non-Western countries.

References

Amanatullah, E. T., & Morris, M. W. (2010). Negotiating gender roles: Gender differences in assertive negotiating are mediated by women's fear of backlash and attenuated when negotiating on behalf of others. *Journal of Personality and Social Psychology*, *98*(2), 256–267. https://doi.org/10.1037/a0017094

Amanatullah, E. T., & Tinsley, C. H. (2013). Punishing female negotiators for asserting too much … or not enough: Exploring why advocacy moderates backlash against assertive female negotiators. *Organizational Behavior and Human Decision Processes*, *120*(1), 110–122. https://doi.org/10.1016/j.obhdp.2012.03.006

Artz, B., Goodall, A. H., & Oswald, A. J. (2018). Do women ask? *Industrial Relations*, *57*(4), 611–636. https://doi.org/10.1111/irel.12214

Babcock, L., Gelfand, M. J., Small, D. A., & Stayn, H. (2006). Gender differences in the propensity to initiate negotiations. In D. De Cremer, M. Zeelenberg, & J. K. Murnighan (Eds.), *Social psychology and economics* (pp. 239–259). Lawrence Erlbaum.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Bear, J. B., & Babcock, L. (2017). Negotiating femininity: Gender-relevant primes improve women's economic performance in gender role incongruent negotiations. *Psychology of Women Quarterly*, *41*(2), 163–174. https://doi.org/10.1177/0361684316679652

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, *55*(3), 789–865. https://doi.org/10.1257/jel.20160995

Borenstein, M. (2009). Effect sizes for continuous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *Handbook of research synthesis and meta-analysis* (2nd ed., pp. 221–235). Sage Publications.

Bosson, J. K., Wilkerson, M., Kosakowska-Berezecka, N., Jurek, P., & Olech, M. (2022). Harder won and easier lost? Testing the double standard in gender rules in 62 countries. *Sex Roles*, *87*(1), 1–19. https://doi.org/10.1007/s11199-022-01297-y

Bowles, H. R., Babcock, L., & Lai, L. (2007). Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask. *Organizational Behavior and Human Decision Processes*, *103*(1), 84–103. https://doi.org/10.1016/j.obhdp.2006.09.001

Bowles, H. R., Babcock, L., & McGinn, K. L. (2005). Constraints and triggers: Situational mechanics of gender in negotiation. *Journal of Personality and Social Psychology*, *89*(6), 951–965. https://doi.org/10.1037/0022-3514.89.6.951

Bowles, H. R., Mazei, J., & Liu, H. H. (2025). "When" versus "whether" gender/sex differences: Insights from psychological research on negotiation, risk-taking, and leadership. *Perspectives on Psychological Science*, *20*(5), 855–873. https://doi.org/10.1177/17456916241231584

Bowles, H. R., Thomason, B., & Bear, J. B. (2019). Reconceptualizing what and how women negotiate for career advancement. *Academy of Management Journal*, *62*(6), 1645–1671. https://doi.org/10.5465/amj.2017.1497

Bowles, H. R., Thomason, B., & Macias-Alonso, I. (2022). When gender matters in organizational negotiations. *Annual Review of Organizational Psychology and Organizational Behavior*, *9*(1), 199–223. https://doi.org/10.1146/annurev-orgpsych-012420-055523

Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., Grange, J. A., Perugini, M., Spies, J. R., & van't Veer, A. (2014). The replication recipe: What makes for a convincing replication?

Journal of Experimental Social Psychology, *50*, 217–224. https://doi.org/10.1016/j.jesp.2013.10.005

Brescoll, V. L. (2011). Who takes the floor and why: Gender, power, and volubility in organizations. *Administrative Science Quarterly*, *56*(4), 622–641. https://doi.org/10.1177/0001839212439994

De Dreu, C. K. W., Weingart, L. R., & Kwon, S. (2000). Influence of social motives on integrative negotiation: A meta-analytic review and test of two theories. *Journal of Personality and Social Psychology*, *78*(5), 889–905. https://doi.org/10.1037/0022-3514.78.5.889

Desai, S. D., & Gunia, B. C. (2023). The interplay of gender and perceived sexual orientation at the bargaining table: A social dominance and intersectionalist perspective. *Organizational Behavior and Human Decision Processes*, *179*, Article 104279. https://doi.org/10.1016/j.obhdp.2023.104279

Dreber, A., Heikensten, E., & Säve-Söderbergh, J. (2022). Why do women ask for less? *Labour Economics*, *78*, Article 102204. https://doi.org/10.1016/j.labeco.2022.102204

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Lawrence Erlbaum.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, *109*(3), 573–598. https://doi.org/10.1037/0033-295X.109.3.573

Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M., & Sczesny, S. (2020). Gender stereotypes have changed: A cross-temporal meta-analysis of U.S. public opinion polls from 1946 to 2018. *American Psychologist*, *75*(3), 301–315. https://doi.org/10.1037/amp0000494

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, *46*(4), 735–754. https://doi.org/10.1037/0022-3514.46.4.735

Eagly, A. H., & Wood, W. (2013). The nature–nurture debates: 25 years of challenges in understanding the psychology of gender. *Perspectives on Psychological Science*, *8*(3), 340–357. https://doi.org/10.1177/1745691613484767

England, P., Levine, A., & Mishel, E. (2020). Progress toward gender equality in the United States has slowed or stalled. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(13), 6990–6997. https://doi.org/10.1073/pnas.1918891117

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

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage Publications.

Giner-Sorolla, R. (2018). *Powering your interaction*. https://approachingblog.wordpress.com/2018/01/24/powering-your-interaction-2/

Giner-Sorolla, R., Amodio, D. M., & van Kleef, G. A. (2018). Three strong moves to improve research and replications alike. *Behavioral and Brain Sciences*, *41*, Article e130. https://doi.org/10.1017/S0140525X18000651

Giner-Sorolla, R., Montoya, A. K., Reifman, A., Carpenter, T., Lewis, N. A., Aberson, C. L., Bostyn, D. H., Conrique, B. G., Ng, B. W., Schoemann, A. M., & Soderberg, C. (2024). Power to detect what? Considerations for planning and evaluating sample size. *Personality and Social Psychology Review*, *28*(3), 276–301. https://doi.org/10.1177/10888683241228328

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). Guilford Press.

Hüffmeier, J., Freund, P. A., Zerres, A., Backhaus, K., & Hertel, G. (2014). Being tough or being nice? A meta-analysis on the impact of hard- and softline strategies in distributive negotiations. *Journal of Management*, *40*(3), 866–892. https://doi.org/10.1177/0149206311423788

Hüffmeier, J., Mazei, J., & Schultze, T. (2016). Reconceptualizing replication as a sequence of different studies: A replication typology. *Journal of Experimental Social Psychology*, *66*, 81–92. https://doi.org/10.1016/j.jesp.2015.09.009

Jarvis, S. N., & Kray, L. (2022). *Fears of backlash underlie gendered participation in question-and-answer sessions*. https://doi.org/10.31234/osf.io/7kftn

Kennedy, J. A., & Kray, L. J. (2015). A pawn in someone else's game? The cognitive, motivational, and paradigmatic barriers to women's excelling in negotiation. *Research in Organizational Behavior*, *35*, 3–28. https://doi.org/10.1016/j.riob.2015.09.002

Kouchaki, M., & Kray, L. J. (2018). "I won't let you down:" Personal ethical lapses arising from women's advocating for others. *Organizational Behavior and Human Decision Processes*, *147*, 147–157. https://doi.org/10.1016/j.obhdp.2018.06.001

Kray, L. J., Kennedy, J. A., & Lee, M. (2024). Now, women do ask: A call to update beliefs about the gender pay gap. *Academy of Management Discoveries*, *10*(1), 7–33. https://doi.org/10.5465/amd.2022.0021

Kray, L. J., Thompson, L., & Galinsky, A. (2001). Battle of the sexes: Gender stereotype confirmation and reactance in negotiations. *Journal of Personality and Social Psychology*, *80*(6), 942–958. https://doi.org/10.1037/0022-3514.80.6.942

Kugler, K. G., Reif, J. A. M., Kaschner, T., & Brodbeck, F. C. (2018). Gender differences in the initiation of negotiations: A meta-analysis. *Psychological Bulletin*, *144*(2), 198–222. https://doi.org/10.1037/bul0000135

Kulik, C. T., & Olekalns, M. (2012). Negotiating the gender divide: Lessons from the negotiation and organizational behavior literatures. *Journal of Management*, *38*(4), 1387–1415. https://doi.org/10.1177/0149206311431307

LeBel, E. P., McCarthy, R. J., Earp, B. D., Elson, M., & Vanpaemel, W. (2018). A unified framework to quantify the credibility of scientific findings. *Advances in Methods and Practices in Psychological Science*, *1*(3), 389–402. https://doi.org/10.1177/2515245918787489

LeBel, E. P., Vanpaemel, W., Cheung, I., & Campbell, L. (2019). A brief guide to evaluate replications. *Meta-Psychology*, *3*, 1–9. https://doi.org/10.15626/MP.2018.843

Leigh, A., & Desai, S. D. (2023). What's race got to do with it? The interactive effect of race and gender on negotiation offers and outcomes. *Organization Science*, *34*(2), 935–958. https://doi.org/10.1287/orsc.2022.1629

Lietz, M., Mazei, J., Mertes, M., & Hüffmeier, J. (2023). Are strategies for women in compensation negotiations more appealing when it is explained how they are meant to impact negotiation outcomes? *Psychology of Women Quarterly*, *47*(1), 80–112. https://doi.org/10.1177/03616843221128484

London, B., Downey, G., Romero-Canyas, R., Rattan, A., & Tyson, D. (2012). Gender-based rejection sensitivity and academic self-silencing in women. *Journal of Personality and Social Psychology*, *102*(5), 961–979. https://doi.org/10.1037/a0026615

Ma, A., Ponce de Leon, R., & Rosette, A. S. (2024). Asking for less (but receiving more): Women avoid impasses and outperform men when negotiators have weak alternatives. *Journal of Applied Psychology*, *109*(7), 1145–1158. https://doi.org/10.1037/apl0001138

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*(3), 709–734. https://doi.org/10.2307/258792

Mazei, J., Bear, J. B., Schaumberg, R., & Hüffmeier, J. (2025). *Study materials, data, and syntax for a close replication of an original study by Amanatullah and Morris (2010)*. https://osf.io/6bfvq/?view_only=5d54007b1dca40b2b21a4eaf1413e7a1

Mazei, J., Hüffmeier, J., Freund, P. A., Stuhlmacher, A. F., Bilke, L., & Hertel, G. (2015). A meta-analysis on gender differences in negotiation outcomes and their moderators. *Psychological Bulletin*, *141*(1), 85–104. https://doi.org/10.1037/a0038184

Molnar, A. (2019). SMARTRIQS: A simple method allowing real-time respondent interaction in Qualtrics surveys. *Journal of Behavioral and*

Experimental Finance, *22*, 161–169. https://doi.org/10.1016/j.jbef.2019.03.005

Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology's renaissance. *Annual Review of Psychology*, *69*(1), 511–534. https://doi.org/10.1146/annurev-psych-122216-011836

Nohe, C., Hüffmeier, J., Bürkner, P., Mazei, J., Sondern, D., Runte, A., Sieber, F., & Hertel, G. (2022). Unethical choice in negotiations: A meta-analysis on gender differences and their moderators. *Organizational Behavior and Human Decision Processes*, *173*, Article 104189. https://doi.org/10.1016/j.obhdp.2022.104189

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., … Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), Article aac4716. https://doi.org/10.1126/science.aac4716

Recalde, M. P., & Vesterlund, L. (2023). Gender differences in negotiation: Can interventions reduce the gap? *Annual Review of Economics*, *15*(1), 633–657. https://doi.org/10.1146/annurev-economics-092022-115353

Rubin, J. Z., & Brown, B. R. (1975). *The social psychology of bargaining and negotiation*. Academic Press.

Rubin, J. Z., Pruitt, D. G., & Kim, S. H. (1994). *Social conflict: Escalation, stalemate, and settlement* (2nd ed.). Mcgraw-Hill Book Company.

Rudman, L. A. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. *Journal of Personality and Social Psychology*, *74*(3), 629–645. https://doi.org/10.1037/0022-3514.74.3.629

Rudman, L. A., & Fairchild, K. (2004). Reactions to counterstereotypic behavior: The role of backlash in cultural stereotype maintenance. *Journal of Personality and Social Psychology*, *87*(2), 157–176. https://doi.org/10.1037/0022-3514.87.2.157

Rudman, L. A., Moss-Racusin, C. A., Phelan, J. E., & Nauts, S. (2012). Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology*, *48*(1), 165–179. https://doi.org/10.1016/j.jesp.2011.10.008

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, *13*(2), 90–100. https://doi.org/10.1037/a0015108

Schoemann, A. M., Boulton, A. J., & Short, S. D. (2017). Determining power and sample size for simple and complex mediation models. *Social Psychological and Personality Science*, *8*(4), 379–386. https://doi.org/10.1177/1948550617715068

Shan, W., Keller, J., & Joseph, D. (2019). Are men better negotiators everywhere? A meta-analysis of how gender differences in negotiation performance vary across cultures. *Journal of Organizational Behavior*, *40*(6), 651–675. https://doi.org/10.1002/job.2357

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 word solution. *Dialogue*, *26*, 4–7.

Simonsohn, U. (2014). *No-way interactions*. https://datacolada.org/2014/03/12/17-no-way-interactions-2/

Simonsohn, U. (2015). Small telescopes: Detectability and the evaluation of replication results. *Psychological Science*, *26*(5), 559–569. https://doi.org/10.1177/0956797614567341

Singelis, T. M. (1994). The measurement of independent and interdependent self-construals. *Personality and Social Psychology Bulletin*, *20*(5), 580–591. https://doi.org/10.1177/0146167294205014

Small, D. A., Gelfand, M., Babcock, L., & Gettman, H. (2007). Who goes to the bargaining table? The influence of gender and framing on the initiation of negotiation. *Journal of Personality and Social Psychology*, *93*(4), 600–613. https://doi.org/10.1037/0022-3514.93.4.600

Stuhlmacher, A. F., & Linnabery, E. (2013). Gender and negotiation: A social role analysis. In M. Olekalns & W. Adair (Eds.), *Handbook of research on negotiation* (pp. 221–248). Edward Elgar. https://doi.org/10.4337/9781781005903.00018

Stuhlmacher, A. F., & Walters, A. E. (1999). Gender differences in negotiation outcome: A meta-analysis. *Personnel Psychology*, *52*(3), 653–677. https://doi.org/10.1111/j.1744-6570.1999.tb00175.x

Thompson, L. L. (2009). *The mind and heart of the negotiator* (4th ed.). Prentice Hall.

Toosi, N. R., Mor, S., Semnani-Azad, Z., Phillips, K. W., & Amanatullah, E. T. (2019). Who can lean in? The intersecting role of race and gender in negotiations. *Psychology of Women Quarterly*, *43*(1), 7–21. https://doi.org/10.1177/0361684318800492

Walters, A. E., Stuhlmacher, A. F., & Meyer, L. L. (1998). Gender and negotiator competitiveness: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *76*(1), 1–29. https://doi.org/10.1006/obhd.1998.2797

Williams, M. J., & Tiedens, L. Z. (2016). The subtle suspension of backlash: A meta-analysis of penalties for women's implicit and explicit dominance behavior. *Psychological Bulletin*, *142*(2), 165–197. https://doi.org/10.1037/bul0000039

Wilson, D. B. (2023). *Practical meta-analysis effect size calculator* (Version 2023.11.27) [Computer software]. https://www.campbellcollaboration.org/calculator/

Received September 9, 2020
Revision received September 25, 2025
Accepted October 17, 2025 ■