Received: 13 June 2022 | Accepted: 18 December 2023

DOI: 10.1002/ejsp.3023

RESEARCH ARTICLE

EASP WILEY

Gender stereotypes may not influence the choice of female leaders: Experimental evidence from a crisis framed as social or economic during the COVID-19 pandemic

Ruri Takizawa[1] | Vincenzo Iacoviello[1] | Michelle K. Ryan[2,3] | Clara Kulich[1]

[1]Faculty of Psychology and Educational Sciences (FPSE), University of Geneva, Geneva, Switzerland

[2]Global Institute for Women's Leadership, The Australian National University, Canberra, Australia

[3]Faculty of Economics and Business, Organisational Behaviour, University of Groningen, Groningen, The Netherlands

Correspondence
Ruri Takizawa, Université de Genève, FPSE, Uni Mail, Boulevard du Pont d'Arve 40, CH-1211 Genève 4, Switzerland.
Email: ruri.takizawa@unige.ch

Funding information
Swiss National Science Foundation, Grant/Award Number: 100019_188934; European Research Council, Grant/Award Number: 725128

Abstract

This research examined whether female (vs. male) leaders are preferred during a pandemic when stereotypically feminine leadership is deemed useful. We hypothesized that citizens prefer female (vs. male) politicians when the crisis is framed as a social (vs. economic) crisis because they believe it requires feminine (vs. masculine) leadership. In a pilot study and three online experiments with US residents (N_{total} = 1675), we manipulated crisis type or a leadership candidate's gender for a task force. While participants indicated that a crisis framed as social (vs. economic) required more feminine leadership, they did not appoint a woman more or rated her as more suitable for the social crisis (vs. economic crisis or a no-crisis situation). Furthermore, the female (vs. male) candidate was not perceived to possess more feminine leadership traits. Overall, participants did not rely on gender stereotypes when explicitly evaluating politicians. We discuss potential explanations for these unexpected results.

KEYWORDS
crisis Leadership, gender stereotypes, glass cliff, politicians, politics, social roles

1 | INTRODUCTION

Prior research has demonstrated a gendered division of tasks in political leadership in line with gender stereotypes and social roles (Eagly & Steffen, 1984). For example, female leaders are perceived to be more people-oriented than male leaders and are allocated tasks accordingly (Alexander & Andersen, 1993). The COVID-19 pandemic provided a recent example where researchers and practitioners commented on and investigated the effectiveness of politicians' leadership in relation to politicians' gender and gendered leadership styles (e.g., Funk, 2020; Grebelsky-Lichtman & Katz, 2020; McGuire et al., 2020; Sergent & Stajkovic, 2020). The present research seeks to understand whether gender stereotypes also shaped citizens' leadership preferences during the COVID-19 crisis. In particular, we focused on the association between candidate gender and the types of crises that are perceived to require skills congruent with gender stereotypes.

1.1 | Evaluations of women (vs. men) in leadership

Research has repeatedly highlighted that political tasks are perceived to be gendered. Whereas male politicians dominate in units addressing the economy, fiscal affairs, foreign affairs or crime, female politicians are more likely found in so-called *domestic* units addressing health and education (Herrnson et al., 2003). Moreover, much research illustrated that voters evaluate female and male politicians according to gender stereotypes (Holman, 2023; Huddy & Terkildsen, 1993a; Koch, 1999, 2000). As in the business and management domains (Eagly et al.,

1992; Koenig et al., 2011; Sczesny et al., 2004), female actors are perceived to be more communal (e.g., people-oriented and honest) and male actors are generally seen as more agentic (e.g., assertive and self-confident, Alexander & Andersen, 1993; Rosenwasser & Dean, 1989). These perceptions may reflect on what leadership style is perceived in, or attributed to, female and male politicians. Thus, female politicians may either generally lead in more stereotypically feminine (i.e., caring) ways than men (e.g., Funk & Philips, 2019) or women may be more likely than men to lead in areas where communal skills are relevant (Eagly & Steffen, 1984).

In particular, in the COVID-19 crisis, a widespread belief suggested that female leaders were managing the crisis better than male leaders when it came to protecting vulnerable group members (Chamorro-Premuzic & Gallop, 2020; Henley & Roy, 2020; Katz, 2021; Taub, 2020). Initial research reported that female-led countries controlled the spread of the virus faster than male-led countries or reported lower COVID-19-related deaths (Coscieme et al., 2020; Chen et al., 2023). Such findings have led to a call for more female leadership in the crisis (e.g., Pan American Health Organization [PAHO], 2021; UN Women, 2020).

1.2 | A female advantage in crisis leadership?

The gender and leadership literature indeed gives accounts of crisis–women associations. The glass cliff metaphor and supporting evidence have emphasized that women are disproportionately appointed to leadership positions in a crisis (Morgenroth et al., 2020; Ryan et al., 2016). Several explanations for the glass cliff have been advanced, with one influential factor being the goal envisioned for the leadership appointment (Ryan et al., 2011). The COVID-19 crisis had a direct and disruptive impact on many people's lives. So, what did people expect from leaders in this context?

Leadership qualities associated with solving a crisis are people-orientation and communion (Mano-Negrin & Sheaffer, 2004). Because female leaders, more than male leaders, are typically perceived to possess these qualities, they are seen as more apt to handle crises effectively (Ryan et al., 2011). This *think crisis–think female* association has been experimentally tested in business scenarios, with results suggesting that people believed that it was more important for managers of struggling companies to have communal (i.e., stereotypically feminine) traits than agentic traits (Study 2; Ryan et al., 2011) and that this belief was linked to a preference for female over male candidates in crisis (Study 2; Bruckmüller & Branscombe, 2010). One may thus reason that women might be seen as more effective leaders in crisis *because* of their ascribed communal leadership traits (see also Oliver et al., 2022).

However, other research demonstrated that agentic leaders were preferred over communal leaders for poorly performing companies compared to strongly performing companies independent of the leader's gender if leadership qualifications motivated the appointment (Kulich et al., 2018). This finding suggests that stereotypically feminine (i.e., communal) traits are not always deemed useful in a crisis. Following up on these disparities, Kulich et al. (2021) proposed that different types of crises may trigger demands for either agentic or communal leadership – independent of the leader's gender. They experimentally illustrated that the preference for communal leaders was stronger in a relational crisis, which implicitly demands stereotypically feminine qualities, than in a financial crisis, which demands masculine traits. The trait–crisis type association aligns with the mechanism suggested by the *think crisis–think female* metaphor.

In the same study, the effect of candidate gender on the candidate choice was smaller, and participants preferred women in a crisis (vs. no crisis), independent of its relational or financial nature. This tendency suggests that (1) gendered leadership traits did not explain candidate choice as regards their actual gender and (2) other mechanisms explaining the preference for women in crisis contexts need to be uncovered. Thus, more research is needed to understand the causalities of gender and gendered traits in different types of crises (Bavik et al., 2021; Hannah et al., 2009). The *think crisis–think female* association has not yet been experimentally investigated in the political domain, with clear differentiation between the types of crises demanding stereotypically feminine or masculine qualities (Morgenroth et al., 2020), even though gender stereotypes have been reported to impact politicians' evaluations significantly (Huddy & Terkildsen, 1993a).

1.3 | Leaders' gender, gendered leadership and the COVID-19 crisis

The COVID-19 pandemic provides an example where politicians' gender and gendered leadership styles were focal discussion points for determining effective leadership in both the press and research (Funk, 2020; Grebelsky-Lichtman & Katz, 2020; McGuire et al., 2020; Sergent & Stajkovic, 2020). Many media posts contrasted female and male politicians' leadership styles and reported that female political leaders handled the health crisis differently from their male counterparts (Bell, 2020; Park, 2022; Waylen, 2021). Researchers reported that female leaders communicated with more empathy and cooperativeness than male leaders and showed more engagement in protecting vulnerable communities (Funk, 2020; Grebelsky-Lichtman & Katz, 2020; McGuire et al., 2020; Sergent & Stajkovic, 2020). The present research seeks to understand whether gender stereotypes about political leaders also shaped citizens' leadership perceptions and whether preferences for female or male leaders depended on how the COVID-19 crisis was framed: either as a social crisis, a domain often associated with women and femininity or an economic crisis, a domain associated with men and masculinity.

The COVID-19 crisis provided an opportunity to test different aspects of a crisis because the pandemic was not only described as a health crisis but also encompassed economic (Béland et al., 2020; Donthu & Gustafsson, 2020; Pak et al., 2020) and social consequences (Fisher & Ryan, 2021; Profeta, 2020; Rimmer, 2020). To gain initial insights into when and why female political leaders were preferred in the COVID-19 crisis, Takizawa et al. (2022) collected observational data with an online questionnaire in five countries during the first year of the pandemic, exploring predictors of female political leader

preferences. They were particularly interested in whether participants would prefer female political leaders if they perceived the social aspects of the crisis to be severe and less so when the economic crisis was perceived to be severe. However, the results were inconclusive. Considering the correlational nature of survey data and the different perceptions people may have in the COVID-19 crisis contexts, we need more evidence from a controlled setting to better understand crisis type–leader gender associations.

2 | THE PRESENT RESEARCH

We experimentally tested whether the *think crisis–think female* association depended on the qualities that were perceived as relevant in a crisis. More specifically, we investigated in the COVID-19 pandemic whether a female political leader was preferred (over a male leader) in a crisis when management required qualities that are stereotypically linked to women and feminine leadership. *Hypothesis 1* predicted that social consequences (compared to economic consequences) of the COVID-19 crisis were perceived to require more stereotypically feminine competencies, namely communal (H_{1a}) and moral (H_{1b}) leadership, and fewer stereotypically masculine competencies (agentic leadership, H_{1c}). We added the dimension of morality, often characterized as being ethical, trustworthy and honest (Leach et al., 2007), to the two universal dimensions of social perception (communion and agency; Fiske et al., 2007) because trust was an important factor for predicting leadership preferences in the COVID-19 pandemic (Bollyky et al., 2023; Everett et al., 2021). Although morality has been shown to be conceptually different from communion (Leach et al., 2007), they are both associated with women and femininity (e.g., Huddy & Terkildsen, 1993b). In politics, moral leadership is also associated with female leaders (Barnes & Beaulieu, 2014; Pew Research Center, 2015). If H_1 is supported and if leaders are associated with gendered stereotypes, then female leaders should be more strongly preferred than male leaders in contexts of the COVID-19 crisis, which require stereotypically feminine qualities. Thus, in *Hypothesis 2*, we predicted that a female (vs. male) candidate would be preferred to lead a unit dealing with social consequences compared to one dealing with economic consequences.

We conducted a pilot study and three online experiments by creating fictitious scenarios about COVID-19 task forces in the United States, which were described as either social or economic units, and participants were asked to match these through different experimental designs with male or female political leaders. National task forces guiding political decisions through the COVID-19 crisis were set up in many countries as one of the first steps in the public health emergency. In the United States, the president established the first task force, which overlooked all measures related to the crisis (WhiteHouse.gov, 2020). The task force even had the power to overturn the Centers for Disease Control and Prevention (e.g., Kaplan, 2020). Based on the recognized importance of such COVID-19 task forces and their much-disputed effectiveness (e.g., Santucci, 2020), a task force provided an excellent framework with high external validity to use in our experiments.

2.1 | Data transparency

Data and the pre-registration of Study 3 are available on the open registries network at https://osf.io/myh9c/?view_only=d0206dc7cfd44895a46e00562b519a71 . We report all measures, manipulations and exclusions of all studies in the paper. Further details on materials, additional measures and explorative analyses are presented in the Supporting Information (SM 1.3 for Study 1, SM 2.2 for Study 2, SM 3.3 for Study 3). We performed data analyses with the sample sizes provided herein. No additional data were sought for any of the studies after the initial data analysis. Participants in all studies were randomly assigned to experimental conditions.

3 | PILOT STUDY

To test whether communal leadership traits and female politicians were associated more with a crisis framed as producing *social* consequences than one framed as producing *economic* consequences, we conducted an initial study in June 2021 with 407 Prolific participants based in the United States. Details of the study are available in the Supporting Information (SM Pilot Study). We manipulated the gender of the political candidate (man vs. woman) who would lead a COVID-19 task force unit, and participants had to appoint the candidate to either the unit dealing with the economic crisis or the one dealing with the social crisis. Additionally, we measured the perceived importance of agentic and communal leadership traits for each position. Results showed that communal leadership traits were rated as more important for the social unit than for the economic unit (H_{1a}). Furthermore, agentic leadership traits were rated as less important for the social unit than the economic unit (H_{1c}). However, we found no effect of candidate gender on choice. Participants presented with a female candidate were not more likely to select the social unit than those with a male candidate. Due to the lack of results in this pilot study, we decided to launch a series of studies to test and exclude alternative explanations for the lack of a link between female leaders and a stereotypically feminine crisis.

4 | STUDY 1

Study 1 tested H_1 and H_2 with improved candidate biographies (such as highlighting expertise in both social and economic issues; see SM for details) and a between-participant manipulation of the mission of the task force (economic vs. social). Moreover, we measured how suitable participants rated a female versus a male politician in the position to lead a unit in economic or social crises (see also Morgenroth et al., 2020). With the suitability rating, we expected to measure more nuanced and honest responses from participants. *Hypothesis 2* was adapted such that we predicted that a female candidate is *rated as more suitable* than a male candidate to lead a unit tackling a social crisis rather than an economic crisis.

4.1 | Method

4.1.1 | Participants

An a priori power analysis for a multiple regression for H_2 ($\alpha = .05$, 80% power) with one tested predictor (Candidate gender × Crisis unit) determined a sample size of 407 for a small effect ($\eta_p^2 = .019$, based on Study 3 from Haslam & Ryan, 2008; between-participants effects from Morgenroth et al., 2020; and the pilot study's suitability ratings for male and female candidates). We recruited 560 participants on Prolific in October 2021 and excluded participants who did not live in the United States ($n = 69$), failed the attention check ($n = 10$), failed the comprehension checks ($n = 47$) or did not specify their gender ($n = 1$). A total of 433 participants were included in the analysis (52.7% women and 47.3% men).

Participants reported their age, $M = 32.75$, $SD = 10.61$, from 18 to 79 years, employment status (67.7% employed) and ethnic identity (74.1% Caucasian or White, 4.8% Asian or Asian-American, 6.2% Hispanic or Latino, 15.7% African-American or Black, 1.2% Native American, Alaska Native, Hawaiian or Pacific Islander, 2.3% Other or Multi-Cultural).

4.1.2 | Study design and procedure

The study had a 2 Candidate gender × 2 Task force unit (social vs. economic) between-participant design. Participants first read a fictitious vignette about a US state where a COVID-19 task force was established. They either learned about the economic unit, 'addressing the economic crisis, which includes the worsened economic situation for many people at present and/or in the future', or the social unit, 'addressing the social crisis, which encompasses social inequalities or social problems that are worse for many people at present and/or in the future' (see SM 1.1). The text explained that the presented task force unit needed a new director to lead the unit's tasks. Next, participants wrote down what they associated with the economic or social aspects of the crisis (matching the unit that they had seen), rated how important they considered the presented unit's responsibilities and how difficult it was to lead. Participants then indicated how important agentic, communal and moral traits were for leading that unit. In the next step, they saw a female or male political candidate shortlisted for the task force director position (SM 1.2). In these biographies, we reported candidates' academic and work-related economic backgrounds as well as their social backgrounds. It was stressed that the candidates were already active government members ('Each unit is led by a director *who is a political representative* appointed by the state government'). Finally, participants rated the candidate's suitability for the position, answered some additional questions about their perceptions and ended with questions about participants' perceptions of the COVID-19 crisis and their demographic information.

4.1.3 | Measures

Comprehension check about the crisis. After the COVID-19 task force information, we asked participants to select the type of crisis that the new unit director in the task force would need to address ('Social inequalities' or 'Economic difficulties').

Agentic, communal and moral leadership traits. We searched perceived ideal traits of *politicians* in the literature and categorized them into the agentic and communal dimensions (Cwalina & Falkowski, 2005; Liberska & Jankowiak, 2016). We compared them to the leadership traits we used in the pilot study and adapted the list with items that seemed more fitting in the political context. Importance ratings were assessed (7-point Likert scales labelled 1 *not at all important* to 7 *very important)* for four agentic leadership traits[1] (*assertive*, *confident*, *determined*, *powerful*; $\alpha = .63$) and six communal leadership traits (*friendly*, *considerate*, *empathetic*, *socially oriented*, *cooperative*, *sensitive to the needs of others*; $\alpha = .82$). Moreover, seven items assessed the perceived importance of moral leadership traits (*honest*, *sincere*, *trustworthy*, *reliable*, *genuine*, *means what they say*, *principled*; $\alpha = .92$; Boukes & Boomgaarden, 2016; Clarke et al., 2018; Leach et al., 2007).

Comprehension check about the candidate. Following the candidate's biography, participants selected the academic backgrounds of the candidate ('social science' and 'economics'). Participants were excluded from analysis if they selected the wrong backgrounds ('biological science' or 'law').

Leader suitability. Six items, based on Haslam and Ryan (2008), measured participant's evaluation of the presented candidate regarding their leadership abilities (e.g., 'This candidate has the required competencies for this job') and their suitability (e.g., 'This candidate fits with this position') for the position on a 7-point Likert scale from 1 *Completely disagree* to 7 *Completely agree* ($\alpha = .82$, $M = 5.49$, $SD = 0.95$).

Comprehension check of candidate gender. Before the demographic questions, participants were asked the name of the candidate they were shown ('Robert Wilson', 'Laura Wilson', 'Neither').

Participant political leaning. Political leaning was measured by an 11-point scale from 0 *left* to 11 *right* ($M = 6.22$, $SD = 3.54$).

4.2 | Results

4.2.1 | Perceptions of economic and social crisis units (H_1)

First, we conducted a repeated-measures analysis of variance (ANOVA) on importance ratings of the leadership traits (agentic, communal, moral) with the unit (social vs. economic) as a between-participant variable. As generally predicted by H_1, we found a Traits

[1] Due to an error in the questionnaire, different agentic items were partially shown depending on the condition. Therefore, we only included the four agentic items that were assessed in both conditions. Moreover, the item *honest* was accidentally used twice as a communal and moral dimension and used for the moral dimension.

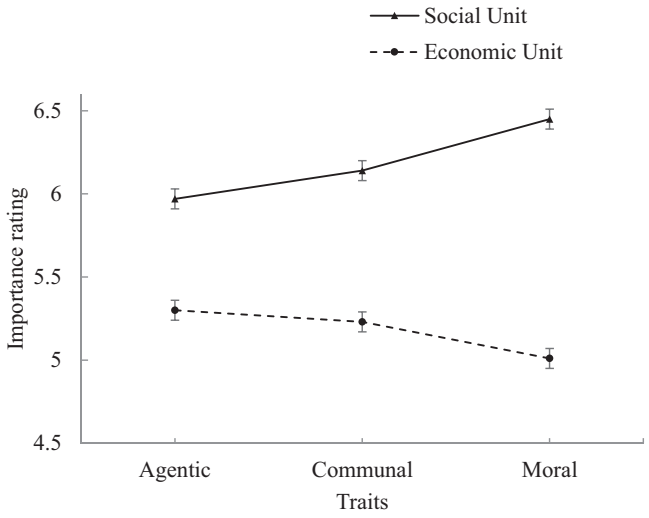

FIGURE 1 Importance ratings of agentic, communal and moral traits for the economic and social units. Error bars indicate standard errors.

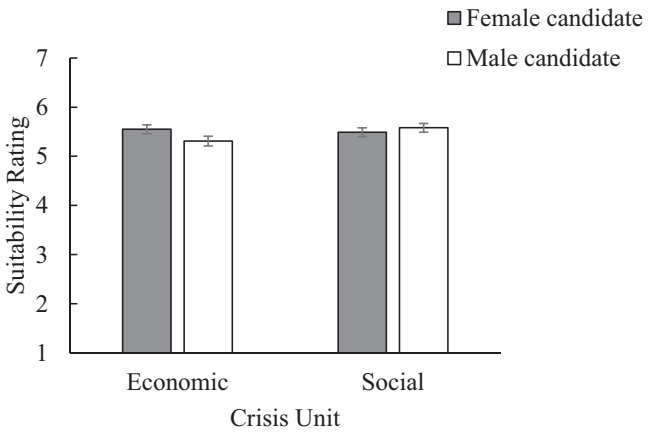

FIGURE 2 Suitability ratings of the candidates depending on their gender and the presented crisis unit. Error bars indicate standard errors.

× Unit interaction, $F(1431) = 96.28$, $p < .001$, $\eta_p^2 = .18$. Post hoc tests revealed that H_{1a} and H_{1b} were supported as communal, 95% CI [−1.08, −0.75], $p < .001$ and moral traits, 95% CI [−1.60, −1.29], $p < .001$, were both perceived to be more important for participants in the social than in the economic unit. Against expectations, agentic traits were also rated as more important by participants in the social (vs. economic) unit, 95% CI [−0.83, −0.51], $p < .001$. Thus, H_{1c} was not supported. It should, however, be noted that the effect size of the difference in unit ratings was smallest for the agentic traits ($MD = 0.67$, $SE = 0.08$, $F(1431) = 65.32$, $p < .001$, $\eta_p^2 = .13$; communal traits: $MD = 0.92$, $SE = 0.08$, $F(1431) = 119.07$, $p < .001$, $\eta_p^2 = .22$; moral traits, $MD = 1.44$, $SE = 0.08$, $F(1431) = 331.47$, $p < .001$, $\eta_p^2 = .44$; see Figure 1). Following this, we tested whether the size of between-unit discrepancies significantly varied between the trait dimensions. We conducted three separate repeated-measures ANOVAs comparing each dimension pair. The analyses revealed that all between-unit discrepancies were different (agentic vs. communal: $F(1431) = 8.66$, $p = .003$, $\eta_p^2 = .02$, agentic vs. moral: $F(1431) = 96.28$, $p < .001$, $\eta_p^2 = .18$, communal vs. moral: $F(1431) = 54.26$, $p < .001$, $\eta_p^2 = .11$). The largest effect size was found for the agentic-moral trait comparison.

4.2.2 | Suitability rating depending on candidate gender (H_2)

In all regression analyses in which we tested H_2, we added participant gender and political leaning as covariates (including all interactions). Previous findings showed that they influenced how both economic and social aspects of the COVID-19 crisis were perceived and whether female political leadership was preferred (Takizawa et al., 2022).

To test whether the female (vs. male) candidate was rated as more suitable for a leadership position if the crisis was social in nature rather than economic, we conducted a regression analysis on suitability with candidate gender (−1 = man, 1 = woman) and crisis unit (−1 = economic, 1 = social) as predictors (see Appendix A for complete results with full statistics). Participants' gender (−1 = man, 1 = woman) and political leaning (centralized), including all interactions, were added as covariates. Regarding our hypothesis, we found no interaction between candidate gender and crisis unit, $B = −0.08$, $SE = 0.05$, 95% CI [−0.17, 0.01], $p = .084$, $\eta_p^2 = .01$ (see Figure 2). All other effects, including those of covariates and interactions, were non-significant, $ps > .051$.

4.3 | Discussion

We found partial support for H_1, showing that communal and moral leadership traits were rated more important for the social than the economic unit, replicating the results of the pilot study. Although smaller, the same pattern occurred for agentic traits, which was at odds with our prediction and with findings from the pilot study. Regarding H_2, suitability ratings were not affected by candidate gender and crisis type, which is consistent with the absence of results for candidate choice in the pilot study.

Several reasons could underline the deviating results for agency. First, a programming error led to participants in the economic and social conditions seeing three different agency items (*strong-minded*, *forceful*, *dedicated* in the economic condition and *skillful*, *task-oriented*, *confident* in the social condition), next to four items that were common to both conditions. Even though we only used the items that were present in both conditions for the analyses, the other items could have involuntarily influenced the ratings of the remaining agency items (Asch, 1946; Yzerbyt & Klein, 2019). This inconsistency might have led to different baseline ratings of the remaining items between the conditions.

Second, we compared the ratings of the units *between* participants in Study 1 (every participant only read about one unit) and *within* participants in the pilot study (every participant saw both units). The absence

of the expected effect of higher importance of agency in the economic (vs. social) unit in Study 1, but its presence in the pilot study, may be due to the direct comparison in the latter. The social unit might be generally perceived to require agentic leadership traits, but in direct comparison to the economic unit, stereotypes linking agency to the economy might become more influential (see Biernat & Manis, 1994). However, complementary analyses illustrated that leading the economic unit is perceived to be more difficult than leading the social unit, which may contradict this potential explanation (see SM 1.3).

So far, both studies suggest that a social (vs. economic) crisis context is perceived to require more stereotypically feminine traits, but this does not result in a preference for female candidates.

5 | STUDY 2

Study 1 left two major questions open, which concern (1) our negligence of testing whether female candidates are actually perceived to possess more communal and less agentic traits than male candidates and (2) we did not consider alternative mechanisms that could explain the choice of women.

5.1 | Gendered stereotypes and candidate gender

Our implicit premise was that participants would perceive the female candidate as endorsing stereotypically feminine (communal) characteristics, therefore choosing her more in a unit requiring such characteristics (i.e., the social unit). However, so far, we have investigated whether perceptions of ideal gendered leadership styles in a given situation would influence candidate preferences (appointments or suitability ratings) regarding their gender. Thus, while our studies informed us about participants' importance ratings of the units of social and economic affairs along the communion-agency trait dimensions, we do not know how the female versus male candidate was perceived along these same dimensions. Study 3 thus sought to investigate gender effects on leadership trait ascriptions, advancing *Hypothesis 3* that participants would perceive the female candidate to possess more communal and moral traits and less agentic traits than the male candidate. Extending this interest, we also explored whether choices are affected by perceptions that women *endorse* communal *values* more than men (Block et al., 2018; Diekman et al., 2010; Konrad et al., 2000).

5.2 | Gender differences concerning change potential

The presented results did not reveal candidate gender effects when comparing crisis types. One reason may be that the effect may have been too weak, as previous findings suggest a stronger crisis-type effect for candidates' gendered traits than for their gender (Kulich et al., 2021). Another theoretically interesting reason could be that alternative mechanisms explain glass cliff choices. A crisis implies that change is required (Pearson & Clair, 1998), and appointing a leader who is different from previous leaders in a crisis signals a change by shifting away from the common, default choice of male leadership (Brown et al., 2011; Bruckmüller & Branscombe, 2010; Reinwald et al., 2023). Therefore, it is not necessarily believed that the female appointee will implement actual changes (Kulich et al., 2015). In the corporate world, the term *window dressing* is used when a woman is appointed to a high-power position to signal to shareholders that the company is doing *something* to address a crisis (Helland & Sykuta, 2004). So, the type of crisis may be unimportant and female candidates could be preferred in any crisis context. If this mechanism also applies in the political domain, the results from our previous studies come with little surprise because we only compared crisis situations.

In political democracies, politicians gain power through the support and votes of the citizens (e.g., Zingher & Farrer, 2016). Therefore, the appointed candidates will need to have a positive impact on public opinion. So, their perceived potential to signal change might be important for being appointed (an effect found for ethnic minority candidates, see Study 2, Aelenei et al., 2020). Regarding gender, the observational COVID-19 crisis study demonstrated that the preference for female leadership increased for those who did not usually endorse female leadership if people were dissatisfied with current political actions, implying that those people might be hoping for some sort of change (Takizawa et al., 2022).

Complementing the gender stereotype hypotheses about leadership styles and the encouraging results from the observational study, we thus aimed to test whether female and male political candidates for high-power positions were perceived to differ in their potential to signal change. In *Hypothesis 4*, we predicted that participants would perceive the female candidate to possess greater potential to signal change but not greater potential to implement actual change than the male candidate.

Moreover, we wanted to confront the *signalling change hypothesis* with the *gender stereotypes hypothesis* by adding a no-crisis comparison condition and analysing candidate appointments across the three conditions. According to the *signalling change hypothesis*, we predicted in *Hypothesis 5a* that the female (vs. male) candidate should more likely be appointed to a crisis context compared to a no-crisis context regardless of crisis type. However, if the *gender stereotypes hypothesis* holds, the alternative *Hypothesis 5b* predicts that the female candidate would be more likely appointed to the social crisis than the economic crisis and no-crisis context, and the male candidate more likely to the economic crisis than the social crisis and no-crisis context.

5.3 | Method

5.3.1 | Participants

As in the pilot study, we ran an a priori power analysis based on Study 3 from Kulich et al. (2021), which conducted a similar analysis. The required sample size (80% power, $\alpha = .05$, odds ratio = 1.80) was 375.

In November 2021, 457 US participants completed our questionnaire via Prolific. We excluded participants who failed the attention check ($n = 5$), the comprehension checks ($n = 26$) or those who did not specify their gender or identified as non-binary because there were too few of them to create an independent group ($n = 6$). Thus, 420 participants were included in the analysis (49.5% women and 50.5% men). A sensitivity analysis for logistic regression revealed that with our sample, $\alpha = .05$, and 80% power, the minimum effect size that we could detect was an odds ratio = 1.74.

Participants reported their age, $M = 42.40$, $SD = 13.85$, from 19 to 79 years, employment status (65.5% employed) and ethnic identity (86.7% Caucasian or White, 4.8% Asian or Asian-American, 4% Hispanic or Latino, 6% African-American or Black, 1.4% Native American, Alaska Native, Hawaiian or Pacific Islander, 1.7% Other or Multi-Cultural).

5.3.2 | Procedure

Because we added a no-crisis context, the scenarios were changed to a COVID-19 crisis task force that sent one political representative to three different city councils (randomized presentation; see SM 2.1) that either increasingly struggled 'with the consequences of the social crisis' or 'with the consequences of the economic crisis' and needed to 'improve the situation and the public opinion'. In the no-crisis city, difficulties resulting from the COVID-19 crisis were presented as having been efficiently dealt with and the focus was set on 'regular affairs' and on 'maintaining the positive public opinion'. Next, participants responded to a comprehension check and continued to rate the political representative's task in each city in terms of importance and difficulty. Participants then saw a female or male political candidate (same materials used in Study 1). They rated the presented candidate regarding agentic, communal and moral leadership traits, perceived values of the candidate and their change potential. Finally, they appointed the candidate to one of the cities (with its situation described again). The questionnaire ended with demographic questions after answering some additional questions.

5.3.3 | Measures

Comprehension check of the task force unit. Following information about the COVID-19 task force and the cities, participants were assigned the correct tasks ('tackle economic difficulties', 'keep things going as they are' or 'tackle social problems') that task force representatives would address in each city (randomized presentation).

Perceived candidate's traits. Four items each measured to what extent the presented candidate was perceived to be agentic (*dominant*, *bold*, *assertive*, *competitive*; $\alpha = .85$), communal (*understanding*, *kind*, *compassionate*, *sympathetic*; $\alpha = .92$) and moral (*sincere*, *honest*, *trustworthy*, *ethical*; $\alpha = .93$) using a 7-point Likert scale labelled 1 *strongly disagree* to 7 *strongly agree*. The agency and communion items were derived from Hentschel et al. (2019), and the morality items were taken from Bongiorno et al. (2021) and Leach et al. (2007).

Perceived candidate's values. Participants indicated to what extent they believed communal and agentic values were important for the presented candidate on a 7-point Likert scale from 1 *not at all important* to 7 *extremely important*. Agentic values ($\alpha = .90$) comprised *power*, *recognition*, *achievement*, *self-promotion*, *independence*, *status* and *competition*, while communal values ($\alpha = .93$) consisted of the items *helping others*, *serving humanity*, *working with people*, *connection with others*, *attending to others*, *caring for others* and *intimacy* (Block et al., 2018; Diekman et al., 2010).

Perceived candidate's change potential. Several items measured to what extent participants believed that the presented candidate possessed change potential using a 7-point Likert scale ranging from 1 *strongly disagree* to 7 *strongly agree*. Four items addressed the candidate's potential to signal change (e.g., 'the assignment of this candidate symbolizes a visible change for the citizens of the city', $\alpha = .87$) and six items were about the candidate's potential to implement actual changes (e.g., the candidate is capable of adequately tackling the consequences that the city has to address, $\alpha = .89$). All items were taken from prior glass cliff studies with some wording adapted to the political context (Kulich et al., 2015, 2018, 2021).

Comprehension check about the candidate, comprehension check of candidate's gender and participant political leaning ($M = 5.79$, $SD = 3.27$) were measured as in Study 1.

5.4 | Results

5.4.1 | Perceived traits of the presented candidate (H_3)

We conducted a repeated-measures ANOVA on the traits (agentic, communal, moral) with candidate gender (woman vs. man) as between-participant variable. A main effect of traits was found, $F(1,418) = 14.22$, $p < .001$, $\eta_p^2 = .03$, such that the attributions of moral traits, $M = 5.23$, $SD = 1.03$, were generally higher than attributions of agentic, $M = 5.03$, $SD = 0.98$ and communal traits, $M = 5.15$, $SD = 0.97$. Other pairwise comparisons of the trait attributions were not significant, $ps > .060$. No Trait × Candidate gender interaction was found, indicating that the attribution of the traits did not differ between the female and male candidate, $F(1,418) = 0.12$, $p = .725$, $\eta_p^2 < .001$, not supporting H_3.

5.4.2 | Perceived values of the presented candidate

We ran a repeated-measures ANOVA on the value ratings (agentic, communal) with candidate gender (woman, man) as between-participant variable. The main effect of value, $F(1,418) = 27.08$, $p < .001$, $\eta_p^2 = .06$, showed that communal values were generally attributed more than agentic values to the candidates. However, no Values × Candidate gender interaction was found, $F(1,418) = 0.13$, $p = .717$, $\eta_p^2 < .001$.

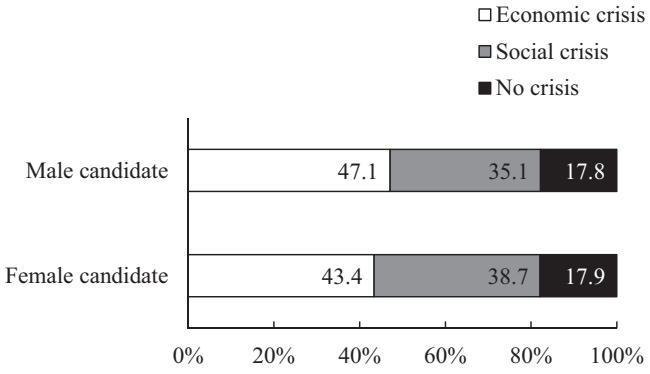

FIGURE 3 The proportion of participants' choices of crisis types depending on candidate gender.

5.4.3 | Perceived change potential of presented candidate (H_4)

We conducted independent samples *t*-tests to compare the evaluations of the female candidate and the male candidate. The female candidate, $M = 4.94$, $SD = 1.12$, was not perceived to possess more potential to signal change to the citizens than the male candidate, $M = 4.94$, $SD = 1.14$, $t(418) = 0.02$, $p = .984$, 95% CI [−0.21, 0.21], not supporting our hypothesis (H_4).

Moreover, no difference in actual change potential was observed (female candidate: $M = 5.24$, $SD = 1.00$; male candidate: $M = 5.31$, $SD = 0.97$), $t(418) = -0.69$, $p = .490$, 95% CI [−0.25, 0.12].

5.4.4 | Crisis choice depending on candidate gender (H_{5a} and H_{5b})

In a preliminary analysis, we checked which city (i.e., crisis type) participants chose for the female candidate and for the male candidate (see Figure 3). According to a chi-square analysis, the general distribution differed significantly from chance, χ^2 (2, $N = 420$) = 49.64, $p < .001$. Overall, the most chosen city was the city with an economic crisis (45.2%), followed by the city with a social crisis (36.9%). The city without any crises was chosen the least (17.9%). A chi-square test of independence revealed that the appointments to different cities did not differ by candidate gender, χ^2 (2, $N = 420$) = 1.7, $p = .709$.

To perform an accurate test for H_{5a} and H_{5b}, we conducted two multinomial logistic regressions on crisis type (1 = city with social crisis, 2 = city with economic crisis, 3 = city without a crisis) with candidate gender (−1 = man, 1 = woman) as predictor and participants' gender (−1 = man, 1 = woman), political leaning (centralized) and all interactions as covariates (see Appendix B for complete results with full statistics).

In the first regression model, the reference category was the city without crisis. This way, we investigated the likelihood of participants choosing the city with a social crisis and the city with an economic crisis compared to the city without a crisis as a function of the predictors. We found that neither the probability of choosing the city with social crisis (vs. no crisis), $B = 0.06$, Wald $X^2(420) = 0.19$, $p = .665$, $e^B = 1.06$, 95% CI [0.80, 1.41], nor the probability of choosing the city with economic crisis (vs. no crisis), $B = -0.03$, Wald $X^2(420) = 0.05$, $p = .828$, $e^B = 0.97$, 95% CI [0.74, 1.27], were higher for the female candidate compared to the male candidate. Therefore, H_{5b} was not supported.

In the second regression model, the reference category was the city with an economic crisis. In addition to the first model, the second model revealed that the female candidate was not more likely than the male candidate to be chosen for the city with social crisis (vs. economic crisis, $B = 0.09$, Wald $X^2(420) = 0.72$, $p = .398$, $e^B = 1.10$, 95% CI [0.89, 1.36]). Thus, H_{5a} was not supported. Overall, the results indicate that the appointment of candidates to different contexts (economic crisis vs. social crisis vs. no crisis) did not differ depending on their gender.

5.5 | Discussion

In Study 2, we incorporated a no-crisis context for our dependent variable and measured how the female candidate was evaluated compared to the male candidate. We tested two sets of hypotheses: One was based on our initial prediction that people's preferences for female and male leaders would be dependent on gendered leadership demands of crisis types (H_3 and H_{5b}). The alternative hypotheses predicted that the type of crisis did not matter in leadership appointments because women would be more likely than men to be chosen to lead a crisis (H_4 and H_{5a}).

Overall, we found no differences in rating between the female and male candidates regarding agentic, communal and moral traits and their potential to signal or lead changes. This finding could explain why our hypotheses on leadership appointment/suitability were not supported in the present and previous studies. If the female candidate is not perceived to be more communal than the male candidate, there is no reason to prefer her over the male candidate to deal with a task that requires more communal skills (i.e., to handle a social crisis). The alternative hypothesis was also not supported because the woman was not appointed to lead a crisis more than a no-crisis situation compared to the man. Hence, we found no glass cliff effect in the context of a COVID-19 crisis task force. Our finding adds to the glass cliff meta-analysis that demonstrated that various factors moderate its emergence (Morgenroth et al., 2020).

6 | STUDY 3

Study 3 was pre-registered and aimed to exclude four potential methodological shortcomings that could have been responsible for failing to observe the expected link between candidate gender and crisis types in the previous studies. Thus, we made four methodological adjustments but retained our initial hypotheses, focusing specifically on H_1 and H_2, which were as follows:

H_1. Social consequences (compared to economic consequences) of the COVID-19 crisis are perceived to require ($H_{1a,b}$) more

stereotypically feminine competencies (i.e., communal and moral leadership traits) and (H_{1c}) fewer stereotypically masculine competencies (agentic leadership traits).

H_2. A female (vs. male) candidate is more likely appointed to lead a unit dealing with social consequences compared to one dealing with economic consequences.

Regarding the methodological alterations, first, the two different crisis types focusing on either social or economic inequalities may have been perceived by participants to be inherently linked. Even though we found that a task force unit dealing with social issues was perceived to require more stereotypically feminine leadership qualities than a unit working on economic issues (Pilot and Study 1), it remains unclear whether participants thought of the social crisis as economic inequalities between social groups, which would also make the economic crisis implicitly present in the context of a social crisis. To address this issue, Study 3 listed concrete tasks in each condition, informing that the task force's mission would be to solve these specific tasks. We also incorporated a manipulation check on top of a comprehension check to ensure that the two types of crisis are distinctly perceived.

Second, the perception of a relatively high level of leader responsibility may have prevented the perception of gender differences in leader suitability. Recent studies suggest that in top management positions, men and women leaders are perceived as equally agentic, indicating that the gender gap in perceived traits may decrease as seniority increases (Gartzia, 2022). We cannot exclude that the level of seniority may also play a role for political leaders, who are generally perceived as more masculine than feminine (Gaffney & Blaylock, 2010). Unlike Ryan et al. (2010), who asked to select candidates for a lower level election, we presented candidates as already active representatives in state government in our studies. This may have triggered stereotypes about politicians overriding gender stereotypes (Eagly & Steffen, 1984). To address this issue, Study 3 utilized an entry-level political role with a lower level of responsibility in the task force.

Third, it is unclear whether the consequences of the COVID-19 crisis were attributed to a controllable or uncontrollable cause. Kulich et al. (2015, Study 1) demonstrated that the perception that management could control the company's performance was a prerequisite for finding a preference for a female manager over a man in a corporate crisis. Indeed, such a tendency was not found in an uncontrollable crisis (i.e., global financial crisis). The causes of the crises might have been ambiguous in our scenario: The pandemic and the health crisis were global and may have been perceived as external. While our vignettes pointed to internal management problems as causes of the current problem in the task force, the real-life presence of the COVID-19 crisis might still have led to interpretations of the situation as uncontrollable. Citizens have experienced much uncertainty during the pandemic (e.g., Maragakis, 2021). Thus, we measured perceptions of controllability to control for it in the analyses in Study 3.

Finally, because we had not pretested our candidate profile, we could not rule out the possibility that the lack of effects was due to their particularities. Therefore, we ran a pretest of new candidate profiles (SM 3.0) that contained more generic information than the previously used profiles and did not differ in perceived agentic, communal and moral leadership traits. In addition, we presented participants with multiple candidates to choose from instead of just one profile per gender to make gender less salient and avoid demand characteristics (Gloor et al., 2020).

6.1 | Method

6.1.1 | Participants

An a priori power analysis for the logistic regression used for H_2 with G*Power (Faul et al., 2009), 80% power to capture a minimum effect size of an odds ratio = 1.80 ($\alpha = .05$) recommended a sample size of 375 participants. From the 448 US participants on Prolific who completed the questionnaire in February 2023, we excluded those who asked us to delete their data after debriefing ($n = 4$), failed the attention check ($n = 4$), the comprehension check ($n = 4$) or those who did not specify their gender or identified as non-binary ($n = 4$). Our final sample consisted of 432 participants (49.8% women, 49.3% men, 18–80 years old, $M = 44.16$, $SD = 14.37$). A sensitivity analysis for logistic regression revealed that, with our sample, $\alpha = .05$ and 80% power, the minimum effect size we could detect was an odds ratio = 1.73.

Participants reported their employment status (65.5% employed) and ethnic identity (87.0% Caucasian or White, 4.4% Asian or Asian-American, 4.2% Hispanic or Latino, 6.0% African-American or Black, 0.7% Native American, Alaska Native, Hawaiian or Pacific Islander).

6.1.2 | Study design and procedure

We manipulated the type of crisis (economic, social) between participants and measured the perceived importance of gendered leadership skills for the crisis (agentic, communal and moral leadership qualities). Participants chose a candidate for a local task force leadership role.

The questionnaire began with a fictitious vignette describing a US city grappling with the long-term consequences of the COVID-19 crisis. These consequences were manipulated: half of the participants read about increased economic instability ('the COVID-19 crisis has accelerated the city's disastrous economic conditions: Since 2020, the city has dealt with increased economic instability'), while the other half read about social inequalities ('the COVID-19 crisis has accelerated the city's disastrous conditions concerning social inequalities: Since 2020, the city has dealt with increased social inequalities'). The text explained that the City Council had established a local task force to tackle those 'local economic difficulties' or 'local social inequalities' in the districts. The tasks of these local task force leaders were specified as working with the City Council to 'oversee financial resources, and implement plans to support at-risk businesses' (economic) or to 'oversee community efforts, and support vulnerable and disadvantaged people to become part of the community' (social).

The text further explained that the council was looking for a new local task force leader for a district where attempts to address the

TABLE 1 Rating of imminence (means and standard deviations) of different types of crises (economic, social and climate) in the experimental conditions (economic vs. social).

	Economic crisis condition		Social crisis condition	
Crisis types	***M***	***SD***	***M***	***SD***
Economic difficulties	6.57	0.94	4.44	2.11
Social inequalities	3.63	2.10	6.41	1.27
Climate crisis	2.60	1.77	2.34	1.74

Abbreviations: M, mean; SD, standard deviation.

economic difficulties/social problems have been met 'with dissatisfaction from the population'. The text said, 'anyone engaged in the community and motivated to help the city is encouraged to apply. This is a good opportunity for anyone interested in a political career'.

Participants answered some comprehension questions about the text and rated the importance of agentic, communal and moral leadership traits for the local task force leader in the scenario. Finally, they saw the pretested descriptions of four pre-selected candidates (two men and two women) in a randomized order. The combination of candidate gender and description was also randomized. These biographies included candidates' names (Robert Miller, Lisa Clark, William Baker, Patricia Smith), birthdates (all candidates born between 1967 and 1971), alma maters (all U.S.-based), relevant experiences (e.g., serving in the Office of Neighborhood Services), three domains of expertise (e.g., financial law, social reform) and hobbies (e.g., swimming; see SM 3.2 for complete profiles).

Participants were asked to select their preferred candidate for the local task force leader role in the presented district, with the question, 'In your opinion, which candidate should the City Council choose as the local task force leader tackling economic difficulties/social inequalities?'.

The study ended with demographic questions.

6.1.3 | Measures

Manipulation check of the crises. After reading the vignette, participants indicated to which degree (from 1 *Not at all* to 7 *A lot*) *economic difficulties*, *social inequalities* and the *climate crisis* (presented in randomized order) needed to be tackled by the new local task force leader (see Table 1 for means).

Comprehension check. They selected the topic of the vignette between finding a new 'local task force leader to tackle the long-term consequences of the COVID-19 crisis' (correct response), 'chef for a local restaurant […]' and 'head of a local hospital […]'.

Perceived controllability. Participants indicated the degree to which they thought *social inequalities* or *economic difficulties* in the described district could be controlled by a new local task force leader, $M = 4.22$, $SD = 1.34$ (from 1 *Not at all* to 7 *Completely*).

Agentic, communal and moral leadership traits. Participants rated how important agentic, $\alpha = .80$, $M = 4.95$, $SD = 1.01$, and communal, $\alpha = .85$, $M = 5.73$, $SD = 0.83$, leadership traits were for the new local task force leader in the shown situation from 1 *Not at all important* to 7 *Very important*. The seven communal (friendly, caring, considerate, understanding, supportive, helpful, emotional) and seven agentic (assertive, competitive, dominant, self-confident, independent, ambitious, aggressive) traits were the same as those used in the pilot study (Abele, 2003; Eagly et al., 2003; Hentschel et al., 2019; Keck, 2019; Scott & Brown, 2006). Moreover, like in Study 1, seven items assessed the perceived importance of moral leadership traits, $\alpha = .90$, $M = 6.41$, $SD = 0.73$. All traits were presented in a randomized order.

Participant political leaning. Political leaning was assessed as in prior studies, $M = 5.95$, $SD = 3.58$.

6.2 | Results

6.2.1 | Preliminary analyses

Manipulation check of the crises. We tested whether participants perceived the type of crisis that they had read about as more imminent than other types of crisis (Table 1). A repeated-measures ANOVA on the perceived crisis type (economic difficulties, social inequalities, climate crisis) with the experimental condition (economic, social) as between-participant variable found a main effect of crisis, $F(2,860) = 461.71$, $p < .001$, $\eta_p^2 = .52$, and a Crisis × Condition interaction, $F(2,860) = 266.47$, $p < .001$, $\eta_p^2 = .38$. As expected, participants in the economic condition rated economic difficulties to be more important than social inequalities, 95% CI [2.60, 3.28], $p < .001$, and the climate crisis, 95% CI [3.69, 4.26], $p < .001$. Participants in the social condition rated social inequalities as more important to tackle than economic difficulties, 95% CI [1.63, 2.30], $p < .001$, and the climate crisis, 95% CI [3.79, 4.33], $p < .001$. Finally, economic difficulties in the economic condition ($M = 6.57$, $SD = 0.94$) and social inequalities in the social condition ($M = 6.41$, $SD = 1.27$) were not perceived as significantly different in imminence according to an independent samples *t*-test, $t(430) = -1.54$, $p = .123$, 95% CI [−0.38, 0.04].

Perceived controllability. We tested whether participants in the economic and social crisis conditions differed in their perception of the controllability of the presented crisis. An independent samples *t*-test, $t(430) = 3.83$, $p < .001$, 95% CI [0.24, 0.74], indicated that the economic crisis, $M = 4.47$, $SD = 1.20$, was perceived to be more controllable by the new local task force leader than the social crisis, $M = 3.98$, $SD = 1.43$.

6.2.2 | Perceptions of economic and social crises (H_1)

To test whether the social crisis was perceived to require more communal and moral leadership traits ($H_{1a,b}$) and less agentic leadership traits (H_{1c}) than the economic crisis, we conducted a repeated-measures ANOVA to compare the importance ratings (agentic, communal, moral) for the crisis (economic, social; see Figure 4). We found a main effect of traits, $F(1,430) = 429.94$, $p < .001$, $\eta_p^2 = .54$, and a Trait × Crisis

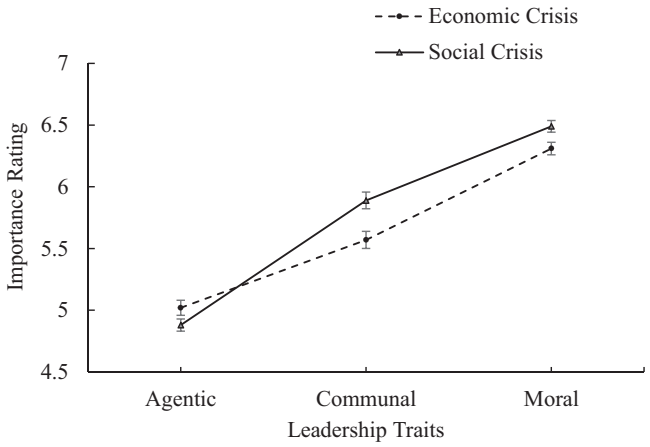

FIGURE 4 Importance ratings of communal, agentic and moral leadership traits for the economic and social units. Error bars indicate standard errors.

interaction, $F(1,430) = 14.59$, $p < .001$, $\eta_p^2 = .03$, indicating that the importance of the traits was dependent on the crisis. As hypothesized ($H_{1a,b}$), communal, $p < .001$, 95% CI [0.17, 0.47], and moral, $p = .010$, 95% CI [0.04, 0.31], leadership traits were rated as more important in the social than economic crisis. However, agentic leadership traits were not perceived to be more important in the economic crisis than the social crisis, $p = .167$, 95% CI [−0.33, 0.06], thus, not supporting H_{1c}.

6.2.3 | Candidate choice depending on crisis (H_2)

To examine whether a female candidate (vs. man) was more likely appointed to lead a local task force dealing with a social crisis (vs. economic; H_2), we conducted a logistic regression on candidate gender (0 = man, 1 = woman) with crisis (−1 = economic, 1 = social) as the predictor and participant gender (−1 = man, 1 = woman) and political leaning (centred continuous variable) as covariates, including all interactions (see Appendix C for complete results). H_2 was not supported because the type of crisis did not predict candidate choice, $B = -0.01$, Wald $X^2(431) = 0.01$, $p = .906$, $e^B = 0.99$, 95% CI [0.81, 1.20]. Politically left-leaning participants were more likely to appoint a woman, $B = -0.10$, Wald $X^2(431) = 0.01$, $p = .001$, $e^B = 0.91$, 95% CI [0.86, 0.96], but participant gender was not predictive of choice, $B = -0.05$, Wald $X^2(431) = 0.21$, $p = .649$, $e^B = 0.96$, 95% CI [0.79, 1.16]. None of the interactions were significant, $ps > .327$. Participants were generally more likely to select the female candidate (56.7%), $B = 0.27$, Wald $X^2(431) = 7.58$, $p = .006$, $e^B = 1.31$.

We repeated the analysis described above while controlling for perceived controllability of the presented crisis (centred, continuous) instead of political leaning. Again, we found no effect of crisis type, $B = -0.02$, Wald $X^2(431) = 0.03$, $p = .861$, $e^B = 0.98$, 95% CI [0.81, 1.20], nor an interaction between crisis type and perceived controllability, $B = 0.12$, Wald $X^2(431) = 2.38$, $p = .123$, $e^B = 1.13$, 95% CI [0.97, 1.31]. No main effect or interactions with participant gender was significant, $ps > .307$. Thus, H_2 was also not supported when controlling for perceived controllability of the presented crisis.

6.3 | Discussion

Study 3 aimed to test our two main hypotheses again by ruling out the methodological shortcomings of our prior studies. We refined the descriptions of the economic and social crises with distinct tasks. The manipulation check confirmed that participants perceived the priorities of each crisis condition to be different. They perceived the crises in the conditions to be imminent to a similar degree. Like in Study 1, our findings partially supported our first hypothesis as communal and moral leadership traits were perceived to be more important in the social than economic crisis ($H_{1a,b}$). However, participants in the economic condition did not rate agentic leadership traits as more important than participants rating the social crisis (H_{1c}). Despite describing the leadership position as an entry-level role, taking into account the perceived controllability of the presented crisis by the task force leader, and presenting multiple pretested candidate profiles to choose from, the candidate choice was not influenced by the candidate's gender and the type of crisis. Like in our prior studies, H_2 was not supported as a woman candidate (vs. a man) was not more likely to be appointed to lead a task force unit dealing with a social crisis than an economic crisis.

7 | GENERAL DISCUSSION

A pilot study and three online experiments were conducted to test whether people in the United States preferred a woman over a man as a political leader during a pandemic, where management required skills that are stereotypically linked to women. Our first hypothesis was partially supported: While the COVID-19 crisis task force dealing with social issues was perceived to require more stereotypically feminine (communal and moral) leadership qualities than the task force working on economic issues, it was not perceived to require less stereotypically masculine (agentic) leadership (Studies 1 and 3). Our second hypothesis was not supported: A female candidate was not more likely to be chosen or rated as more suitable than a male candidate to lead a COVID-19 task force unit dealing with social consequences (Studies 1–3). Study 2 provided a potential explanation for these results: Female and male political candidates were not perceived to differ in their possession of agentic, communal and moral traits and were equally preferred in a crisis compared to a no-crisis context. Overall, our findings suggest that in the context of the COVID-19 crisis, participants did not solely rely on gender and gender stereotypes for rating and appointing political leaders. While we cannot conclude that gender stereotypes did not influence a preference for female leaders in crisis because we did not find a higher preference for female than male leaders, we found that participants adapted preferred gendered leadership traits depending on crisis type. In the following sections, we will discuss the results.

First, we should consider the particularity of the COVID-19 crisis. The multifaceted consequences of the pandemic impacted many people's lives. Thus, participants likely cared more about this crisis and its related political decisions than when experiments addressed fictitious companies or hypothetical elections on a regional level. This difference matters as people are more likely to rely on heuristics and stereotypes when not motivated to think deeply (Bodenhausen, 1993). The context of our study made clear that actual improvements in the situations were expected from the task force leaders in the crisis conditions. Studies from the corporate context demonstrated that agentic leaders were generally preferred in struggling companies regardless of their gender (Kulich et al., 2018). Overall, it is possible that participants in our studies have actively dismissed heuristics and gender stereotypes when given a chance to decide on leadership that is supposed to improve a precarious situation that impacts people's lives.

Second, Studies 1 and 2 left the question open of whether we did not find the expected gender and crisis link because of the candidate's seniority. Successful politicians might be primarily perceived as prototypical politicians rather than stereotypical exemplars of their gender category (Eagly & Steffen, 1984). However, in Study 3, we still found no female preference for the social crisis despite lowering the candidates' seniority level and describing the tasks as entry level. We can deduct several possibilities from this finding. One is that female politicians constitute a subtype with their own qualities within the group of women. Some evidence suggests that female politicians do not share the qualities ascribed to women in general while also not being ascribed to male stereotypes (Schneider & Bos, 2014). Thus, traditional gender stereotypes might not be activated when rating women in politics (see also Hargrave & Blumenau, 2022).

It is also possible that with more female representation in politics (Blazina & Desilver, 2021), gender stereotypes are generally becoming less salient. Prior research in politics suggested that candidate gender is a low-information heuristic if other information is present (Crowder-Meyer et al., 2020; Dolan & Lynch, 2014; McDermott, 1997). Moreover, it has been suggested that female stereotypes about politicians are only activated if they are communicated and highlighted in political campaigns (Bauer, 2014). Alternatively, gender stereotypes in politics might be changing. According to social role theory (Eagly, 1987), gender stereotypes decrease if social roles are equally distributed between men and women. Thus, our finding that female politicians are perceived to be more stereotypically masculine could support the notion that gender stereotypes are dynamic and malleable (Bosak et al., 2018).

While the exact reasons for the lack of effect are not evident based on our studies, it is likely that other variables than gender are more important for participants when appointing political leaders to deal with different tasks efficiently (see also Gartzia, 2022; Kulich et al., 2021). For instance, while we controlled for participants' political leaning, we did not manipulate the party affiliation of the task force units or the candidate. In the United States, the Democratic and Republican parties demonstrated distinct attitudes and approaches to the COVID-19 crisis (Kerr et al., 2021). In our studies, participants might have made no distinction between male and female candidates because the context of the crisis and the desired outcomes were perceived to be more important (see also Devroe, 2020).

7.1 | Implications

During the pandemic, gendered leadership was prevalent in the news (Chamorro-Premuzic & Gallop, 2020; Henley & Roy, 2020) and research (Chen et al., 2023; Coscieme et al., 2020). Female leaders seemed to display a more communal leadership style than male leaders (Grebelsky-Lichtman & Katz, 2020; McGuire et al., 2020; Sergent & Stajkovic, 2020). This impression led to an increased call for more female leadership during the health crisis (Katz, 2021; PAHO, 2021; UN Women, 2020). Past research illustrated that media reports are often biased when depicting female and male politicians (Larson, 2001), often endorsing traditional gender stereotypes (see also Atkeson & Krebs, 2008; Kittilson & Fridkin, 2008; Hayek & Russmann, 2020).

However, our findings suggest that these COVID-19-related media posts might not have necessarily reflected or changed laypeople's opinions (see also Joseph et al., 2021). Our participants did not seem to perceive a significant difference in leadership style and change potential based on candidate gender nor did they generally prefer female over male leaders for high-status political positions in the COVID-19 crisis. Thus, we suggest that the evaluation or selection of political candidates can be made without gender bias under certain conditions if the goal is to find an effective leader.

Our findings in the COVID-19 crisis show that people can choose political leaders without being influenced by gender stereotypes, and they raise the question of why women leaders are sometimes preferred over men leaders in today's political arena, especially in a crisis (Morgenroth et al., 2020). Even though prior literature helps us understand why our hypothesis regarding gender and gendered tasks was not supported, our findings raise new questions. Because we can only speculate about why the female and male candidates were perceived similarly, more research is needed to investigate the link between gender stereotypes, perceptions of political tasks and voter preferences.

7.2 | Methodological limitations

All studies were conducted online in the United States, which may have influenced the conditions under which participants completed the questionnaires and the particular COVID-19-related contexts they were situated in. Consequently, it is essential to acknowledge that the generalizability of our findings to different contexts might be limited (Henrich et al., 2010).

Furthermore, although our experimental studies were designed to complement a previous correlational study (Takizawa et al., 2022), we recognize that scenario experiments have their own limitations (Aguinis & Bradley, 2014; Gloor et al., 2020). A more comprehensive approach could be adopted in future research, combining various

quantitative methods, including experimental, correlational and archival, to address these research questions.

8 | CONCLUSION

Prior research on the glass cliff phenomenon has shown the complexity underlying perceived female leadership advantage and the appointment of female leaders in crises. Our results provide an insight into citizens' perceptions and attitudes towards women and political leadership tasks in the context of a real-life crisis. Citizens did not choose crisis leaders based on their gender. These results might indicate a step towards less gendered perceptions of political leaders and underline that media might exaggerate gender-based leadership valuations. Future research should address which cues are more salient in choosing crisis leaders to deepen our understanding of citizens' needs during precarious times.

ACKNOWLEDGEMENTS

This research was financed by the Geneva–Exeter Joint Seed Funding. The work was further supported by grant no. 100019_188934 from the Swiss National Science Foundation awarded to the last author and a European Research Council Consolidator grant (no. 725128) awarded to the third author.

Open access publishing facilitated by Australian National University, as part of the Wiley - Australian National University agreement via the Council of Australian University Librarians.

CONFLICT OF INTEREST STATEMENT

We have no known conflict of interest to disclose.

DATA AVAILABILITY STATEMENT

Data are available on the open registries network at https://osf.io/myh9c/?view_only=d0206dc7cfd44895a46e00562b519a71.

ETHICS APPROVAL STATEMENT

This manuscript adheres to the APA ethical guidelines. Studies were approved by the University of Exeter's Psychology Ethics Committee and by the University Commission for Ethical Research at the University of Geneva.

TRANSPARENCY STATEMENT

All results are reported honestly, the studies were conducted ethically and the submitted work is original. All data, variables and supplementary materials are publicly available.

ORCID

Ruri Takizawa https://orcid.org/0000-0003-4764-6274
Vincenzo Iacoviello https://orcid.org/0000-0002-6654-8330
Michelle K. Ryan https://orcid.org/0000-0003-1091-9275
Clara Kulich https://orcid.org/0000-0002-9483-2128

REFERENCES

Abele, A. E. (2003). The dynamics of masculine-agentic and feminine-communal traits: findings from a prospective study. *Journal of Personality and Social Psychology*, *85*(4), 768–776. https://doi.org/10.1037/0022-3514.85.4.768

Aelenei, C., Assilaméhou-Kunz, Y., Iacoviello, V., & Kulich, C. (2020). The political glass cliff: When left-wing orientation leads to minority candidate choices for hard-to-win seats. *European Review of Applied Psychology*, *70*(3), 100–514.

Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods*, *17*(4), 351–371. https://doi.org/10.1177/1094428114547952

Alexander, D., & Andersen, K. (1993). Gender as a factor in the attribution of leadership traits. *Political Research Quarterly*, *46*(3), 527–545. https://doi.org/10.1177/106591299304600305

Asch, S. E. (1946). Forming impressions of personality. *The Journal of Abnormal and Social Psychology*, *41*(3), 258–290. https://doi.org/10.1037/h0055756

Atkeson, L. R., & Krebs, T. B. (2008). Press coverage of mayoral candidates: The role of gender in news reporting and campaign issue speech. *Political Research Quarterly*, *61*(2), 239–252. https://doi.org/10.1177/1065912907308098

Barnes, T. D., & Beaulieu, E. (2014). Gender stereotypes and corruption: How candidates affect perceptions of election fraud. *Politics & Gender*, *10*(3), 365–391. https://doi.org/10.1017/S1743923X14000221

Bauer, N. M. (2014). Emotional, sensitive, and unfit for office? Gender stereotype activation and support female candidates. *Political Psychology*, *36*(6), 691–708. https://doi.org/10.1111/pops.12186

Bavik, Y. L., Shao, B., Newman, A., & Schwarz, G. (2021). Crisis leadership: A review and future research agenda. *The Leadership Quarterly*, *32*, 101518. https://doi.org/10.1016/j.leaqua.2021.101518

Bell, D. A. (2020, August 20). "Why female leaders are faring better than "wartime presidents" against COVID-19." Fortune, https://fortune.com/2020/08/20/women-female-leaders-vs-wartime-president-trump-jacinda-ardern-angela-merkel-covid-19-coronavirus/

Béland, L. P., Brodeur, A., & Wright, T. (2020). *The short-term economic consequences of Covid-19: Exposure to disease, remote work and government response* (IZA Discussion Paper No. 13159). https://ssrn.com/abstract=3584922

Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, *66*(1), 5–20. https://doi.org/10.1037/0022-3514.66.1.5

Blazina, C., & Desilver, D. (2021, January 15). A record number of women are serving in the 117th Congress. Pew Research Center, https://www.pewresearch.org/fact-tank/2021/01/15/a-record-number-of-women-are-serving-in-the-117th-congress/

Block, K., Croft, A., & Schmader, T. (2018). Worth less?: Why men (and women) devalue care-oriented careers. *Frontiers in Psychology*, *9*, 1353. https://doi.org/10.3389/fpsyg.2018.01353

Bodenhausen, G. V. (1993). Emotions, arousal, and stereotypic judgments: A heuristic model of affect and stereotyping. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, cognition, and stereotyping* (pp. 13–37). Academic Press. https://doi.org/10.1016/B978-0-08-088579-7.50006-5

Bollyky, T. J., Kickbusch, I., & Petersen, M. B. (2023, January 30). The trust gap. How to fight pandemics in a divided country. Foreign Affairs. https://www.foreignaffairs.com/united-states/trust-gap-fight-pandemic-divided-country

Bongiorno, R., Bain, P. G., Ryan, M., Kroonenberg, P. M., & Leach, C. W. (2021). *Think leader-think (immoral, power-hungry) man: An expanded framework for understanding stereotype content and leader gender bias*. PsyArXiv. https://doi.org/10.31234/osf.io/p5uya

Bosak, J., Eagly, A. H., Diekman, A., & Sczesny, S. (2018). Women and men of the past, present, and future: Evidence of dynamic gender stereotypes in

Ghana. *Journal of Cross-Cultural Psychology*, *49*(1), 115–129. https://doi.org/10.1177/0022022117738750

Boukes, M., & Boomgaarden, H. G. (2016). Politician seeking voter: How interviews on entertainment talk shows affect trust in politicians. *International Journal of Communication*, *10*, 1145–1166. https://ijoc.org/index.php/ijoc/article/view/2849/1584

Brown, E. R., Diekman, A. B., & Schneider, M. C. (2011). A change will do us good: Threats diminish typical preferences for male leaders. *Personality and Social Psychology Bulletin*, *37*(7), 930–941. https://doi.org/10.1177/0146167211403322

Bruckmüller, S., & Branscombe, N. R. (2010). The glass cliff: When and why women are selected as leaders in crisis contexts. *British Journal of Social Psychology*, *49*(3), 433–451. https://doi.org/10.1348/014466609X466594

Chamorro-Premuzic, T., & Gallop, C. (2020, April 7). 7 Leadership lessons men can learn from women. Harvard Business Review. https://hbr.org/2020/04/7-leadership-lessons-men-can-learn-from-women

Chen, C. R., Chen-Edinboro, L., & Xu, C. (2023). Does economic power or leadership matter? The difference between male and female world leaders in managing pandemics. *Applied Economics Letters*, *30*, 27–32. https://doi.org/10.1080/13504851.2021.1971610

Clarke, N., Jennings, W., Moss, J., & Stoker, G. (2018). *The good politician: Folk theories, political interaction, and the rise of anti-politics*. Cambridge University Press. https://doi.org/10.1017/S1537592720001516

Coscieme, L., Fioramonti, L., Mortensen, L. F., Pickett, K. E., Kubiszewski, I., Lovins, H., McGlade, J., Ragnarsdóttir, K. V., Roberts, D., Costanza, R., De Vogli, R., & Wilkinson, R. (2020). Women in power: female leadership and public health outcomes during the COVID-19 pandemic. *MedRxiv*. https://doi.org/10.1101/2020.07.13.20152397 medRxiv

Crowder-Meyer, M., Gadarian, S. K., Trounstine, J., & Vue, K. (2020). A different kind of disadvantage: Candidate race, cognitive complexity, and voter choice. *Political Behavior*, *42*(2), 509–530. https://doi.org/10.1007/s11109-018-9505-1

Cwalina, W., Falkowski, A., & Kaid, L. L. (2005). Advertising and the image of politicians in evolving and established democracies: Comparative study of the Polish and the US presidential elections in 2000. *Journal of Political Marketing*, *4*(2-3), 19–44. https://doi.org/10.1300/J199v04n02_02

Devroe, R. (2020). Voters' evaluation of (contra-) prototypical political candidates. An experimental test of the interaction of candidate gender and policy position cues in Flanders (Belgium). *Electoral Studies*, *68*, 102240.https://doi.org/10.1016/j.electstud.2020.102240

Diekman, A. B., Brown, E. R., Johnston, A. M., & Clark, E. K. (2010). Seeking congruity between goals and roles: A new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychological Science*, *21*(8), 1051–1057. https://doi.org/10.1177/0956797610377342

Dolan, K., & Lynch, T. (2014). It takes a survey: Understanding gender stereotypes, abstract attitudes, and voting for women candidates. *American Politics Research*, *42*(4), 656–676. https://doi.org/10.1177/1532673X13503034

Donthu, N., & Gustafsson, A. (2020). Effects of COVID-19 on business and research. *Journal of Business Research*, *117*, 284–289. https://doi.org/10.1016/j.jbusres.2020.06.008

Eagly, A. H. (1987). *Sex differences in social behavior: A Social-role interpretation* (1st ed.). Psychology Press. https://doi.org/10.4324/9780203781906

Eagly, A. H., Johannesen-Schmidt, M. C., & van Engen, M. L. (2003). Transformational, transactional, and laissez-faire leadership styles: A meta-analysis comparing women and men. *Psychological Bulletin*, *129*(4), 569–591. https://doi.org/10.1037/0033-2909.129.4.569

Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, *111*(1), 3–22. https://doi.org/10.1037/0033-2909.111.1.3

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

Everett, J. A., Colombatto, C., Awad, E., Boggio, P., Bos, B., Brady, W. J., Chawla, M., Chituc, V., Chung, D., Drupp, M. A., Goel, S., Grosskopf, B., Hjorth, F., Ji, A., Kealoha, C., Kim, J. S., Lin, Y., Ma, Y., Maréchal, M. A., … Crockett, M. J. (2021). Moral dilemmas and trust in leaders during a global health crisis. *Nature Human Behaviour*, *5*(8), 1074–1088. https://doi.org/10.1038/s41562-021-01156-y

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fisher, A. N., & Ryan, M. K. (2021). Gender inequalities during COVID-19. *Group Processes & Intergroup Relations*, *24*(2), 237–245. https://doi.org/10.1177/1368430220984248

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Funk, K. D. (2020). Local responses to a global pandemic: Women mayors lead the way. *Politics & Gender*, *16*(4), 968–974. https://doi.org/10.1017/S1743923X20000410

Funk, K. D., & Philips, A. Q. (2019). Representative budgeting: Women mayors and the composition of spending in local governments. *Political Research Quarterly*, *72*(1), 19–33. https://doi.org/10.1177/1065912918775237

Gaffney, A. M., & Blaylock, D. L. (2010). Hillary Clinton's race: Did she match the presidential prototype? *Advancing Women in Leadership Journal*, *30*, 1–15. https://doi.org/10.21423/awlj-v30.a294

Gartzia, L. (2022). Self and other reported workplace traits: A communal gap of men across occupations. *Journal of Applied Social Psychology*, *52*, 568–587. https://doi.org/10.1111/jasp.12848

Gloor, J. L., Gazdag, B., & Reinwald, M. (2020). Overlooked or undercooked? Critical review & recommendations for experimental methods in diversity research. In A. Risberg, S. Just, & F. Villeseche (Eds.), *Routledge companion to organizational diversity research methods* (pp. 91–106). Routledge.

Grebelsky-Lichtman, T., & Katz, R. (2020). Gender effect on political leaders' nonverbal communicative structure during the COVID-19 crisis. *International Journal of Environmental Research and Public Health*, *17*(21), 7789. https://doi.org/10.3390/ijerph17217789

Hannah, S. T., Uhl-Bien, M., Avolio, B. J., & Cavarretta, F. L. (2009). A framework for examining leadership in extreme contexts. *The Leadership Quarterly*, *20*(6), 897–919. https://doi.org/10.1016/j.leaqua.2009.09.006

Hargrave, L., & Blumenau, J. (2022). No longer conforming to stereotypes? Gender, political style and parliamentary debate in the UK. *British Journal of Political Science*, *52*(4), 1584–1601. https://doi.org/10.1017/S0007123421000648

Haslam, S. A., & Ryan, M. K. (2008). The road to the glass cliff: Differences in the perceived suitability of men and women for leadership positions in succeeding and failing organizations. *The Leadership Quarterly*, *19*(5), 530–546. https://doi.org/10.1016/j.leaqua.2008.07.011

Hayek, L., & Russmann, U. (2020). Those who have the power get the coverage—Female politicians in campaign coverage in Austria over time. *Journalism*, *23*(1), 224–242. https://doi.org/10.1177/1464884920916359

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not weird. *Nature*, *466*, 29. https://doi.org/10.1038/466029a

Helland, E., & Sykuta, M. (2004). Regulation and the evolution of corporate boards: Monitoring, advising, or window dressing? *The Journal of Law and Economics*, *47*(1), 167–193. https://doi.org/10.1086/380473

Henley, J., & Roy, E. A. (2020, April 25), Are female leaders more successful at managing the coronavirus crisis? The Guardian. https://www.theguardian.com/world/2020/apr/25/why-do-female-

leaders-seem-to-be-more-successful-at-managing-the-coronavirus-crisis

Hentschel, T., Heilman, M. E., & Peus, C. V. (2019). The multiple dimensions of gender stereotypes: A current look at men's and women's characterizations of others and themselves. *Frontiers in Psychology*, *10*, 11. https://doi.org/10.3389/fpsyg.2019.00011

Herrnson, P. S., Lay, J. C., & Stokes, A. K. (2003). Women running "as women": Candidate gender, campaign issues, and voter-targeting strategies. *The Journal of Politics*, *65*(1), 244–255. https://doi.org/10.1111/1468-2508.t01-1-00013

Holman, M. R. (2023). Gender stereotyping questions accurately measure beliefs about the traits and issue strengths of women and men in politics. *Journal of Women, Politics & Policy*, *44*(1), 90–104. https://doi.org/10.1080/1554477X.2023.2162285

Huddy, L., & Terkildsen, N. (1993a). Gender stereotypes and the perception of male and female candidates. *American Journal of Political Science*, *37*, 119–147. https://doi.org/10.2307/2111526

Huddy, L., & Terkildsen, N. (1993b). The consequences of gender stereotypes for women candidates at different levels and types of office. *Political Research Quarterly*, *46*(3), 503–525. https://doi.org/10.1177/2F106591299304600304

Joseph, K., Shugars, S., Gallagher, R., Green, J., Mathé, A. Q., An, Z., & Lazer, D. (2021). *(Mis) alignment between stance expressed in social media data and public opinion surveys*. arXiv. https://doi.org/10.48550/arXiv.2109.01762

Kaplan, S. (2020, October 9). *White House blocked C.D.C. from requiring masks on public transportation*. The New York Times. https://www.nytimes.com/2020/10/09/health/coronavirus-covid-masks-cdc.html

Katz, M. (2021, March 17). The COVID crisis shows why we need more female leadership. Fortune. https://fortune.com/2021/03/17/covid-female-women-leadership-jacinda-ardern/

Keck, S. (2019). Gender, leadership, and the display of empathic anger. *Journal of Occupational and Organizational Psychology*, *92*(4), 953–977. https://doi.org/10.1111/joop.12264

Kerr, J., Panagopoulos, C., & van der Linden, S. (2021). Political polarization on COVID-19 pandemic response in the United States. *Personality and Individual Differences*, *179*, 110892. https://doi.org/10.1016/j.paid.2021.110892

Kittilson, M. C., & Fridkin, K. (2008). Gender, candidate portrayals and election campaigns: A comparative perspective. *Politics & Gender*, *4*(3), 371–392. https://doi.org/10.1017/S1743923X08000330

Koch, J. W. (1999). Candidate gender and assessments of Senate candidates. *Social Science Quarterly*, *80*, 84–96. http://www.jstor.org/stable/42863875

Koch, J. W. (2000). Do citizens apply gender stereotypes to infer candidates' ideological orientations? *Journal of Politics*, *62*, 414–429. https://doi.org/10.1111/0022-3816.00019

Koenig, A. M., Eagly, A. H., Mitchell, A. A., & Ristikari, T. (2011). Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin*, *137*(4), 616–642. https://doi.org/10.1037/a0023557

Konrad, A. M., Ritchie, J. E., Jr., Lieb, P., & Corrigall, E. (2000). Sex differences and similarities in job attribute preferences: A meta-analysis. *Psychological Bulletin*, *126*, 593–641. https://doi.org/10.1037//0033-2909.126.4.593

Kulich, C., Gartzia, L., Komarraju, M., & Aelenei, C. (2021). Contextualizing the think crisis–think female stereotype in explaining the glass cliff: Gendered traits, gender, and type of crisis. *PLoS One*, *16*(3), e0246576. https://doi.org/10.1371/journal.pone.0246576

Kulich, C., Iacoviello, V., & Lorenzi-Cioldi, F. (2018). Solving the crisis: When agency is the preferred leadership for implementing change. *The Leadership Quarterly*, *29*(2), 295–308. https://doi.org/10.1016/j.leaqua.2017.05.003

Kulich, C., Lorenzi-Cioldi, F., Iacoviello, V., Faniko, K., & Ryan, M. K. (2015). Signaling change during a crisis: Refining conditions for the glass cliff. *Journal of Experimental Social Psychology*, *61*, 96–103. https://doi.org/10.1016/j.leaqua.2017.05.003

Larson, S. G. (2001). American women and politics in the media: A review essay. *PS: Political Science and Politics*, *34*(2), 227–230. https://www.jstor.org/stable/1350209

Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, *93*(2), 234–249. https://doi.org/10.1037/0022-3514.93.2.234

Liberska, H., & Jankowiak, J. (2016). Psychological gender, the perfect politician stereotype and assessment of a female presidential candidate. In H. Liberska (Ed.), *Current psychosocial problems in traditional and novel approaches: Three worlds of human life* (pp. 141–156). Wydawnictwo Uniwersytetu Kazimierza Wielkiego.

Mano-Negrin, R., & Sheaffer, Z. (2004). Are women "cooler" than men during crises? Exploring gender differences in perceiving organisational crisis preparedness proneness. *Women in Management Review*, *19*(2), 109–122. https://doi.org/10.1108/09649420410525315

Maragakis, L. (2021, October 21). *Coronavirus second wave, third wave and beyond: what causes a COVID surge*. John Hopkins Medicine. https://www.hopkinsmedicine.org/health/conditions-and-diseases/coronavirus/first-and-second-waves-of-coronavirus

McDermott, M. L. (1997). Voting cues in low-information elections: Candidate gender as a social information variable in contemporary United States elections. *American Journal of Political Science*, *41*, 270–283. https://doi.org/10.2307/2111716

McGuire, D., Cunningham, J. E., Reynolds, K., & Matthews-Smith, G. (2020). Beating the virus: An examination of the crisis communication approach taken by New Zealand Prime Minister Jacinda Ardern during the Covid-19 pandemic. *Human Resource Development International*, *23*(4), 361–379. https://doi.org/10.1080/13678868.2020.1779543

Morgenroth, T., Kirby, T. A., Ryan, M. K., & Sudkämper, A. (2020). The who, when, and why of the glass cliff phenomenon: A meta-analysis of appointments to precarious leadership positions. *Psychological Bulletin*, *146*, 797–829. https://doi.org/10.1037/bul0000234

Oliver, A. G., Pfarrer, M. D., & Neville, F. (2022). Grand Challenges and female leaders: An exploration of relational leadership during the COVID-19 pandemic. Business & Society. Advance online publication. https://doi.org/10.1177/00076503221141880

Pan American Health Organization (PAHO). (2021, March 8). *PAHO Director calls for more women in leadership of the fight against COVID-19*. https://www.paho.org/en/news/8-3-2021-paho-director-calls-more-women-leadership-fight-against-covid-19

Pak, A., Adegboye, O. A., Adekunle, A. I., Rahman, K. M., McBryde, E. S., & Eisen, D. P. (2020). Economic consequences of the COVID-19 outbreak: the need for epidemic preparedness. *Frontiers in Public Health*, *8*, 241. https://doi.org/10.3389/fpubh.2020.00241

Park, S. (2022). Gendered leadership during the COVID-19 pandemic: how democracy and representation moderate leadership effectiveness. *Public Management Review*, *24*(11), 1802–1823. https://doi.org/10.1080/14719037.2021.1937294

Pearson, C. M., & Clair, J. A. (1998). Reframing crisis management. *Academy of Management Review*, *23*(1), 59–76. https://doi.org/10.5465/amr.1998.192960

Pew Research Center. (2015). *Women and leadership: Public says women are equally qualified, but barriers persist*. Pew Research Center. https://www.pewresearch.org/social-trends/2015/01/14/women-and-leadership/

Profeta, P. (2020). Gender equality and public policy during COVID-19. *CESifo Economic Studies*, *66*(4), 365–375. https://doi.org/10.1093/cesifo/ifaa018

Reinwald, M., Zaia, J., & Kunze, F. (2023). Shine bright like a diamond: When signaling creates glass cliffs for female executives. *Journal of Management*, *49*(3), 1005–1036. https://doi.org/10.1177/01492063211067518

Rimmer, A. (2020). Covid-19: Disproportionate impact on ethnic minority healthcare workers will be explored by government. *BMJ*, *369*, m1562. https://doi.org/10.1136/bmj.m1562

Rosenwasser, S. M., & Dean, N. G. (1989). Gender role and political office: Effects of perceived masculinity/femininity of candidate and political office. *Psychology of Women Quarterly*, *13*(1), 77–85. https://doi.org/10.1111/j.1471-6402.1989.tb00986.x

Ryan, M. K., Haslam, S. A., Hersby, M. D., & Bongiorno, R. (2011). Think crisis–think female: The glass cliff and contextual variation in the think manager–think male stereotype. *Journal of Applied Psychology*, *96*(3), 470–484. https://doi.org/10.1037/a0022133

Ryan, M. K., Haslam, S. A., & Kulich, C. (2010). Politics and the glass cliff: Evidence that women are preferentially selected to contest hard-to-win seats. *Psychology of Women Quarterly*, *34*(1), 56–64. https://doi.org/10.1111/j.1471-6402.2009.01541.x

Ryan, M. K., Haslam, S. A., Morgenroth, T., Rink, F., Stoker, J., & Peters, K. (2016). Getting on top of the glass cliff: Reviewing a decade of evidence, explanations, and impact. *The Leadership Quarterly*, *27*(3), 446–455. https://doi.org/10.1016/j.leaqua.2015.10.008

Santucci, J. (2020, February 27). *What we know about the White House coronavirus task force now that Mike Pence is in charge*. USA Today. https://www.usatoday.com/story/news/politics/2020/02/27/coronavirus-what-we-know-mike-pence-and-task-force/4891905002/

Schneider, M. C., & Bos, A. L. (2014). Measuring stereotypes of female politicians. *Political Psychology*, *35*(2), 245–266. https://doi.org/10.1111/pops.12040

Sczesny, S., Bosak, J., Neff, D., & Schyns, B. (2004). Gender stereotypes and the attribution of leadership traits: A cross-cultural comparison. *Sex Roles*, *51*(11), 631–645. https://doi.org/10.1007/s11199-004-0715-0

Scott, K. A., & Brown, D. J. (2006). Female first, leader second? Gender bias in the encoding of leadership behavior. *Organizational Behavior and Human Decision Processes*, *101*(2), 230–242. https://doi.org/10.1016/j.obhdp.2006.06.002

Sergent, K., & Stajkovic, A. D. (2020). Women's leadership is associated with fewer deaths during the COVID-19 crisis: Quantitative and qualitative analyses of United States governors. *Journal of Applied Psychology*, *105*(8), 771–783. https://doi.org/10.1037/apl0000577

Takizawa, R., Robinson, S., Aelenei, C., Iacoviello, V., & Kulich, C. (2022). A five-nation study of the impact of political leaning and perception of crisis severity on the preference for female and minority leaders during the COVID-19 pandemic. *Current Research in Ecological and Social Psychology*, *3*, 100055. https://doi.org/10.1016/j.cresp.2022.100055

Taub, A. (2020, May 15). *Why are women-led nations doing better with Covid-19?* The New York Times. https://www.nytimes.com/2020/05/15/world/coronavirus-women-leaders.html

United Nations Entity for Gender Equality and the Empowerment of Women (UN Women). (2020). *COVID-19 and women's leadership: From an effective response to building back better*. https://www.unwomen.org/en/digital-library/publications/2020/06/policy-brief-covid-19-and-womens-leadership#view

Waylen, G. (2021). Gendering political leadership: hypermasculine leadership and Covid-19. *Journal of European Public Policy*, *28*(8), 1153–1173. https://doi.org/10.1080/13501763.2021.1942160

WhiteHouse.gov. (2020, January 29). *Statement from the Press Secretary Regarding the President's Coronavirus Task Force*. https://trumpwhitehouse.archives.gov/briefings-statements/statement-press-secretary-regarding-presidents-coronavirus-task-force/

Yzerbyt, V., & Klein, O. (2019). *Psychologie sociale*. De Boeck Supérieur.

Zingher, J. N., & Farrer, B. (2016). The electoral effects of the descriptive representation of ethnic minority groups in Australia and the UK. *Party Politics*, *22*(6), 691–704. https://doi.org/10.1177/1354068814556895

SUPPORTING INFORMATION

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

How to cite this article: Takizawa, R., Iacoviello, V., Ryan, M. K., & Kulich, C. (2024). Gender Stereotypes may not Influence the Choice of Female Leaders: Experimental Evidence from a Crisis Framed as Social or Economic During the COVID-19 Pandemic. *European Journal of Social Psychology*, *54*, 558–576. https://doi.org/10.1002/ejsp.3023

APPENDIX A: STUDY 1

Complete regression model testing H_2

Suitability Ratings of the Candidate as a Function of Candidate Gender (−1 = man, 1 = woman), Crisis Unit (−1 = economic, 1 = social), Participant Gender (−1 = man, 1 = woman), Participant Political Leaning (from left to right) and all Interactions

	B	*SE*	95% CI	*p*	η_p^2
(Intercept)	**5.47**	**.05**	**[5.37, 5.56]**	**.000**	**.970**
Candidate Gender	0.05	.05	[−0.05, 0.14]	.336	.002
Crisis Unit	0.06	.05	[−0.03, 0.16]	.186	.004
Participant Political Leaning	0.01	.01	[−0.01, 0.04	.363	.002
Participant Gender	−0.02	.05	[−0.11, 0.07]	.685	.000
Candidate Gender × Crisis Unit	−0.08	.05	[−0.17, 0.01]	.084	.007
Crisis Unit × Political Leaning	0.01	.01	[−0.02, 0.03]	.590	.001
Candidate Gender × Political Leaning	−0.03	.01	[−0.05, 0.00]	.051	.009
Political Leaning × Participant Gender	−0.02	.01	[−0.05, 0.01]	.165	.005
Crisis Unit × Participant Gender	−0.04	.05	[−0.14, 0.05]	.369	.002
Candidate Gender × Participant Gender	−0.02	.05	[−0.11, 0.08]	.741	.000
Candidate Gender × Crisis Unit × Political Leaning	0.01	.01	[−0.02, 0.03]	.716	.000
Candidate Gender × Crisis Unit × Participant Gender	0.01	.05	[−0.09, 0.10]	.881	.000
Crisis Unit × Political Leaning × Participant Gender	0.02	.01	[−0.01, 0.05]	.157	.005
Candidate Gender × Political Leaning × Participant Gender	0.01	.01	[−0.02, 0.04]	.472	.001
Candidate Gender × Crisis Unit × Political Leaning × Participant Gender	−0.00	.01	[−0.03, 0.02]	.761	.000

Note: Values in boldface are statistically significant.

APPENDIX B: STUDY 2

Complete regression models testing H_{5a} *and* H_{5b}

Multinomial Logistic Regression on Choice (1 = city with social crisis, 2 = city with economic crisis, 3 = city without a crisis) as a Function of Candidate Gender (−1 = man, 1 = woman), Participant Gender (−1 = man, 1 = woman), Political Leaning (centred) and Their Interactions

Crisis	Predictors	*B*	*SE*	Wald X^2	*p*	e^B	95% CI
1 = Social	**(Intercept)**	**0.74**	**.14**	**26.44**	**<.000**		
	Candidate Gender	0.06	.14	0.19	.665	1.10	[0.80, 1.41]
	Participant Gender	−0.10	.14	0.48	.488	0.91	[0.68, 1.20]
	Political Leaning	0.05	.04	1.36	.244	1.05	[0.97, 1.15]
	Candidate Gender × Participant Gender	−0.07	.14	0.24	.625	0.93	[0.70, 1.24]
	Candidate Gender × Political Leaning	−0.02	.04	0.31	.579	0.98	[0.90, 1.06]
	Participant Gender × Political Leaning	−0.05	.04	1.08	.298	0.96	[0.88, 1.04]
	Candidate Gender × Participant Gender × Political Leaning	−0.05	.04	1.34	.248	0.95	[0.87, 1.04]
2 = Economic	**(Intercept)**	**0.95**	**.14**	**46.60**	**<.001**		
	Candidate Gender	−0.03	.14	0.05	.828	0.74	[0.74, 1.27]
	Participant Gender	−0.04	.14	0.07	.792	0.73	[0.73, 1.27]
	Political Leaning	0.02	.04	0.12	.725	0.93	[0.93, 1.10]
	Candidate Gender × Participant Gender	−0.08	.14	0.31	.579	0.71	[0.71, 1.22]
	Candidate Gender × Political Leaning	−0.01	.04	0.10	.758	0.91	[0.91, 1.07]
	Participant Gender × Political Leaning	−0.02	.04	0.15	.704	0.91	[0.91, 1.07]
	Candidate Gender × Participant Gender × Political Leaning	−0.05	.04	1.50	.220	0.87	[0.87, 1.03]

Note: The reference category is 3 = No crisis. Values in boldface are statistically significant.

Crisis	Predictors	*B*	*SE*	Wald X^2	*p*	e^B	95% CI
1 = Social	(Intercept)	−0.21	.11	3.59	.058		
	Candidate Gender	0.09	.11	0.72	.398	1.10	[0.89, 1.36]
	Participant Gender	−0.06	.11	0.33	.565	0.94	[0.76, 1.16]
	Political Leaning	0.04	.03	1.16	.281	1.01	[0.97, 1.11]
	Candidate Gender × Participant Gender	0.01	.11	0.00	.948	1.01	[0.81, 1.25]
	Candidate Gender × Political Leaning	−0.01	.03	0.11	.736	0.99	[0.93, 1.06]
	Participant Gender × Political Leaning	−0.03	.03	0.78	.378	0.97	[0.91, 1.04]
	Candidate Gender × Participant Gender × Political Leaning	−0.00	.03	0.00	.968	1.00	[0.94, 1.07]
3 = No Crisis	**(Intercept)**	**−0.95**	**.14**	**46.60**	**<.001**		
	Candidate Gender	0.03	.14	0.05	.828	1.03	[0.79, 1.35]
	Participant Gender	0.04	.14	0.07	.792	1.04	[0.79, 1.36]
	Political Leaning	−0.02	.04	0.12	.725	0.99	[0.91, 1.07]
	Candidate Gender × Participant Gender	0.08	.14	0.31	.579	1.08	[0.82, 1.42]
	Candidate Gender × Political Leaning	0.01	.04	0.10	.758	1.01	[0.93, 1.10]
	Participant Gender × Political Leaning	0.02	.04	0.15	.704	1.02	[0.94, 1.11]
	Candidate Gender × Participant Gender × Political Leaning	0.05	.04	1.50	.220	1.05	[0.97, 1.15]

Note: The reference category is 2 = Economic crisis. Values in boldface are statistically significant.

APPENDIX C: STUDY 3

Complete regression models testing H_2

Logistic Regression on Candidate gender (0 = man, 1 = woman) with crisis (−1 = economic, 1 = social) as the predictor and participant gender (−1 = man, 1 = woman) and political leaning (centred continuous variable) as covariates, including all interactions.

Predictors	*B*	*SE*	Wald X^2	*p*	e^B	95% CI
(Intercept)	**0.27**	**.10**	**7.58**	**.006**	**1.31**	
Crisis	−0.01	.10	0.01	.906	0.99	[0.81, 1.20]
Participant Gender	−0.05	.10	0.21	.649	0.96	[0.79, 1.16]
Political Leaning	**−0.10**	**.03**	**11.51**	**.001**	**0.91**	**[0.86, 0.96]**
Crisis × Participant Gender	−0.10	.10	0.96	.327	0.91	[0.75, 1.10]
Crisis × Political Leaning	−0.01	.03	0.10	.758	0.99	[0.94, 1.05]
Participant Gender × Political Leaning	0.02	.03	0.36	.546	1.02	[0.96, 1.07]
Crisis × Participant Gender × Political Leaning	−0.00	.03	0.00	.985	1.00	[0.95, 1.06]