The Leadership Quarterly 36 (2025) 101910

Contents lists available at ScienceDirect

The Leadership Quarterly

journal homepage: www.elsevier.com/locate/leaqua

Full Length Article

Gender differences in dishonesty when leaders make decisions on behalf of their team☆

Kerstin Grosch [a,1], Stephan Müller [b,2], Holger A. Rau [c,*], Lilia Wasserka-Zhurakhovska [d,e,3]

[a] WU Vienna University of Economics and Business, Austria
[b] Bard College Berlin, Germany
[c] University of Duisburg–Essen, University of Göttingen, Germany
[d] University of Galway; J.E. Cairnes School of Business and Economics, Ireland
[e] Portland State University, USA

ARTICLE INFO

Dataset link: https://osf.io/my7ek/?view_only=cfd97dd2d321432c913f94f2c2758e58

JEL classification:
C91
D91
J16
M14
H26

Keywords:
Leadership
Decisions for others
Experiments
Gender differences
Dishonesty

ABSTRACT

This study examines the ethical dilemma faced by leaders, balancing financial gains and ethical considerations, with a focus on gender differences. We experimentally study such a dilemma in which leaders can benefit their teams at the expense of moral costs from dishonest reporting. We measure, first, individual dishonesty preferences and, second, reporting decisions for teams in a leadership role using outcome-reporting games in a laboratory setting. Individual dishonesty preferences predict men's propensity to apply for leadership. We further find that women have lower initial dishonesty preferences compared to men but increase dishonesty when assuming leadership roles. A follow-up study indicates that women leaders act dishonestly when they expect that most team members also report dishonestly. When leadership roles are randomly assigned rather than self-selected, we find no statistically significant difference in how women and men respond to them.

Introduction

Leaders face pressing global challenges, such as climate change, health crises, and inclusive growth, all of which demand ethical considerations. Higher ethical conduct of leaders can foster a culture of honesty, reciprocity, trust, and less toxic competition among coworkers, leading to positive social and economic outcomes (Alan et al., 2022; d'Adda et al., 2017; Liu et al., 2022). Honesty is, therefore, a crucial element for effective leadership and long-term group success (Ashton & Lee, 2008; Johnson et al., 2011; Knack & Keefer, 1997; Somanathan & Rubin, 2004). However, in the short term, leaders may exploit dishonesty to gain advantages for their group members. This behavior can lead to increased support for dishonest leaders, perpetuating a cycle of deceit and undermining trust in leadership, commonly observed in the public and the private sector (Huppert & Levine, 2023).

However, research has yet to establish how the assumption of leadership roles itself affects individuals' ethical behavior, particularly whether potential gender differences in honesty persist when making decisions that affect others. To address this gap, our study contributes to ethical leadership theory (Brown et al., 2005), builds on moral leadership theory (Treviño et al., 2003), and the concept of honesty humility (Ashton & Lee, 2008; Johnson et al., 2011), by examining how assuming leadership roles impacts individuals' ethical behavior. We also contribute to research on the "dark side" of leadership (Hogan & Hogan, 2001; Schyns & Schilling, 2013), which has implications for

☆ We are grateful to Martin G. Kocher and Marie Claire Villeval for their valuable feedback and to the audiences at the University of Duisburg–Essen, the Gender Economics and Workplace conference in Nuernberg, and the VCEE seminar in Vienna for their insightful comments. We acknowledge the financial support provided by the University of Goettingen, and the University of Duisburg-Essen. Kerstin Grosch acknowledges funding by the Austrian Science Fund (FWF) [T 1263-G].

* Correspondence to: Platz der Göttinger Sieben 3, 37073 Göttingen, Germany.
E-mail addresses: kerstin.grosch@wu.ac.at (K. Grosch), s.mueller@berlin.bard.edu (S. Müller), holger.rau@uni-goettingen.de (H.A. Rau), lilia.wasserka-zhurakhovska@universityofgalway.ie (L. Wasserka-Zhurakhovska).
[1] Welthandelsplatz 1, 1020 Vienna (Austria).
[2] Platanenstrasse 24, 13156 Berlin (Germany).
[3] 49 Upper Newcastle, Galway, H91 YK8V (Ireland).

https://doi.org/10.1016/j.leaqua.2025.101910
Received 23 March 2024; Received in revised form 26 July 2025; Accepted 22 August 2025
Available online 13 October 2025

leader–member exchange (LMX) theory (Graen & Uhl-Bien, 1995), by specifically investigating gender differences in leadership aspirations and how men and women's ethical conduct changes when assuming leadership roles.

Increasing the number of women in leadership positions could be a suitable strategy to foster honest decision-making in leadership positions, as various experimental studies demonstrate that women exhibit a higher tendency towards ethical behavior, particularly in the realm of honesty (Conrads et al., 2014; Dreber & Johannesson, 2008; Grosch & Rau, 2017; Houser et al., 2016, 2012), altruism (Brañas-Garza et al., 2018b; Eckel & Grossman, 1998), and prosociality (Grosch & Rau, 2017; Soutschek et al., 2017). Despite the growing demand for ethical leadership and women's higher ethical standards, they remain underrepresented in leadership positions, especially in industries where ethical conduct is deemed important (Bear et al., 2010; European Institute for Gender Equality, 2012; Fernandez-Mateo & Fernandez, 2016; Flabbi et al., 2019; Gobillon et al., 2015; Matsa & Miller, 2013; Zenger & Folkman, 2019). The underrepresentation of women in leadership positions limits the available empirical evidence on their influence on ethical leadership, resulting in relatively few empirical studies on the topic (Bear et al., 2010; Chattopadhyay & Duflo, 2004; Haucap & Heldman, 2023; Matsa & Miller, 2013; Swamy et al., 2001). [4] Several of these papers report correlational evidence and cannot disentangle selection effects on women's ethical behavior in leadership roles. That is, they cannot establish a causal link between women's personal honesty preferences and their decision to assume leadership. Moreover, there is no evidence indicating whether women with personal honesty preferences maintain such ethical behavior in leadership positions. It is crucial to consider these implications, particularly concerning affirmative action policies aimed at increasing the number of women in leadership positions, due to the growing demand for enhanced ethical conduct (United Nations, 2019).

In this paper, we address these research gaps. Our analysis extends role congruity theory (Eagly & Karau, 2002) by investigating how women's ethical behavior changes in leadership positions. It complements research on gender differences in leadership styles (Bass & Avolio, 1994), focusing specifically on ethical decision-making in contexts where leaders make decisions affecting others. To address this scenario, we conducted a controlled experiment to investigate how women and men report payoffs individually and as team leaders reflecting (dis-)honest behavior. We also examine the role of promotion procedures. Thus, our paper pays special attention to the inclusion of economic perspectives and methods in leadership research, as postulated by Garretsen et al. (2020) and Kosfeld (2020). Our experiment studies the crucial characteristic of many leadership decisions of making ethical decisions with monetary consequences for others. We focus on the ethical dimension of honesty as a fundamental value in both business and politics (Caselli & Morelli, 2004; Schwartz, 2005). Specifically, leaders may have two main reasons for behaving dishonestly. First, leaders often have personal incentives to misreport outcomes because their compensation and advancement depend on performance (Burns & Kedia, 2006; Necker & Paetzel, 2022). Second, leaders' decisions affect stakeholder payoffs, such as shareholders for managers and staff members for politicians (Berman et al., 1999). Because leaders are often evaluated based on stakeholder satisfaction, their decisions may be influenced by stakeholder preferences, reflecting the role of social norms and beliefs in decision-making.

We designed an economic experiment on leadership because the field of economics can advance leadership research not only in concepts, but also in causality and context (Garretsen et al., 2020). Thus, our design builds on other economic leadership experiments where subjects first make individual decisions and, thereafter, make decisions that also affect the outcomes of others when acting as leaders (e.g., Alan et al., 2020; Ertac & Gurdal, 2012). [5] We adapt this design to study whether ethical women continue to behave ethically when making decisions as leaders. Another related research question focuses on the role of the promotion procedure in changing ethical behavior, with a particular focus on gender differences in dishonesty under different promotion procedures. Moreover, we investigate whether women in leadership positions behave more ethically than men. We also examine whether dishonesty preferences can predict the decision to apply for a leadership position.

Although our experiment is stylized, it encompasses characteristics that model dishonest behavior in business situations. Specifically, in our setting participants have to report the realization of a private signal, the outcome of a die-roll (Fischbacher & Föllmi-Heusi, 2013). In this setup, reporting higher numbers leads to increased payoffs, thus creating an incentive for misreporting. For instance, this setup resembles situations in which managers know the real outcome and may intentionally inflate company returns (e.g., Bollen & Pool, 2009; Burns & Kedia, 2006), by misreporting sale figures of teams (Church et al., 2012), the quality of products (Belot & Van De Ven, 2017, 2019), or figures to evade taxes (Joulfaian, 2000). The reporting setup of Fischbacher and Föllmi-Heusi (2013) has been demonstrated to predict real-life behavior in the fields of corrupt behavior in India (Hanna & Wang, 2017) and Denmark (Barfort et al., 2019), wearing masks improperly (Tobol et al., 2020), and fare evasion in public transportation (Dai et al., 2018; Potters & Stoop, 2016). The task has also been employed in a military field experiment (Ruffle & Tobol, 2017).

A main feature of our experiment is that subjects participate twice in the reporting setup. First, they make the reporting decisions for their own payoff only, which enables us to measure individual dishonesty preferences. Second, participants make their reporting decision as a team leader, which affects their own payoff as well as the payoffs of two other team members. In our experiment, we also measure subjects' willingness to take up the leadership role by asking them whether they want to become a leader or not. Eliciting subjects' dishonesty preferences and their willingness to become a leader allows us to analyze how an institutional context with self-chosen (*endogenous*) leadership predicts dishonesty as a leader. Our focus is on the analysis of the behavioral change from the first decision (individual dishonest behavior) to the second decision (dishonest behavior in the role of a leader). We compare the behavioral change and analyze if women and men respond differently to assuming the leadership role (Eagly & Karau, 2002). To study the role of the promotion procedure, we ran a control treatment without the possibility to apply for the leadership position, that is, leaders are externally appointed by a random mechanism (*exogenous leadership*). By applying economic methodologies to questions of ethical leadership, our study bridges economic and leadership research. This interdisciplinary approach allows us to isolate the causal effects of leadership roles on ethical behavior, contributing to a nuanced understanding of how organizational contexts influence leadership ethics.

Our results corroborate lab-experimental evidence that women behave less dishonestly than men when deciding on individual outcomes, for instance, in lying situations where lying only benefits the person who lies and hurts somebody else (e.g., Dreber & Johannesson, 2008; Grosch & Rau, 2017; Houser et al., 2016, 2012; Muehlheusser et al., 2015). The gender difference in reporting decisions is no longer statistically significant when subjects act as leaders, because women become more often dishonestly in this role, whereas men's behavior

[4] For instance, the cited studies indicate that women in leadership roles have the potential to mitigate corruption, decrease the likelihood of cartel formation, enhance the delivery of public goods in the political sphere, improve social responsibility ratings, and display increased sensitivity to workers' vulnerability to unemployment risk.

[5] We follow the argument in Alan et al. (2020) that a *"major component of a leader's job is to hold the power and responsibility for making decisions on behalf of others"*.

does not differ significantly between the two domains. Moreover, our results show that it is particularly dishonest men who tend to sort into leadership and misreport in both domains. We find no such relation for women. Our control treatment reveals that women only increase their misreporting from individual to team outcomes when they can apply for leadership, but it is statistically non-significant under an external appointment. The findings highlight that women's increase in dishonesty is not primarily driven by the team context *per se*. It is caused by both having the option to choose to become a leader *and* making decisions for the team. We ran a follow-up study to learn more about potential underlying predictors. The findings suggest that the belief about other team members' dishonesty is the key predictor for women in leadership positions to increase misreporting on behalf of the team. We discuss the implications of the results in the conclusion.

Our study contributes to the scarce experimental evidence on gender differences in leadership behavior and advances various streams of leadership research (Ashton & Lee, 2008; Brown et al., 2005; Graen & Uhl-Bien, 1995; Hogan & Hogan, 2001; Johnson et al., 2011; Schyns & Schilling, 2013) by examining the role of gender and dishonesty preferences on leaders' ethical decision-making and selection into leadership. Additionally, we expand on existing explanations for the lack of women in leadership positions, including discrimination (Kübler et al., 2018), gender-role attitudes (Alesina et al., 2013), and differences in preferences (Azmat & Petrongolo, 2014; Croson & Gneezy, 2009). Our findings reveal that the underrepresentation of women in leadership is not linked to individual dishonesty preferences. Importantly, we demonstrate that gender differences in honesty are context-dependent. These differences vary depending on whether one is an employee or a leader, and they are influenced by the process of attaining leadership positions. This comprehensive approach enhances our understanding of how individual preferences, gender, and institutional contexts shape ethical decision-making in leadership, bridging economic perspectives with leadership theory.

Study 1: Experimental design

In this section, we describe the design of our within-subjects experiment. In the beginning, we elicit data on economic preferences in several consecutive parts. We use these preference data as pilot data for another experiment on unincentivized vs. incentivized elicitation of preferences (Grosch et al., 2023). [6] Afterwards, we collect the main data for this experiment, that is, we apply modifications of the die-rolling game introduced by Fischbacher and Föllmi-Heusi (2013) to measure dishonest behavior when misreporting individual outcomes ("individual preferences for dishonesty") and team outcomes ("preferences for dishonesty as a team leader"). We inform subjects step by step about each part by giving them new instructions for each part. Subjects are told that, at the end of the session, the computer will randomly select one of the parts for the payout, and that the main parts described below will always be paid. Each session ends with a questionnaire on socio-demographics.

Individual preferences for dishonesty

To measure subjects' individual dishonesty preferences, we implement a modification of the method by Fischbacher and Föllmi-Heusi (2013). In this part, subjects have to report the payoff associated with the outcome of a die-roll. This initial stage allows us to assess individual dishonesty preferences. To have control over individual misreporting behavior, we apply a computerized version of the die-rolling game that records the real die outcome. This approach is similar to Kocher et al. (2017). Although subjects are anonymous per design, they cannot disguise their lies and, therefore, we expect subjects to be less dishonest than in the original die-rolling game (Kajackaite & Gneezy, 2017). To demonstrate to subjects that the die is fair, they can repeatedly press a button for 20 s that randomly displays one side of a six-sided die whenever they press the button on the computer screen. At the end of the 20 s, subjects are asked to press the button one more time and to report the payoff associated with the outcome of the actual die-roll. They know that the report determines their payment in this task. In the task, each associated payoff corresponds to the die number times three. For instance, a one is associated with €3, a two is associated with €6, .., a five is associated with €15. The only exception is the number six, which is associated with no payment to mitigate the risk of introducing a focal point (Fischbacher & Föllmi-Heusi, 2013). This first part of our within-subjects experiment allows us to compare individual dishonesty preferences to the situation, where subjects can misreport team outcomes, explained in the following.

Preferences for dishonesty as a team leader

In this part, we measure dishonest behavior when subjects decide as team leaders. The leadership stage aligns with ethical leadership theory (Brown et al., 2005) and provides insights into how individuals navigate moral dilemmas when responsible for others. To investigate this, we play a die-rolling game similar to the previous one. That is, subjects again roll a six-sided die and report the outcome. We apply the same payoff structure as in the previous part (e.g., a die-roll of a four is associated with a payoff of €12). The crucial difference to the previous part is that subjects learn that they have been randomly matched in teams of three and that each team member's payoff is determined by the team leader's report. The experimental instructions point out that each of the other two team members receives the same payoff as the one reported by the leader. We do not use the word "leader" in the instructions, and call the person who determines the team outcome "person A". The following paragraph describes the two promotion procedures that we employ, which are randomized across sessions.

Main treatment "endogenous leadership": Before subjects roll the die, they can choose whether they want to be in the role of "person A" (leader) or not. This choice element connects to the concept of moral leadership (Treviño et al., 2003) by examining how ethical preferences might influence the pursuit of leadership positions. The selection process operates as follows: when only one team member expresses willingness to assume the leadership role, they are automatically appointed as the team leader. When more than one person says "yes", a random draw selects one of the applicants for leadership. When no one applies, the random draw selects one person among the three team members. [7] To ensure a team decision could be reached even if no one applied for leadership, all participants were required to state a team decision, that is, we applied the strategy method (Selten, 1967). Because of the option to assume leadership, we call this treatment *"endogenous leadership"*. The choice mechanism enables us to relate the subjects' individual dishonesty preferences to their willingness to act as a leader. Moreover, we can analyze whether a subject's choice to assume responsibility predicts dishonesty in the leadership position. In our study, the costs of applying for and acting in a leadership role, as well as the potential disutility from misreporting or the burden of making decisions on behalf of others, are negligible. This may result in a relatively high number of applicants. We are, however, not interested in the exact extent to which subjects are willing to act as leaders but in the behavioral changes across contexts. In this regard, the application of

[6] Note that we do not provide any feedback before the end of the experiment. Furthermore, these parts are identical across treatments and can, therefore, not induce any treatment differences. We provide the entire instructions of the experiment, including these parts in Supplementary Material.

[7] This case has not occurred in our data.

the strategy method is important, as it allows us to isolate the effect of changing decision environments on behavior while avoiding selection-driven changes. After subjects decide whether they want to act as a leader or not, we elicit their beliefs about how many of the other team members wanted to become leaders to learn about their perceived chance to become a leader. Subjects receive €1 for a correct guess.

Control treatment "exogeneous leadership": To analyze the role of the endogenous selection promotion procedure for a potential shift in dishonesty, we conducted a control treatment called *exogenous leadership*. In contrast to the *endogenous* treatment, where subjects can apply for the leadership position, employees cannot apply for a leadership position in the *exogenous leadership* treatment. Instead, they are automatically included in the applicant pool and appointed to a leadership position exogenously (e.g., by a third party, see also Erkal et al. 2022). Whereas we keep the sequence of actions similar to the main treatment, we disable the leadership choice and a random draw determines leadership in the control treatment. [8] To account for the possibility that subjects in our main treatment may hold different beliefs on the likelihood of ending up as a leader, we apply different probabilities of becoming a leader in the *exogenous* treatment. The probabilities vary between one-third, that is, we tell all three team members that their probability of becoming a leader is one-third, and one-half, that is, we tell one team member that this person cannot become a leader for sure, whereas the other two team members are told that the probability of becoming a leader is one half. [9]

Next, all subjects are told to roll the die once and to simultaneously enter the outcome they want to report should they become person A (leader). [10] They know that this decision only becomes relevant if they are selected as the leader. Subjects are given the anonymous ID (Subject 1, 2 or 3) of the selected leader, as well as the leader's report at the end of the experiment. However, they are not informed of the leader's actual die roll.

Our experimental design models the selection of leaders and their hierarchical decision-making. In other words, subjects become leaders, and as such, they are responsible for the team's payoffs, which they report in a non-strategic situation. In contrast, the decisions in Kocher et al. (2017) and Lohse and Simon (2021) are not hierarchical and focus on strategic contexts to analyze dishonest decisions in groups. In Kocher et al. (2017) group members have to achieve payoff commonality, that is, subjects only earn a positive amount if they report the same die outcome as the other group members. [11] Whereas, in Lohse and Simon (2021) coordinating on the same reported own outcomes guarantees that a dishonest report of subjects' own outcomes is feasible. Another difference compared to these studies is that we analyze subjects' decisions and the consequences of selection into leadership. In this respect, our leaders decide independently on behalf of others, whereas subjects in Kocher et al. (2017) and Lohse and Simon (2021) meet in a chat before they make their reporting decisions. We deliberately refrain from a strategic team context to avoid confounds in answering our research question on leaders' decision to misreport outcomes on behalf of their team members. The treatments offer valuable insights from a managerial perspective, as they allow us to shed light on the implications of different appointment procedures (see e.g., Bohnet et al., 2016; Murciano-Goroff, 2022).

Procedure

The experiment was conducted at a German university, and it was programmed with the software z-Tree (Fischbacher, 2007). All subjects were students and recruited with the subject-pool software ORSEE (Greiner, 2015). A total of 282 subjects participated, with 144 in the main treatment group (75 women and 69 men) and 118 in the control group (57 women and 61 men). The sample size was determined by budgetary constraints and orientated on the study by Kocher et al. (2017), which included 117 participants per main treatment condition. After subjects made their decisions, we ran a questionnaire to verbally elicit their preferences. This is part of a pilot study for a project on the relationship between non-incentivized and incentivized elicitation of preferences (Grosch et al., 2023). To control for order effects, we conducted some sessions with the questionnaire at the beginning. At the very end of the experiment, we asked for the subjects' socio-demographics. Participants were from various disciplines with a mean age of 24 years. In our sample, 50.4% of the subjects are women. Sessions lasted approximately 70 min. We paid subjects in cash at the end of the experiment, and earnings were on average €10.81, including a show-up fee of €5. In the following, we present the hypotheses [12] and results of our main study (Study 1).

Study 1: Hypotheses

In this study, we focus on misreporting behavior where no other party can be betrayed or deceived. Therefore, we refrain from the term "lying" and use the term "dishonest behavior".

The experimental literature on gender differences in individual dishonest behavior finds predominantly that men behave more dishonestly than women for selfish black lies, that is, when being dishonest benefits oneself and harms another person/a third party in the lab (Conrads et al., 2014; Dreber & Johannesson, 2008; Grosch & Rau, 2017; Houser et al., 2016, 2012), in face-to-face interactions (Lohse & Qari, 2021), and in the field (Azar et al., 2013; Bucciol et al., 2013). These findings inform our first hypothesis on individual reporting behavior.

Hypothesis 1. Men are more likely than women to be dishonest when reporting outcomes that affect only their own payoffs

Misreporting as a team leader generates a benefit for the team members and can therefore be seen as a Pareto improvement over telling the truth for the participants. Thus, they may receive an extra utility from being dishonest as a leader than when reporting individual outcomes. In line with that, Gino et al. (2013) demonstrate that the more other people benefit from misreporting, the more people are willing to be dishonest. This behavior could be interpreted through the lens of the "dark side" of leadership (Hogan & Hogan, 2001), where leaders may justify unethical actions for perceived team benefits. Hence, we expect that subjects are more likely to misreport team outcomes compared to individual outcomes.

[8] Just as in the main treatment, we collect some additional variables at the end of the study, such as risk preferences. To make sure that our subject pools do not differ between the two treatments, we test whether there are differences between the samples on these variables. This is not the case ($p > 0.340$ for all comparisons).

[9] We do not have teams in which only one team member becomes the leader for sure as only in one of 144 cases a subject applied for leadership and at the same time believed that no other team member would in the *endogenous* treatment. Additionally, in 4 out of 144 cases, subjects abstained from applying for leadership while simultaneously believing their team members would also refrain from applying.

[10] This approach is different from experiments with sequential designs where leaders move first, and other subjects may imitate their behavior to analyze the role of "leading-by-example" (e.g., Amore et al., 2022; Gächter et al., 2012; Güth et al., 2007).

[11] In a control treatment, Kocher et al. (2017) also analyze simultaneous team decisions of reporting individual outcomes when subjects do not receive an incentive for reporting similar outcomes.

[12] We refrain from setting up a hypothesis about the main treatment versus the control treatment condition because the main purpose was to learn more about the role of the endogenous promotional procedure for the shift in dishonesty from the individual to the leader decision, rather than formulating a clear theory with a directed hypothesis for *comparing* dishonesty under the endogenous versus the *exogenous* treatment condition.

Moreover, we expect gender differences in the dishonesty of leaders, based on the following reasoning. Compared to men, women are expected to demonstrate higher prosociality at work (Brañas-Garza et al., 2018a) and have been found, on average, to be more prosocial (e.g., Andreoni & Vesterlund, 2001; Croson & Gneezy, 2009; Eckel & Grossman, 1998; Rand et al., 2016). The leader's decision to be dishonest for the team could be perceived as a prosocial act. Because women's level of prosociality is more pronounced than men's, the increase in misreporting between the individual and the team context may be stronger for women than for men. This expectation aligns with role congruity theory (Eagly & Karau, 2002), which suggests that women may adapt their behavior to meet perceived leadership expectations.

Hypothesis 2. (a) Subjects are more likely to be dishonest when reporting outcomes that affect their team's payoff rather than just their own (b) Compared to men, women are more likely to engage in dishonest behavior when misreporting outcomes that affect team payoffs than when misreporting outcomes that affect their own payoff.

People's attitudes towards dishonesty may be vital for leadership applications when leadership is understood to implicitly demand unethical behavior. Other studies have shown that people choose leadership positions based on individual characteristics that resonate with the characteristics of the decision environment, for instance, risk preferences, overconfidence, competitive preferences, preferences for free-riding behavior (e.g., Alan et al., 2020; Barber & Odean, 2001; Cappelen et al., 2016; Eckel & Grossman, 2002; Ertac & Gurdal, 2012; Niederle et al., 2013; Niederle & Vesterlund, 2007; Reuben et al., 2012). Moreover, a lab experiment by Brassiolo et al. (2021) shows that institutions with a cheating opportunity attract (repel) dishonest (honest) subjects. Similar evidence is found by Konrad et al. (2021). This self-selection into leadership roles based on dishonesty preferences could be seen as a manifestation of the "dark side" of leadership (Hogan & Hogan, 2001).

In our setting, we expect that individual dishonesty preferences predict subjects' decisions to become leaders because dishonesty pays off in our decision context. Therefore, subjects with an individual dishonesty preference may apply for leadership to ensure that they maximize their individual profit when deciding to misreport team outcomes. Honest subjects, on the contrary, may not apply for leadership to enforce honest behavior in the team domain because they do not have monetary incentives to do so and they can shift responsibility to team members who applied. In contrast, dishonest subjects have strong monetary incentives for becoming a leader and misreporting team outcomes. This hypothesis relates to the concept of moral leadership (Treviño et al., 2003) and how individual ethical preferences may influence the pursuit of leadership positions.

Hypothesis 3. Subjects with an individual dishonesty preference are more likely to apply for leadership

Results: Study 1

In this section, we present the main results of Study 1 that compare women's and men's misreporting behavior of individual and team outcomes. We then present Study 2, an online experiment examining the underlying predictors of behavioral changes among individuals who endogenously selected into leadership roles. [13]

[13] All regressions presented in these sections were estimated using standard maximum likelihood estimation without robust standard errors, as our experimental design with random assignment typically satisfies the assumptions for this approach. We verified that using robust standard errors yields similar point estimates and only slightly different standard errors, confirming the robustness of our results.

Main results

First, we focus on subjects' misreporting behavior in our main treatment, where subjects can apply for leadership (*endogenous leadership*). We categorize cases as "profitable dishonest reports" when subjects increased their earnings by misreporting the associated payoffs of the real outcome of the die roll. In this case, the dummy variable "dishonest behavior" is one, otherwise, the variable is zero. [14] This measure captures potential dishonesty for die-roll outcomes between one and four, where subjects inflated their reports by claiming outcomes associated with higher payoffs (i.e., reporting a result between two and five instead of the actual lower outcome).

We show the share of misreporting individual outcomes using white bars and the share of misreporting team outcomes with black bars in Fig. 1. The figure conditions on men (left panel) and women (right panel). [15] When reporting individual outcomes, our data confirm commonly found gender differences in dishonesty (e.g., Conrads et al., 2014; Grosch & Rau, 2017; Jacobsen et al., 2018; Kocher et al., 2017). That is, men (26%) are five times more frequently dishonest than women (5%) (Fisher's exact test, $p = 0.001$), supporting Hypothesis 1.

We turn to our first research question and analyze misreporting behavior when deciding as team leaders. We do not find a statistically significant gender difference when subjects report team outcomes (Fisher's exact test, $p = 0.353$). Due to limited statistical power to detect small effects, we must treat non-significant results with caution. [16] In the team domain, women significantly increase dishonest behavior by more than four times from 5% to 24% (Wilcoxon matched-pairs test, $p < 0.001$). By contrast, there is no statistically significant difference in the dishonest behavior of men in the two contexts (individual outcomes: 26%; team outcomes: 32%) (Wilcoxon matched-pairs test, $p = 0.346$). Although the difference for men is not statistically significant, our study lacks the power to detect small differences. [17] We find support for Hypothesis 2b since women are more likely than men to increase their dishonest behavior in the team domain. In general, we find that subjects behave more dishonestly when reporting team outcomes (28%) than individual outcomes (15%) (Wilcoxon matched-pairs test, $p = 0.002$). This finding is in line with the results from Kocher et al. (2017) and Lohse and Simon (2021) who analyze group decisions in a strategical setting with pre-play communication and without leaders. [18] This result supports Hypothesis 2a.

[14] The dummy variable is set to "0" when subjects reported the real outcome, or when they reported an outcome that was to their disadvantage. However, we did not observe the latter case in Study 1.

[15] We also find that all results in the *endogenous leadership* and *exogenous leadership* treatments are robust to excluding the data where subjects were shown a five and therefore could not behave dishonestly. All descriptives are very much in line with the results and the results of the statistical tests do not change significantly.

[16] To assess whether the non-significant findings may be due to insufficient statistical power, we conducted a post hoc power analysis using G*Power (two-tailed Fisher's exact test, $\alpha = 0.05$, n = 69 for group 1 and n = 75 for group 2, p1 = 0.24, p2 = 0.32). The achieved power was only 0.1601, indicating a high risk of Type II error. This suggests that our study was underpowered to detect effects of this size, and the non-significant results should be interpreted with caution.

[17] To examine whether the non-significant findings might be attributable to insufficient statistical power, we conducted a post hoc power analysis using G*Power (two-tailed Fisher's exact test, $\alpha = 0.05$, n = 69 per group, p1 = 0.26, p2 = 0.32). The resulting power was only 0.0896, indicating a substantial risk of Type II error. This suggests that the study was underpowered to detect effects of this magnitude, and the non-significant results should therefore be interpreted with caution.

[18] Castillo et al. (2022) replicate (Kocher et al., 2017) and show that groups are not more dishonest than individuals when a local charity is hurt by subjects' dishonesty.

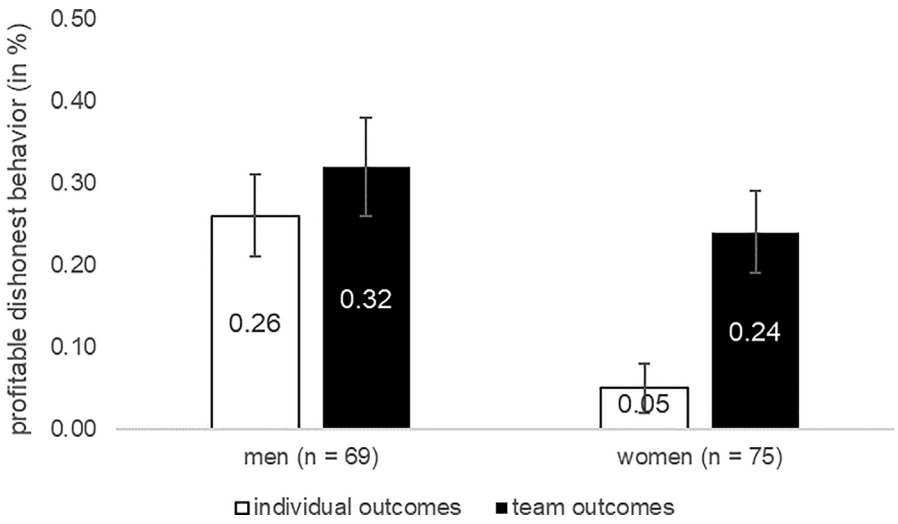

Fig. 1. Percentage of profitable dishonest reports in *endogenous leadership*. White (black) bars present misreporting for individual outcomes (team outcomes). Standard error bars included.

Result 1. Dishonest Behavior for Individual and Team Outcomes
(a) Women behave less dishonestly than men when reporting individual outcomes
(b) Women are more likely to increase their dishonest behavior when making decisions about team outcomes compared to men.

Potential predictors and correlates of misreporting team outcomes

To better understand leaders' motivation to behave dishonestly when deciding about team outcomes, we focus on potential behavioral factors that affect dishonest behavior in *endogenous leadership*. We start with regression analyses to investigate the association of individual characteristics and individual motives with misreporting team outcomes.

The association with individual determinants

Table 1 presents probit regressions on subjects' likelihood to misreport team outcomes. Precisely, the regressions focus on the association with subjects' gender (*woman*), and their individual preference for dishonesty, captured by a dummy (*misreported ind. outcome)* which is one (zero) when subjects misreported (truthfully reported) individual outcomes. We include the variable *(perceived) chance of becoming a leader* in models (2), (4), and (6). This variable is derived from two factors: the subject's own willingness to become a leader (yes or no), and their guess about how many other team members (0, 1, or 2) want to become a leader. The variable ranges from 0 to 1, with possible values of 0, 1/3, 1/2, and 1. For example, if someone wants to become a leader and estimates that no other team member wants to lead, then their perceived chance of becoming a leader is 1. If this person wants to lead but thinks that 2 others also want to become leaders, then their perceived chance drops to 1/3. If a person does not want to become a leader and thinks that at least another person wants to become a leader, their perceived chance to become a leader is 0. This approach captures both the subject's own leadership aspirations and their perception of competition for the leadership role. Models (1)–(2) present the results from all subjects, whereas models (3)–(4) focus on women only, and the last two models analyze the data from men only. This approach allows us to determine whether the findings in our data are gender-specific and gain insights into the factors that contribute to increased dishonest behavior among women in the team domain. [19]

Finally, we include subjects' age, measured by the year of their birth (*age*), a dummy whether they study economics (*econ*), and an order dummy for the timing of the verbal elicitation of preferences (beginning vs. end of the sessions) in models (2), (4), and (6). Table 7 in the Supplementary Material gives an overview of the descriptives and our control variables in all of our treatments and both studies. Accordingly, we present correlation matrices of these variables in Tables 8, 9, and 10. All regressions report marginal effects with standard errors in parentheses. We report regressions with standard coefficients in Table 4 of the Supplementary Material. [20]

All models highlight that subjects' individual dishonesty preferences positively correlate with their dishonest behavior as leaders. Precisely, the highly significant positive coefficient of *misreported ind. outcome* shows that subjects who misreported individual outcomes are also more likely to misreport team outcomes. Moreover, the models do not indicate that this result is gender-specific. Results are robust to the inclusion of controls (models (2), (4), and (6)). This robustness holds, if we additionally include subjects' risk aversion, their ambiguity aversion, and their (Fehr & Schmidt, 1999) parameters as controls. Finally, models (1) and (2) confirm that there is no statistically significant difference between men and women in terms of how they misreport team outcomes.

Note that, as described in detail in Section "Preferences for dishonesty as a team leader", subjects had to make a reporting decision

[19] An obvious solution to test for a potential gender-specific effect of individual dishonesty preferences or their perceived chance of becoming a leader on subjects' propensity to behave dishonestly as a leader would be to include the interaction of *woman* and *misreported ind. outcome* or of *woman* and *(perceived) chance of becoming a leader*. However, including interaction terms in probit models is problematic (see Ai & Norton, 2003). We test the robustness of all our probit results in OLS regressions, and we also test for the respective interaction terms in OLS regressions. The results from the OLS regressions do not deviate from the results reported in the paper.

[20] As a robustness check, we estimate OLS regressions of the profit from dishonest behavior in the group context on the profit from dishonest behavior in the individual context, with and without the full set of control variables used in the main specification. In this analysis, we additionally control for the potential gain from misreporting in both the individual and group contexts—measured as the difference between the reported payoff and the actual payoff from the die roll. This accounts for the fact that the incentive to report dishonestly depends on the value of the die roll (e.g., a roll of 5 offers little room for profitable misreporting, while a roll of 2 offers substantial room). The results remain qualitatively similar: for both genders, we find a positive and statistically significant correlation between the realized profit from dishonest behavior in the individual context and in the group context.

Table 1
Probit regressions on misreporting team outcomes (Study 1: *endogenous leadership*).

	misreporting team outcomes					
	all		*women*		*men*	
	(1)	(2)	(3)	(4)	(5)	(6)
misreported ind. outcome	0.351***	0.322***	0.429**	0.485**	0.338***	0.235**
	(0.083)	(0.085)	(0.187)	(0.190)	(0.090)	(0.100)
woman	0.009	0.022				
	(0.074)	(0.074)				
(perceived) chance of becoming a leader		0.252		0.238		0.302
		(0.194)		(0.268)		(0.277)
controls[a]	no	yes	no	yes	no	yes
obs.	144	144	75	75	69	69

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1.
Notes: The regressions report average marginal effects. [a] Controls: year of birth, whether subjects study economics, and an order dummy.
The results do not change, if we also include subjects' risk aversion, their ambiguity aversion, and their (Fehr & Schmidt, 1999) parameters as controls. *Misreported ind. outcome* remains significant at the 1%-level in model (2) and at the 5%-level in models (4) and (6). Moreover, *women* remains statistically non-significant.

for the team irrespective of their willingness to assume leadership. Thus, the observed relation between misreporting of individual and team outcomes cannot be caused by selection effects into leadership. As additional robustness checks, we also control for subjects' (perceived) chance of becoming a leader, which includes their stated willingness to become a leader. It turns out to be always statistically non-significant. Finally, when we restrict the analyses to subjects who indicated their willingness to become a leader, the estimation results presented in Table 1 hardly change.

The role of the promotion procedure: the possibility to apply for leadership

Our analysis of participants' reporting decisions as team leaders in our main treatment shows a strong correlation with their individual dishonesty preferences. The decision-maker faces two changes in the team domain compared to the individual domain. First, they can apply for leadership, second, their reporting decision affects other persons' payoffs. To isolate the effect of payoff externalities on subjects' misreporting of team outcomes, we conducted the control treatment *exogenous leadership*, in which all subjects in a team are automatically included in the applicant pool and one subject is randomly selected as leader. To account for differing beliefs of becoming the leader in the main treatment, we varied the probabilities in this control treatment.

The data show that the probability (1/3 vs. 1/2) does not have a statistically significant influence on the fraction of misreported team outcomes (Fisher's exact test, $p = 0.323$) or increase misreporting from individual to team outcomes (a dummy, which is positive when subjects misreported team but not individual outcomes) (Fisher's exact test, $p = 0.439$). [21] We also run two Probit regressions on misreporting team outcomes and the increase in misreporting from individual to team outcomes, which confirm the non-parametric test results. The regressions show that the probability of becoming a leader in the *exogenous leadership* treatment does not affect the probability for misreporting team outcomes ($p = 0.757$) and the increase in misreporting from individual to team outcomes ($p = 0.339$). [22] Thus, we merge these data.

Fig. 2 presents the share of misreports for individual and team outcomes when leadership is exogenously determined. As expected, we confirm that men behave significantly more often dishonestly (36%) than women (12%) (Fisher's exact test, $p = 0.003$). [23] Focusing on team outcomes, we find a moderate but statistically non-significant increase for men (from 36% to 46%) (Wilcoxon matched-pairs test, $p = 0.180$). [24] In contrast to the *endogenous* treatment, women show a less pronounced increase of dishonest behavior from the individual (12%) to the team domain (18%) which is also statistically non-significant (Wilcoxon matched-pairs test, $p = 0.257$). [25] Hence, the gender difference in leaders' dishonest behavior remains when leaders are exogenously determined (Fisher's exact test, $p = 0.001$).

Our control treatment highlights that the increase in dishonesty among women in leadership positions is no longer statistically significant when they cannot apply for leadership. We do not find such an effect for men. The finding suggests that the driver for Result 1b is the opportunity to apply for leadership positions.

Result 2. Dishonest Behavior as Leaders under Exogenous Leadership

Under exogenous leadership, women's increase in dishonest behavior is not statistically significant. Consequently, the gender difference in individual dishonesty preferences remains when deciding as leaders.

[21] In the *endogenous* treatment, we run a similar analysis and find that subjects' perceived chance of becoming a leader does not have a statistically significant effect on misreporting team outcomes (Fisher's exact test, $p = 0.256$) or the increase in misreporting from individual to team outcomes (Fisher's exact test, $p = 0.527$). In the *endogenous* treatment, the perceived chance to become a leader depends on the decision-maker's willingness to become a leader and on her belief about the number of team members willing to become a leader. Thus, the perceived chance to become a leader takes on values in $\{0, \frac{1}{3}, \frac{1}{2}, 1\}$.

[22] In the two regressions, we include a dummy controlling for the two probabilities of ending up as a leader. We also include a gender dummy and the same controls as in Table 1.

[23] Although baseline dishonesty rates appear higher in the *endogenous* treatment, these differences are not statistically significant by gender (male: $p = 0.256$; female: $p = 0.206$, Fisher's exact tests). Probit regressions confirm that individual dishonesty rates do not significantly differ between treatments for women (−0.466, SE = 0.325; with controls: −0.525, SE = 0.328). Crucially, our main conclusions are based on within-subject analyses that examine how individual participants change their dishonesty behavior when transitioning from individual to team decision-making contexts, making baseline differences between treatments less relevant for our core findings.

[24] To examine whether the non-significant effects might be attributable to insufficient statistical power, we conducted a post hoc power analysis using G*Power (two-tailed Fisher's exact test, $\alpha = 0.05$, n = 61 per group, p1 = 0.36, p2 = 0.46). The resulting power was only 0.1599, indicating a substantial risk of Type II error. This suggests that the study was underpowered to detect effects of this magnitude, and the non-significant results should therefore be interpreted with caution.

[25] To assess whether the non-significant findings could be attributed to limited statistical power, we conducted a post hoc power analysis using G*Power (two-tailed Fisher*s exact test, $\alpha = 0.05$, n = 57 per group, p1 = 0.12, p2 = 0.18). The analysis yielded a power of only 0.0905, indicating a considerable risk of Type II error. This suggests that the study lacked sufficient power to detect effects of this magnitude, and the non-significant results should be interpreted with caution.

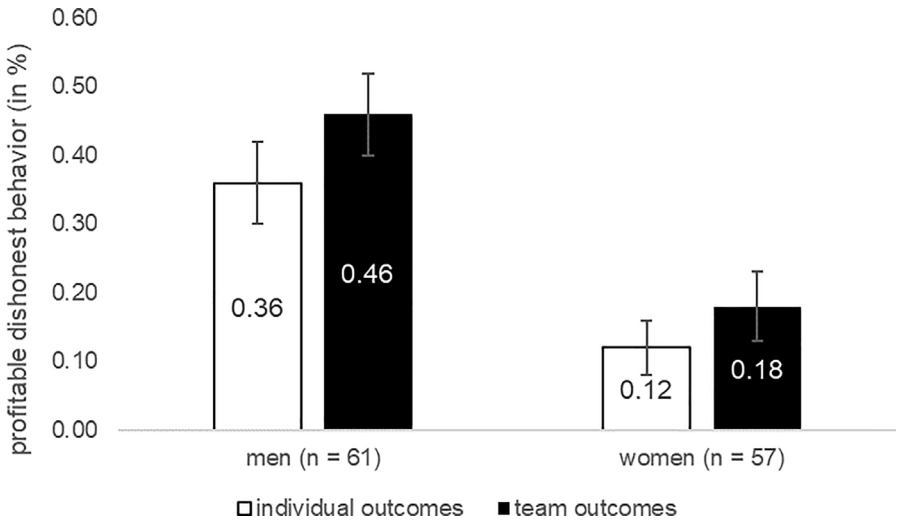

Fig. 2. Percentage of misreports under *exogenous leadership*. White (black) bars present the reports for individual outcomes (team outcomes). Standard error bars included.

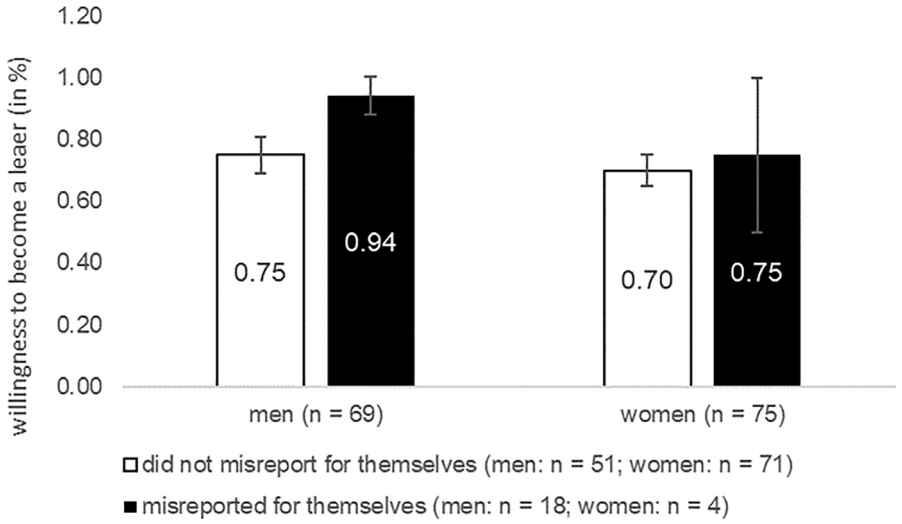

Fig. 3. Percentage of subjects who want to become a leader, conditional on gender in *endogenous leadership*. White (black) bars present subjects who reported truthfully for themselves (misreported). Standard error bars included.

The willingness to lead: The role of individual dishonesty and gender

In the following, we test Hypothesis 3, which expects that subjects with individual dishonesty preferences are more likely to apply for leadership positions. Motivated by our previous results, we also want to explain why women compared to men misreport team outcomes more often than individual outcomes. Therefore, we investigate the willingness to become a leader for men and women separately.

Fig. 3 provides an overview of women's and men's share of applying for the leadership role by individual dishonesty preferences. For men, it can be seen that a higher fraction (94%, or 17 out of 18 men) decides to become a leader when they behaved dishonestly at the individual decision, as compared to men who behaved honestly in the individual domain (75%, or 38 out of 51 men). We find that this difference in the willingness to become a leader is statistically significant on a 10%-level (Fisher's exact test: $p = 0.094$). For women, we find that the fractions of women who want to become a leader are similar when they behaved dishonestly (75%, or 3 out of 4 women) and honestly (70%, or 50 out of 71 women) in the individual domain (Fisher's exact test: $p = 1.000$). This result should be interpreted with caution, as only four women misreported. Overall, dishonest men are more likely to take on leadership roles, while dishonesty appears to be less relevant for women. Thus, our data supports Hypothesis 3 for men while we find no statistically significant evidence in our data for women. [26]

Study 2: Channels of changes in dishonest behavior

In Study 1, we observed that women increased their dishonest behavior as leaders under the *endogenous leadership* condition. There are two possible reasons for this increase in dishonesty: guilt sharing, where individuals feel diminished moral responsibility for their actions in group settings (Conrads et al., 2013; Gino et al., 2013; Wiltermuth, 2011), or efficiency concerns (Engelmann et al., 2004) because dishonesty by a team leader generates greater overall welfare. However, these explanations are insufficient considering our control condition results. Importantly, the data does not show a statistically significant

[26] Generally, there is no statistically significant difference between men and women in the willingness to apply for leadership (Fisher's exact test: $p = 0.250$).

increase in dishonest behavior among women in the *exogenous leadership* condition, despite the presence of guilt-sharing opportunities and identical payoff structures in the team domain across the two treatments. Specifically, our control treatment shows that women only statistically significantly increase dishonest behavior when they can apply for leadership positions. Although these findings highlight the importance of the *endogenous leadership* choice, we know little about the channels through which women's behavior changes. [27] Therefore, we pre-registered and conducted a follow-up study (Study 2). [28] Based on the findings in Study 1, we pre-registered the hypotheses that men behave more dishonestly than women when deciding for themselves (H4) and that under *endogenous leadership* women increase dishonesty from the individual to the team domain more strongly than men (H5). In Study 1, we also found indicative evidence that women who wanted to become a leader more often switched from individual honesty preferences to dishonest behavior for teams (23%) than women who did not want to become a leader (14%). We do not find any evidence of such an association in our data for men. Thus, we pre-registered a third hypothesis in Study 2 (H6), which expects that women who assume leadership show a stronger increase in misreporting from individual to team outcomes than women who do not apply for leadership. The sample size orientates on Study 1 allowing to detect similar effect sizes as in Study 1.

Study 2 aims to achieve two goals. First, we examine the underlying predictors of behavioral changes among individuals who endogenously selected into leadership roles. Given the payoff externalities of leaders' reporting decisions on team members' outcomes, leaders' dishonest behavior might be influenced by their prosociality. Furthermore, leaders might not just care about payoff consequences for their team members but also about making reporting decisions that align with their team members' individual dishonesty preferences. To analyze these predictors of behavioral change, we added two additional measures. As a measure of subjects' prosociality, we elicit their social value orientation (SVO). Additionally, we measure subjects' beliefs about their team members' individual dishonesty preferences to analyze how these beliefs influence behavioral change when acting as a leader. The new dishonesty task in Study 2 is binary, which simplifies the measurement of these beliefs.

Second, Study 2 attempts to replicate our findings on endogenous selection into leadership roles from Study 1 by using a different dishonesty task and subject pool in an online experiment, thereby testing the robustness and increasing the generalizability of our results. For Study 2, we employ the dots task (Gino et al., 2010), which requires participants to report whether they observe more dots on the "left" or "right" side of a quadratic area. This methodological change eliminates the risk of participants being dishonest because they perceive the task as gambling, which was possible in Study 1's die-rolling paradigm. [29]

Experimental design

The experimental setup is almost identical to the *endogenous leadership* condition of the first study. The main difference is the use of a different dishonesty measure than in Study 1. In Study 2, we use the dots task (Gino et al., 2010), in which we ask participants to report on which half of a quadratic area ("left" or "right") they see more dots. Reporting "right" corresponds to misreporting and leads to a higher payoff than reporting "left" which is a truthful report.

The study comprises four parts, and one of them is randomly determined to be payoff-relevant. In part one, we elicit subjects' SVO with the slider measure introduced by Murphy et al. (2011). Here, subjects are repeatedly confronted with two possible payoff allocations between them and another subject. In each decision set, the allocations vary the payoff differences and subjects have to trade off different money allocations. Based on their choices, we calculate an SVO angle for each subject (see Supplementary Material for instructions as well as for a screenshot of one of the allocation decisions; for the angle's calculation see Murphy et al., 2011). In part two, we measure dishonest behavior and subjects reported individual outcomes. A truthful report leads to a payoff of £0.20, and a dishonest report to a payoff of £2.00. Part three is similar to the team-dishonesty measure in the former experiment, except for the different dishonesty game. A truthful report leads to a payoff of £0.20 for each team member, and a dishonest report £2.00 for each team member. Thereafter, in part four, we elicit subjects' beliefs on the team members' individual dishonesty preferences in an incentivized way. They are asked about their belief of how many other team members indicated "right" in part two of the experiment. A correct guess yields a payoff of £2.00. Finally, we asked several survey questions to gather additional evidence on women's and men's motivation to act as a leader. We asked them if they wanted to influence payoffs and/or have the power of decision-making. [30]

We recruited 156 subjects (76 men and 80 women) using Prolific (Palan & Schitter, 2018) and surveyed them using Qualtrics. [31] We used ex-post matching, which is standard in online experiments that do not require real-time interaction (Rand, 2012). Participants were informed that they would be matched with others in the study, without any specifying of timing. Decisions were made independently, and matching occurred after data collection. Payments, including bonuses based on decisions and random draws (as explained in the instructions), were issued together a few days later, following Prolific standards. The sample is limited to UK citizens with a high school degree or a higher education. The average participant was 32.57 years old. The average time spent on the experiment was 6.81 min. Participants earned £1.89 on average, including a show-up fee of £1.00.

[27] To prepare Study 2, we conducted a pilot lab experiment in a similar within-subjects setting as in the *endogenous leadership* treatment of Study 1 with 219 subjects (90 men and 129 women). Afterwards, we elicited leaders' beliefs on the dishonest behavior of a randomly selected team member in the individual domain. A disadvantage of this approach is that we have to apply the strategy method and that we have to compute mean beliefs of the guesses, as the die-rolling task can result in six different outcomes. The results of the pilot study suggest that women who wanted to become leaders increase dishonesty for teams when holding an above-median belief on the reported die number of their team members in the individual domain. Based on the findings of the pilot study, Study 2 was designed to improve the analysis of subjects' beliefs regarding team members' behavior when reporting high payoffs in the individual domain. Therefore, we conducted a well-powered pre-registered online experiment, applying a simpler dishonesty task and easier belief elicitation (see above).

[28] The pre-registration can be found here: https://aspredicted.org/gm9v3.pdf.

[29] As an additional robustness check, we conducted a pre-registered experiment (https://aspredicted.org/re92p.pdf) to analyze whether our findings are confounded by the repetition of decisions. In this experiment, participants make a single reporting decision which corresponds to the group context of the *endogenous* treatment of Study 2. This design allows us to analyze whether the observed absence of the gender gap in the group context can also be found in isolation, that is, when no prior decision in an individual context has been made. In our robustness study, men and women behave dishonestly in 51% (men) and 47% (women) of the cases. Firstly, their decisions are not statistically significantly different from each other (Fisher's exact test: $p = 0.753$). Secondly, the behavior of men (Fisher's exact test:$p = 1.000$) and women (Fisher's exact test:$p = 0.753$) is not statistically significantly different to the group decisions in Study 2.

[30] Survey responses do not differ between men and women. Therefore, we do not report them in the paper.

[31] A limitation of our study is the absence of an a priori power calculation for Study 2. While the sample size was based on Study 1, where we observed significant effects, this approach may limit the generalizability and statistical precision of our findings. Future research should aim to replicate the results with a larger sample and a formally pre-registered power analysis.

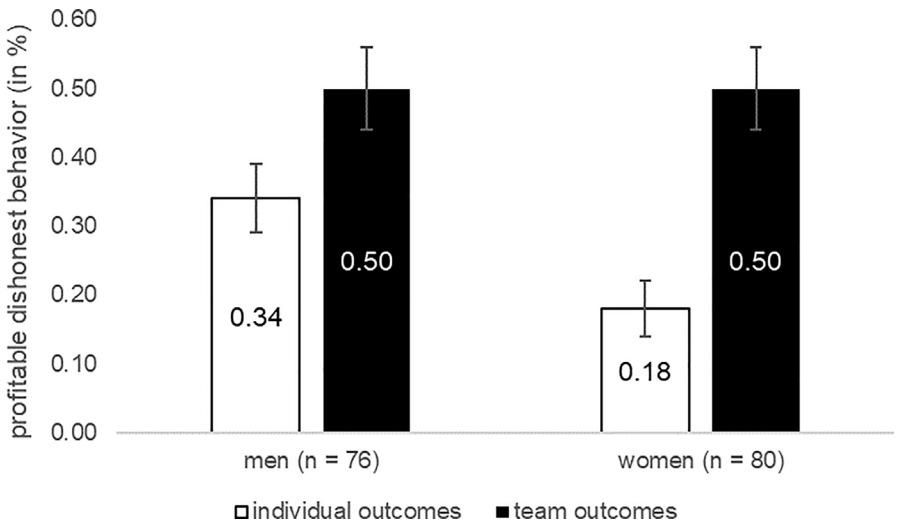

Fig. 4. Percentage of misreporting under *endogenous leadership* in Study 2. White (black) bars present misreporting individual outcomes (team outcomes). Standard error bars included.

Replication of the results

We start with our results on gender differences in dishonest behavior. We condition dishonest behavior on men (left panel) and women (right panel). Fig. 4 presents women's and men's percentages of dishonest reports when misreporting individual outcomes (white bars) and team outcomes (black bars) in Study 2.

As can be seen in Fig. 4, the results in Study 2 look similar to the results in Study 1. Again, we find a gender difference in individual dishonesty preferences. Men behave significantly more dishonestly than women (Fisher's exact test: $p = 0.018$). The frequency of men's dishonest reports is almost two times higher (34%) than women's (18%). The result supports H4, and it is in line with the result from Study 1. By contrast, the gender difference is no longer statistically significant when subjects act as leaders (Fisher's exact test: $p = 1.000$), as both genders misreport team outcomes in 50% of the cases. In Study 2, the increase in dishonest behavior is positive and statistically significant for both genders (Wilcoxon signed-rank test, men: $p = 0.007$; women: $p < 0.001$). Importantly, a higher percentage of women switch from truthfully reporting individual outcomes to misreporting team outcomes compared to men. This difference is statistically significant at the 10% level (Fisher's exact test: $p = 0.075$). [32] The finding supports H5 and is in line with the results of Study 1, which suggest that in *endogenous leadership* particularly women increase dishonest behavior when deciding as leaders. A closer look reveals that the percentage of women switching from honest to dishonest behavior is not statistically significantly different between those who apply for leadership positions (36%) and those who do not (32%). (Fisher's exact test: $p = 0.797$). [33] Thus, we do not find support for H6.

Following the structure in Study 1, we now analyze the behavioral predictors behind leaders' motivations to report dishonestly. In Table 2, we present similar regressions to those in Table 1 of Study 1, examining subjects' likelihood to misreport team outcomes. We control for the influence of individual dishonesty preferences, gender, and subjects' beliefs on how many other subjects want to become a leader. Thus, we include the same controls as in Study 1, report marginal effects, and present standard errors in parentheses. [34] We report regressions with standard coefficients in Table 5 in the Supplementary Material.

The results highlight that only individual dishonesty preferences, captured by the highly significant positive coefficient, *misreported ind. outcome*, predict whether subjects misreport team outcomes. All other variables are statistically non-significant. In sum, we replicate the findings of Table 1 (Study 1).

Next, we focus on the subjects' willingness to become a leader. The percentage of men who want to become leaders is higher by 12 percentage points (88% vs. 76%) when they show an individual preference for dishonesty. We find that this difference is less pronounced for women (79% vs. 71%). We do not find that these differences are significant in Study 2 (Fisher exact tests, men: $p = 0.238$; women: $p = 0.747$). [35] In sum, we can replicate most of the results of Study 1. Importantly, in Study 2 we again observe that women switch from individual truthful reporting to dishonest behavior as leaders more often than men. As a consequence, the gender gap in misreporting behavior closes in the team domain, similar to Study 1. This finding supports the idea to focus on further analyses of the belief about other team members' dishonesty preferences.

Potential predictors of switching behavior

In this section, we look deeper into the potential behavioral factors predicting the main result. First, we analyze whether our two additional measures are different across genders. Potential differences may help to explain why the observed increase in dishonesty is particularly strong among women. In line with the literature (e.g., Grosch & Rau, 2017), we find women to be more prosocial according to the social value orientation measure as compared to men (Mann–Whitney test: $p = 0.059$). We find no gender difference in average beliefs about their team members' individual dishonesty preferences (Mann–Whitney test: $p = 0.804$). Next, we turn to the relevance of the two potential behavioral factors for leaders' reporting decisions.

As we find that both genders increase their dishonest behavior significantly as leaders, we now focus on subjects' decisions to switch from an individual truthful report to misreporting team outcomes. We examine how subjects' decisions to switch are associated with their

[32] The variable *switch* is coded as one if persons are honest when deciding for themselves and dishonest when deciding as leaders. Otherwise, the variable is coded as zero.

[33] Overall, we find that 44 subjects (28 women; 16 men) of 156 subjects (80 women; 76 men) switch from honest to dishonest behavior.

[34] The results are also robust for OLS regressions with and without included interaction terms.

[35] Similar as in Study 1 we do not find gender differences in the willingness to become a leader (Fisher's exact test: $p = 0.266$).

Table 2
Probit regressions on misreporting team outcomes (Study 2).

	misreporting team outcomes					
	all		*women*		*men*	
	(1)	(2)	(3)	(4)	(5)	(6)
misreported ind. outcome	0.477***	0.501***	0.456***	0.449***	0.471***	0.559***
	(0.069)	(0.071)	(0.136)	(0.137)	(0.064)	(0.066)
woman	0.080	0.076				
	(0.074)	(0.074)				
(perceived) chance of becoming a leader		0.094		0.292		−0.137
		(0.198)		(0.286)		(0.271)
controls[a]	no	yes	no	yes	no	yes
obs.	156	156	80	80	76	76

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1.
Notes: The regressions report average marginal effects.
[a] Controls: age, and whether subjects hold a university degree.

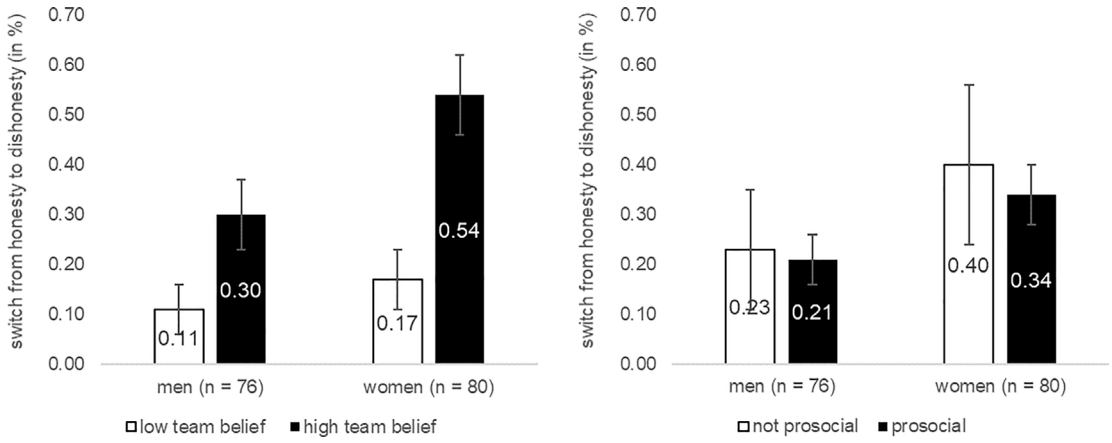

Fig. 5. Percentage change from an individual truthful report to misreporting team outcomes conditional on gender and behavioral factor (left panel: beliefs, right panel: prosocial behavior) in Study 2. White (black) bars present the reports for low levels (high levels). Standard error bars included.

beliefs about team members' dishonesty preferences and their own SVO.

Fig. 5 presents an overview of the percentage change from reporting truthful individual outcomes to misreporting team outcomes. We condition on gender and the potential behavioral factor (left panel: beliefs, right panel: prosocial behavior). We define a dummy that subjects hold a "high team belief" ("low team belief") when they believed that the strict majority, that is, two (none or one) other team members reported high payoffs at the individual stage.

Fig. 5 highlights that men (n=40) and women (n=39) with a high team belief switch more often than men (n=36) and women with a low team belief (n=41). Importantly, the effect size, as well as the statistical significance of this difference, is larger for women (54% vs. 17%) than for men (30% vs. 11%) (Fisher's exact test, women: $p = 0.001$; men: $p = 0.053$). Thus, we find a statistically significant gender difference when subjects hold a high team belief, that is, women are significantly more likely to switch than men (Fisher's exact test, $p = 0.041$). The gender difference is not statistically significant among leaders with a low team belief (Fisher's exact test, $p = 0.528$). By contrast, the right panel of the diagram demonstrates that being prosocial is not associated with the switching behavior of either gender (Fisher's exact tests, men: $p = 1.000$, n=13 "not prosocial", n=63 "prosocial"; women: $p = 0.734$, n=10 "not prosocial", n=70 "prosocial"). [36] The results indicate that leaders' beliefs but not their prosociality predict their behavioral change. Moreover, beliefs play a greater role for women's dishonest behavior as leaders than for men's. [37] We confirm these results using probit regression analyses in Table 3.

In the regressions, we included our two main variables of interest. First, we include the exact number (0, 1, or 2) of other team members whom the subjects believed to have reported dishonestly (team belief). Second, we include the measure of subjects *social value orientation* (i.e., the SVO angle). Models (1) and (2) additionally include a gender dummy (*woman*), whereas models (3)–(4) and (5)–(6) estimate models (1) and (2) separately for women and men. Regressions (2), (4), and (6) include subjects' age in years, and whether they have a university degree as control variables. The regressions report marginal effects

[36] Given that relatively few subjects are "not prosocial" in our data, the statistical power is very limited and we need to treat the non-significant result with caution.

[37] We acknowledge potential limitations in belief elicitation, as prior research has shown discrepancies between stated and revealed beliefs; for example, about others' prosociality, the tendencies for selfish individuals to strategically misreport, or underestimate others' prosociality (Andreoni & Sanchez, 2020; Molnár & Heintz, 2016; Serra-Garcia & Szech, 2022). Importantly, whereas motivated beliefs may influence individual and group lying decisions, they cannot account for the switch in dishonest behavior we observe in our study.

Table 3
Probit regressions on subjects' likelihood to switch from an individual truthful report to misreporting team outcomes (Study 2)

	switch from honest to dishonest behavior					
	all		*women*		*men*	
	(1)	(2)	(3)	(4)	(5)	(6)
team belief	0.221***	0.222***	0.301***	0.318***	0.139**	0.141**
	(0.049)	(0.049)	(0.064)	(0.066)	(0.069)	(0.061)
social value orientation	0.000	0.000	−0.002	−0.002	0.002	0.000
	(0.004)	(0.004)	(0.006)	(0.006)	(0.005)	(0.004)
woman	0.138**	0.129*				
	(0.066)	(0.066)				
controls[a]	no	yes	no	yes	no	yes
obs.	156	156	80	80	76	76

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1.
Note: The regressions report average marginal effects.
[a] Controls: age, and whether subjects hold a university degree.

(we report regressions with standard coefficients in Table 6 of the Supplementary Material). The results are also robust when using OLS regressions.

The analyses show that leaders who believe that a higher number of their team members reported a high payoff are more likely to switch, whereas subjects' social value orientation does not predict switching behavior. Models (3)–(6) show that these findings hold for both genders. The subsample regressions show that the coefficient of *team belief* is more than twice as large for women than for men ((3)–(4) vs. (5)–(6)). This suggests that team beliefs are more important predictors for women than for men, confirming previous results. We estimate a linear probability model which includes an interaction between *woman* and *team belief*. The coefficient of the interaction is positive and significant at a 10%-level ($p = 0.078$). Finally, in line with our previous results, models (1)–(2) confirm that women are generally more likely than men to increase dishonesty from the individual to the team domain.

Our findings indicate that leaders' behavioral shifts are predicted by their beliefs about team members' preferences for high payoff reports, but not by gender or prosociality levels. The regression analysis shows that the link between beliefs and behavior is more pronounced for women, partly accounting for their stronger behavioral shift compared to men's. A closer examination of the data supports this finding: whereas women's team belief coefficient is higher, their actual team belief (1.36) is statistically indistinguishable from men's (1.38; Mann–Whitney test: $p = 0.804$). The observed gender difference in prosociality, although noteworthy, fails to account for women's more pronounced behavioral shift.

Result 3. Behavioral factors predicting the switch from honesty to dishonesty
(a) The likelihood that women (and men) switch from an individual truthful report to misreporting team outcomes is highly positively correlated with their belief about their team members' behavior of reporting high payoffs in the individual domain. Moreover, this relation is more pronounced for women.
(b) Leader's prosociality does not predict switching behavior in our data.

Discussion and conclusion

In this paper, we analyze gender differences in ethical decision-making (dishonesty) of *leaders* in a laboratory (Study 1) and an online study (Study 2). Our experiments are based on within-subjects settings with two stages in which subjects first make a report that determines their individual outcomes and then a report that determines the payoffs for their teams as leaders. We also model subjects' deliberate choice to apply for the leadership role. We can analyze whether this choice is related to the individual dishonesty preferences of the subjects and whether it is associated with the misreporting behavior of leaders for their teams.

Our laboratory experiment (Study 1) demonstrates that men behave more dishonestly than women in the individual domain, corroborating the prevailing findings (e.g., Dreber & Johannesson, 2008; Grosch & Rau, 2017; Houser et al., 2016, 2012; Muehlheusser et al., 2015). One novel finding of this study is that women change their behavior in leadership roles when they can actively apply for them. However, we did not find statistically significant differences for men in our data, suggesting that they act more similarly in individual and team domains. This result corroborates the idea that women change their behavior once they assume responsibility for others. A related finding is provided by the experiment of Fornwagner et al. (2023), which shows that women act more competitively when they decide on behalf of others than individually. We acknowledge that in our study we ruled out such a competitive element to ensure clean identification. In the context of payoff reporting decisions, one could envision a setting in which teams compete against each other. It is possible that the gender differences observed in our results might manifest differently in such a competitive, inter-group environment. Overall, our findings on the shifts in women's ethical behavior when assuming leadership roles are consistent with role congruity theory (Eagly & Karau, 2002), indicating that women align their behavior with perceived leadership expectations more strongly than men do.

In a treatment, we alter the promotion procedure and disable the active application for the leadership role, that is, all employees are in the applicant pool for the leadership role by default (similar to Erkal et al., 2022). The control treatment shows that, under this promotion procedure, women's dishonesty does not differ significantly across the individual and team domains, suggesting a smaller or potentially non-existent effect, which we refrain from overstating given the low power for this specific analysis. Hence, the dishonesty of women in leadership positions does not substantially increase *per se*, but depends on the type of promotion procedure.

To learn more about the behavioral change in women under the promotion procedure with active applications, we conducted Study 2. We focus on the endogenous promotion procedure, motivated by our finding that women increase their dishonest behavior statistically significantly as leaders. We analyze two potential predictors. First, we examine if gender differences in prosociality predict the behavioral change in women. Second, we examine if beliefs about others' reporting behavior are associated with this result. We find that prosociality does not predict women's increase in dishonesty when acting as a leader. However, women leaders are more likely to behave in a way that reflects their belief that their team members prefer to report dishonestly than men are. That is, women behave dishonestly in the leadership role when they believe that their team members prefer to report dishonestly.

It is important to note that we cannot be sure whether all high-payoff reports in the dot task represent dishonest behavior, because participants may make mistakes when counting the numbers of dots. However, there is no reason to expect such errors would systematically favor profitable payoffs.

We further find that men with an individual preference for dishonesty tend to apply more often for leadership, whereas we lack the statistical power to make conclusions for women. This finding contributes to the "dark side" leadership concept (Hogan & Hogan, 2001), suggesting that these tendencies are prevalent among men in leadership positions. Our investigation enriches the literature on gender disparities in leadership representation by examining how differences in attitudes and preferences, such as risk tolerance, overconfidence, and gender role expectations, contribute to the under-representation of women in leadership positions (Eagly & Karau, 2002; Ertac & Gurdal, 2012; Reuben et al., 2012). [38] Our findings of this study also have implications for moral leadership theory (Treviño et al., 2003), particularly in understanding how individual ethical preferences may influence the pursuit of leadership roles.

An integral part of a leader's job is to assume responsibility for making decisions on behalf of others. We show that the choice to assume team responsibility is associated with behaving dishonestly as leaders for women, independent of their individual honesty preferences. Thus, this promotion procedure leads to a change in women's behavior, that is, they give up their individual ethical preferences. In our study, women significantly increase their dishonesty as leaders only when they actively choose to apply for leadership, but not — or only to a lesser extent — when they are automatically included in the applicant pool. Although this is an interesting finding, we can only speculate why this is the case. Through an active application, people can express their willingness to assume responsibility. The literature on endogenous institutions demonstrates that subjects are more willing to adhere to a certain behavior after they have deliberately voted for institutions. For example, subjects contribute more to public goods after they have voted for a punishment institution that intends to prevent free-riding (Kosfeld et al., 2009; Sutter et al., 2010). [39] Thus, women who actively apply ("vote") for a position involving unethical decisions may feel more obliged to adhere to the social norm compared to a situation with a different promotion procedure. This finding also highlights the importance of promotion procedures in which women maintain their ethical preferences and contribute to higher overall ethical conduct in their companies (Alan et al., 2022). Our findings suggest that companies aiming to strengthen ethical conduct should adopt internal promotion procedures in which suitable candidates are included in the applicant pool by default (Erkal et al., 2022). This approach enables women to act in accordance with their individual ethical preferences when assuming leadership roles. Such a promotion procedure has also been shown to facilitate breaking the glass ceiling (Erkal et al., 2022). High ethical conduct among leaders, in turn, impact the ethical conduct of employees throughout the company because they tend to align their ethical behavior with that of their leaders, as demonstrated by Alan et al. (2022). Thus, companies and other institutions can promote ethical behavior among their employees through the type of promotion procedure in place. When we further extrapolate our results, the type of promotion procedure may have further unintended consequences. Women in executive positions who adapt their leadership behaviors to the perceived social norm of their team may face further long-term consequences. That is, changing preferences and adapting to the preferences of others may lead to higher mental stress (Gardiner & Tiggemann, 1999). To reduce higher mental stress, women may eventually give up managerial positions or work part-time, further contributing to the gender gap in leadership positions (Manning & Petrongolo, 2008). There is little causal evidence on how individual worker behavior changes when the role transforms from employee to manager. Our study is novel in this regard, has several important workplace policy implications, and may stimulate further research in the future, contributing to basic knowledge about women's and men's behavioral change when assuming leadership positions.

[38] Our regression analysis in Study 1 reveals that risk tolerance is significantly and positively associated with dishonest behavior within the group domain, as evidenced by our control variable.

[39] Similarly, people cooperate more in prisoner's dilemmas after they were elected to make the decisions (Schories, 2022).

Limitations of the study

Whereas our study provides valuable insights into gender differences in leadership and ethical decision-making, it is important to acknowledge several limitations. First, we make use of a simple experimental design to ensure clean identification, which may not fully capture the complexities of real-world leadership contexts. For instance, our study lacks competitive elements that might be present in corporate environments, potentially affecting gender differences in behavior. Second, our belief measure has limitations in its ability to distinguish between perceptions of others' dishonest behavior and perceptions of others' ability to identify the actual outcomes presented to subjects in our honesty task (the dots task). Third, our study provides evidence of women's behavioral changes when they assume leadership roles, particularly when actively applying for these positions. Although we identify behavioral factors predicting this behavior, such as beliefs about team members' preferences, there may be additional psychological mechanisms at play. Further research could expand on our findings to examine how women and men approach leadership roles differently. Lastly, our study focuses on short-term behavioral changes in a controlled environment. Long-term consequences of adapting leadership behaviors to perceived team norms, such as increases in mental stress or impacts on career trajectories, are beyond the scope of this study and warrant further research. Despite these limitations, our findings provide valuable insights into gender differences in leadership ethics and have important implications for workplace policies.

CRediT authorship contribution statement

Kerstin Grosch: Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Stephan Müller:** Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Holger A. Rau:** Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Lilia Wasserka-Zhurakhovska:** Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.leaqua.2025.101910.

Data availability

We have uploaded our data on the Open Science Framework (OSF) at https://osf.io/my7ek/?view_only=cfd97dd2d321432c913f94f2c2758e58.

References

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters*, *80*(1), 123–129.

Alan, S., Corekcioglu, G., & Sutter, M. (2022). Improving workplace climate in large corporations: A clustered randomized intervention. *Quarterly Journal of Economics*.

Alan, S., Ertac, S., Kubilay, E., & Loranth, G. (2020). Understanding gender differences in leadership. *Economic Journal*, *130*(626), 263–289.

Alesina, A., Giuliano, P., & Nunn, N. (2013). On the origins of gender roles: Women and the plough. *Quarterly Journal of Economics*, *128*(2), 469–530.

Amore, M. D., Garofalo, O., & Guerra, A. (2022). How leaders influence (un) ethical behaviors within organizations: A laboratory experiment on reporting choices. *Journal of Business Ethics*, 1–16.

Andreoni, J., & Sanchez, A. (2020). Fooling myself or fooling observers? avoiding social pressures by manipulating perceptions of deservingness of others. *Economic Inquiry*, *58*(1), 12–33.

Andreoni, J., & Vesterlund, L. (2001). Which is the fair sex? Gender differences in altruism. *Quarterly Journal of Economics*, *1*(116), 293–312.

Ashton, M. C., & Lee, K. (2008). The hexaco model of personality structure and the importance of the h factor. *Social and Personality Psychology Compass*, *2*(5), 1952–1962.

Azar, O. H., Yosef, S., & Bar-Eli, M. (2013). Do customers return excessive change in a restaurant?: A field experiment on dishonesty. *Journal of Economic Behavior & Organization*, *93*, 219–226.

Azmat, G., & Petrongolo, B. (2014). Gender and the labor market: What have we learned from field and lab experiments? *Labour Economics*, *30*, 32–40.

Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics*, *116*(1), 261–292.

Barfort, S., Harmon, N. A., Hjorth, F., & Olsen, A. L. (2019). Sustaining honesty in public service: The role of selection. *American Economic Journal: Economic Policy*, *11*(4), 96–123.

Bass, B. M., & Avolio, B. J. (1994). Transformational leadership and organizational culture. *the International Journal of Public Administration*, *17*(3–4), 541–554.

Bear, S., Rahman, N., & Post, C. (2010). The impact of board diversity and gender composition on corporate social responsibility and firm reputation. *Journal of Business Ethics*, *97*(2), 207–221.

Belot, M., & Van De Ven, J. (2017). How private is private information? the ability to spot deception in an economic game. *Experimental Economics*, *20*(1), 19–43.

Belot, M., & Van De Ven, J. (2019). Is dishonesty persistent? *Journal of Behavioral and Experimental Economics*, *83*, Article 101464.

Berman, S. L., Wicks, A. C., Kotha, S., & Jones, T. M. (1999). Does stakeholder orientation matter? The relationship between stakeholder management models and firm financial performance. *Academy of Management JOurnal*, *42*(5), 488–506.

Bohnet, I., Geen, A. Van, & Bazerman, M. (2016). When performance trumps gender bias: Joint vs. *Separate Evaluation. Management Science*, *62*(5), 1225–1234.

Bollen, N. P., & Pool, V. K. (2009). Do hedge fund managers misreport returns? Evidence from the pooled distribution. *Journal of Finance*, *64*(5), 2257–2288.

Brañas-Garza, P., Capraro, V., & Rascon-Ramirez, E. (2018a). Gender differences in altruism on mechanical turk: Expectations and actual behaviour. *Economics Letters*, *170*, 19–23.

Brañas-Garza, P., Capraro, V., & Rascon-Ramirez, E. (2018b). Gender differences in altruism on mechanical turk: Expectations and actual behaviour. *Economics Letters*, *170*, 19–23.

Brassiolo, P., Estrada, R., Fajardo, G., & Vargas, J. (2021). Self-selection into corruption: Evidence from the lab. *Journal of Economic Behavior & Organization*, *192*, 799–812.

Brown, M. E., Treviño, L. K., & Harrison, D. A. (2005). Ethical leadership: A social learning perspective for construct development and testing. *Organizational Behavior and Human Decision Processes*, *97*(2), 117–134.

Bucciol, A., Landini, F., & Piovesan, M. (2013). Unethical behavior in the field: Demographic characteristics and beliefs of the cheater. *Journal of Economic Behavior & Organization*, *93*, 248–257.

Burns, N., & Kedia, S. (2006). The impact of performance-based compensation on misreporting. *Journal of FInancial EConomics*, *79*(1), 35–67.

Cappelen, A. W., Reme, B.-A., Sørensen, E. Ø., & Tungodden, B. (2016). Leadership and incentives. *Management Science*, *62*(7), 1944–1953.

Caselli, F., & Morelli, M. (2004). Bad politicians. *Journal of Public Economics*, *88*(3–4), 759–782.

Castillo, G., Choo, L., & Grimm, V. (2022). Are groups always more dishonest than individuals? the case of salient negative externalities. *Journal of Economic Behavior & Organization*, *198*, 598–611.

Chattopadhyay, R., & Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in india. *Econometrica*, *72*(5), 1409–1443.

Church, B. K., Hannan, R. L., & Kuang, X. J. (2012). Shared interest and honesty in budget reporting. *Accounting, Organizations and Society*, *37*(3), 155–167.

Conrads, J., Irlenbusch, B., Rilke, R. M., Schielke, A., & Walkowitz, G. (2014). Honesty in tournaments. *Economics Letters*, *123*(1), 90–93.

Conrads, J., Irlenbusch, B., Rilke, R. M., Walkowitz. Lying, G., & incentives, team. (2013). Journal of Economic Psychology. *34*, 1–7.

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, *47*(2), 448–474.

d'Adda, G., Darai, D., Pavanini, N., & Weber, R. A. (2017). Do leaders affect ethical conduct? *Journal of the European Economic Association*, *15*(6), 1177–1213.

Dai, Z., Galeotti, F., & Villeval, M. C. (2018). Cheating in the lab predicts fraud in the field: An experiment in public transportation. *Management Science*, *64*(3), 1081–1100.

Dreber, A., & Johannesson, M. (2008). Gender differences in deception. *Economics Letters*, *99*(1), 197–199.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, *109*(3), 573.

Eckel, C. C., & Grossman, P. J. (1998). Are women less selfish than men?: Evidence from dictator experiments. *the Economic Journal*, *108*(448), 726–735.

Eckel, C. C., & Grossman, P. J. (2002). Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution and Human Behavior*, *23*(4), 281–295.

Engelmann, D., Strobel. Inequality aversion, M., & efficiency (2004). And maximin preferences in simple distribution experiments. *American Economic Review*, *94*(4), 857–869.

Erkal, N., Gangadharan, L., & Xiao, E. (2022). Leadership selection: Can changing the default break the glass ceiling? *the Leadership Quarterly*, *33*(2), Article 101563.

Ertac, S., & Gurdal, M. Y. (2012). Gender, leadership and risk-taking in groups. *Journal of Economic Behavior & Organization*, *83*(1), 24–30.

European Institute for Gender Equality (2012). European Institute for Gender Equality. Gender equality and climate change. https://eige.europa.eu/publications/gender-equality-and-climate-change-report.

Fehr, E., & Schmidt, A. (1999). Does social capital have an economic payoff? a cross-country investigation. *The Quarterly Journal of Economics*, *114*(3), 817–868.

Fernandez-Mateo, I., & Fernandez, R. M. (2016). Bending the pipeline? executive search and gender inequality in hiring for top management jobs. *Management Science*, *62*(12), 3636–3655.

Fischbacher, U. (2007). Z-tree: Zurich toolbox for ready-made economic experiments. *Experimental EConomics*, *10*(2), 171–178.

Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies in disguise - an experimental study on cheating. *Journal of the European Economic Association*, *11*(3), 525–547.

Flabbi, L., Macis, M., Moro, A., & Schivardi, F. (2019). Do female executives make a difference? The impact of female leadership on gender gaps and firm performance. *Economic Journal*, *129*(622), 2390–2423.

Fornwagner, H., Pompeo, M., & Serdarevic, N. (2023). Choosing competition on behalf of someone else. *Management Science*, *69*(3), 1555–1574.

Gächter, S., Nosenzo, D., Renner, E., & Sefton, M. (2012). Who makes a good leader? cooperativeness, optimism, and leading-by-example. *Economic Inquiry*, *50*(4), 953–967.

Gardiner, M., & Tiggemann, M. (1999). Gender differences in leadership style, job stress and mental health in male-and female-dominated industries. *Journal of OCcupational and ORganizational PSychology*, *72*(3), 301–315.

Garretsen, H., Stoker, J. I., & Weber, R. A. (2020). Economic perspectives on leadership: Concepts, causality, and context in leadership research. , journal=The Leadership Quarterly. *31*(3), Article 101410.

Gino, F., Ayal, S., & Ariely, D. (2013). Self-serving altruism? The lure of unethical actions that benefit others. *Journal of EConomic BEhavior & ORganization*, *93*, 285–292.

Gino, F., Norton, M. I., & Ariely, D. (2010). The counterfeit self: The deceptive costs of faking it. *Psychological Science*, *21*(5), 712–720.

Gobillon, D. M., Laurent, & Roux, S. (2015). Estimating gender differences in access to jobs. *Journal of Labor Economics*, *33*(2), 317–363.

Graen, G. B., & Uhl-Bien, M. (1995). Relationship-based approach to leadership: Development of leader-member exchange (lmx) theory of leadership over 25 years: Applying a multi-level multi-domain perspective. *the Leadership Quarterly*, *6*(2), 219–247.

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, *1*(1), 114–125.

Grosch, K., Mittal, N., Stephan, M., Rau, H., Vollmer, S., & Wasserka-Zhurakhovska, L. (2023). Simplifying the elicitation of preferences: A lab-in-the-field approach in India. Mimeo.

Grosch, K., & Rau, H. A. (2017). Gender differences in honesty: The role of social value orientation. *Journal of Economic Psychology*, *62*, 258–267.

Güth, W., Levati, M. V., Sutter, M., & Van De. Heijden, E. (2007). Leading by example with and without exclusion power in voluntary contribution experiments. *Journal of Public Economics*, *91*(5–6), 1023–1042.

Hanna, R., & Wang, S.-Y. (2017). Dishonesty and selection into public service: Evidence from India. *American Economic Journal: Economic Policy*, *9*(3), 262–290.

Haucap, J., & Heldman, C. (2023). On the sociology of cartels. *European Journal of Law and Economics*, *56*(2), 289–323.

Hogan, R., & Hogan, J. (2001). Assessing leadership: A view from the dark side. *International Journal of Selection and Assessment*, *9*(1–2), 40–51.
Houser, D., List, J. A., Piovesan, M., Samek, A., & Winter, J. (2016). Dishonesty: From parents to children. *European Economic Review*, *82*, 242–254.
Houser, D., Vetter, S., & Winter, J. (2012). Fairness and cheating. *European Economic Review*, *56*(8), 1645–1655.
Huppert, E., & Levine, E. (2023). The rise of dishonest leaders: Causes and solutions. *Academy of Management Perspectives*, *37*(3), 239–251.
Jacobsen, C., Fosgaard, T. R., & Pascual-Ezama, D. (2018). Why do we lie? a practical guide to the dishonesty literature. *Journal of Economic Surveys*, *32*(2), 357–387.
Johnson, M. K., Rowatt, W. C., & Petrini, L. (2011). A new trait on the market: Honesty–humility as a unique predictor of job performance ratings. *Personality and Individual Differences*, *50*(6), 857–862.
Joulfaian, D. (2000). Corporate income tax evasion and managerial preferences. *Review of Economics and Statistics*, *82*(4), 698–701.
Kajackaite, A., & Gneezy, U. (2017). Incentives and cheating. *Games and Economic Behavior*, *102*, 433–444.
Knack, S., & Keefer, P. (1997). Does social capital have an economic payoff? a cross-country investigation. *the Quarterly Journal of Economics*, *112*(4), 1251–1288.
Kocher, M. G., Schudy, S., & Spantig, L. (2017). I lie? We lie! why? Experimental evidence on a dishonesty shift in groups. *Management Science*, *64*(9), 3971–4470.
Konrad, K. A., Lohse, T., & Simon, S. A. (2021). Pecunia non olet: on the self-selection into (dis) honest earning opportunities. *Experimental Economics*, *24*(4), 1105–1130.
Kosfeld, M. (2020). The role of leaders in inducing and maintaining cooperation: The cc strategy. *the Leadership Quarterly*, *31*(3), Article 101292.
Kosfeld, M., Okada, A., & Riedl, A. (2009). Institution formation in public goods games. *American Economic Review*, *99*(4), 1335–1355.
Kübler, D., Schmid, J., & Stüber, R. (2018). Gender discrimination in hiring across occupations: a nationally-representative vignette study. *Labour Economics*, *55*, 215–229.
Liu, D., Meagher, K. J., & Wait, A. (2022). Market conditions and firm morality: Employee trust in the honesty of their managers. *Journal of Economic Behavior & Organization*, *204*, 89–106.
Lohse, T., & Qari, S. (2021). Gender differences in face-to-face deceptive behavior. *Journal of Economic Behavior & Organization*, *187*, 1–15.
Lohse, T., & Simon, S. A. (2021). Compliance in teams–implications of joint decisions and shared consequences. *Journal of Behavioral and Experimental Economics*, *94*, Article 101745.
Manning, A., & Petrongolo, B. (2008). The part-time pay penalty for women in britain. *Economic Journal*, *118*(526), F28–F51.
Matsa, D. A., & Miller, A. R. (2013). A female style in corporate leadership? evidence from quotas. *American Economic Journal: Applied Economics*, *5*(3), 136–169.
Molnár, A., & Heintz, C. (2016). Beliefs about people's prosociality: Eliciting predictions in dictator games. (p. 19). CEU: Department of Economics-Working Paper.
Muehlheusser, G., Roider, A., & Wallmeier, N. (2015). Gender differences in honesty: Groups versus individuals. *Economics Letters*, *128*, 25–29.
Murciano-Goroff, R. (2022). Missing women in tech: The labor market for highly skilled software engineers. *Management Science*, *68*(5), 3262–3281.
Murphy, R. O., Ackermann, K. A., & Handgraaf, M. (2011). Measuring social value orientation. *Judgment and Decision Making*, *6*(8), 771–781.
Necker, S., & Paetzel, F. (2022). The effect of losing and winning on cheating and effort in repeated competitions. CESifo Working Paper.
Niederle, M., Segal, C., & Vesterlund, L. (2013). How costly is diversity? affirmative action in light of gender differences in competitiveness. *Management Science*, *59*(1), 1–16.
Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*, *122*(3), 1067–1101.
Palan, S., & Schitter, C. (2018). Prolific. ac-a subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27.
Potters, J., & Stoop, J. (2016). Do cheaters in the lab also cheat in the field? *European Economic Review*, *87*, 26–33.
Rand, D. G. (2012). The promise of mechanical turk: How online labor markets can help theorists run behavioral experiments. *Journal of Theoretical Biology*, *299*, 172–179.
Rand, D. G., Brescoll, V. L., Everett, J. A., Capraro, V., & Barcelo, H. (2016). Social heuristics and social roles: Intuition favors altruism for women but not for men. *Journal of Experimental Psychology: General*, *145*(4), 389.
Reuben, E., Rey-Biel, P., Sapienza, P., & Zingales, L. (2012). The emergence of male leadership in competitive environments. *Journal of Economic Behavior & Organization*, *83*(1), 111–117.
Ruffle, B. J., & Tobol, Y. (2017). Clever enough to tell the truth. *Experimental Economics*, *20*, 130–155.
Schories, F. E. (2022). The influence of indirect democracy and leadership choice on cooperation. *Experimental Economics*, *25*(4), 1173–1201.
Schwartz, M. S. (2005). Universal moral values for corporate codes of ethics. *Journal of Business Ethics*, *59*(1–2), 27–44.
Schyns, B., & Schilling, J. (2013). How bad are the effects of bad leaders? a meta-analysis of destructive leadership and its outcomes. *the Leadership Quarterly*, *24*(1), 138–158.
Selten, R. (1967). Die strategiemethode zur erforschung des eingeschränkt rationalen verhaltens im rahmen eines oligopolexperiments. In H. Sauermann (Ed.), *Beiträge zur experimentellen wirtschaftsforschung*. Mohr, Tübingen.
Serra-Garcia, M., & Szech, N. (2022). The (in) elasticity of moral ignorance. *Management Science*, *68*(7), 4815–4834.
Somanathan, E., & Rubin, P. H. (2004). The evolution of honesty. *Journal of Economic Behavior & Organization*, *54*(1), 1–17.
Soutschek, A., Burke, C. J., Raj. Beharelle, A., Schreiber, R., Weber, S. C., Karipidis, I. I., Te. Velden, J., Weber, B., Haker, H., Kalenscher, T., et al. (2017). The dopaminergic reward system underpins gender differences in social preferences. *Nature Human Behaviour*, *1*(11), 819–827.
Sutter, M., Haigner, S., & Kocher, M. G. (2010). Choosing the carrot or the stick? Endogenous institutional choice in social dilemma situations. *Review of Economic Studies*, *77*(4), 1540–1566.
Swamy, A., Knack, S., Lee, Y., Azfar. Gender, O., & corruption (2001). Journal of Development Economics. *64*(1), 25–55.
Tobol, Y., Siniver, E., & Yaniv, G. (2020). Dishonesty and mandatory mask wearing in the covid-19 pandemic. *Economics Letters*, *197*, Article 109617.
Treviño, L. K., Brown, M., & Hartman, L. P. (2003). A qualitative investigation of perceived executive ethical leadership: Perceptions from inside and outside the executive suite. *Human Relations*, *56*(1), 5–37.
United Nations (2019). More women in power vital for creating sustainable world, deputy secretary-general says, noting no country on track to fully achieve gender equality by 2030. https://www.un.org/press/en/2019/dsgsm1305.doc.htm.
Wiltermuth, S. S. (2011). Cheating more when the spoils are split. *Organizational Behavior and Human Decision Processes*, *115*(2), 157–168.
Zenger, J., & Folkman, J. (2019). Women score higher than men in most leadership skill. In *Harvard business review*.