TI 2024-002/VII
Tinbergen Institute Discussion Paper

Social desirability bias in attitudes towards sexism and DEI policies in the workplace

Anne Boring[1]
Josse Delfgaauw[2]

1 Erasmus University Rotterdam, Tinbergen Institute, and Sciences Po (LIEPP)

2 Erasmus University Rotterdam and Tinbergen Institute

Social desirability bias in attitudes towards sexism and DEI policies in the workplace*

Anne Boring† & Josse Delfgaauw‡

10th January 2024

ABSTRACT

Do workers speak their mind about sexism and about diversity, equity, and inclusion (DEI) policies in the workplace? We measure social desirability bias regarding sexism and DEI policies using a list experiment survey among workers from five male-dominated industries in France and in the US. In both countries and, remarkably, among both men and women, we document substantial social desirability bias. Managers exhibit a larger bias than non-managerial employees. This difference between voiced and real attitudes may make organizations overestimate support for DEI policies in their workforce, rendering such policies less effective.

Keywords: Sexism, Diversity, Social desirability bias, List experiment survey.
JEL Classification Numbers: J16, J71, M14, M5.

*María José González Fuentes provided excellent research assistance on this project. For very useful comments and suggestions while developing this project, we thank Martin Abel, Billur Aksoy, Iris Bohnet, Jen Brown, Edward Chang, Katie B. Coffman, Zoë Cullen, Will Dobbie, Morgane Laouenan, Suanna Oh, and Dario Sansone. The authors would also like to thank conference participants at the EEA's Annual Meeting, the KVS Nederlandse Economendag, UAE Discrimination and Diversity Workshop, Women in Finance Workshop, Women in Politics Workshop, and seminar participants at University of Exeter, Erasmus University Rotterdam, Sorbonne University, and the University of Florida. Anne Boring would further like to thank the Women and Public Policy Program at Harvard Kennedy School for their support. The Women in Business Chair at Sciences Po provided financial support on this project.

†Erasmus University Rotterdam, Tinbergen Institute, and Sciences Po (LIEPP), boring@ese.eur.nl or anne.boring@sciencespo.fr.

‡Erasmus University Rotterdam and Tinbergen Institute, delfgaauw@ese.eur.nl.

1 Introduction

Sexism in the workplace can take many forms. It ranges from subtle forms of implicit bias and microaggressions, to sexist or obscene jokes, degrading comments about women's ability, and explicit quid pro quo forms of sexual harassment (Basford et al., 2014; Hersch, 2011). Sexism contributes to a hostile work environment that can negatively impact the productivity and well-being of workers. Exposure to sexist behaviors, especially more hostile forms of sexual harassment, is one reason why women exit male-dominated industries or refrain from seeking leadership positions (Adams-Prassl et al., 2022; Batut et al., 2021; Folke and Rickne, 2022; Folke et al., 2020; Wu, 2020).[1]

Firms are increasingly implementing diversity, equity and inclusion (DEI) policies to address sexist behaviors (Azmat and Boring, 2020). DEI policies include any interventions aimed at increasing the representation of members of under-represented groups at all levels of a firm and improving firm culture such that workers from various backgrounds feel treated equitably and included.[2] In some countries, firms have primarily adopted gender based quota policies, such as in France where gender quotas are also enforced by government regulation (Azmat and Boring, 2020). In other countries like the US, firms have extensively implemented training programs that inform workers about sexism and the types of behaviors that contribute to a hostile work environment (Dobbin and Kalev, 2022).[3] However, the effectiveness of these policies remains unclear. While quota policies for recruitment and promotions enable more women to reach leadership positions and can reduce sexual harassment in firms (Niederle et al., 2013; Dobbin and Kalev, 2017), they can also lead to a competence signalling problem for women (Bijkerk et al., 2021; Avery, 2023) and a push-back from men who feel that they are not getting the promotions they deserve. Studies on DEI training programs have found little, if any, significant effects, especially in terms of behavior change from men (Chang et al., 2019). Moreover, training programs can also elicit push-back from workers, especially when training is mandatory (Dobbin and Kalev, 2016, 2019; Leslie, 2019; Forscher et al., 2019; Kalev et al., 2006).[4]

[1] More broadly, workers sort out of occupations to avoid prejudice from colleagues and hostile workplace cultures (Plug et al., 2014).

[2] Our paper focuses on gender-related DEI policies, while DEI generally concerns other minority groups as well.

[3] Training programs can take many forms, such as unconscious bias training, diversity training, sexual harassment training, and bystander training. DEI policies can also include initiatives such as commitments to diversity charters or labels, signalling non-discrimination in hiring, etc.

[4] By making sexism and sexual harassment more salient, DEI training can also have unintended consequences. Gertsberg (2022) finds that after #MeToo, junior female economists were less likely to start new collaborations with

The success of DEI policies requires support from a firm's workforce (Scarborough et al., 2019). However, employees and managers may not always speak their mind about these policies due to the sensitive nature of issues surrounding sexism and sexual harassment. Various factors can drive individuals to hide their true opinions, such as social image concerns or stigma (Bénabou and Tirole, 2006), career concerns (Morris, 2001), and fear of reprisal (Fehr and Gächter, 2000; Fehr and Fischbacher, 2004). As a result, firms may overestimate support for DEI policies among their staff.[5] Such firms may experience unanticipated pushback against DEI policies and may also fail to install proper checks and balances to ensure actual compliance with policies (Iyer, 2022).

In this paper, we study whether workers openly voice their opinions about sexism in the workplace and DEI policies or whether their reported views contain social desirability bias (SDB). SDB arises when individuals bias their answers to survey questions in the direction of a social norm and provide what they perceive to be socially acceptable responses instead of revealing their true beliefs (Zerbe and Paulhus, 1987; Stantcheva, 2022). We employ a survey conducted among five male-dominated and high-paying industries in both France and the US with over 1,600 respondents. We estimate the extent of SDB using the list experiment methodology (Miller, 1984; Glynn, 2013; Coffman et al., 2017). By randomly varying whether respondents' answers to four potentially sensitive questions[6] are exposed or remain hidden, the list experiment provides a measure of SDB when responses are exposed.

List experiments have been used to assess the extent of SDB regarding a wide variety of topics, such as LGBT sentiment in the workplace (Coffman et al., 2017; Aksoy et al., 2022), social distancing during the COVID-19 pandemic (Allen et al., 2021), attitudes towards affirmative action for African-Americans (Kuklinski et al., 1997), and prejudice against immigrants (Janus, 2010; Igarashi and Nagayoshi, 2022). Our aim is to determine whether SDB is a relevant concern with respect to workplace sexism and DEI policies. It is beyond the scope of our study to uncover the underlying mechanism driving SDB in the context of workplace sexism and DEI policies.

For perceived effectiveness of DEI policies, we document substantial SDB. In both France and the US, we estimate that the fraction of respondents that agrees with the statement "DEI policies

senior male colleagues, leading to fewer new projects. The paper argues that men responded to higher perceived risk of being accused of sexual harassment by sorting out of interactions with women.

[5]Bursztyn et al. (2023) finds that misperceptions about gender norms are widespread. In more gender-equal countries, the paper finds that people overestimate the support for affirmative action favoring women.

[6]Two are related to perceived effectiveness of DEI policies, and two are related to perceptions of sexism.

in the workplace do more good than harm" is 18 percentage points higher among respondents whose answer is exposed than among respondents whose answer remains hidden. Hence, actual support for DEI policies in the workplace is significantly lower than one would infer from asking directly. This holds for both men and women. We find only small gender differences in both stated and actual support for DEI policies.

For perceptions of sexism in the workplace, we document significant levels of SDB in France, but not in the US. However, in both countries, men tend to display more SDB than women, particularly in response to the statement "There is good reason for women to complain about sexism and sexual harassment in the workplace."

Remarkably, we estimate that in both countries, managers exhibit more SDB than employees in non-managerial positions. This is particularly so regarding DEI policies, where stated support among managers is higher than among non-managers, but our estimates show that actual support is lower among managers. Lastly, we find that in the US, both stated and actual attitudes toward DEI policies differ strongly along political lines.

The next section describes the survey and the sampling of respondents. Section 3 provides some descriptive statistics and describes the empirical approach. In Section 4, we present the results. We end with a discussion of implications for researchers and organisations.

2 The survey

2.1 Survey design

The main part of the survey is the list experiment. Survey respondents are randomly assigned to either the veiled condition or the direct condition. In the veiled condition, respondents receive lists of five statements. Each list contains a (potentially) sensitive statement. The respondents must indicate with how many statements they agree, without indicating with which statements they agree. In the direct condition, respondents must indicate whether they agree with the sensitive statement directly, and with how many of the four non-sensitive statements they agree. Hence, where the respondents in the direct condition may be inclined to give a socially desirable answer to the sensitive question, those in the veiled condition can express their true attitude as it remains hidden (provided the non-sensitive questions are chosen appropriately, as discussed below). Com-

parison between the two groups gives an estimate of the true attitude at the population level as well as an estimate of the degree of SDB regarding the sensitive issue.

We use four potentially sensitive items in the list experiment:

1. "I dislike sexist jokes in the workplace."
2. "There is good reason for women to complain about sexism and sexual harassment in the workplace."
3. "Nowadays, when a woman gets a promotion instead of a man, it is generally because she is more competent than the man."
4. "Diversity, equity, and inclusion policies in the workplace do more good than harm."

The first two items address sexism and its prevalence in the workplace. Item 1 elicits respondents' personal preference for sexist humor. Research suggests that the use of humor in sexist remarks reduces perceptions of sexist intentions (Mallett et al., 2016), but that individuals who rate higher on scales of sexism tend to appreciate sexist humor more than those who rate lower (Greenwood and Isbell, 2002; Thomas and Esses, 2004). Research also suggests that the use of humor can increase workers' social status in the workplace (Bitterly et al., 2017), although interpretation of what is humorous (benign or hostile) depends on the prevailing social norm (Gray and Ford, 2013; Warren et al., 2021). Item 2 elicits respondents' general perception of prevalence and effects of sexism in the workplace. Workers who disagree with these two items will see limited value in efforts to confront sexism.

The other two items elicit perceptions of DEI policies. Item 3 addresses a recurring concern regarding positive discrimination including quota policies (Foley and Williamson, 2019). Workers who believe that current policies undeservedly favour women may not openly state this view, as it may signal that they hold a negative view of women in higher positions (Bijkerk et al., 2021). Item 4 asks for a general perception as to whether DEI policies have a positive influence at work. Here, respondents' perception can be influenced by their experience with such policies. While some respondents may oppose DEI policies generically, others may support them in principle but find the specific policies they have experienced ineffective or counterproductive. Our design allows us to assess the extent to which respondents hide their critique of DEI-policies.

Coffman et al. (2017) finds that list experiments work better at eliciting SDB if the sensitive

answer to a question is "no" rather than "yes".[7] Two recent meta-analyses (Ehler et al., 2021; Li and Van den Noortgate, 2022) also support this framing choice in list experiments. In particular, Li and Van den Noortgate (2022) finds that list experiments are more effective when measuring SDB in attitudes that fit with mainstream social norms (in our case, supporting DEI policies and non-sexist behaviors in the workplace). This is because respondents are more willing to present themselves in a favorable manner. Finally, online survey respondents tend to provide faster, more automatic responses to items involving desirable traits (Andersen and Mayerl, 2019). Therefore, we follow best practice based on this literature and phrase all four items such that the socially desirable answer is to agree with the item. We also include a list with a placebo item: "I currently have a Netflix subscription." Hence, in total we have five lists in the list experiment.

List experiments work best when the four non-sensitive items in each set are such that one item is low prevalence, one item is high prevalence, and the two final items are negatively correlated (Glynn, 2013; Coffman et al., 2017). This prevents floor and ceiling effects, which could expose respondents' views on the sensitive items even in the veiled condition. With these criteria, we selected items from the General Social Survey. We ran a pilot survey on 167 respondents, which led to some adjustments in the set of non-sensitive items.[8] The complete survey can be found in Appendix C.

Before the list experiment part, the survey elicits several personal and work characteristics. After the list experiment part, we include questions on job satisfaction and workplace culture. We also include several questions on perceived sexism and DEI policies. These direct questions may induce socially desirable answers. If so, we would expect the differences between respondents from the direct and the veiled conditions on these questions to be smaller than the differences in the list experiment. Finally, we add a question on respondents' political views. Respondents place themselves on a 10-point scale from the political left to the political right. We group respondents into category Right if they answered 7 or higher, into Center if they answered 5 or 6, and into Left if they answered 4 or lower.

Careless or inattentive respondents reduce the quality of survey data (Stantcheva, 2022). Following Haaland et al. (2023), we include an attention check in the survey directly after the list

[7] See section 6 of Coffman et al. (2017) for a complete discussion about this finding.

[8] The same items were used in France and the US. The only difference was that the item 'I would move within the US for work' in the US survey was phrased as 'I would move within France for work' in the French survey.

experiment questions. Before running the survey, we set criteria for including respondents into our sample: (i) respondents should answer the attention check question correctly, (ii) spend at least 5 (7) seconds per list question in the direct (veiled) condition, and (iii) spend at least 5 minutes on the complete survey.

2.2 Survey respondents

An online panel company distributed the survey in the summer of 2022 to employed workers in five male-dominated industries: information technology, finance (which we defined as including banks, credit unions, financial advisors, tax and accounting firms, insurance companies, private equity and venture capital firms), law, business consultancy, and engineering. Individuals who were not employed by a firm in one of these five sectors were screened-out. We selected workers holding at least a high school degree and we ensured that the fraction of women in our sample would lie between 40 and 60 percent in order to be able to document gender differences. We chose to conduct the survey in France and the US, because the two countries have had a different approach to DEI policies. While France has been one of the leading countries in terms of public policies involving quotas and creating incentives for firms to promote more women to leadership positions, the US has a longer history of training programs for workers (Azmat and Boring, 2020).[9] Our sample contains 703 respondents for France and 921 for the US.[10]

3 Descriptive statistics and empirical approach

Table 1 provides the descriptive statistics of respondents. Women are 54% of respondents in the French survey and 58% in the US survey.[11] In France, 66% of respondents are younger than 45, in the US this is 43%. In the US, 82% of respondents are white and 18% are people of color (POC).[12] Most respondents work for the finance industry: 59% of respondents in France and 44%

[9]Firms in France started implementing training programs only recently. In 2017, a new law (*loi du 27 janvier 2017 "Égalité et Citoyenneté"*) made it mandatory for recruiters of firms with more than 300 employees to follow a training program against discrimination in hiring. In our sample, 45% of French respondents vs. 76% of US respondents declare having followed some type of DEI training.

[10]We initially aimed for 1,000 respondents in France and in the US. The survey ran from the end of June 2022 to August 2022, after which we had 703 respondents who met our inclusion criteria for France and 921 for the US. Table B1 in Appendix B shows the fraction of respondents that did not meet our criteria.

[11]We gave respondents the option to select female, male or other. Only one respondent in the US selected other. This respondent is not included in our final sample.

[12]It is not legal to collect information about race in French surveys.

of respondents in the US. About half of our respondents hold a management position.

Concerning randomization, Table 2 shows no statistically significant differences on observable characteristics between respondents assigned to the veiled condition and respondents assigned to the direct condition, neither in France nor in the US samples. For both countries, the fraction of respondents in the direct condition is slightly higher than in the veiled condition.

Our empirical approach follows Coffman et al. (2017). In the veiled condition, respondent i reports the number of statements he or she agrees with out of the five statements in list question q. Denote this by y_{qi}^V. In the direct condition, let d_{qi} be equal to one if the respondent agrees with the direct sensitive item and zero otherwise, and let c_{qi} be the respondent's answer to the number of remaining four statements she agrees with. We construct $y_{qi}^D = d_{qi} + c_{qi}$, which gives the number of statements in list question q that respondent i reports to agree with in the direct condition.

Given randomization and truthful reporting, we would expect that $E(y_{qi}^D) = E(y_{qi}^V)$. A significant difference between $E(y_{qi}^D)$ and $E(y_{qi}^V)$ points to SDB regarding the directly asked item. We denote this difference as $\mu_q = E(y_{qi}^V) - E(y_{qi}^D)$. Given that the sensitive answer to a question is 'no', the presence of SDB would imply $\mu_q < 0$. Using OLS, we examine the degree of SDB by estimating

$$y_{qi} = \mu_q V_i + \beta X_i + \epsilon_{qi}, \tag{1}$$

where $V_i \in \{0, 1\}$ indicates whether respondent i was in the veiled condition, X_i is a set of control variables, and ϵ_{qi} is a noise term. As control variables, we include dummy variables for gender, age, race (US only), education (having at least the equivalent of a Bachelor's degree), being a manager, having children, working part time, and industry and region fixed effects.

Besides randomization, two assumptions underlie the list experiment methodology (Blair and Imai, 2012) that allows interpreting $\mu_q < 0$ as SDB. First, the difference between the veiled and the direct condition does not affect respondents' answers to the non-sensitive items ('no design effect'). Blair and Imai (2012) propose a statistical test based on comparing conditional probabilities between the two conditions. Intuitively, as the lists contain one more item in the veiled condition than in the direct condition, the proportion of respondents in the direct condition that answers at least c_q on the (short) list should lie in between the proportions of respondents in the veiled treatment that answers at least c_q and at least $c_q + 1$. The test does not reject the null hypothesis of

no design effects for any of our lists. Hence, there is no indication that respondents' answers to the control items are affected by the inclusion of the sensitive item to the lists in the veiled condition.

The second assumption is that respondents in the veiled condition give a truthful response to the sensitive item ('no liars'). This assumption cannot be tested, as true attitudes are unobserved. However, the no-liars assumption is likely violated in the presence of significant floor and ceiling effects (Blair and Imai, 2012). Figure A1 in the Appendix shows the histograms of the responses to the list questions in both countries. We find very few respondents that agree with either zero or all statements. Furthermore, across the sensitive list questions, the average fraction of respondents in the direct condition that agrees with zero out of the four non-sensitive items is less than two percent. Hence, our list questions succeed in allowing nearly all respondents in the veiled condition to answer the sensitive question truthfully without exposing their answer.

4 Results

Before presenting the results of estimating Equation (1), Figure 1 presents the difference μ_q for each list question. First, the placebo question shows small and non-significant differences between the veiled and the direct condition for both France and the US, reassuring that the differences on the other lists are driven by the sensitive items. In contrast, we find substantial differences between the two conditions for all four sensitive items in France and for two sensitive items in the US. French respondents in the direct condition are considerably more likely than respondents in the veiled condition to claim that they dislike sexist jokes, that women rightfully complain about sexism, that women receive promotions based on competence, and that DEI policies do more good than harm. For the US, we obtain a similar result for the last two items.

Table 3 presents the results of estimating Equation (1). In France, 74.6 percent of respondents in the direct condition claim to dislike sexist jokes in the workplace. In the veiled condition, the estimate is 24 percentage points lower, which implies that only 50.6 percent of respondents actually dislike sexist jokes. We find similarly significant differences between the veiled and the direct condition on the other three sensitive items. For the US, we find no differences between the two conditions for disliking sexist jokes and for women's rightful complaints about sexism. For the two items related to DEI policies, we do estimate significant differences, comparable in magnitude to

the estimates for France. We cannot determine the drivers behind the differences between France and the US. One possible explanation is that diversity training programs are more widespread in the US as compared to France. Even though training can generate backlash in the short-run, it is possible that informing workers about sexism has led to a real shift in norms over time in the US.

Given the design of the list experiment, we can interpret the differences between the veiled and the direct conditions as the extent of SDB. Figure A2 in the Appendix presents further evidence in line with this interpretation. After the list questions, we asked respondents in both conditions several direct questions concerning sexism and DEI policies. Figure A2 shows that we find no significant differences in the average response to these questions between respondents in the direct and the veiled conditions. In other words, removing the ability to answer truthfully without exposing one's views makes the responses in the veiled condition equally biased as responses in the direct condition. This finding also mitigates two potential concerns. The first concern is that the differences between the conditions could have affected the SDB in the direct condition. For instance, if respondents in the direct condition inferred our research interest from the directly-asked sensitive questions, this could lead to inflated SDB in their responses. However, Figure A2 provides no indication of inflated SDB in the direct condition. A second concern is that our randomization may have failed on unobservable characteristics, leading to differences in attitudes across respondents in the two conditions. If so, we would expect such differences to also surface in Figure A2. Hence, we are confident that our findings in Figure 1 and Table 3 can be interpreted as the extent of SDB on sexism and DEI-policies in our sample.

We have estimated Equation (1) for subgroups. Figures 2 to 4 depict the stated and estimated actual agreement with each item.[13] One caveat of these analyses is that the number of observations in some cells is quite low, yielding large confidence intervals. We therefore also provide Table 4, which shows the significance of the interaction terms for different characteristics (gender, age, race, education, and management position) and the veiled condition.

Figure 2 presents the results of the subgroup estimations by gender, age, education, and race (for the US). In the direct condition, the fractions of men and women that agree with the sensitive statements are quite similar. However, for the items on sexist jokes and women's complaints about sexism, we find substantially more SDB among men than among women. This also implies that the

[13]Table A1 provides the corresponding estimated SDB μ_q per sensitive item.

fraction of men that actually agrees with these statements is substantially lower than the fraction of women. For the two items on DEI policies, we find a similar pattern for men and women. Hence, the support for DEI policies among women in the direct condition also overstates their actual support.

Regarding age, we find an interesting difference between France and the US for the DEI policy questions. In France, young employees display more SDB than older employees and have less actual support for DEI policies, whereas in the US the reverse holds. We conjecture that this is driven by differences in approach to DEI policies between the two countries. The emphasis on quota and promotion policies in France affect relatively young employees most, whereas older employees in the US may see limited value in the focus on DEI training.

In France, lower-educated employees display more SDB than higher-educated employees. This difference is less pronounced in the US. Lastly, we find more SDB among white employees in the US than among people of color.

Do managers display more or less SDB than employees? Figure 3 presents the results of estimating Equation (1) separately for managers and for non-managerial employees. In the direct condition, managers appear to agree at least as often with the sensitive statements as non-managerial employees. However, managers display more SDB, particularly so in France. In both countries, our findings indicate that managers are less concerned about sexism and have less faith in DEI policies than non-managerial employees.[14] We can only speculate as to why managers display more SDB. One explanation is that managers want to influence workplace norms by setting a good example. Another explanation is that managers are more often subjected to and/or restricted by DEI policies.

Figure 4 shows that both stated and estimated actual agreement with the sensitive items tend to be lower among respondents that have right-wing political views compared to respondents with left-wing political views. The estimated levels of SDB are quite similar across the political spectrum.[15] Most striking are the stark differences in beliefs (both stated and actual) regarding the effects of DEI policies between left-wing and right-wing respondents in the US.

[14] As discussed in Section 3, we include a set of controls in estimating the SDB, implying that the findings in Figure 3 are not driven by differences in demographic characteristics between managers and non-managerial employees.
[15] Past research has found stronger SDB among Democrats for issues related to women in politics (Setzler, 2019) and to LGBTs in the workplace (Coffman et al., 2017).

5 Discussion

Studying male-dominated industries in France and the US, we document widespread SDB in workers' reported attitudes regarding sexism and DEI policies. This has implications for researchers and for firms. Researchers should be aware that many people do not speak their minds on these topics. Moreover, it is possible that particular interventions affect stated but not actual views. Chang et al. (2019) finds that a short online diversity training positively affected men's supportive attitudes towards women, but did not affect men's actual support for women in the workplace. One interpretation is that the change in attitudes reflects a change in SDB more than a change in actual attitudes. Our findings imply that disentangling changes in actual attitudes (and behavior) from changes in stated views is important to evaluate the effectiveness of DEI policies.

Given that respondents exhibit SDB in a survey, we expect such bias to also affect communication in the workplace. This need not be altogether negative. Insofar as it reduces sexist behavior, SDB may improve women's productivity and well-being at work. At the same time, worker support for reducing sexism and implementing DEI policies may be lower than voiced. Employees who perceive sexism as non-prevalent or unimportant are reluctant to share these views. Similarly, employees may oppose DEI policies without expressing their opposition, and push back once such policies are implemented. If more workers view DEI policies as unnecessary or harmful than firms expect, implementation of such policies may lead to lower compliance, more push-back, and, as a result thereof, fewer positive effects than firms anticipate. Furthermore, our findings suggest that a sizable fraction of managers may pay lip-service, but does not actually support DEI policies. As managers play a key role in addressing sexism and implementing DEI policies (Dobbin et al., 2015; Scarborough et al., 2019), this lack of genuine managerial support further hampers policy effectiveness.

Our findings suggest that improved communication within organisations may contribute towards more effective policies. Alan et al. (2023) finds that a management training program emphasizing communication and pro-social behavior improves workplace climate and productivity. In our survey, we asked questions about perceptions of the quality of communication and trust with colleagues and managers, satisfaction with workplace culture, and whether sexism is a serious problem in the respondents' workplace. We find that employees who are more positive about communication with

colleagues and managers are also more positive about workplace culture (Figure A3) and indicate to perceive less sexism (Figure A4). An interesting avenue for future research is to analyse whether effective workplace communication training reduces sexism and whether this reduction mediates the effect of the training on workplace climate and productivity.

References

Adams-Prassl, A., Huttunen, K., Nix, E., and Zhang, N. (2022). Violence against women at work. *CEPR Discussion Paper No. 17504.*

Aksoy, B., Carpenter, C. S., and Sansone, D. (2022). Understanding labor market discrimination against transgender people: Evidence from a double list experiment and a survey. *NBER working paper No. 30483.*

Alan, S., Corekcioglu, G., and Sutter, M. (2023). Improving workplace climate in large corporations: A clustered randomized intervention. *The Quarterly Journal of Economics*, 138(1):151–203.

Allen, J., Mahumane, A., Riddell, J., Rosenblat, T., Yang, D., and Yu, H. (2021). Correcting perceived social distancing norms to combat covid-19. *NBER working paper No. 28651.*

Andersen, H. and Mayerl, J. (2019). Responding to socially desirable and undesirable topics: Different types of response behaviour? *Methods, data, analyses (mda)*, 13(1):7–35.

Avery, M. (2023). The self-limiting dynamics of affirmative action. *Mimeo*, Monash University.

Azmat, G. and Boring, A. (2020). Gender diversity in firms. *Oxford Review of Economic Policy*, 36(4):760–782.

Basford, T. E., Offermann, L. R., and Behrend, T. S. (2014). Do you see what I see? Perceptions of gender microaggressions in the workplace. *Psychology of Women Quarterly*, 38(3):340–349.

Batut, C., Coly, C., and Schneider-Strawczynski, S. (2021). It's a man's world: Culture of abuse, #metoo and worker flows. *Dondena Working Papers No. 149.*

Bénabou, R. and Tirole, J. (2006). Incentives and prosocial behavior. *American Economic Review*, 96(5):1652–1678.

Bijkerk, S. H., Dominguez-Martinez, S., Kamphorst, J., and Swank, O. H. (2021). Labor market quotas when promotions are signals. *Journal of Labor Economics*, 39(2):437–460.

Bitterly, T. B., Brooks, A. W., and Schweitzer, M. E. (2017). Risky business: When humor increases and decreases status. *Journal of personality and social psychology*, 112(3):431.

Blair, G. and Imai, K. (2012). Statistical analysis of list experiments. *Political Analysis*, 20(1):47–77.

Bursztyn, L., Cappelen, A. W., Tungodden, B., Voena, A., and Yanagizawa-Drott, D. H. (2023). How are gender norms perceived? *NBER working paper No. 31049.*

Chang, E. H., Milkman, K. L., Gromet, D. M., Rebele, R. W., Massey, C., Duckworth, A. L., and Grant, A. M. (2019). The mixed effects of online diversity training. *Proceedings of the National Academy of Sciences*, 116(16):7778–7783.

Coffman, K. B., Coffman, L. C., and Ericson, K. M. M. (2017). The size of the LGBT population and the magnitude of antigay sentiment are substantially underestimated. *Management Science*, 63(10):3168–3186.

Dobbin, F. and Kalev, A. (2016). Why diversity programs fail. *Harvard Business Review*, 94(7).

Dobbin, F. and Kalev, A. (2017). Training programs and reporting systems won't end sexual harassment. Promoting more women will. *Harvard Business Review*, 70(4):687–702.

Dobbin, F. and Kalev, A. (2019). The promise and peril of sexual harassment programs. *Proceedings of the National Academy of Sciences*, 116(25):12255–12260.

Dobbin, F. and Kalev, A. (2022). *Getting to diversity: What works and what doesn't*. Harvard University Press.

Dobbin, F., Schrage, D., and Kalev, A. (2015). Rage against the iron cage: The varied effects of bureaucratic personnel reforms on diversity. *American Sociological Review*, 80(5):1014–1044.

Ehler, I., Wolter, F., and Junkermann, J. (2021). Sensitive questions in surveys: A comprehensive meta-analysis of experimental survey studies on the performance of the item count technique. *Public Opinion Quarterly*, 85(1):6–27.

Fehr, E. and Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, 25(2):63–87.

Fehr, E. and Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, 90(4):980–994.

Foley, M. and Williamson, S. (2019). Managerial perspectives on implicit bias, affirmative action, and merit. *Public Administration Review*, 79(1):35–45.

Folke, O. and Rickne, J. (2022). Sexual harassment and gender inequality in the labor market. *The Quarterly Journal of Economics*, 137(4):2163–2212.

Folke, O., Rickne, J., Tanaka, S., and Tateishi, Y. (2020). Sexual harassment of women leaders. *Daedalus*, 149(1):180–197.

Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., and Nosek, B. A. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology*, 117(3):522–559.

Gertsberg, M. (2022). The unintended consequences of #MeToo: Evidence from research collaborations. *mimeo, University of Melbourne*.

Glynn, A. N. (2013). What can we learn with statistical truth serum? Design and analysis of the list experiment. *Public Opinion Quarterly*, 77(S1):159–172.

Gray, J. A. and Ford, T. E. (2013). The role of social context in the interpretation of sexist humor. *Humor*, 26(2):277–293.

Greenwood, D. and Isbell, L. M. (2002). Ambivalent sexism and the dumb blonde: Men's and women's reactions to sexist jokes. *Psychology of Women Quarterly*, 26(4):341–350.

Haaland, I., Roth, C., and Wohlfart, J. (2023). Designing information provision experiments. *Journal of Economic Literature*, 61(1):3–40.

Hersch, J. (2011). Compensating differentials for sexual harassment. *American Economic Review*, 101(3):630–34.

Igarashi, A. and Nagayoshi, K. (2022). Norms to be prejudiced: List experiments on attitudes towards immigrants in Japan. *Social Science Research*, 102:102647.

Iyer, A. (2022). Understanding advantaged groups' opposition to diversity, equity, and inclusion (dei) policies: The role of perceived threat. *Social and Personality Psychology Compass*, 16(5):e12666.

Janus, A. L. (2010). The influence of social desirability pressures on expressed immigration attitudes. *Social Science Quarterly*, 91(4):928–946.

Kalev, A., Dobbin, F., and Kelly, E. (2006). Best practices or best guesses? Assessing the efficacy of corporate affirmative action and diversity policies. *American Sociological Review*, 71(4):589–617.

Kuklinski, J. H., Sniderman, P. M., Knight, K., Piazza, T., Tetlock, P. E., Lawrence, G. R., and Mellers, B. (1997). Racial prejudice and attitudes toward affirmative action. *American Journal of Political Science*, 41(2):402–419.

Leslie, L. M. (2019). Diversity initiative effectiveness: A typological theory of unintended consequences. *Academy of Management Review*, 44(3):538–563.

Li, J. and Van den Noortgate, W. (2022). A meta-analysis of the relative effectiveness of the item count technique compared to direct questioning. *Sociological Methods & Research*, 51(2):760–799.

Mallett, R. K., Ford, T. E., and Woodzicka, J. A. (2016). What did he mean by that? Humor decreases attributions of sexism and confrontation of sexist jokes. *Sex Roles*, 75:272–284.

Miller, J. D. (1984). *A new survey technique for studying deviant behavior*. The George Washington University.

Morris, S. (2001). Political correctness. *Journal of Political Economy*, 109(2):231–265.

Niederle, M., Segal, C., and Vesterlund, L. (2013). How costly is diversity? Affirmative action in light of gender differences in competitiveness. *Management Science*, 59(1):1–16.

Plug, E., Webbink, D., and Martin, N. (2014). Sexual orientation, prejudice, and segregation. *Journal of Labor Economics*, 32(1):123–159.

Read, B., Wolters, L., and Berinsky, A. J. (2022). Racing the clock: Using response time as a proxy for attentiveness on self-administered surveys. *Political Analysis*, 30(4):550–569.

Scarborough, W. J., Lambouths, D. L., and Holbrook, A. L. (2019). Support of workplace diversity policies: The role of race, gender, and beliefs about inequality. *Social Science Research*, 79:194–210.

Setzler, M. (2019). Measuring bias against female political leadership. *Politics & Gender*, 15(4):695–721.

Stantcheva, S. (2022). How to run surveys: A guide to creating your own identifying variation and revealing the invisible. *Annual Review of Economics*, 15.

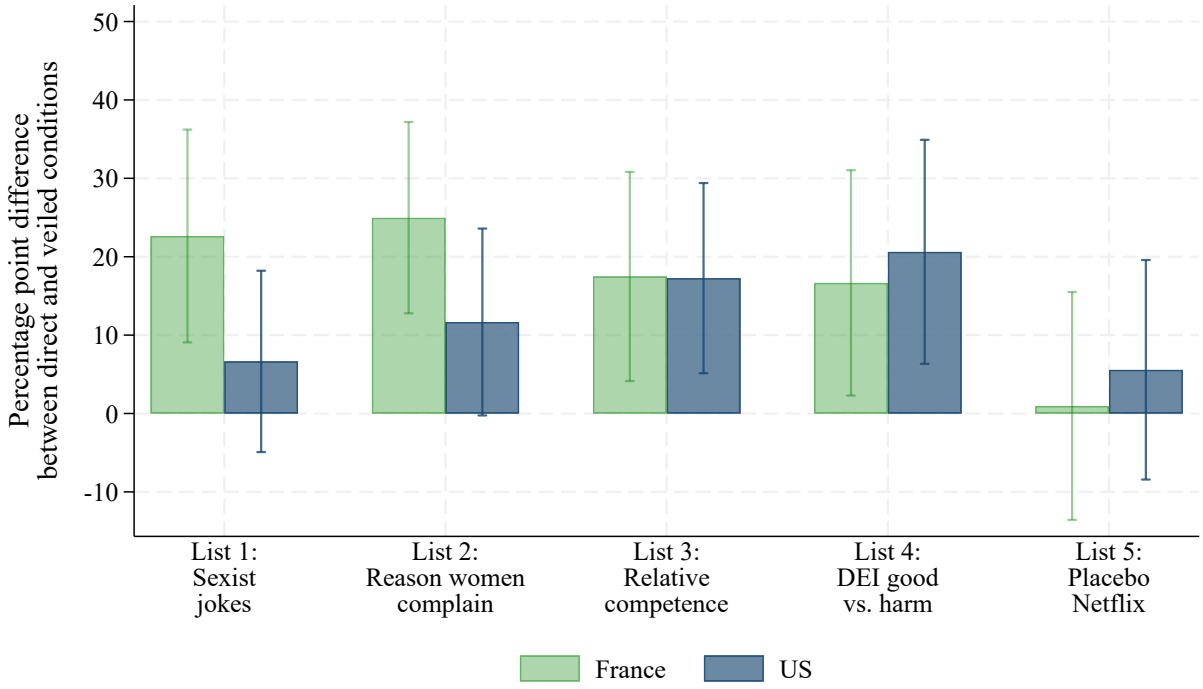
Thomas, C. A. and Esses, V. M. (2004). Individual differences in reactions to sexist humor. *Group Processes & Intergroup Relations*, 7(1):89–100.

Warren, C., Barsky, A., and McGraw, A. P. (2021). What makes things funny? An integrative review of the antecedents of laughter and amusement. *Personality and Social Psychology Review*, 25(1):41–65.

Wu, A. H. (2020). Gender bias in rumors among professionals: An identity-based interpretation. *Review of Economics and Statistics*, 102(5):867–880.

Zerbe, W. J. and Paulhus, D. L. (1987). Socially desirable responding in organizational behavior: A reconception. *The Academy of Management Review*, 12(2):250–264.

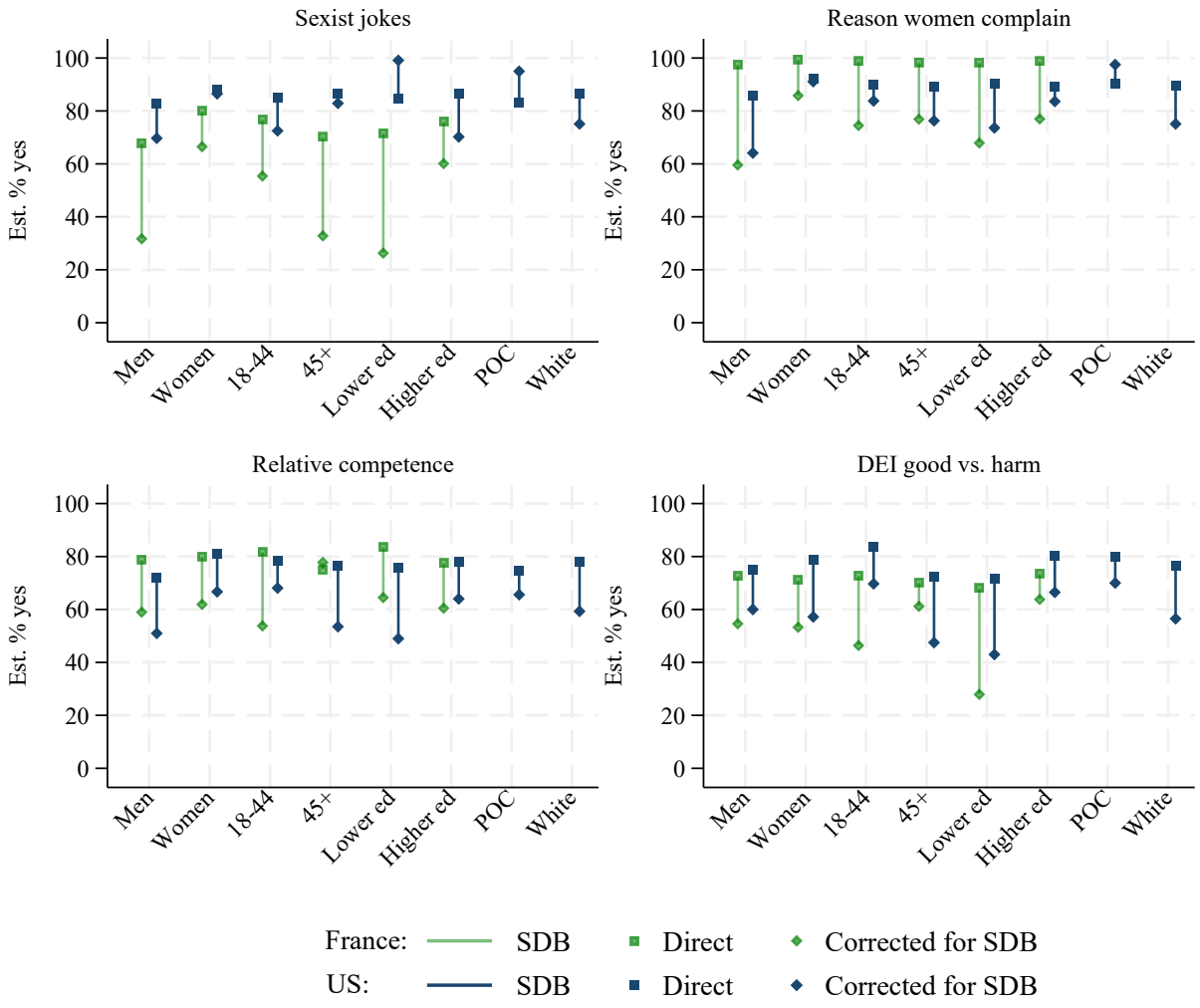
Figure 1. Raw Differences between Direct and Veiled Conditions in Total Items per List



Notes: A positive difference indicates social desirability bias. The items that are asked directly in the direct condition are the following.

1. Sexist jokes: "I dislike sexist jokes in the workplace".
2. Reason women complain: "There is good reason for women to complain about sexism and sexual harassment in the workplace".
3. Relative competence: "Nowadays, when a woman gets a promotion instead of a man, it is generally because she is more competent than the man".
4. DEI good vs. harm: "Diversity, equity, and inclusion policies at the workplace do more good than harm".
5. Placebo Netflix: "I currently have a Netflix subscription".

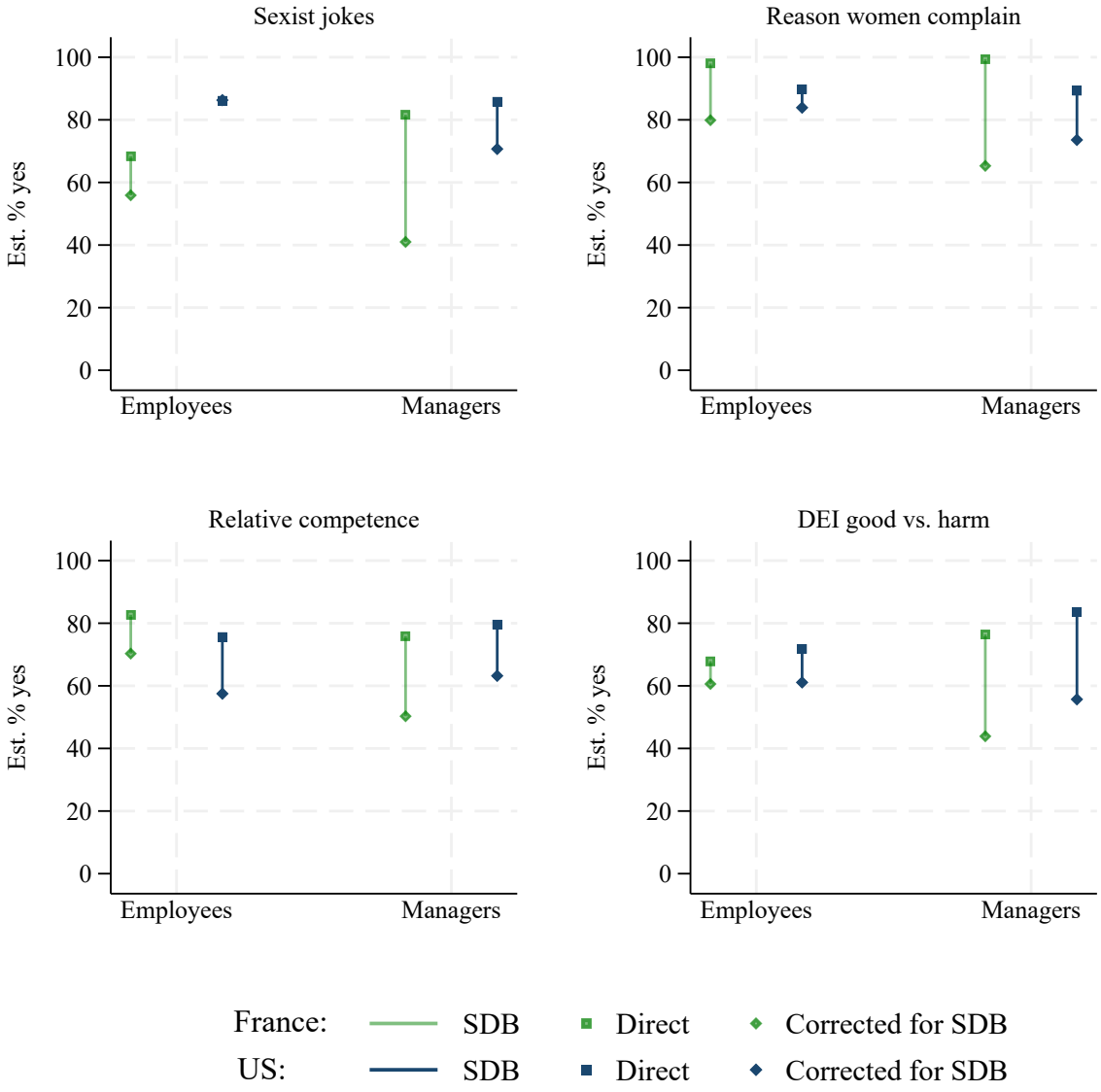
Figure 2. Social Desirability Bias by Group



Notes: 'Direct' indicates the fraction of respondents in the direct condition that agreed with the item. The SDB is estimated based on equation (1), as presented in Table A1. The items that are asked directly in the direct condition are the following.

1. Sexist jokes: "I dislike sexist jokes in the workplace".
2. Reason women complain: "There is good reason for women to complain about sexism and sexual harassment in the workplace".
3. Relative competence: "Nowadays, when a woman gets a promotion instead of a man, it is generally because she is more competent than the man".
4. DEI good vs. harm: "Diversity, equity, and inclusion policies at the workplace do more good than harm".

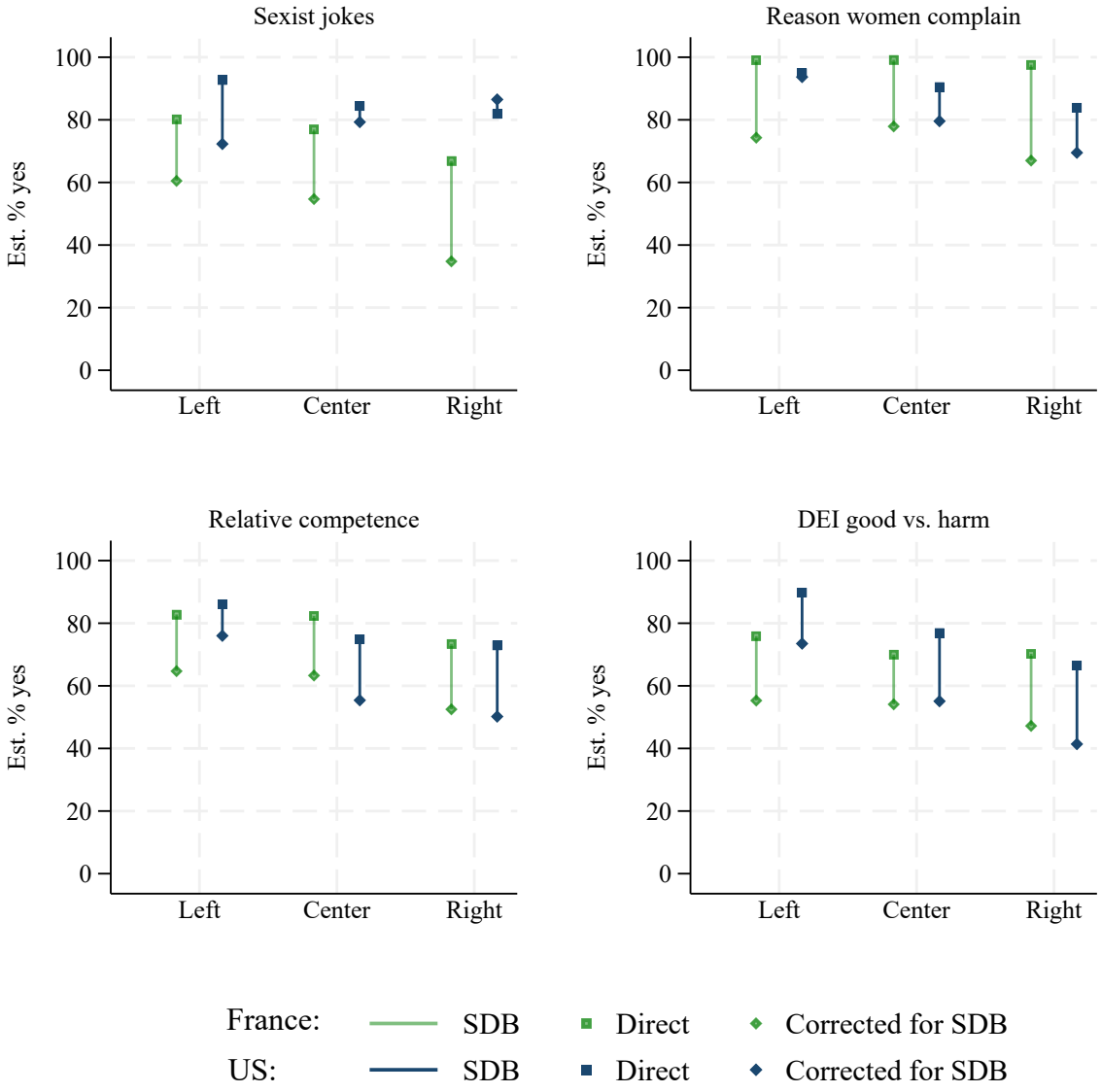
Figure 3. Social Desirability Bias among Managers



Notes: 'Direct' indicates the fraction of respondents in the direct condition that agreed with the item. The SDB is estimated based on equation (1), as presented in Table A1.

1. Sexist jokes: "I dislike sexist jokes in the workplace".
2. Reason women complain: "There is good reason for women to complain about sexism and sexual harassment in the workplace".
3. Relative competence: "Nowadays, when a woman gets a promotion instead of a man, it is generally because she is more competent than the man".
4. DEI good vs. harm: "Diversity, equity, and inclusion policies at the workplace do more good than harm".

Figure 4. Social Desirability Bias by Political Affiliation



Notes: 'Direct' indicates the fraction of respondents in the direct condition that agreed with the item. The SDB is estimated based on equation (1), as presented in Table A1.

1. Sexist jokes: "I dislike sexist jokes in the workplace".
2. Reason women complain: "There is good reason for women to complain about sexism and sexual harassment in the workplace".
3. Relative competence: "Nowadays, when a woman gets a promotion instead of a man, it is generally because she is more competent than the man".
4. DEI good vs. harm: "Diversity, equity, and inclusion policies at the workplace do more good than harm".

Table 1. Characteristics of Survey Respondents

	France		US	
	Mean	S.D.	Mean	S.D.
Gender: Female	0.54	0.50	0.58	0.49
Age: 18 to 44	0.66	0.47	0.43	0.50
Race:				
Asian			0.04	0.19
Black			0.07	0.25
Hispanic			0.06	0.23
White			0.82	0.39
Has children	0.64	0.48	0.64	0.48
Higher education	0.68	0.47	0.66	0.47
Management position	0.49	0.50	0.46	0.50
Employed part-time	0.10	0.30	0.12	0.32
Industry:				
ITC	0.22	0.42	0.26	0.44
Finance	0.59	0.49	0.44	0.50
Law	0.02	0.13	0.20	0.40
Consulting	0.12	0.32	0.08	0.27
Engineering	0.05	0.23	0.02	0.16
Political opinions:				
Right	0.35	0.48	0.31	0.46
Center	0.32	0.47	0.40	0.49
Left	0.32	0.47	0.28	0.45

Notes: N=703 for the French sample and N=921 for the US sample. Higher education includes respondents with at least a Bachelor's or equivalent degree.

Table 2. Randomization Test

	France			US		
	Mean		t-test	Mean		t-test
	Veiled	Direct	p-value	Veiled	Direct	p-value
Gender: Female	0.53	0.55	0.498	0.57	0.59	0.488
Age: 18 to 44	0.66	0.66	0.973	0.43	0.43	0.924
Race:						
Asian				0.03	0.05	0.195
Black				0.07	0.07	0.712
Hispanic				0.06	0.05	0.523
White				0.83	0.81	0.449
Has children	0.64	0.64	0.980	0.65	0.64	0.713
Bachelor's degree or higher	0.68	0.69	0.891	0.67	0.66	0.970
Management position	0.51	0.47	0.253	0.47	0.46	0.675
Employed part-time	0.11	0.09	0.477	0.11	0.12	0.471
Industry:						
Information, technology, communication	0.23	0.22	0.841	0.27	0.25	0.664
Finance	0.58	0.59	0.747	0.44	0.44	0.910
Law	0.02	0.02	0.690	0.19	0.21	0.426
Consulting	0.12	0.12	0.873	0.08	0.07	0.647
Engineering	0.06	0.05	0.739	0.02	0.03	0.755
Political opinions:						
Right	0.38	0.34	0.266	0.31	0.32	0.776
Center	0.29	0.35	0.107	0.40	0.41	0.927
Left	0.33	0.31	0.634	0.29	0.28	0.696

Notes: This table shows the share of respondents with different characteristics in the veiled and the direct conditions.

Table 3. The Effect of Veiled Condition on Reports of Sensitive Beliefs

	(1) Sexist jokes	(2) Reason women complain	(3) Relative competence	(4) DEI good vs. harm	(5) Placebo Netflix
Panel A: France					
% Direct condition	74.6	98.6	79.5	71.9	71.1
	(2.3)	(0.6)	(2.1)	(2.3)	(2.4)
Δ Veiled condition	-24.0***	-24.4***	-18.7***	-18.5***	-2.4
	(6.8)	(6.4)	(6.8)	(7.4)	(7.2)
Estimated true %	50.6	74.2	60.8	53.4	68.7
Panel B: US					
% Direct condition	85.9	89.6	77.4	77.2	69.9
	(1.6)	(1.4)	(1.9)	(1.9)	(2.1)
Δ Veiled condition	-6.7	-10.2*	-17.3***	-18.2***	-5.8
	(5.9)	(6.1)	(6.0)	(7.2)	(6.9)
Estimated true %	79.2	79.4	60.1	59.0	64.1

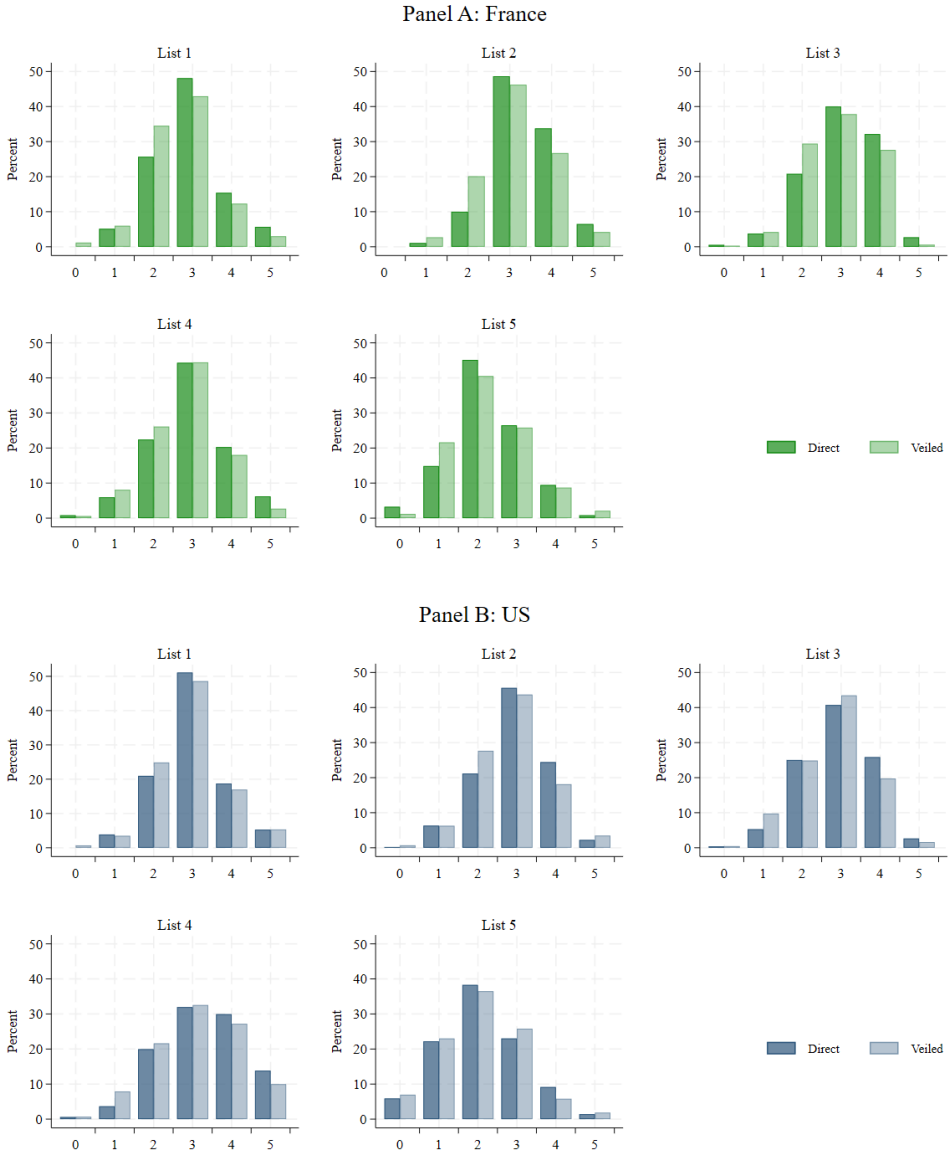
Notes: N=921 (491 in the direct condition) for the US sample and N=703 (370 in the direct condition) for the French sample. In the first four columns, we examine responses to four sensitive statements: (1) Sexist jokes: "I dislike sexist jokes in the workplace"; (2) Reason women complain: "There is good reason for women to complain about sexism and sexual harassment in the workplace"; (3) Relative competence: "Nowadays, when a woman gets a promotion instead of a man, it is generally because she is more competent than the man"; and (4) DEI good vs. harm: "Diversity, equity, and inclusion policies at the workplace do more good than harm". In the final column, we present the result for the placebo statement we included: (5) Placebo Netflix: "I currently have a Netflix subscription". % Direct condition is the sample mean of respondents in the direct condition. Δ Veiled condition is the coefficient on Veiled estimated using Equation 1 with the following controls: female, age (respondent is between 18 and 44 years old), race (US only), education, being a manager, having children, working part time, and industry and region fixed effects. Estimated true % is the sum of % Direct condition and Δ Veiled condition. Standard deviations are in parentheses for the Direct condition statistics. Heteroskedasticity-robust standard errors are in parentheses for regression results. Significance levels: *** p<0.01 and * p<0.1.

Table 4. The Effect of Veiled Condition and Interaction Terms

	(1) Sexist jokes	(2) Reason women complain	(3) Relative competence	(4) DEI good vs. harm
Panel A: France				
Veiled	-0.357***	-0.379***	-0.192*	-0.200*
	(0.100)	(0.100)	(0.0992)	(0.105)
Veiled × Female	0.214	0.250*	0.00908	0.0270
	(0.136)	(0.128)	(0.136)	(0.145)
Veiled	-0.316***	-0.237**	0.0122	-0.0597
	(0.118)	(0.116)	(0.115)	(0.124)
Veiled × Age:18 to 44	0.112	-0.0105	-0.299**	-0.189
	(0.145)	(0.138)	(0.141)	(0.156)
Veiled	-0.423***	-0.274**	-0.234*	-0.390***
	(0.128)	(0.115)	(0.126)	(0.132)
Veiled × Higher education	0.264*	0.0441	0.0685	0.299*
	(0.148)	(0.137)	(0.149)	(0.158)
Veiled	-0.121	-0.166*	-0.122	-0.0808
	(0.0897)	(0.0901)	(0.0960)	(0.105)
Veiled × Management	-0.246*	-0.159	-0.131	-0.212
	(0.136)	(0.128)	(0.135)	(0.147)
Panel B: US				
Veiled	-0.150	-0.224**	-0.192**	-0.144
	(0.0916)	(0.0991)	(0.0971)	(0.112)
Veiled × Female	0.143	0.209*	0.0324	-0.0654
	(0.120)	(0.126)	(0.124)	(0.146)
Veiled	-0.0219	-0.120	-0.232***	-0.239**
	(0.0758)	(0.0791)	(0.0818)	(0.0979)
Veiled × Age:18 to 44	-0.103	0.0424	0.137	0.129
	(0.121)	(0.125)	(0.122)	(0.143)
Veiled	0.151	0.0687	-0.104	-0.124
	(0.145)	(0.145)	(0.140)	(0.160)
Veiled × White	-0.265*	-0.208	-0.0833	-0.0710
	(0.159)	(0.161)	(0.155)	(0.179)
Veiled	0.131	-0.194*	-0.248**	-0.285**
	(0.112)	(0.108)	(0.102)	(0.125)
Veiled × Higher education	-0.297**	0.139	0.113	0.155
	(0.132)	(0.132)	(0.128)	(0.152)
Veiled	0.0122	-0.0574	-0.184**	-0.0994
	(0.0792)	(0.0785)	(0.0842)	(0.102)
Veiled × Management	-0.170	-0.0961	0.0238	-0.179
	(0.119)	(0.123)	(0.122)	(0.142)

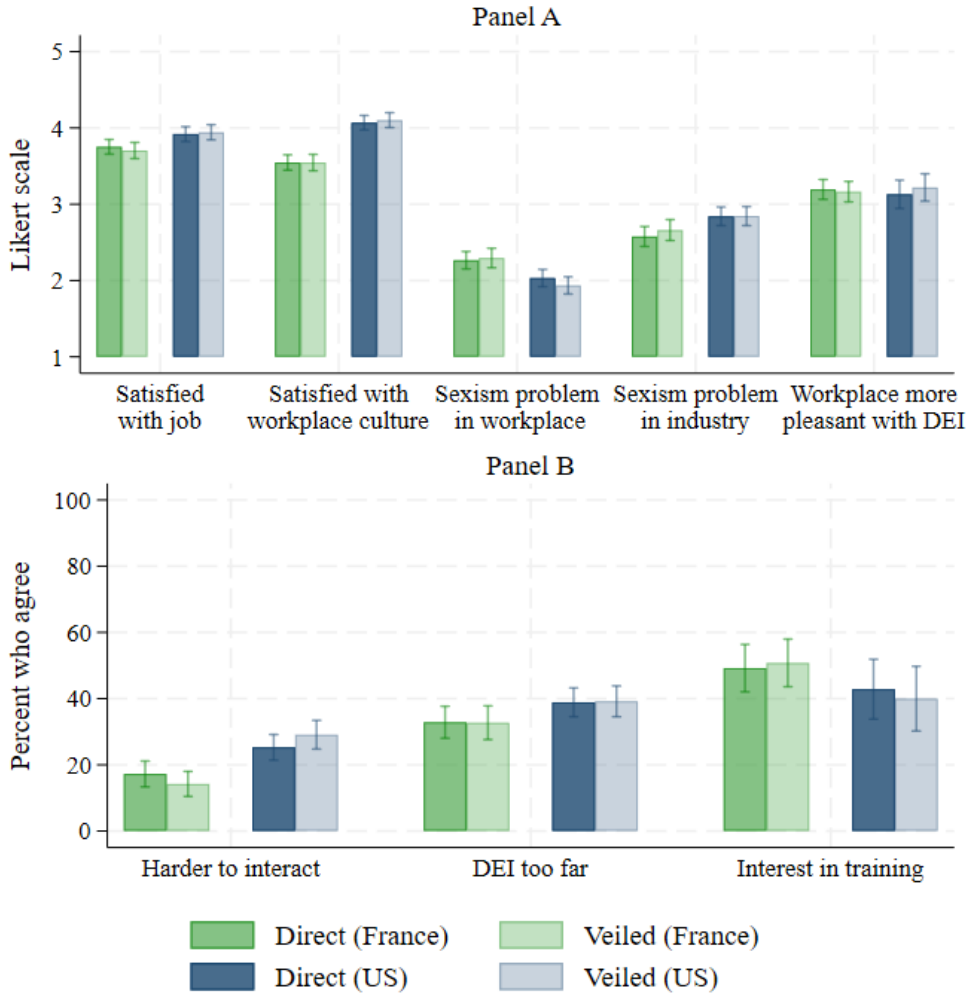
Notes: This table presents results of regressions including interaction terms between Veiled condition and different individual characteristics of respondents. We examine responses to four sensitive statements: (1) Sexist jokes: "I dislike sexist jokes in the workplace"; (2) Reason- women complain: "There is good reason for women to complain about sexism and sexual harassment in the workplace"; (3) Relative competence: "Nowadays, when a woman gets a promotion instead of a man, it is generally because she is more competent than the man"; and (4) DEI good vs. harm: "Diversity, equity, and inclusion policies at the workplace do more good than harm". Regressions include the following controls (not reported): female, age, race (US only), education, being a manager, having children, working part time, and industry and region fixed effects. Standard errors are in parentheses. Significance levels: *** p<0.01, ** p<0.05 and * p<0.1.

Figure A1. Distribution of agreement across lists



Notes: This figure plots the distribution of total number of statements with which respondents agreed, in the direct and veiled conditions, for each list and for each country. The y-axis shows the percentage of respondents who agreed with zero up to five statements.

Figure A2. Means of Perceptions of Workplace Culture, Sexism, and DEI Training, by Condition



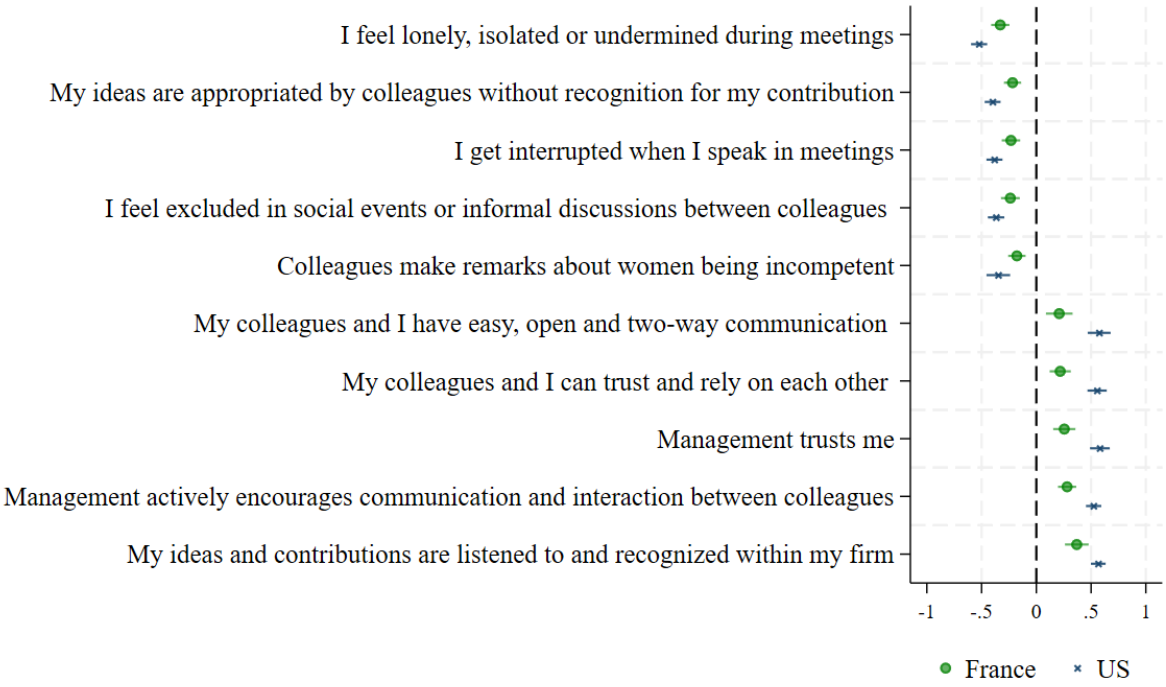
Notes: The questions on Panel A are on a five-point Likert scale:

- "How satisfied are you with your current job?" (Extremely dissatisfied to extremely satisfied.)
- "How satisfied are you with the workplace culture at your current job?" (Extremely dissatisfied to extremely satisfied.)
- "Sexism is a serious problem in my workplace." (Strongly disagree to strongly agree.)
- "Sexism is a serious problem in my industry." (Strongly disagree to strongly agree.)
- "If all my colleagues would get diversity, equity, and inclusion training, my workplace would be more pleasant for me." (Strongly disagree to strongly agree.)

The questions on Panel B are summarized as those who agree with the following statements:

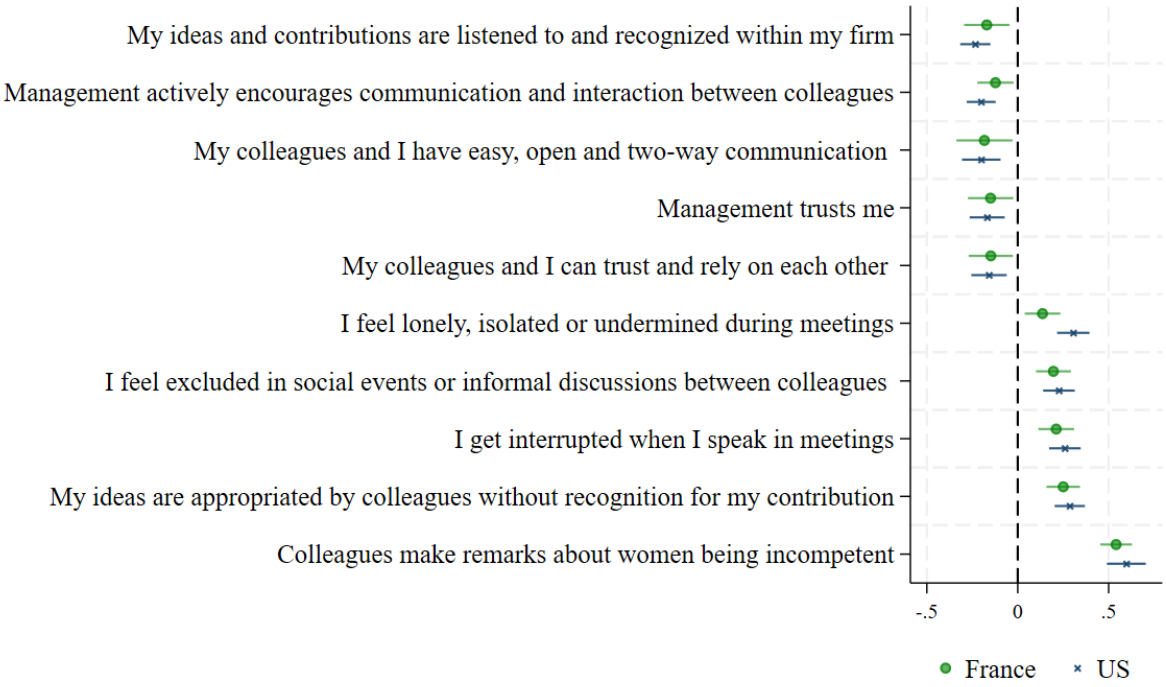
- "The increased focus on sexual harassment in the workplace has made it harder for men and women to know how to interact with each other in the workplace."
- "Actions taken in favor of diversity, equity, and inclusion in the workplace have gone too far".
- "Would you be interested in [DEI] training?" (if had not followed one yet).

Figure A3. Satisfied with workplace culture



Notes: This figure shows the results of a series of regressions correlating respondents' answers to the question "How satisfied are you with the workplace culture at your current job?" (five-point Likert scale) with their agreement with our ten measures of quality of communication and management (see questions 27 and 28 of our questionnaire in Section Appendix C). We run each regression separately. Regressions include the following controls (not reported): female, age, race (US only), education, being a manager, having children, working part time, and industry and region fixed effects. Horizontal lines represent 95% confidence intervals.

Figure A4. Sexism is a serious problem in my workplace



Notes: This figure shows the results of a series of regressions correlating respondents' answers to the question "Please indicate whether you agree or disagree with the following statement: Sexism is a serious problem in my workplace" (five-point Likert scale), with their agreement with our ten measures of quality of communication and management (see questions 27 and 28 of our questionnaire in Section Appendix C). We run each regression separately. Regressions include the following controls (not reported): female, age, race (US only), education, being a manager, having children, working part time, and industry and region fixed effects. Horizontal lines represent 95% confidence intervals.

Table A1. The Effect of Veiled Conditions on Reports of Sensitive Beliefs, by Subgroups

	France					US				
	N	Sexist jokes	Reason women complain	Relative competence	DEI good vs. harm	*N*	Sexist jokes	Reason women complain	Relative competence	DEI good vs. harm
Female	381	-13.5	-13.7	-18.1*	-18.0	538	-1.5	-1.1	-14.5*	-21.6**
		(9.6)	(8.8)	(9.8)	(10.9)		(7.9)	(7.5)	(7.6)	(10.0)
Male	322	-36.2***	-37.9***	-19.8**	-18.1*	383	-13.2	-21.9**	-20.9**	-14.9
		(10.0)	(9.2)	(10.2)	(10.0)		(10.1)	(9.4)	(9.4)	(11.4)
Age: 18-44	467	-21.4**	-24.3***	-27.9***	-26.4***	399	-12.4	-6.3	-10.2	-13.8
		(8.6)	(7.6)	(9.4)	(9.5.3)		(9.3)	(8.6)	(9.1)	(10.1)
Age: 45+	236	-37.4***	-21.5	-2.8	-9.0	522	-3.9	-12.9	-23.2***	-24.9***
		(11.4)	(13.2)	(13.6)	(13.1)		(8.0)	(8.4)	(7.8)	(9.3)
White						753	-11.5*	-14.3**	-18.8***	-20.1**
							(6.3)	(7.2)	(5.9)	(8.3)
POC						168	12.0	7.1	-8.9	-9.8
							(14.6)	(16.5)	(12.3)	(15.8)
Lower education	222	-45.3***	-30.3**	-19.1	-40.2***	309	14.2	-16.7	-26.8**	-28.5**
		(13.7)	(11.3)	(14.3)	(13.9)		(11.6)	(10.5)	(10.8)	(13.2)
Higher education	481	-15.9**	-21.8***	-17.1**	-9.8	612	-16.3**	-5.7	-14.2*	-13.5
		(7.8)	(8.1)	(7.3)	(9.0)		(7.1)	(7.5)	(7.5)	(8.7)
Manager	345	-40.6***	-34.1***	-25.6***	-32.5***	428	-15.1	-15.7*	-16.3*	-27.9***
		(9.5)	(9.6)	(11.8)	(10.0)		(9.8)	(9.4)	(8.7)	(8.6)
Employee	358	-12.5	-18.1**	-12.4	-7.2	493	0.3	-5.9	-18.1**	-10.7
		(8.5)	(8.2)	(8.2)	(11.6)		(7.2)	(8.8)	(7.9)	(10.3)
Left	226	-19.7*	-24.8***	-18.1*	-20.6*	262	-20.4**	-1.1	-10.2	-16.3
		(11.2)	(11.0)	(10.5)	(15.7)		(10.0)	(10.5)	(9.7)	(11.6)
Center	228	-22.3**	-21.3*	-19.0*	-15.9	372	-5.1	-10.9	-19.4**	-21.8**
		(10.8)	(11.2)	(10.9)	(14.1)		(8.1)	(9.5)	(8.8)	(11.0)
Right	249	-32.1***	-30.6**	-20.9*	-23.0*	287	4.6	-14.4	-22.7**	-25.1**
		(12.1)	(12.8)	(12.7)	(12.3)		(12.0)	(11.7)	(10.6)	(12.0)

Notes: This table presents the effect of the Veiled condition estimated using Equation 1 for each subgroup, including the following controls (not reported): female, age, race (US only), education, being a manager, having children, working part time, and industry and region fixed effects. We examine responses to four sensitive statements: (1) Sexist jokes: "I dislike sexist jokes in the workplace"; (2) Women complain: "There is good reason for women to complain about sexism and sexual harassment in the workplace"; (3) Relative competence: "Nowadays, when a woman gets a promotion instead of a man, it is generally because she is more competent than the man"; and (4) DEI good vs. harm: "Diversity, equity, and inclusion policies at the workplace do more good than harm". Standard errors are in parentheses. Significance levels: *** p<0.01, ** p<0.05 and * p<0.1.

Appendix B Sample description [Online Only]

This section describes how we collected the sample of respondents for the list experiment. In May 2022, we filed an application with the Internal Review Board (IRB) for experimental research at the Erasmus School of Economics. We obtained IRB approval on May 24th, 2022. In our IRB application, we specified that our target population was 2,200 respondents from France and the US. We planned a pilot study of 200 participants (100 in France and 100 in US), before launching our final data collection of 1,000 participants in each country. The purpose of the pilot study was to calibrate the attention checks and to test whether the items that we included in our lists led to floor or ceiling effects.

For the pilot study, the survey company collected information from 143 respondents for France and 102 respondents for the US. Some respondents (55 for France and 23 for the US) were screened-out because they did not fulfil the criteria to participate in the study (they were either self-employed, did not work in one of the five industries or they did not meet the education level requirement). We were therefore left with 88 respondents for France and 79 for the US. We had included a timer for the list questions, which enabled us to measure how much time respondents spent on each list. We noticed that some respondents spent so little time on some lists that it was clear that they were not properly reading the questions. The timer helped to eliminate the answers from respondents who are fast and inattentive. However, some respondents may be slow and inattentive (Read et al., 2022). We therefore added another attention check, which has been suggested in the economics literature by Haaland et al. (2023). The question is the following: "The next question is about the following problem. In questionnaires like ours, sometimes there are participants who do not carefully read the questions and just quickly click through the survey. This means that there are a lot of random answers which compromise the results of research studies. To show that you read our questions carefully, please enter turquoise as your answer to the following question."

At the end of the pilot study, and before the launch of the final study, we agreed with the survey agency that we would not take into account any respondent who did not meet the following attention criteria:

- at least seven seconds on five-item list questions;
- at least five seconds on four-item list questions;
- at least five minutes on the survey;
- and answered the favorite color question correctly.[16]

The agency agreed to replace respondents who did not meet these criteria, for free. The cost-per-incidence we agreed to was €3.57. Data collection for the final survey started on June 24th and ended on August 30th, 2022. At that point, the survey agency decided to stop data collection, because the replacement of respondents who did not meet the attention check criteria had become too costly for them. Indeed, only 32% of US respondents and 30% of French respondents passed

[16] We accepted any spelling close to "turquoise" as being a valid answer.

all tests (Table B1). Our final sample of respondents is therefore lower than intended. Indeed, our data set includes 703 respondents for France and 921 respondents for the US, which is below the 1,000 we had planned for each country.

Table B1. Percentage of respondents who fail the attention checks

Test	France	US
Five or seven seconds on lists	41.4%	36.4%
Five minutes on survey	25.5%	17.6%
Favorite color test	66.6%	62.8%
N	*2,361*	*2,895*

Note: Respondents who fail at least one test: 70.2% for France and 68.2% for the US. Many respondents fail several tests.

Appendix C Questionnaire [Online Only]

This survey is part of a research project conducted by a team of researchers at Erasmus University Rotterdam and Sciences Po.

The survey takes approximately 10 minutes to complete.

The survey asks your opinion on a variety of topics. Your answers are stored anonymously.

Thank you very much for your contribution to the advancement of scientific knowledge!

Section 1

In this first section, we ask several questions related to your personal characteristics.

1. What is your gender?

   ◯ Female
   ◯ Male
   ◯ Other

2. What is your age?

   ◯ 18-24 years old
   ◯ 25-34 years old
   ◯ 35-44 years old
   ◯ 45-54 years old
   ◯ 55-64 years old
   ◯ 65 years old or more

   Note: The following question was asked in the US, but not in France.

3. How would you describe yourself?

   ◯ American Indian or Alaska Native
   ◯ Asian
   ◯ Black or African American
   ◯ Hispanic, Latino or of Spanish Origin
   ◯ Native Hawaiian or Other Pacific Islander
   ◯ White
   ◯ Other

4. Do you have children?

   ◯ Yes
   ◯ No

   Note: The following question was asked to respondents who answered "Yes" to the children question.

5. How old is your youngest child?

   ◯ 0 to 18 months old
   ◯ 19 months to 3 years old
   ◯ 4 to 6 years old
   ◯ 7 to 12 years old
   ◯ 13 to 18 years old
   ◯ 19 or older

6. In which state do you currently reside? *Note: List of 50 states plus DC, Puerto Rico, and "I do not live in the United States" for the US. List of 13 regions for France and "I do not live in France".*
7. What is the highest level of education you've completed?
   - ◯ Less than high school degree
   - ◯ High school degree or equivalent (e.g. GED)
   - ◯ Some college, no degree
   - ◯ Associate degree (e.g. AA, AS)
   - ◯ Bachelor's degree (e.g. BA, BS)
   - ◯ Master's degree (e.g. MA, MS, MEd)
   - ◯ Professional degree (e.g. MD, DDS, DVM)
   - ◯ Doctorate (e.g. PhD, EdD)
8. What industry do you work for?
   - ◯ Information Technology
   - ◯ Finance *Includes banks, credit unions, financial advisors, tax and accounting firms, insurance companies, private equity and venture capital firms*
   - ◯ Law
   - ◯ Business Consultancy
   - ◯ Engineering
   - ◯ Other
9. What is your current employment status?
   - ◯ Employed full time (40 or more hours per week)
   - ◯ Employed part time (up to 39 hours per week)
   - ◯ Self-employed
   - ◯ Other
10. Do you hold a management position?
    - ◯ Yes
    - ◯ No

Section 2: List experiment questions

Veiled treatment

In this section we will present a series of five statements and ask you to tell us with how many of them you agree. We do not need to know which ones, just how many.

12. Below are five statements. Please tell us with how many of them you agree. We do not need to know which ones, just how many.
    - If something can go wrong for me, it will.
    - Last year, I was employed for at least six months.
    - The government should remove strict environmental regulations on businesses.
    - The government should finance projects to create new jobs.
    - I dislike sexist jokes in the workplace.

    Please indicate the total number of statements that you agree with.

0	1	2	3	4	5
◯	◯	◯	◯	◯	◯

13. Below are five statements. Please tell us with how many of them you agree. We do not need to know which ones, just how many.

    - Nuclear power stations are completely safe.
    - I am not a hunter.
    - The government should reduce income differences.
    - I am in favor of the death penalty.
    - There is good reason for women to complain about sexism and sexual harassment in the workplace.

    Please indicate the total number of statements that you agree with.

0	1	2	3	4	5
◯	◯	◯	◯	◯	◯

14. Below are five statements. Please tell us with how many of them you agree. We do not need to know which ones, just how many.

    - I am a regular user of Flickr.
    - I enjoy listening to music.
    - I would move within the US for work. *Note: In France, this statement was "I would move within France for work."*
    - I usually count on good things happening to me.
    - Nowadays, when a woman gets a promotion instead of a man, it is generally because she is more competent than the man.

    Please indicate the total number of statements that you agree with.

0	1	2	3	4	5
◯	◯	◯	◯	◯	◯

15. Below are five statements. Please tell us with how many of them you agree. We do not need to know which ones, just how many.

    - I hardly ever expect things to go my way.
    - I would accept a position with a lower pay.
    - I regularly spend time with my relatives.
    - I would accept a longer commute to get to work.
    - I currently have a Netflix subscription.

    Please indicate the total number of statements that you agree with.

0	1	2	3	4	5
◯	◯	◯	◯	◯	◯

16. Below are five statements. Please tell us with how many of them you agree. We do not need to know which ones, just how many.

    - Pesticides used in farming are dangerous.
    - Social benefits from the government encourage people to work more.
    - Marijuana should be legal.

- I enjoy some aspects about my job.
- Diversity, equity, and inclusion policies at the workplace do more good than harm.

Please indicate the total number of statements that you agree with.

0	1	2	3	4	5
○	○	○	○	○	○

Direct treatment

In this section we will present a series of four statements and ask you to tell us with how many of them you agree. We do not need to know which ones, just how many.

12. Below are four statements. Please tell us with how many of them you agree. We do not need to know which ones, just how many.

    - If something can go wrong for me, it will.
    - Last year, I was employed for at least six months.
    - The government should remove strict environmental regulations on businesses.
    - The government should finance projects to create new jobs.

    Please indicate the total number of statements that you agree with.

0	1	2	3	4
○	○	○	○	○

13. Do you agree with the following statement?
    I dislike sexist jokes in the workplace.

    - Yes
    - No

14. Below are four statements. Please tell us with how many of them you agree. We do not need to know which ones, just how many.

    - Nuclear power stations are completely safe.
    - I am not a hunter.
    - The government should reduce income differences.
    - I am in favor of the death penalty.

    Please indicate the total number of statements that you agree with.

0	1	2	3	4
○	○	○	○	○

15. Do you agree with the following statement?
    There is good reason for women to complain about sexism and sexual harassment in the workplace.

    - Yes
    - No

16. Below are four statements. Please tell us with how many of them you agree. We do not need to know which ones, just how many.

    - I am a regular user of Flickr.

- I enjoy listening to music.
- I would move within the US for work. *Note: In France, this statement was "I would move within France for work."*
- I usually count on good things happening to me.

Please indicate the total number of statements that you agree with.

0	1	2	3	4
◯	◯	◯	◯	◯

17. Do you agree with the following statement?
Nowadays, when a woman gets a promotion instead of a man, it is generally because she is more competent than the man.

    - Yes
    - No

18. Below are four statements. Please tell us with how many of them you agree. We do not need to know which ones, just how many.

    - I hardly ever expect things to go my way.
    - I would accept a position with a lower pay.
    - I regularly spend time with my relatives.
    - I would accept a longer commute to get to work.

    Please indicate the total number of statements that you agree with.

0	1	2	3	4
◯	◯	◯	◯	◯

19. Do you agree with the following statement?
I currently have a Netflix subscription.

    - Yes
    - No

20. Below are four statements. Please tell us with how many of them you agree. We do not need to know which ones, just how many.

    - Pesticides used in farming are dangerous.
    - Social benefits from the government encourage people to work more.
    - Marijuana should be legal.
    - I enjoy some aspects about my job.

    Please indicate the total number of statements that you agree with.

0	1	2	3	4
◯	◯	◯	◯	◯

21. Do you agree with the following statement?
Diversity, equity, and inclusion policies at the workplace do more good than harm.

    - Yes
    - No

Section 3: Attention test

The next question is about the following problem. In questionnaires like ours, sometimes there are participants who do not carefully read the questions and just quickly click through the survey. This means that there are a lot of random answers which compromise the results of research studies. To show that you read our questions carefully, please enter turquoise as your answer to the following question.

22. What is your favorite color? _____

Section 4: Workplace culture and job satisfaction

In the following section, we will ask you questions related to your work and workplace culture.

23. How likely are you to still be working in the same industry five years from now?

    ◯ Extremely unlikely
    ◯ Somewhat unlikely
    ◯ Neither likely nor unlikely
    ◯ Somewhat likely
    ◯ Extremely likely

24. How satisfied are you with your current job?

    ◯ Extremely dissatisfied
    ◯ Somewhat dissatisfied
    ◯ Neither satisfied nor dissatisfied
    ◯ Somewhat satisfied
    ◯ Extremely satisfied

25. How many direct colleagues do you have? *(Approximate number) Direct colleagues are those you interact with regularly, whether they are a member of your team or of another part of the firm.* _____

26. How many of your direct colleagues are men? *(Approximate number)* _____

27. Please indicate whether you agree or disagree with each of the following statements, which relate to how you perceive your interactions with your direct colleagues in your firm.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
My colleagues and I have easy, open and two-way communication	◯	◯	◯	◯	◯
My ideas and contributions are listened to and recognized within my firm	◯	◯	◯	◯	◯
My colleagues and I can trust and rely on each other	◯	◯	◯	◯	◯
Management actively encourages communication and interaction between colleagues	◯	◯	◯	◯	◯
Management trusts me	◯	◯	◯	◯	◯

28. How prevalent are the following situations within your work environment?

	This never happens	This rarely happens	This occasionally happens	This is common	This happens all the time
I feel lonely, isolated or undermined during meetings	◯	◯	◯	◯	◯
I feel excluded in social events or informal discussions between colleagues	◯	◯	◯	◯	◯
My ideas are appropriated by colleagues without recognition for my contribution	◯	◯	◯	◯	◯
I get interrupted when I speak in meetings	◯	◯	◯	◯	◯
Colleagues make remarks about women being incompetent	◯	◯	◯	◯	◯

29. How satisfied are you with the workplace culture at your current job?

- ◯ Extremely dissatisfied
- ◯ Somewhat dissatisfied
- ◯ Neither satisfied nor dissatisfied
- ◯ Somewhat satisfied
- ◯ Extremely satisfied

30. What share of your colleagues do you think enjoys sexist jokes? *Your estimate (%).*

0	10	20	30	40	50	60	70	80	90	100
◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯

31. Please indicate whether you agree or disagree with the following statement.
Sexism is a serious problem in my workplace.

- ◯ Strongly disagree
- ◯ Somewhat disagree
- ◯ Neither agree nor disagree
- ◯ Somewhat agree
- ◯ Strongly agree

32. Please indicate whether you agree or disagree with the following statement.
Sexism is a serious problem in my industry.

- ◯ Strongly disagree
- ◯ Somewhat disagree
- ◯ Neither agree nor disagree
- ◯ Somewhat agree
- ◯ Strongly agree

33. Do you think the increased focus on sexual harassment in the workplace has made it easier or harder for men and women to know how to interact with each other in the workplace?

- ◯ Has made it easier
- ◯ Has made it harder
- ◯ Hasn't made much difference

34. Do you think that actions taken in favor of diversity, equity, and inclusion in the workplace have gone too far or not far enough?

    ◯ Too far
    ◯ Not far enough
    ◯ Neither too far nor not far enough

35. Have you ever followed a training related to diversity, equity, and inclusion?

    ◯ Yes, it was mandatory
    ◯ Yes, it was voluntary
    ◯ No

36. Have you found this training useful? *Note: Question displayed to those who answered "Yes" to question 36.*

    ◯ Yes
    ◯ No

37. Would you be interested in such a training? *Note: Question displayed to those who answered "No" to question 36.*

    ◯ Yes
    ◯ No

38. Please indicate whether you agree or disagree with the following statement.
    If all my colleagues would get diversity, equity, and inclusion training, my workplace would be more pleasant for me.

    ◯ Strongly disagree
    ◯ Somewhat disagree
    ◯ Neither agree nor disagree
    ◯ Somewhat agree
    ◯ Strongly agree

Section 5

This final section includes three general questions.

39. In political matters, people talk of "the left" and "the right". Generally speaking, how would you place your views on this scale?

Left				Center					Right
1	2	3	4	5	6	7	8	9	10
◯	◯	◯	◯	◯	◯	◯	◯	◯	◯

40. Please do your best to answer the following question correctly. A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball.
    How much does the ball cost? _____
41. Was there anything unclear or confusing about this survey? _____