

Behavioral measures of bias-reducing, inclusive behaviors:

1) Mentorship

Two weeks to three months after participants complete the initial training, they will be invited by their employer to participate in a mentorship/connectivity program to foster inclusiveness in the workplace. Participants will be allowed to select up to five colleagues to connect with. We will compare the average number of women selected as mentees per participant in each condition as our primary DV. We will also examine the average number of non-white individuals selected as mentees per participant in each condition. We will use an intend-to-treat analysis where participants who do not choose to participate will be counted as nominating zero people.

2) Recognition for Awards

Two weeks to three months after participants complete the initial training, they will separately be invited to nominate colleagues for an excellence award. Participants will be allowed to nominate up to five colleagues. We will compare the average number of women nominated per participant in each condition as our primary DV. We will also examine the average number of non-white individuals nominated per participant in each condition. We will use an intend-to-treat analysis where participants who choose not to participate will be counted as nominating zero people.

3) Willingness to Provide Guidance to Others

One to five months after participants complete the initial training, participants will be asked to take time to speak with a new female or male employee to provide guidance about working at the company. Half of participants will be asked to speak with a female; half of participants will be asked to speak with a male. We will look at what percent of participants in each condition was willing to speak with a female and compare that to the percent willing to speak with a male.

Engagement with the issue of workplace bias

1) One to three weeks after completing the initial training, participants will receive an email from the research team asking if they would be willing to volunteer their time to complete a short survey that will help address inequalities that women and minorities face in the workplace. We will measure what percent of participants in each condition is will to start this survey and what percent of participants in each condition is willing to finish the survey.

2) Once a week following the training up to 12 weeks total, participants will be texted follow-up messages that will ask questions about stereotypes, bias, and inclusion and provide links to articles and videos about these topics. We will measure the average number of interactions per participant we receive across conditions (an interaction being a reply to a text or a click through to one of the articles or videos).

Attitudinal change regarding bias:

1) Modern Sexism.

We will use Swim et al. (1995) measure of Modern Sexism to see whether our intervention changed sexist attitudes.

2) Self-other perceptions of bias.

We assess the extent to which participants believe that they themselves exhibit gender and racial bias and stereotyping and to what extent they believe that the average person exhibits gender and

racial bias and stereotyping. We will look at how the self rating changes across conditions, how the other (i.e. average person) rating changes across conditions, as well as how the difference between the self rating and the other rating changes across conditions.

### 3) Situational Judgment Test.

We will seek to detect bias in participants' judgments via a situational judgment test to assess decision-making in bias-prone situations. In the situational judgment test, participants will be presented with different scenarios that can occur in the workplace. From a list of options, they will select how they would choose to behave in that situation. The situations described (and the response options) have been created based on interviews with employees of our partner organization to provide construct validity. Because the response options in this type of assessment are written to offer the participant multiple options – thus masking the socially desirable response – situational judgment tests are more difficult to fake than other personality assessments. We will use the situational judgment test to assess whether people choose behaviors that would interrupt or reinforce the influence of bias in different situations that commonly arise in the workplace. Scores on the situational judgment test are the sum of the scores of the individual items on the test with total scores ranging from -20 to 20.

### Analysis

Our main analytic strategy will be to compare participants between the treatment conditions and the control condition on each of these measures using t-tests and ordinary least squares regressions. As a baseline, we will do pairwise t-tests between conditions for each of the dependent variables. Next, to show robustness, we will run regressions to see the effects of the treatment conditions on the dependent variables including relevant control variables such as participant demographics (e.g., gender, race, office location). Finally, to see if there are differential effects based on demographics (specifically the gender of the participant and the office location of the participant), we will run OLS regressions with interaction terms to see how the intervention may vary in effectiveness between, for example, men and women or between participants based in the United States versus those based outside of the United States.

For measures that require nominating colleagues, we will include office location fixed effects and cluster standard errors by offices to account for the fact that levels of female representation and minority representation may vary depending on the office location.