# Supplemental Online Material

# Study 1

**Link to Anonymized Data and R Script (OSF)**

<https://osf.io/z8c9k/?view_only=d8b5901086d14bdeaf0c7a2eb68daabe>

**Creation of the Training and Testing Set (MBA Students’ Recalled Pro-Diversity Arguments)**

***Data Collection***

To build a set of reference instrumental and non-instrumental pro-diversity arguments, we recruited 394 participants at a European business school to complete a survey in the weeks prior to starting their MBA program (*Mage* = 28.25, *SD* = 2.39, range = [23; 37]; 240 men, 136 women, 3 third gender, 18 non-specified; 13 industries; 272 organizations). All students completed this survey in preparation for a mandatory diversity orientation workshop, and the sample size was determined by the number of participants who consented to their responses being used in academic research. Participants were first asked to recall their most recent employer prior to the MBA, and then informed as follows: “Different organizations use different types of arguments to justify a commitment to diversity (in terms of gender, race, sexual orientation), if they express one. In your most recent organization, can you describe what the top 3 arguments in favor of diversity were?” Participants listed up to 3 arguments used by their most recent employer in a free response forma.

***Hand-Labeling of the Arguments***

We recruited two independent coders – blind to the study’s hypotheses – for a first round of coding of the 1,182 free responses collected. In this first round, the coders identified pro-diversity arguments, coded as 1, vs. blank, incomplete, inarticulate or anti-diversity responses, coded as 0. The kappa in this first round was acceptable (percent agreement = 89.7%, kappa = 0.79). Disagreements were resolved in a reconciliation session in which one of the authors was present. The first round of coding found 503 pro-diversity arguments, provided by 220 participants.

In the second round of coding, we provided the same two coders with descriptions of business case and fairness case pro-diversity arguments, labeled as “Category A” and “Category B” to preserve coders’ blindness to the research question (see next section for coding guidelines). The descriptions provided to coders specified that an argument could only pertain to “Category A” if (1) it explicitly referred to at least one benefit that diversity would yield to the organization, and (2) it did not refer to any values/principles, or any diversity outcome, irrelevant to the organization’s interests. In contrast, an argument could only pertain to “Category B” if (1) it presented diversity as an intrinsic value, and (2) it did not explicitly refer to any benefits that the company could derive from supporting diversity. If an argument could not exclusively fit into either “Category A” or “Category B” because it had features relevant to both categories, then coders were to code the argument as “mixed”. After reading these descriptions, coders categorized the pro-diversity arguments into one of four categories: Category A (business case arguments only), Category B (fairness case arguments only), both categories (mixed arguments), or neither of them (arguments classified as “other”). Inter-rater agreement was acceptable (percent agreement = 76.2%, kappa = 0.55; Cohen, 1960), and disagreements were resolved in a reconciliation session attended by one of the authors.

***Results***

The first round of coding found that 220 participants collectively provided 503 pro-diversity arguments. Of the other 174 participants, 121 (i.e., 30.5% of participants) did not recall any type of argument used by their most recent employer to promote diversity (as indicated either by an explicit statement that their most recent employer did not use any such argument, or by a blank response), and 53 gave incomplete, inarticulate, or anti-diversity responses. Of the 503 pro-diversity arguments recorded, 66.2% were coded as business case, 28.0% as fairness case, 3.8% as mixed, and 2.0% as other (see Table S1 for examples). A one-sample t-test showed that business case arguments were significantly more prevalent than fairness case arguments in employees’ recollections of their former employers’ pro-diversity arguments (*M* = 0.66, *SD* = 0.47, *t*(502)=7.68, *p* < 0.001, 95% CI = [0.12, 0.20]).

**Verbatim Guidelines for the Second Round of Coding**

**Category A:**

Some comments will describe arguments used by organizations that focus on the direct or indirect **benefits of diversity for organizations' financial performance or market share**.

For example, claiming that diversity should be promoted because a diverse workforce yields:

* superior (financial) performance (*direct* benefit)
* more thought diversity, more creative thinking, or better decision-making (*indirect* benefit)
* greater team performance (*indirect* benefit)
* better rapport with specific customer segments (*indirect* benefit)
* access to a larger talent pool (*indirect* benefit)
* better corporate reputation among customers (*indirect* benefit), etc.

would qualify as comments that relate to the benefits of diversity for the organization’s financial performance or market share.

Arguments in this category refer to:

* the **instrumental benefits that organizations receive from diversity**
* and not to any values/principles, or any positive diversity outcomes that do not directly or indirectly benefit the organization.

**Category B:**

Some comments will describearguments used by organizations that focus on the **moral, fair and inclusive aspects of diversity**.

For example, claims that diversity should be promoted because organizations ought to:

* treat people fairly
* offer equal opportunities to all, regardless of gender, race, or sexual orientation
* include different people, and avoid excluding on the basis of group membership
* promote diversity because it is in line with their corporate values
* promote diversity because it makes people/employees feel better or happier

would qualify as comments that relate the moral, fair and inclusive aspects of diversity.

Arguments in this category refer to:

* the promotion of **diversity as an intrinsic value**
* and not to any instrumental benefits that the company could derive from behaving morally, fairly, or inclusively

**Important**

Arguments can also have elements of both categories, in which case you should enter a “1” in both categories.

Arguments can have none of the characteristics corresponding to Category A or B, in which case you should put a “0” in each column.

|  |  |  |
| --- | --- | --- |
| **Argument type** | **Argument sub-category** | **Original quote** |
| Instrumental – coded as “business case” | Diversity as instrumental to the financial bottom-line | “It’s better for the firm’s profits and bottom line.” |
| Diversity as instrumental to thought, creativity, decision-making | “Diversity of opinions allows to create/deliver a better product by introducing different ideas.” |
| Diversity as instrumental to talent recruitment | “More diverse companies are better able to win top talent.” |
| Diversity as instrumental to customer relationships | “Our clients are diverse, so I need to look like them to win business.” |
| Diversity as instrumental to corporate reputation | “It’s is good for branding.” |
| Non-instrumental – coded as “fairness case” | Diversity as motivated by equality and fairness (without any explicit link between equality and fairness and potential benefits for organizational performance) | “People deserve to be treated equality regardless of their sexual orientation, race, or gender.” |
| Diversity as motivated by inclusion (without any explicit link between inclusion and potential benefits for organizational performance) | “Cultivate an inclusive environment” |
| Diversity as motivated by corporate values (without any explicit link between corporate values and potential benefits for organizational performance) | “Combat the notion that engineering is not for women.” |
| Diversity as motivated by employees’ well-being (without any explicit link between employees’ well-being and potential benefits for organizational performance) | “Happier teams” |
| Mixed | Mix of instrumentality and non-instrumentality | “It’s the right thing to do and it’s good for our business” |
| Neither | Compliance to laws or norms | “Just to comply with diversity metrics” |

Table S1. *Study 1 Sample responses to the prompt: “In your most recent organization, can you describe what the top three arguments in favor of diversity were?”*

**Secondary Measures in the Survey Used to Collect MBAs’ Recalled Pro-Diversity Arguments**

***Participants’ Ratings of Type of Case***

After reporting up to three pro-diversity arguments, MBA students were asked to indicate the extent to which each argument fitted definitions of the business and fairness case. To facilitate participants’ understanding of these constructs, we provided them with definitions (adapted from Eagly, 2016) of the business case: “advocat[ing] in favor of diversity on the basis of its expected direct or indirect effects on the company's financial bottom-line” and of the fairness case: “advocat[ing] in favor of diversity on the basis of ethical and social justice principles.” Participants then rated on 2 separate items the extent to which each previously-listed argument fitted the business and the fairness case definitions, using a scale ranging from 1 “Not at all” to 7 “Extremely”.

***Inclusion Perceptions***

For each arguments that they had listed, participants completed 2 items measuring the extent to which they perceived each argument as conducive to inclusion in the workplace: “Please indicate how much you think this argument fosters an inclusive culture?” and “Please indicate how much you think this argument makes each employee feel valued?” (scale: 1 “Not at all” to 7 “Extremely”). These items (*r* = 0.57, *p* < 0.001) were aggregated to form a measure of inclusion perceptions.

***Performance Perceptions***

For each of the arguments that they had listed, participants also responded to the following item: “Please indicate how much you think this argument has a positive impact on performance?” (scale: 1 “Not at all” to 7 “Extremely”).

***Preferred Argument***

Participants who had listed at least two arguments were asked to imagine that they were leaders in a company, and to indicate which argument they would choose to make if they could only

use one argument to promote diversity among their employees

***Personal Case for Diversity***

Participants were asked to imagine that they were leaders in a company, and that they had the opportunity to promote diversity among their employees. Participants were asked to write their personal preferred arguments in favor of diversity.

## Machine Learning Procedure for Computerized Text Analysis

***Step 1: Training the Classifier***

**Creating the Training & Testing Sets.**

***Why Arguments Rather Than Organizational Diversity Cases for the Training Set?*** Let us first start by clarifying how arguments differ from diversity cases. “Diversity cases” here describe the text collected on the websites of real organizations (here, the Fortune 500), in which they justify why diversity matters to them in the form of several sentences, which together form a collection of long and complex paragraphs. In contrast, “arguments” refer to single sentences or “bullet point” phrases that MBA students listed upon being asked to report up to 3 separatearguments that they remembered their latest employer using to justify diversity (i.e., these were *not* MBA students’ personal arguments). We made the assumption that these recalled arguments would be accurate enough reflections the rhetoric used in the Fortune 500 companies’ diversity cases, because MBAs at this particular school mostly come from large corporate professional organizations (e.g., banks, consulting).

Bearing in mind that an algorithm’s classification performance is only as good as the training set from which it learns the characteristics of the classification categories, we opted to use arguments recalled by MBA students as the training set (rather than diversity cases) for two reasons. Our first goal was therefore to maximize the *quality* of the training set, that is, to ensure that the training set contained the purest possible examples of each category. Given this, there was a meaningful advantage to training the algorithm on the arguments recalled by MBAs rather than on full organizational diversity cases. The selected dataset of “arguments” indeed consists of separate units of text (single sentences or “bullet point” phrases) that all contain a single diversity justification – a “pure” instrumental or non-instrumental one (e.g., “It’s better for the firm’s profits and bottom line”, vs. “[Diversity] cultivate[s] an inclusive environment”). The advantage of using these arguments for the training set is that having the algorithm learn from these “pure” examples of the “business case” and “fairness case” increased the likelihood that it would accurately learnthe characteristics of each category, and thus subsequently perform well as a classifier. In contrast, using companies’ long and complex diversity cases as a training set would have provided the algorithm with a “noisier” training set to learn from, which would likely have decreased its subsequent performance as a classifier.

Our second goal was to maximize the *quantity* of examples of each type of case in the training set, and the quantity of fairness case examples in particular. We predicted that the fairness case would be less represented among the Fortune 500 companies’ diversity cases than the business case. We also expected the same to be true of companies outside the Fortune 500, thus again making organizational diversity cases a sub-optimal option for the training set. In contrast, the 141 fairness case arguments collected among MBAs allowed us to provide the algorithm with a much greater number of fairness case examples than might have possible to find naturalistically among organizations’ diversity cases, which allowed for better performance.

**Building the LASSO Classifier.** The LASSO procedure allows to keep in the logit model only the words that have the highest predictive power for the classification outcome, and forcing to zero the coefficients of the predictors with relatively lower predictive power (a process called “regularization”; see next section for further details). Therefore, a key advantage of the LASSO procedure relative to alternative approaches is that it allows for greater interpretability of which specific predictors drive the classification (in contrast to Ridge classification, which keeps myriads of predictors with close-to-zero coefficients in the model).

***The Regularization Process.*** The mechanism by which a LASSO classifier forces to zero the coefficients of those predictors with low predictive power (a process called “regularization”) is determined by a tuning parameter, lambda, whose value can vary from 0 to infinity. The choice of the value for lambda has a direct influence on the set of words that contribute to the classification outcome (e.g., “productive”, “creative”, “business” for the business case, versus “fair”, “inclusive”, “equal” for the fairness case; see Tables S2 and S3 for a complete list of words). This is because the LASSO coefficient estimates associated with each of these words (i.e., their predictive power) are determined by minimizing the sum of squared errors with a penalty for complexity, which depends on lambda:

with:

the probability for a given Argument i to be of a “fairness case” type

α the intercept

the predictive power of Word k to predict the “fairness case” category

*M* the number of arguments in the corpus

*N* the number of words in the entire corpus that are appear more than 5 times

λ the tuning hyperparameter

For each value of , a LASSO regression will thus produce a different set of coefficient estimates, . For example, when lambda is equal to 0, the coefficient estimates in the LASSO classifier are identical to least squares estimates, whereas when lambda tends to infinity, all coefficients in the LASSO classifier are constrained to zero, which yields a null model.

To select the “optimal” lambda, the LASSO procedure tests a wide range of values for lambda, selects the value of lambda that yields the best classification results on the training set, and builds a classifier using this specific value of lambda. The selection of this “optimal” lambda was done through a 10-fold cross-validation procedure.

***The 10-Fold Cross-Validation Procedure.*** Fitting a classifier on a subset of the MBA students’ arguments (the training set, rather than the entire sample of MBA arguments) could increase bias in the model, because the classifier has fewer cases to learn from. To minimize this risk, we implemented a 10-fold process of cross-validation on the training set for determining the lambda yielding the lowest cross-validation error rate.

For each given value of lambda tested by the LASSO classifier, the arguments in the training set were randomly divided into 10 folds, 9 of which were used to train a LASSO classifier (i.e., to build a model as described above based on the arguments contained in these folds), and 1 of which was used to validate it (i.e., to test the error rate of said LASSO classifier when classifying arguments in the remaining validation fold, which was done by comparing the classifier’s predicted label for each argument in this fold against its actual label provided by the two coders). For each value of lambda tested, the cross-validation procedure was thus performed 10 times iteratively, each time changing the fold of arguments that was used for cross-validation amongst the 10 folds available. The test error rates across all 10 iterations were then averaged into a cross-validation error rate, which was attributed to the value of lambda being tested. After running the 10-fold cross-validation procedure for all tested values of lambda, the computer determined the value of lambda that yielded the lowest average cross-validation error rate (in combination with the values of other hyperparameters over which we iterated through a grid search: weights, decision threshold, and proportion of MBA arguments assigned to the training set – see below).

***Grid search.*** Concomitantly to optimizing the lambda hyperparameter and the proportion of MBA arguments assigned to the training vs. testing set, we optimized two other hyperparameters: weights, and cut-off (or “decision threshold”), which aimed at correcting for the imbalance that characterized our training set. Indeed, 70% of MBA arguments in the training set pertained to the “business case” category, whereas only 30% pertained to the “fairness case” category (the same was true of the overall set of MBA arguments, i.e., the training and testing sets were stratified). This imbalance meant that the classifier would have fewer examples of the “fairness case” category to learn from in the training set relative to the “business case” category, which increased the risk that the classifier may fail to correctly identify “fairness case” arguments, and misclassify them as “business case” arguments (Fernández et al., 2018). To mitigate this risk, we first used weights to associate a greater “penalty” to the misclassification of “fairness case” (vs. “business case”) arguments, as advised in unbalanced contexts (Analytics Vidhya, 2020). In addition, we tuned the algorithm’s decision threshold for classifying arguments as “business” vs. “fairness case” – another practice recommended in unbalanced contexts (Provost, 2000). Indeed, a typical decision threshold (or “cut-off” value) to classify text instances into two categories based on their predicted probability is 0.5. However, in contexts where the training set is unbalanced (i.e., contain more instances of one category than the other), failing to adjust this standard threshold has been deemed to be “a critical mistake” in the machine learning field (Provost, 2000), because it fails to address the risk that the classifier may misclassify instances of the minority category (here, “fairness case” arguments) as pertaining to the majority category (here, “business case”; Fernández et al., 2018). Adjusting the cut-off has thus become the standard recommendation in the field (Provost, 2000), as it allows to optimize the algorithm’s classification performance by optimizing the trade-off between the its recall performance (the proportion of arguments labelled as “fairness case” by coders that is also classified as such by the algorithm) and its precision performance (the proportion of arguments classified as “fairness case” by the algorithm that were indeed labelled as “fairness case” by coders). This trade-off is captured by the F1-score metric, which consists in the harmonic mean of recall and precision performances. This measure of algorithmic performance is preferred to accuracy in unbalanced contexts (Lever et al., 2016), as it assesses the algorithm’s performance in light of its performance on the minority category, which is most at risk of misclassification in unbalanced contexts – here the “fairness case”.

We used a grid search to determine in a non-arbitrary way which values of weights and cut-off (as well as lambda and proportion of MBA arguments to assign to the training set) would maximize the classifier’s F1-score on the training set (i.e., yielded the best classification results, by minimizing the misclassification of “fairness case” arguments). The grid search revealed that the following values maximized the classifier’s F1-score on the training set: lambda = 0.0027, weights = 0.7, decision threshold = 0.75, proportion of MBA arguments in training set = 0.8. Finally, we built the classifier using this set of values for the hyperparameters. Tables S2 and S3 show the words selected as predictors in the logit model of the classifier (the coefficients of all other words in the text corpus were forced to zero by the lambda hyperparameter).

Table S2. *Study 1 Logit coefficients of the first LASSO classifier for the words detected as predictive of a business case for diversity (negative coefficients)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Words predictive of the business case** | **Logit coefficient (predictive power)** |  |  |  |  |
| set | -41.55761389 | nationality | -9.507314384 | promotes | -4.919229969 |
| productive | -25.41800185 | results | -9.410777851 | challenges | -4.762132878 |
| local | -25.41041165 | expand | -9.287411936 | benefit | -4.669676234 |
| encourage | -24.24460819 | market | -9.20523345 | group | -4.479102022 |
| future | -22.35550241 | impact | -9.02176564 | customer | -4.446445941 |
| grow | -21.82728212 | different | -8.982402606 | organization | -4.319267231 |
| hiring | -21.52418995 | complex | -8.96578393 | understand | -4.087245005 |
| beneficial | -21.09889053 | cultures | -8.83897495 | drives | -3.999698489 |
| creates | -20.09080606 | solutions | -8.676155206 | need | -3.777710385 |
| open | -20.07719064 | broader | -8.649401239 | help | -3.767495578 |
| make | -19.22406502 | views | -8.587174506 | population | -3.645296094 |
| stronger | -17.68306969 | view | -8.45627236 | challenge | -3.581912345 |
| client | -17.64818403 | perspectives | -8.220701682 | many | -3.483821933 |
| creating | -17.63799176 | learn | -8.062636137 | collaboration | -3.282209659 |
| passions | -16.98443383 | given | -7.978484276 | diversity | -3.212123217 |
| nationalities | -16.64852584 | recruitment | -7.961402914 | viewpoints | -3.092655307 |
| companies | -16.62231827 | thinking | -7.945336228 | firm | -3.04611599 |
| organisation | -16.00726169 | base | -7.793553669 | decisions | -2.915892079 |
| etc | -16.00623589 | experience | -7.769974693 | beliefs | -2.861643431 |
| greater | -15.91256893 | performance | -7.746310816 | top | -2.758674069 |
| quality | -15.04879738 | groups | -7.596022788 | respond | -2.57210698 |
| skill | -14.66400449 | economic | -7.439806458 | pool | -2.31074355 |
| talents | -14.38357213 | creative | -7.3747724 | across | -2.136594139 |
| talent | -14.21115704 | breadth | -7.322070454 | well | -1.897885334 |
| delivery | -14.1483696 | decision | -7.27448738 | team | -1.84254998 |
| retains | -13.88394232 | creativity | -7.143171791 | way | -1.601255414 |
| regardless | -13.85374921 | wider | -7.140309513 | effective | -1.445748736 |
| ways | -13.60217702 | advantage | -7.135578974 | communication | -1.327663862 |
| clients | -13.42706834 | teams | -7.063693703 | get | -1.304315618 |
| necessary | -13.19856933 | can | -7.046994929 | helps | -1.188025325 |
| dynamic | -12.63464817 | experiences | -7.005404288 | awareness | -1.102473592 |
| power | -12.5463941 | important | -6.901117686 | better | -0.917297783 |
| company | -12.45269158 | productivity | -6.795170646 | collective | -0.81116192 |
| makes | -12.32359724 | lead | -6.394152079 | approach | -0.695333201 |
| tech | -12.17597928 | increase | -6.162160739 | needed | -0.601534643 |
| skills | -12.02357198 | american | -6.117015006 | learning | -0.558964679 |
| management | -11.80720634 | diverse | -5.940043587 | attract | -0.546479369 |
| looking | -11.74216304 | therefore | -5.689626816 | fundamental | -0.409271046 |
| status | -11.71850637 | thought | -5.550273699 | choice | -0.205370172 |
| ideas | -11.2324636 | good | -5.483396533 | contributing | -0.106188191 |
| business | -11.21085838 | new | -5.440755168 | comes | -0.061938572 |
| best | -11.08507008 | attracting | -5.376100147 | variety | -0.036670927 |
| workforce | -10.45162878 | opinions | -5.339396585 | idea | -0.024076633 |
| global | -10.42782265 | differently | -5.32423567 | international | -0.022299997 |
| key | -10.36360098 | services | -5.283280667 | think | -0.000610756 |
| operate | -10.23800097 | able | -5.12106958 | priority | -6.85E-05 |
| industry | -9.846576616 | problem | -5.116170161 | behaviors | -5.87E-07 |
| society | -9.846362982 | enriches | -5.08753026 | enable | -8.16E-14 |
| every | -9.822333777 | bring | -5.034264668 | supply | -5.23E-14 |
| religion | -9.674807152 | innovation | -4.95835073 | take | -2.09E-15 |

Table S3. *Study 1 Logit coefficients of the first LASSO classifier for the words detected as predictive of a fairness case for diversity (positive coefficients)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Words predictive of the fairness case** | **Logit coefficient (predictive power)** |  |  |  |  |
| among | 28.70163794 | giving | 7.880525765 | inclusiveness | 3.198152602 |
| provide | 24.25696735 | committed | 7.880524876 | making | 2.981453859 |
| allow | 20.51039349 | ethical | 7.880501397 | discrimination | 2.626842029 |
| number | 18.82293495 | atmosphere | 7.809627359 | inclusivity | 2.626838191 |
| sexual | 16.01708344 | employee | 7.608808652 | create | 2.078390795 |
| come | 15.32290651 | equal | 7.076824575 | unique | 2.043862466 |
| us | 15.25255109 | right | 6.53960323 | gender | 1.99189968 |
| office | 12.21587919 | workplace | 6.070260463 | social | 1.689319871 |
| outside | 12.11770116 | equally | 5.695318863 | openness | 1.616123334 |
| internally | 11.42466667 | colleagues | 5.533670104 | provides | 1.210721395 |
| cultural | 11.11577124 | equality | 4.595304026 | place | 1.112777974 |
| women | 10.79917831 | equity | 4.016808897 | (Intercept) | 0.551273037 |
| capabilities | 10.5073742 | others | 4.010441931 | feels | 0.437890767 |
| shared | 10.50736475 | respect | 3.870081738 | genders | 0.307758117 |
| work | 10.22737898 | opportunities | 3.808429219 | orientations | 0.015763273 |
| fair | 9.787919442 | promote | 3.637560577 | feel | 1.38E-13 |
| ceo | 9.173851618 | everyone | 3.556801191 | educational | 7.89E-14 |
| inclusive | 9.13951349 | opportunity | 3.511215686 | leadership | 2.05E-14 |
| acceptance | 7.950873803 | inclusion | 3.446007928 | empowered | 8.54E-15 |
| unconscious | 7.88053005 | fairness | 3.270806698 | support | 6.15E-15 |
| personal | 7.880526835 | individual | 3.209875989 |  |  |

***Step 2: Testing the Classifier***

**Measures of Algorithm Performance.** Below, we present the definitions of the measures of algorithmic performance which we discuss in detail in the next section:

**Performance on the Testing Set – First LASSO Classifier.** Obtaining the predicted labels (0 for “business case”, 1 for “fairness case”) for all arguments in the testing set allowed us to assess the performance of the LASSO classifier on unseen data, by comparing the predicted labels to the actual labels provided by the two coders for arguments in the testing set (see Figure S1 for confusion matrix).

The classifier had a 78.4% F1-score, which is a satisfactory result for accuracy[[1]](#footnote-1) (James et al., 2017). To understand the F1-score (which focuses on the algorithm’s classification performance on the category most at risk of misclassification – the minority category, here the “fairness case”), it is useful to look at the algorithm’s recall vs. precision performance – two facets of algorithmic performance for between which a tradeoff exists, and which the F1-score captures. This algorithm had a higher precision (83%) than recall performance (74%). The 83% precision level suggests that among cases classified as “fairness cases” by the algorithm, only 17% were misclassified business cases – i.e., the set of cases identified by the algorithm as “fairness cases” was rather “pure”. However, due to the precision-recall tradeoff, this high precision performance came undermined recall performance. Specifically, the purity achieved in the set of cases classified as “fairness case” by the algorithm came at the expense of exhaustiveness in the algorithm’s detection of fairness cases in the testing set – only 74% of cases labeled as “fairness case” by the coders were also classified as “fairness cases” by the algorithm.

*Figure S1.* Confusion matrix for the first LASSO algorithm on the testing set. In green: correct predictions. In red: erroneous predictions.

|  |  |  |
| --- | --- | --- |
| **Actual labels attributed**  **by the human coders**  **Labels predicted**  **by the LASSO classifier** | **1**  **(Fairness case)** | **0**  **(Business case)** |
| **1**  **(Fairness case)** | **20** | **4** |
| **0**  **(Business case)** | **7** | **63** |

*Note.* F1-score: 78.4%

Precision: 20/(20+4) = 83%

Recall: 20/(20+7) = 74%

Overall test error rate: (7+4)/(20+63+4+7) = 11.7%

Miss rate for “fairness case” arguments (1-“recall”): 7/(7+20) = 26%

False alarm rate for “fairness case” arguments (1-“precision”): 4/(4+20) = 17%

Complementary insights into the algorithm’s performance come from signal detection theory (Macmillan & Creelman, 1991), which provides a framework to better understand the test error rate by decomposing it into a “miss” rate (or “false negative” rate), and a “false alarm” rate (or “false positive” rate). We found that the classifier had a higher “miss” rate (26%) than “false alarm” rate (17%) when detecting fairness case arguments in the testing set, suggesting that the classifier may also have a tendency to underestimate the number of fairness cases (and thus to overestimate the number of business cases) represented among the Fortune 500 diversity cases[[2]](#footnote-2). To test the robustness of these results, we therefore built and used a second LASSO classifier.

**Robustness checks – Second LASSO Classifier.** To assess the robustness of the results obtained above with the first LASSO classifier, we built a second LASSO classifier using a well-established, alternative classification approach (James et al., 2017), whereby the classifier uses word stems (e.g., “perform”) rather than entire words (e.g., “performance”, “performing”, “performed”) as independent variables in the logit model. The methodology used to build this second classifier was identical to that of our first classifier – that is, we used a grid search to determine the hyperparameters that would maximize its performance in the cross-validation phase, and accordingly built the classifier using those hyperparameters (lambda = 0.0062, weights = 0.7, decision threshold = 0.55, proportion of MBA arguments in training set = 0.6). Tables S4 and S5 show the word stems selected as predictors in the logit model of the classifier (the coefficients of all other word stems in the text corpus were forced to zero through the lambda hyperparameter).

This second classifier had a 75.2% F1-score on the testing set (see Figure S2 for confusion matrix). In contrast to the first algorithm, this second algorithm had a higher recall (80%) than precision performance (71%). The 80% recall level suggests that among those cases in the testing set that had been labeled as “fairness case” by the coders, 80% were also detected as “fairness cases” by the algorithm (an increase relative to the first algorithm’s recall performance). However, due to the precision-recall tradeoff, this high recall performance undermined precision performance (71%). Specifically, improved exhaustiveness in the algorithm’s detection of fairness cases was achieved at the expense of “purity” in the set of cases classified as “fairness cases” by the algorithm – among those, 29% were in fact misclassified business cases (a deterioration relative to the first algorithm). In sum, the set of cases classified by the second algorithm as “fairness cases” was less “pure” than the first classifier’s set (i.e., it contained more misclassified business cases), but it *also* included a greater proportion of the fairness cases actually present in the testing set.

*Figure S2.* Confusion matrix for the second Lasso algorithm on the testing set. In green: correct predictions. In red: erroneous predictions.

|  |  |  |
| --- | --- | --- |
| **Actual labels attributed**  **by the human coders**  **Labels predicted**  **by the LASSO classifier** | **1**  **(Fairness case)** | **0**  **(Business case)** |
| **1**  **(Fairness case)** | **44** | **18** |
| **0**  **(Business case)** | **11** | **116** |

*Note.* F1-score: 75.2%

Precision: 44/(44+18) = 71%

Recall: 44/(44+11) = 80%

Overall test error rate: (11+18)/(44+116+11+18) = 15.3%

Miss rate for “fairness case” arguments (1-“recall”): 11/(11+44) = 20%

False alarm rate for “fairness case” arguments (1-“precision”): 18/(18+44) = 29%

Decomposing the test error rate using the framework provided by signal detection theory provides complementary insights into the performance of the second classifier. We found that in contrast to the first classifier, this second classifier had a higher “false alarm” rate (29%) than “miss” rate (20%) when detecting fairness case arguments in the testing set, suggesting that the classifier may also have the tendency to overestimate the number of fairness cases (and thus to underestimate the number of business cases) represented among the Fortune 500 diversity cases[[3]](#footnote-3). In other words, given the first classifier underestimates the number of fairness cases among the Fortune 500 diversity cases whereas the second overestimates it (and vice versa for the number of business cases), considering the results of both classifiers in tandem provides a lower and upper bound for the representation of each type of cases among the Fortune 500 diversity cases.

***Results.*** We ran the second classifier on the Fortune 500 diversity cases. Setting aside the 90 companies of the Fortune 500 companies (18%) that did not have any organizational diversity case, this second classifier estimated that among the 410 remaining organizational diversity cases collected, 388 (i.e., about 78% of the Fortune 500) pertained to the business case category, whereas 22 (under 5% of the Fortune 500) pertained to the fairness case category. The order of magnitude of these estimated proportions is consistent with those of the first classifier, thus corroborating the conclusion that the world’s largest corporations overwhelmingly appeal to the business case (rather than fairness case) to convey their commitment to diversity. Interestingly, this second classifier offers an upper bound for the prevalence of the fairness case (resp., a lower bound for the prevalence of the business case), whereas the first classifier offers a lower bound for the prevalence of the fairness case (resp., an upper bound for the prevalence of the business case). We thus conclude that 78 to 80% of the Fortune 500 make the business case for diversity, whereas 1 to 5% only make the fairness case.

Table S4. *Study 1 Logit coefficients of the second LASSO classifier for the word stems detected as predictive of a business case for diversity (negative coefficients)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Word stems predictive of the business case** | **Logit coefficient (predictive power)** |  |  |  |  |
| qualiti | -21.43614315 | nation | -6.691870512 | opinion | -3.451498492 |
| take | -21.42468155 | result | -6.531679398 | set | -2.839961338 |
| incorpor | -14.25693218 | wider | -6.507347057 | intern | -2.834119399 |
| critic | -14.24187226 | innov | -6.477102381 | design | -2.777345911 |
| client | -13.84926273 | differ | -6.426193146 | execut | -2.766569843 |
| best | -13.42281477 | solut | -6.394828621 | group | -2.764719587 |
| compani | -13.40687503 | communiti | -6.310417122 | key | -2.633695013 |
| futur | -13.25387729 | custom | -6.221674788 | collabor | -2.619473644 |
| can | -11.4806985 | experi | -6.028549187 | enhanc | -2.472844816 |
| busi | -11.37583015 | think | -5.886364392 | enabl | -2.416869043 |
| talent | -11.28606832 | approach | -5.798415673 | expand | -2.219543912 |
| allow | -11.18866407 | perspect | -5.692945414 | status | -2.208844201 |
| make | -10.10529491 | base | -5.618392735 | abl | -2.058802016 |
| encourag | -9.992143525 | brand | -5.505278476 | recruit | -1.831210135 |
| etc | -9.893366564 | advantag | -5.483314597 | employe | -1.453668265 |
| idea | -9.807930696 | core | -5.314173492 | thought | -1.31618608 |
| repres | -9.563783533 | unit | -5.239312471 | abil | -1.053906618 |
| manag | -9.531902094 | peopl | -4.976898225 | get | -0.899573428 |
| skill | -9.491227327 | import | -4.930504298 | broader | -0.801551607 |
| thrive | -9.361406449 | polici | -4.758320609 | member | -0.55375209 |
| workforc | -9.342512049 | product | -4.742760749 | help | -0.552084147 |
| deliveri | -9.169607257 | perform | -4.72599481 | competit | -0.491523398 |
| market | -8.541259312 | dynam | -4.241502667 | tech | -0.45547146 |
| learn | -8.377593346 | accept | -4.168535965 | expertis | -0.392288944 |
| industri | -7.859193082 | attract | -4.068784544 | increas | -0.350966021 |
| understand | -7.497826281 | viewpoint | -4.036094572 | strength | -0.076065438 |
| organis | -7.402292287 | way | -3.8470426 | local | -2.81E-13 |
| creativ | -7.361956714 | new | -3.726343783 | leader | -2.11E-13 |
| view | -7.30536551 | sourc | -3.72441511 | limit | -1.41E-13 |
| team | -7.020955606 | like | -3.64572472 | compon | -1.01E-13 |
| benefici | -6.957481457 | top | -3.590248122 | american | -5.21E-14 |
| global | -6.926955508 | lead | -3.564303501 | live | -3.72E-14 |
| problem | -6.869851092 | divers | -3.520299667 | globe | -2.25E-15 |

Table S5. *Study 1 Logit coefficients of the second LASSO classifier for the word stems detected as predictive of a fairness case for diversity (positive coefficients)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Word stems predictive of the fairness case** | **Logit coefficient (predictive power)** |  |  |  |  |
| outsid | 35.16700765 | equal | 4.348346878 | potenti | 1.835565064 |
| treat | 16.63978178 | atmospher | 4.14283877 | social | 1.476454166 |
| women | 9.927795084 | hire | 4.127098323 | promot | 1.16162958 |
| strive | 9.140253466 | inclus | 4.10179111 | organ | 0.889730882 |
| aim | 8.509075511 | us | 3.913832654 | leadership | 0.83120173 |
| ensur | 8.067332439 | welcom | 3.440315951 | environ | 0.790778148 |
| day | 7.135104699 | fair | 3.365300655 | offic | 0.744905193 |
| overal | 7.12376633 | candid | 3.166981213 | (Intercept) | 0.656777704 |
| capabl | 6.445750879 | individu | 2.997768879 | uniqu | 0.560301851 |
| ceo | 6.325908755 | provid | 2.802637062 | cultiv | 0.444560193 |
| around | 6.302029327 | discrimin | 2.754811194 | lgbtq | 0.438043095 |
| opportun | 5.740928067 | work | 2.554065875 | everyon | 0.135923276 |
| interest | 5.695667533 | respect | 2.442158818 | role | 0.006788013 |
| institut | 5.492116285 | balanc | 2.424598826 | support | 1.07E-13 |
| ethic | 5.33722716 | open | 2.158966173 | educ | 7.98E-14 |
| unconsci | 5.337116772 | equiti | 2.080156016 | bias | 4.76E-14 |
| right | 4.684294111 | lgbt | 2.003858321 | number | 3.88E-14 |

**Examples of Fortune 500 Diversity Cases Classified as Instrumental (“Business Case”) Versus Non-Instrumental (“Fairness Case”) by the LASSO Classifiers**

In bold are the phrases most characteristic of each type of organizational diversity case.

***Examples of an Instrumental Diversity Case: JP Morgan, Levi Strauss, and Intel***

“Employees are our greatest **asset**, and we strive to attract talent from the broadest pool **to foster innovation, creativity and productivity**. There is **tremendous power** that results from this kind of diversity. In fact, creating a diverse and inclusive environment is **critical to our success**, and we are deeply committed to hiring and retaining employees from different backgrounds, experiences and locations. Diversity brings together people with unique perspectives, and inclusion creates opportunities for all individuals to contribute and work together to **achieve success** as a whole. We believe working in an inclusive environment **motivates exceptional effort**, or—put more simply—it **makes us all better at what we do**. Our diverse workforce **helps our customers and business partners achieve their business goals**. By recruiting the highest quality people who **reflect the customers and communities that we serve**, we **increase our ability to deliver the best possible solutions**.”

“We know that **diverse teams make** **stronger companies** - and we encourage you to be yourself, add your voice and know that we **celebrate everything that you bring to the table**. Valuing diverse perspectives has been **integral to our success**. While we know that creating pay equity is simply the right thing to do, research shows **workplace diversity is also good for business**. We consistently strive to bridge pay gaps in all our employee markets.”

“Recognizing that diversity and inclusion are **critical to our success** is just one step. It's our action and efforts that lift goals into reality.”

***Examples of a Non-Instrumental Diversity Case: Kohl’s, Olin, and Tenet Health***

“At Kohl's we are committed to a **fair** and **equitable** workplace where everyone is **respected** as a valued member of the team. This extends to our suppliers as well as our Associates. As an extension of our **values**, Kohl's aspires to provide **equal access** to business opportunities to a diverse supply base, including minority-owned, women-owned, veteran-owned, LGBT-owned, and other diversity suppliers. This **inclusive** strategy enables us to **support** our customers, Associates and community.”

“Olin provides **equal opportunities** to employees and ensures the ongoing **safety** and **livelihood** of our people and communities. Our robust employee learning and development programs further build our culture of **safety**, **integrity** and **inclusion**.”

“At Tenet, we believe in a diverse and inclusive environment, one that is grounded in our dedication to the health and **well-being of all people**. **Respecting, nurturing and encouraging diversity** of thought, background and experience contribute to **positive work environments** that result in exceptional patient care. Tenet’s **equal opportunity** employment and employee development programs embrace the unique characteristics of our people and our communities and **encourage continuous individual improvement**. We embrace the diversity of our co-workers, physicians, suppliers and patients. **We do not harass or discriminate** on the basis of race, ethnicity, religion, gender, sexual orientation, gender identity or expression, national origin, age, disability, veteran status or any other characteristic protected by law. We embrace diversity **because it's our culture, and because it's the right thing to do**.”

# Study 2

**Link to Anonymized Pre-Registration on the Open Science Framework (OSF)**

<https://osf.io/b7ean/?view_only=19c3dd1bdbfa4ac781d50818bf036bb2>

We inadvertently forgot to pre-register online that we would exclude heterosexual allies from analyses. This exclusion criterion was inadvertently omitted at the time of pre-registration. However, given that the goal of Study 2 was to investigate the consequences of the business (vs. fairness) case for diversity among LGBTQ+ individuals, our decision to nonetheless apply this exclusion criterion cannot be deemed arbitrary or data-driven.

**Link to Anonymized Study Materials, Data, and SPSS Syntax (OSF)**

<https://osf.io/7657d/?view_only=273d93aa1f094690885f4713d8df9620>

**Verification of Demographic Balance Across Conditions**

Our lab’s policy is to verify that the demographics measured in experiments are balanced across conditions, to ensure that the core assumption of baseline comparability across cells in experimental designs is valid (Fives et al., 2013). We found no significant different across conditions for any of the demographic variables measured in this study.

**Additional Results Not Reported in the Main Text**

***Analyses of the Main Effect of Condition on Attraction to the Organization***

We explored the main effect of the manipulation on Attraction to the organization using an independent-sample t-test. The results are reported in Table S6 below.

Table S6. *Study 2 Descriptive Statistics for, and Results of the T-Test on Attraction to the Organization.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Outcome variable** | **Predictor** | **BC**  ***M***  **(*SD*)** | **FC**  ***M***  **(*SD*)** | ***t*** | ***p*** | ***d*** | **95% CI** | |
|  | **Lower bound** | **Upper bound** |
| Attraction to the organization | Condition | 3.71  (0.94) | 3.90  (0.78) | 1.31 | 0.19 | 0.21 | -0.11 | 0.53 |

*Note. N* = 151*.* Degrees of freedom were equal to 149.

**Secondary Measures**

***Secondary Outcome***

**Anticipated Performance.** Participants also indicated the extent to which they felt they would be able to perform well in this organization by responding to a single item: “I feel I would be able to perform well in this organization”, using a scale ranging from 1 “Strongly disagree” to 6 “Strongly agree”.

***Alternative Process***

**Stigma Consciousness.** Participants additionally indicated the extent to which they anticipated being aware of the stigma and stereotypes attached to their sexual-orientation minority identity when interacting with others in the organization by responding to one item adapted from Pinel's (1999) Stigma Consciousness Scale: “When interacting with people in this organization, I feel like they would interpret all my behaviors in terms of the fact that I am an LGBT+ individual”, on a scale ranging from 1 “Strongly disagree” to 6 “Strongly agree”.

**Secondary Results**

***Secondary Outcome***

We explored potential indirect effects of type of organizational diversity case (X) via anticipated sense of belonging (M) on the two additional outcomes (Y) (Model 4 in Hayes 2013). To test this, we effect-coded type of organizational diversity case (Fairness case condition = -1; Business case condition = 1). The coefficients reported below are indirect effects and their bias-corrected, bootstrapped 95% CIs, computed with 10,000 resamples using the PROCESS macro (Hayes, 2013).

**Anticipated Performance.** We found a significant indirect effect of type of diversity case on anticipated ability to perform well in the prospective organization, through anticipated sense of belonging (*b* = -0.069, *SE* = 0.037, 95% CI [-0.149; -0.006]). That is, LBGT+ participants exposed to a business (versus fairness) diversity case reported significantly lower anticipated sense of belonging to the organization making this case, which in turn was associated with feeling significantly less able to perform well in the organization.

***Alternative Process***

**Stigma Consciousness.** We explored the possibility that stigma consciousness would drive the detrimental effects of the business case for diversity on anticipated sense of belonging among underrepresented group members. We tested for an indirect effect of type of organizational diversity case (X), via stigma consciousness (M) on anticipated sense of belonging (Y) (Model 4 in Hayes 2013). We effect-coded Type of organizational diversity case (Fairness case condition = -1; Business case condition = 1). We computed the indirect effect and its bias-corrected, bootstrapped 95% CIs, computed with 10,000 resamples using the PROCESS macro in SPSS (Hayes, 2013). The indirect effect of type of diversity case through stigma consciousness on anticipated sense of belonging was not significant (*b* = -0.01, *SE* = 0.01, 95% CI [-0.042; 0.010]).

# Study 3

**Link to Anonymized Pre-Registration on the Open Science Framework (OSF)**

<https://osf.io/wdsma/?view_only=883f3f85d4224025bbb12a8f654e69a7>

**Link to Anonymized Study Materials, Data, SPSS Syntax, and R Script (OSF)**

https://osf.io/awmey/?view\_only=1554369df02a47bf802c1434827b0949

**Full Scale for Social Identity Threat (*α* = 0.95; adapted from Cohen & Garcia, 2005, and Rattan et al., 2018)**

1. How much would you worry that people in this company might draw conclusions about you based on gender stereotypes (\*)?
2. How much do you think gender stereotypes (\*) would affect your peers’ impressions of your ability?
3. How much do you think you might face biased evaluations from managers in this company because of your gender (\*)?

Scale: 1 “Not at all” to 6 “Extremely”

(\*) “gender stereotypes” was replaced by “racial stereotypes” in Study 5, and by “racial and/or gender stereotypes” in Study 6; “your gender” was replaced by “your race” in Study 5, and by “your race and/or gender” in Study 6.

**Confirmatory Factor Analysis**

To test whether the hypothesized process (anticipated sense of belonging) and the outcome variables (Attraction to the organization, Desire to join the firm) represent distinct constructs, we conducted a confirmatory factor analysis (CFA). A five-factor model fitted the data best (RMSEA = 0.08, CFI = 0.95, SRMR = 0.08, χ2(179) = 570.95). The factors were: attraction to the organization, desire to join the firm, anticipated sense of belonging – membership, anticipated sense of belonging – acceptance, anticipated sense of belonging – rejection. The fit was significantly better than a one-factor model (RMSEA = 0.23, CFI = 0.56, SRMR = 0.19, χ2(189) = 3796.05), a two-factor model with attraction to the organization and desire to join the firm on one factor, and all facets of anticipated sense of belonging on a second factor (RMSEA = 0.22, CFI = 0.58, SRMR = 0.18, χ2(188) = 3570.41), a three-factor model with attraction to the organization, desire to join the firm, and all facets of anticipated sense of belonging on one factor (RMSEA = 0.18, CFI = 0.71, SRMR = 0.17, χ2(186) = 2528.37), or a four-factor model with attraction to the organization, desire to join the firm, anticipated sense of belonging – rejection, and the membership and acceptance facets of anticipated sense of belonging on one factor (RMSEA = 0.08, CFI = 0.95, SRMR = 0.08, χ2(183) = 604.15).

Thus, attraction to the organization, desire to join the firm, and the different facets of anticipated sense of belonging assess distinct constructs. We thus followed our pre-registration plan of analyzing the outcome variable measures tapping potential consequences of anticipated sense of belonging separately, and with the mediator – anticipated sense of belonging – split by subscale.

**Verification of Demographic Balance Across Conditions and Genders**

Our lab’s policy is to verify the balance across conditions and genders of all the demographics measured in experiments, to ensure that the core assumption of baseline comparability across cells in experimental designs is valid (Fives et al., 2013). In the course of these analyses, we found that five covariates significantly differed across conditions: Political ideology (*t*(369) = -2.05, *p* = 0.041), Level of managerial responsibility in their current/most current job (*t*(369) = 3.05, *p* = 0.002), Number of subordinates in one’s current/most recent job (*t*(353.68) = -4.07, *p* < 0.001), Seriousness of job search (*t*(369) = -2.89, *p* = 0.004), and Number of job interviews done (*t*(369) = -2.24, *p* = 0.025). Specifically, compared to participants randomly assigned to the fairness case condition, participants in the business case condition were significantly more conservative (*MBusiness* = 4.16, *SE* = 0.13, 95% CI = [3.89, 4.43], *MFairness* = 3.79, *SE* = 0.13, 95% CI = [3.55, 4.03], *t*(369) = -2.05, *p* = 0.041, *d* = -0.21), had a significantly lower level of managerial responsibilities in their current/most recent job (*MBusiness* = 3.26, *SE* = 0.09, 95% CI = [3.08, 3.44], *MFairness* = 3.67, *SE* = 0.09, 95% CI = [3.48, 3.86], *t*(369) = 3.05, *p* = 0.002, *d* = 0.32), had a significantly greater number of subordinates/direct reports in their current/most current job (*MBusiness* = 3.28, *SE* = 0.14, 95% CI = [2.98, 3.58], *MFairness* = 2.47, *SE* = 0.14, 95% CI = [2.23, 2.71], *t*(353.68) = -4.07, *p* < 0.001, *d* = -0.43), were significantly more serious in the process of job search (*MBusiness* = 3.11, *SE* = 0.06, 95% CI = [2.99, 3.23], *MFairness* = 2.85, *SE* = 0.06, 95% CI = [2.72, 2.98], *t*(369) = -2.89, *p* = 0.004, *d* = -0.30), and had gone through a significantly greater number of job interviews (*MBusiness* = 1.01, *SE* = 0.07, 95% CI = [0.87, 1.15], *MFairness* = 0.78, *SE* = 0.07, 95% CI = [0.64, 0.92], *t*(369) = -2.24, *p* = 0.025, *d* = -0.23).

In addition, we found significant differences across genders on ten demographic variables: Political ideology (*t*(369) = -2.22, *p* = 0.027), Level of managerial responsibility in their current/most recent job (*t*(369) = 3.91, *p* < 0.001), Number of subordinates in one’s current/most recent job (*t*(369) = -2.95, *p* = 0.003), Seriousness of job search (*t*(369) = -2.03, *p* = 0.043), Number of job interviews done (*t*(314.31) = -3.20, *p* = 0.002), as well as Current employment status (χ2(3, 371) = 10.55, *p* = 0.014), Type of STEM occupation currently or previously held (χ2(10, 371) = 22.67, *p* = 0.012), Targeted industry for job search (χ2(3, 371) = 20.28, *p* < 0.001), Racial representation (χ2(1, 371) = 4.55, *p* = 0.033), and Native English speaker status (χ2(1, 371) = 3.00, *p* = 0.083). Specifically, compared to men, women were significantly more conservative (*MWomen* = 3.78, *SE* = 0.13, 95% CI = [3.54, 4.03], *MMen* = 4.19, *SE* = 0.13, 95% CI = [3.93, 4.44], *t*(369) = -2.22, *p* = 0.027, *d* = -0.23), had a significantly higher level of managerial responsibilities in their current/most current job (*MWomen* = 3.71, *SE* = 0.09,95% CI = [3.53, 3.89], *MMen* = 3.19, *SE* = 0.10,95% CI = [3.00, 3.38], *t*(369) = 3.91, *p* < 0.001, *d* = 0.41), had a significantly lower number of subordinates/direct reports in their current/most current job (*MWomen* = 2.59, *SE* = 0.14, 95% CI = [2.32, 2.86], *MMen* = 3.18, *SE* = 0.14, 95% CI = [2.90, 3.46], *t*(369) = -2.95, *p* = 0.003, *d* = -0.31), were significantly less serious in the process of job search (*MWomen* = 2.89, *SE* = 0.06, 95% CI = [2.77, 3.02], *MMen* = 3.08, *SE* = 0.07, 95% CI = [2.95, 3.21], *t*(369) = -2.03, *p* = 0.043, *d* = -0.21), and had gone through a significantly lower number of job interviews (*MWomen* = 0.74, *SE* = 0.07, 95% CI = [0.60, 0.87], *MMen* = 1.06, *SE* = 0.07, 95% CI = [0.92, 1.20], *t*(314.31) = -3.20, *p* = 0.002, *d* = -0.34). In addition, women represented 89% of participants not currently working, but only 33% of unemployed participants, 51% of participants working full-time, and 63% of participants working part-time. Women represented the minority of participants currently or previously holding an engineer role (42%) or a computer occupation (43%), but the majority of participants in all other occupations: STEM-related post-secondary teacher positions (73% women), Mathematical science occupations (72%), Life and physical science technician role (72%), “Other” occupation (69%), STEM-related sales position (67%), Life scientist (60%), STEM-related management role (53%), and Physical scientist (53%). Men and women were equally represented among participants currently or previously working as a Drafter, engineering technician, or mapping technician (50%). There were more women than men among participants looking for a job in Math (80%), in the Physical Sciences (69%), but fewer women than men among participants looking for a job in Engineering (45%) and in Computer Science (46%). There were also more women than men among non-White participants (60%), but fewer women among White participants (48%). Finally, there were more women than men among non-native English speakers (67%), but women and men were almost equally represented among native English-speakers (51%).

This imbalance across conditions (which appeared in spite of random assignment of participants to the two experimental conditions) and across genders could thus have confounded the results of the Condition x Gender analyses without covariates that we pre-registered on OSF. In the main text, we thus controlled in all Condition x Gender analyses for all ten unbalanced covariates (Political ideology, Level of managerial responsibility, Number of subordinates, Seriousness of job search, Number of job interviews done, Current employment status, Type of STEM occupation currently or previously held, Targeted industry for job search, Racial representation, and Native English speaker status).

**Additional Results Not Reported in the Main Text**

***Analyses of the Effect of Condition x Gender on Overall Sense of Belonging***

We conducted a two-way ANCOVA to investigate the effect of condition on the overall anticipated sense of belonging measure among women vs. men, controlling for the unbalanced demographic variables listed above. There was a significant effect of the Condition x Gender interaction on overall anticipated sense of belonging (*F*(1, 344) = 6.75, *p* = 0.010, = 0.019, 90% CI [0.003, 0.050]). In support of H2, pairwise comparisons revealed that female job seekers in the business case condition anticipated significantly lower sense of belonging (*MBusiness* = 4.66, *SE* = 0.15, 95% CI = [4.36, 4.96]) than their counterparts in the fairness case condition (*MFairness* = 4.93, *SE* = 0.14, 95% CI = [4.65, 5.21], *t*(344) = -2.13, *p* = 0.034, *d* = -0.33, 95% CI [-0.61, -0.04]). In contrast, and in support of H5, male job seekers were unaffected by the manipulation (*MBusiness* = 4.74, *SE* = 0.15, 95% CI = [4.44, 5.03], *MFairness* = 4.54, *SE* = 0.16, 95% CI = [4.22, 4.85], *t*(344) = 1.54, *p* = 0.12, *d* = 0.24, 95% CI [-0.06, 0.54]).

***Analyses of the Effect of Condition x Gender on Attraction to, and Desire to Join, the Organization***

We explored the effect of Condition x Gender on Attraction to the organization and Desire to join the organization using ANCOVAs controlling for the unbalanced demographic variables listed above. The results are reported in Table S7 below.

Table S7. *Study 3 Descriptive Statistics and Results of the 2 (Condition: Business case vs. Fairness case) x 2 (Gender: Women vs. Men) ANCOVAs on Attraction to the Organization and Desire to Join the Organization.*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Outcome variables** | **Predictors** | ***M***  **(*SE*)** | | ***M***  **(*SE*)** | | ***F*** | ***p*** | ***η2*** | **90% CI** | |
| **Lower bound** | **Upper bound** |
| 1. Attraction to the organization |  | **BC** | | **FC** | |  |  |  |  |  |
| Condition | 4.19  (0.12) | | 4.13  (0.12) | | 0.58 | 0.45 | 0.002 | 0.0000 | 0.016 |
|  |  | **Women** | | **Men** | |  |  |  |  |  |
|  | Gender | 4.18  (0.11) | | 4.14  (0.12) | | 0.28 | 0.60 | 0.001 | 0.0000 | 0.013 |
|  | Interaction |  |  |  |  | 1.21 | 0.27 | 0.004 | 0.0000 | 0.021 |
| *Pairwise comparisons* | |  | | BC | FC | ***t*** | ***p*** | **Cohen’s d** | **95% CI** | |
| Women | | 4.16  (0.13) | 4.19  (0.12) | 0.24 | 0.81 | 0.04 | -0.25 | 0.32 |
|  |  | Men | | 4.21  (0.13) | 4.07  (0.13) | -1.31 | 0.19 | -0.20 | -0.50 | 0.09 |
| 2. Desire to join the organization |  | **BC** | | **FC** | |  |  |  |  | |
| Condition | 5.51  (0.18) | | 5.46  (0.18) | | 0.15 | 0.70 | <0.001 | 0.0000 | 0.011 |
|  | **Women** | | **Men** | |  |  |  |  |  |
|  | Gender | 5.66  (0.18) | | 5.32  (0.19) | | 7.25 | 0.007 | 0.021 | 0.003 | 0.052 |
|  | Interaction |  |  |  |  | 4.02 | 0.046 | 0.012 | 0.0001 | 0.037 |
| *Pairwise comparisons* | |  | | BC | FC | ***t*** | ***p*** | **Cohen’s d** | **95% CI** | |
| Women | | 5.55  (0.12) | 5.77  (0.11) | 1.16 | 0.25 | 0.18 | -0.11 | 0.46 |
|  |  | Men | | 5.45  (0.12) | 5.20  (0.13) | -1.68 | 0.09 | -0.26 | -0.56 | 0.04 |

*Note*. *N* = 371. For ANCOVAs, all between-groups degrees of freedom are equal to 1, and all within-groups degrees of freedom are equal to 344. For pairwise comparisons, all degrees of freedom are equal to 344.

**Secondary Measures**

***Secondary Outcomes***

We measured two additional outcome variables, which we thought might be indirectly affected by the type of organizational diversity case, through anticipated sense of belonging.

**Anticipated Performance.** Participants completed the same measure as in Study 2.

**Perceptions of Genuineness.** Participants indicated the extent to which they perceived the organization’s commitment to diversity as genuine on 2 items (e.g., “This organization’s desire to support diverse individuals is genuine”; *α* = 0.53; scale: 1 “Strongly disagree” to 6 “Strongly agree”).

***Alternative Process***

We included an additional variable, which we wanted to explore as a potential mediator of the relationship between the type of organizational diversity case condition and anticipated sense of belonging. While social identity threat focuses on underrepresented individuals’ concern about being judged based on their group members, an alternative mechanism for the effect of type of diversity case on anticipated sense of belonging reported in the main text could be the extent to which participants typically expect to be stereotyped by others (Pinel, 1999).

**Stigma Consciousness.**To explore this alternative theoretical account, we measured participants’

Stigma Consciousness using an adapted version of 7 items from Pinel's (1999) scale[[4]](#footnote-4) that were matched to participant gender (e.g., “I would almost never think about the fact that I am female/male when I would interact with men/women in this organization”; *α* = 0.82, 1 “Strongly disagree” to 6 “Strongly agree”).

**Secondary Results**

***Conditional Indirect Effects Through Anticipated Sense of Belonging on Secondary Outcomes***

We tested for indirect effects of type of organizational diversity case (X), via the three facets of sense of belonging (membership (M1), acceptance (M2), and rejection (M3)), on each of the outcomes below (Y) dependent on participant gender (W), which could moderate the X-Mi links or the X-Y link (Model 8 in Hayes 2013), and controlling for the unbalanced demographic variables identified above. The conditional indirect effects through the membership and acceptance facets of sense of belonging were not supported (see Table S8). Thus, we only present below the conditional indirect effects through the rejection facet of sense of belonging.

***Anticipated Performance*.** Women’s greater anticipated rejection in the business (versus fairness) case condition in turn predicted significantly lower anticipation to perform well at work (*b* = -0.02, *SE* = 0.01, 95% CI [-0.044; -0.0003]). However, this indirect effect was not significantly different from the non-significant indirect effect among men (*b* = 0.003, *SE* = 0.01, 95% CI [-0.011; 0.021]), as indicated by the non-significant index of moderated mediation (*b* = 0.02, *SE* = 0.02, 95% CI [-0.001; 0.057]).

***Perceptions of Genuineness.*** Gender significantly moderated the indirect effect of type of diversity case on perceptions of genuineness, through anticipated rejection (index of moderated mediation: *b* = 0.08, *SE* = 0.04, 95% CI [0.010; 0.157]. Women’s greater anticipated rejection in the business (versus fairness) case condition in turn predicted significantly lower perceptions of the organization as genuine (*b* = -0.06, *SE* = 0.03, 95% CI [-0.119; -0.019]), but this indirect effect was not significant among male job seekers (*b* = 0.01, *SE* = 0.03, 95% CI [-0.036; 0.066]).

Table S8. *Study 3* *Secondary results for the moderated mediations analyses, IV = Type of organizational diversity case, DVs = Anticipated performance and Perceptions of genuineness, W = Gender, M1 = Anticipated membership, M2 = Anticipated acceptance, M3 = Anticipated rejection, Covariates = Seriousness of job search, Number of job interviews done, Level of managerial responsibility, Number of subordinates, Political ideology, Current employment status, Type of STEM occupation currently or previously held, Targeted industry for job search, Race, and Native English-speaker status.*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Outcome variables** | **Sub-sample** |  | **Total effect** | **Indirect effect 1** | **Index of moderated mediation 1** | **Indirect effect 2** | **Index of moderated mediation 2** | **Indirect effect 3** | **Index of moderated mediation 3** | **Direct effect** |
| Anticipated performance | Women | *b*  *(SE)*  95% CI  *p* | -0.08  (0.13)  [-0.324; 0.172]  0.55 | -0.01  (0.03)  [-0.065; 0.049] | 0.06  (0.04)  [-0.021; 0.146] | -0.02  (0.02)  [-0.059; 0.011] | 0.04  (0.03)  [-0.006; 0.097] | -0.02  (0.01)  [-0.044;  -0.0003] | 0.02  (0.02)  [-0.001; 0.057] | 0.003  (0.05)  [-0.094; 0.099]  0.96 |
| Men | *b*  *(SE)*  95% CI  *p* | 0.23  (0.13)  [-0.022; 0.483]  0.073 | 0.05  (0.03)  [-0.001; 0.113] | 0.02  (0.02)  [-0.009; 0.057] | 0.003  (0.01)  [-0.011; 0.021] | 0.04  (0.05)  [-0.058; 0.135]  0.43 |
| Perceptions of genuineness | Women | *b*  *(SE)*  95% CI  *p* | -0.14  (0.13)  [-0.383; 0.107]  0.27 | -0.002  (0.01)  [-0.031; 0.026] | 0.02  (0.02)  [-0.011; 0.078] | -0.02  (0.02)  [-0.055; 0.009] | 0.04  (0.03)  [-0.004; 0.095] | -0.07  (0.03)  [-0.119;  -0.019] | 0.08  (0.04)  [0.010; 0.157] | 0.016  (0.05)  [-0.084; 0.116]  0.75 |
| Men | *b*  *(SE)*  95% CI  *p* | 0.17  (0.13)  [-0.084; 0.414]  0.19 | 0.02  (0.02)  [-0.002; 0.063] | 0.02  (0.02)  [-0.007; 0.057] | 0.01  (0.03)  [-0.036; 0.066] | 0.03  (0.05)  [-0.073; 0.128]  0.60 |

*Note. N* = 371. Indirect effects 1, 2 and 3 represent the indirect effects of Type of diversity case on each of the outcome variables through, respectively, M1 = Anticipated membership, M2 = Anticipated acceptance and M3 = Anticipated rejection. Confidence intervals (CIs) were computed with the bias-corrected bootstrap method with 10,000 resamples.

## *Conditional Serial Indirect Effects Through Social Identity Threat & Anticipated Sense of Belonging on Outcomes Reported In-Text*

Given the support for H2, H3, H4, and H5 in this study (see analyses reported in-text), we also explored a serial indirect effect of type of organizational diversity case (X) through social identity threat (M1) to Anticipated rejection (M2), on each of the outcomes (Y), dependent on participant gender (W), which could moderate the X-M1 links or the X-Y link (Model 86 in Hayes, 2018; see Figure S3), and controlling for the unbalanced demographic variables identified above.

X

Type of organizational diversity case

(Business case = 1,

Fairness case = -1)

Y

* Attraction to the organization
* Desire to join the organization

M2

Anticipated

sense of belonging

– Rejection

M1

Social identity threat

W

Gender

(Male = 1,

Female = -1)

*Figure S3.* Conditional serial indirect effects analyses conducted in Study 3, corresponding to Model 86 in Hayes (2018).

*Note.* The path represents the indirect effect of Type of organizational diversity case (X) on respectively, Attraction to the Organization and Desire to join the organization (Y), serially through Social identity threat (M1) and Anticipated rejection (M2), conditional on Gender (W), and controlling for unbalanced demographics across conditions.

**Attraction to the Organization.** Consistent with our theory, female job seekers in STEM experienced significantly greater social identity threat in the business (vs. fairness) case condition, which in turn predicted greater anticipated rejection, and in turn, lower attraction to the organization (*b* = -0.02, *SE* = 0.01, 95% CI [-0.052; -0.002]). In contrast, this serial indirect effect was not supported among men (*b* = 0.001, *SE* = 0.01, 95% CI [-0.024; 0.030]) – though we note that the moderated mediation did not reach significance (index of moderated mediation: *b* = 0.03, *SE* = 0.02, 95% CI [-0.006; 0.068]).

**Desire to Join the Organization.** Consistent with our theory, female job seekers in STEM experienced significantly greater social identity threat in the business (vs. fairness) case condition, which in turn predicted greater anticipated rejection, and in turn, lower desire to join the organization (*b* = -0.04, *SE* = 0.02, 95% CI [-0.089; -0.005]). In contrast, this serial indirect effect was not supported among men (*b* = 0.002, *SE* = 0.02, 95% CI [-0.043; 0.052]) – though we note that the moderated mediation did not reach significance (index of moderated mediation: *b* = 0.05, *SE* = 0.03, 95% CI [-0.012; 0.118]).

***Alternative Process***

**Stigma Consciousness.** We explored the possibility that stigma consciousness would drive the detrimental effects of the business case for diversity on sense of belonging among underrepresented group members. We tested for an indirect effect of type of organizational diversity case (X), via stigma consciousness (M) on each of the three facets of anticipated sense of belonging (Y), dependent on participant gender (W), which could moderate the X-M links or the X-Y link (Model 8 in Hayes 2013), and controlling for the unbalanced demographic variables identified above. We effect-coded Type of organizational diversity case (Fairness case condition = -1; Business case condition = 1) and Gender (Female = -1; Male = 1). We computed indirect effects and their bias-corrected, bootstrapped 95% CIs, computed with 10,000 resamples using the PROCESS macro in SPSS (Hayes, 2013). None of the conditional indirect effects of type of diversity case through stigma consciousness on the membership, acceptance or rejection facets of anticipated sense of belonging were supported for either gender.

# Pilot Study for the Manipulation in Study 3

# *Method*

## Participants. We recruited 100 participants on MTurk (*Mage* = 34.79, *SD* = 9.62, range = [20; 71]; 63 men, 36 women, 1 nonbinary; 73 European Americans, 8 African Americans, 5 Latina/o Americans, 4 East Asian Americans, 4 South-East Asian Americans, 6 identified as Multiracial; 67 employed full-time, 20 employed part-time, 6 unemployed, 1 retired, 6 not currently working).

## Procedure. We used a between-subjects experimental design (Condition: Business case vs. Fairness case). After providing informed consent, participants read: “You are going to read an organization's diversity statement. Please read it carefully, as we will subsequently be asked to evaluate it”. Participants were then randomly assigned to one of two conditions: business or fairness case.

## Organizational Diversity Case Manipulation. Participants were then randomly assigned to read a business [or a *fairness*] case for diversity, each crafted based on diversity cases published on the websites of the world’s top 10 STEM companies (Dill, 2016) to increase the external validity of our manipulation:

Diversity and inclusion are part of our company's commitment to *performance [equality]*. Behind this focus is a simple but powerful idea: That diversity simply *makes good business sense [is the right thing to do]*. Our diversity and inclusion initiatives drive *positive business results by advancing our reputation to attract, retain, and engage diverse talents [a sense of community by advancing our values of respecting, supporting, and nurturing diverse talents].* We also strive to create an environment in which our company can *leverage the unique contributions of our diverse employees to develop innovative solutions for our diverse customer base [empower our diverse employees to grow and thrive as human beings whose ideas are heard and appreciated].* In sum, we firmly believe that diversity and inclusion can help our organization *meet and exceed our business goals [foster respect and opportunity for all]*.

After reading the manipulation, participants responded to two manipulation checks, and to the measures below.

## Measures.

***Manipulation Check - Perceived Instrumentality*.** Participants were asked to indicate the extent to which they agreed with the following statement: “The company sees diversity as core to the organization’s financial bottom-line.” They responded using a Likert scale ranging from 1 “Strongly disagree” to 7 “Strongly agree”.

***Manipulation Check - Perceived Non-Instrumentality*.** Participants indicated the extent to which they agreed with the following statement: “The company sees diversity as core to fairness and equality.”, using a Likert scale ranging from 1 “Strongly disagree” to 7 “Strongly agree”.

***Affective Reactions.*** Participants were asked to complete a PANAS scale (Watson et al., 1988) to measure the extent to which they felt positive and negative affect following the experimental manipulation. Sample items to measure positive affect were: “Enthusiastic”, “Interested” (α = 0.92), and for negative affect were: “Distressed”, “Nervous” (α = 0.80). Participants responded using a Likert scale ranging from 1 “Very slightly or not at all” to 5 “Extremely”.

***Liking of the Case.*** Participants indicated the extent to which they liked the diversity case, using a Likert scale ranging from 1 “Strongly dislike” to 7 “Strongly like”.

***Interest in the Case.*** Participants indicated the extent to which they found the diversity case interesting, using a Likert scale ranging from 1 “Not at all interesting” to 5 “Very interesting”.

***Perceived Length.*** Participants indicated how long they perceived the diversity case to be, using a Likert scale ranging from 1 “Not at all long” to 5 “Very long”.

***Perceived Difficulty.*** Participants indicated how difficult they perceived the diversity case to be, using a Likert scale ranging from 1 “Not at all difficult” to 5 “Very difficult”.

# *Results*

To ensure that all results presented below are robust, we controlled for the demographic characteristic that varied across conditions despite random assignment to conditions (native English-speaker status), as per our lab policy. The results remain unchanged without controls.

**Main Effects of Condition.**

***Manipulation Check – Perceived Instrumentality*.** We conducted a one-way ANCOVA to investigate the effect of condition on perceived instrumentality, controlling for native English-speaker status. We found a significant main effect of condition (*F*(1, 97) = 29.29, *p* < 0.001, = 0.232, 90% CI [0.118, 0.341]). Specifically, participants randomly assigned to the business case condition perceived the diversity case as significantly more instrumental (*MBusiness* = 5.70, *SE* = 0.20, 95% CI [5.31, 6.09]) than their counterparts in the fairness case condition (*MFairness* = 4.18, *SE* = 0.20, 95% CI [3.79, 4.57]).

***Manipulation Check – Perceived Non-Instrumentality*.** We conducted a one-way ANCOVA to investigate the effect of condition on perceived non-instrumentality, controlling for native English-speaker status. We found a significant main effect of condition (*F*(1, 97) = 8.45, *p* = 0.005, = 0.080, 90% CI [0.015, 0.175]). Specifically, participants randomly assigned to the business case condition perceived the diversity case as significantly less non-instrumental (*MBusiness* = 5.36, *SE* = 0.20, 95% CI [4.96, 5.76]) than their counterparts in the fairness case condition (*MFairness* = 6.20, *SE* = 0.20, 95% CI [5.80, 6.60]).

***Affective Reactions.*** A one-way ANCOVA, controlling for native English-speaker status, showed no significant main effect of condition on positive affect (*F*(1, 97) = 0.005, *p* = 0.94, < 0.001, 90% CI [0.000, 0.002], *MBusiness* = 2.95, *SE* = 0.14, 95% CI [2.68, 3.22], *MFairness* = 2.93, *SE* = 0.14, 95% CI [2.66, 3.21]), or on negative affect (*F*(1, 97) = 0.53, *p* = 0.47, = 0.005, 90% CI [0.000, 0.053], *MBusiness* = 1.13, *SE* = 0.04, 95% CI [1.05, 1.22], *MFairness* = 1.18, *SE* = 0.04, 95% CI [1.09, 1.26]).

***Liking of the Case.*** We found no significant main effect of condition on the extent to which participants liked the diversity case (*F*(1, 97) = 0.08, *p* = 0.78, = 0.001, 90% CI [0.000, 0.030], *MBusiness* = 6.21, *SE* = 0.24, 95% CI [5.74, 6.67], *MFairness* = 6.11, *SE* = 0.24, 95% CI [5.65, 6.58]).

***Interest in the Case.*** There was no significant main effect of condition on the extent to which participants found the diversity case to be interesting (*F*(1, 97) = 0.10, *p* = 0.76, = 0.001, 90% CI [0.000, 0.032], *MBusiness* = 2.99, *SE* = 0.19, 95% CI [2.62, 3.36], *MFairness* = 2.91, *SE* = 0.19, 95% CI [2.54, 3.28]).

***Perceived Length.*** There was no significant main effect of condition on the extent to which participants found the diversity case to be long (*F*(1, 97) = 0.01, *p* = 0.92, < 0.001, 90% CI [0.000, 0.004], *MBusiness* = 2.02, *SE* = 0.15, 95% CI [1.72, 2.33], *MFairness* = 2.00, *SE* = 0.15, 95% CI [1.69, 2.31]).

***Perceived Difficulty.*** We found no significant main effect of condition on the extent to which participants found the diversity case to be difficult (*F*(1, 97) = 0.09, *p* = 0.76, = 0.001, 90% CI [0.000, 0.031], *MBusiness* = 1.35, *SE* = 0.09, 95% CI [1.17, 1.53], *MFairness* = 1.39, *SE* = 0.09, 95% CI [1.21, 1.57]).

**Condition x Gender Interactions.**

To ensure that all results presented below are robust, we controlled for the demographic characteristic that varied across conditions despite random assignment to conditions (native English-speaker status) and across gender groups (Age, Education level, and Political views), as per our lab policy. The results remain unchanged without controls.

***Manipulation Check – Perceived Instrumentality.*** We conducted a two-way ANCOVA to test the effect of Condition x Gender on perceived instrumentality, controlling for native English-speaker status, Age, Education level, and Political views. We found a significant interaction effect (*F*(1, 91) = 6.24, *p* = 0.014, = 0.064, 90% CI [0.007, 0.157]). However, a simple effects analysis showed that relative to their same-gender counterparts in the fairness case condition, both women in the business case condition (*MWomen, Business* = 6.14, *SE* = 0.31, 95% CI [5.52, 6.75], *MWomen, Fairness* = 3.66, *SE* = 0.35, 95% CI [2.96, 4.36], *F*(1, 91) = 28.33, *p* < 0.001, = 0.237, 90% CI [0.119, 0.350]) and men in the business case condition perceived the diversity case to be significantly more instrumental (*MMen, Business* = 5.41, *SE* = 0.26, 95% CI [4.89, 5.93], *MMen, Fairness* = 4.39, *SE* = 0.24, 95% CI [3.92, 4.87], *F*(1, 91) = 8.30, *p* = 0.005, = 0.084, 90% CI [0.015, 0.182]).

***Manipulation Check – Perceived Non-Instrumentality*.** We conducted a two-way ANCOVA to test the effect of Condition x Gender on perceived non-instrumentality, controlling for native English-speaker status. We found no significant interaction effect (*F*(1, 91) = 0.24, *p* = 0.62, = 0.003, 90% CI [0.000, 0.045]).

***Affective Reactions.*** A two-way ANCOVA, controlling for native English-speaker status, showed no significant effect of Condition x Gender on positive affect (*F*(1, 91) = 2.32, *p* = 0.13, = 0.025, 90% CI [0.000, 0.098]), or on negative affect (*F*(1, 91) = 1.68, *p* = 0.20, = 0.018, 90% CI [0.000, 0.085]).

***Liking of the Case.*** We found no significant effect of Condition x Gender on the extent to which participants liked the diversity case (*F*(1, 91) = 0.81, *p* = 0.37, = 0.009, 90% CI [0.000, 0.065]).

***Interest in the Case.*** There was no significant effect of Condition x Gender on the extent to which participants found the diversity case to be interesting (*F*(1, 91) = 0.35, *p* = 0.55, = 0.004, 90% CI [0.000, 0.050]).

***Perceived Length.*** There was no significant effect of Condition x Gender on the extent to which participants found the diversity case to be long (*F*(1, 91) = 0.71, *p* = 0.40, = 0.008, 90% CI [0.000, 0.062]).

***Perceived Difficulty.*** We found no significant effect of Condition x Gender on the extent to which participants found the diversity case to be difficult (*F*(1, 91) = 1.36, *p* = 0.25, = 0.015, 90% CI [0.000, 0.079]).

***Discussion***

The manipulation did not have any significant effect on the variables measured, beyond the manipulation checks. This suggests that the two cases did not trigger significantly different affective reactions, that participants equally liked the two cases, and found them of comparable interest, length and difficulty. In addition, there was no evidence that the manipulation affected women and men differently.

# Study 4

**Link to Anonymized Study Materials, Data, and R Script (OSF)**

<https://osf.io/gk93e/?view_only=7bc840364b8e48a8a4e296d64a09a795>

**Full Scale for Identity Conflict (*α* = 0.90)**

1. In this organization, I would be seen as less of a “real” engineer because of my gender.
2. In this organization, people would see me as a female engineer, rather than just as an engineer.
3. In this organization, I would feel trapped in the role of the female engineer.
4. In this organization, there would be no constraints on my ability to be seen as an engineer.
5. In this organization, I would feel typecast as the “woman” on my engineering team.

Scale: 1 “Strongly disagree” to 7 “Strongly agree”

Future work should fully validate this scale.

**Full Scale for Feelings of Exploitation (*α* = 0.91)**

1. I feel the organization would use me to claim a good image regarding gender diversity.
2. I would feel treated as a gender diversity token in this organization.
3. I feel that the organization would try to exploit my gender to enhance its reputation with regards to diversity.
4. I feel the organization would use my gender to its advantage.

Scale: 1 “Strongly disagree” to 6 “Strongly agree”

Future work should fully validate this scale.

**Confirmatory Factor Analysis**

To test whether the hypothesized process (social identity threat), the alternative processes (identity conflict, feelings of being exploited, feelings of being seen as interchangeable, perceptions of the organization as externally vs. internally motivated to control prejudice) and the outcome variable (sense of belonging) represent distinct constructs, we conducted a confirmatory factor analysis (CFA). A nine-factor model fitted the data best (RMSEA = 0.05, CFI = 0.95, SRMR = 0.06, χ2(629) = 1494.58). The factors were: sense of belonging – membership, sense of belonging – acceptance, sense of belonging – rejection, social identity threat, identity conflict, feelings of being exploited, feelings of being seen as interchangeable, perceptions of the organization as externally motivated, and perceptions of the organization as internally motivated.

The fit was significantly better than a one-factor model (RMSEA = 0.16, CFI = 0.49, SRMR = 0.13, χ2(665) = 8941.93), a two-factor model with all facets of sense of belonging, social identity threat, identity conflict, feelings of being exploited and feelings of being seen as interchangeable on a first factor, and perceptions of the organization as externally and internally motivated on a second factor (RMSEA = 0.15, CFI = 0.52, SRMR = 0.16, χ2(664) = 8537.01), a three-factor model with all facets of sense of belonging, social identity threat and identity conflict on a first factor, feelings of being exploited and feelings of being seen as interchangeable on a second factor, and perceptions of the organization as externally and internally motivated on a third factor (RMSEA = 0.14, CFI = 0.62, SRMR = 0.17, χ2(662) = 6939.09), a four-factor model with all facets of sense of belonging on a second factor, social identity threat and identity conflict on a second factor, feelings of being exploited and feelings of being seen as interchangeable on a third factor, and perceptions of the organization as externally and internally motivated on a fourth factor (RMSEA = 0.11, CFI = 0.73, SRMR = 0.14, χ2(659) = 5005.17), a five-factor model with all facets of sense of belonging on a first factor, social identity threat and identity conflict on a second factor, feelings of being exploited and feelings of being seen as interchangeable on a third factor, perceptions of the organization as externally motivated on a fourth factor, and perceptions of the organization as internally motivated on a fifth factor (RMSEA = 0.10, CFI = 0.79, SRMR = 0.09, χ2(655) = 4106.70), a six-factor model with all facets of sense of belonging on a second factor, social identity threat and identity conflict on a second factor, feelings of being exploited on a third, feelings of being seen as interchangeable on a fourth factor, perceptions of the organization as externally motivated on a fifth factor, and perceptions of the organization as internally motivated on a sixth factor (RMSEA = 0.09, CFI = 0.82, SRMR = 0.09, χ2(650) = 3530.22), a seven-factor model with all facets of sense of belonging on a first factor, social identity threat on a second factor, identity conflict on a third factor, feelings of being exploited on a fourth factor, feelings of being seen as interchangeable on a fifth factor, perceptions of the organization as externally motivated on a sixth factor, and perceptions of the organization as internally motivated on a seventh factor (RMSEA = 0.09, CFI = 0.84, SRMR = 0.09, χ2(644) = 3231.86), and an eight-factor model with the membership and acceptance facets of sense of belonging on one factor, sense of belonging – rejection on a second factor, social identity threat on a third factor, identity conflict on a fourth factor, feelings of being exploited on a fifth factor, feelings of being seen as interchangeable on a sixth factor, perceptions of the organization as externally motivated on a seventh factor, and perceptions of the organization as internally motivated on an eighth factor (RMSEA = 0.06, CFI = 0.93, SRMR = 0.06, χ2(637) = 1794.37). We thus analyzed the indirect effects of type of organizational diversity case through each of the potential mediators on sense of belonging – split by subscale.

**Verification of Demographic Balance Across Conditions**

Our lab’s policy is to verify that the demographic variables measured in experiments are balanced across conditions, to ensure that the core assumption of baseline comparability across cells in experimental designs is valid (Fives et al., 2013). We found no significant difference across conditions for any of the demographics measured in Study 4.

**Additional Results Not Reported in the Main Text**

***Analyses of the Main Effect of Condition on Overall Sense of Belonging***

We conducted a one-way ANOVA to investigate the effect of condition on the overall anticipated sense of belonging measure. Replicating findings in Study 3, there was a significant main effect of condition on overall sense of belonging (*F*(2, 506) = 4.63, *p* = 0.010, = 0.018, 90% CI [0.002, 0.039]). As predicted, STEM women in the business case condition anticipated significantly lower belonging (*MBusiness* = 4.66, *SD* = 0.07, 95% CI = [4.52, 4.79]**)** relative to their counterparts in the control condition (*MControl* = 4.96, *SD* = 0.07, 95% CI = [4.82, 5.09], *t*(506) = -3.03, *p* = 0.003, *d* = -0.33, 95% CI = [-0.54, -0.11])**.** Unexpectedly however, there was no significant difference between the business and fairness conditions (*MFairness* = 4.79, *SD* = 0.07, 95% CI = [4.65, 4.92], *t*(506) = -1.32, *p* = 0.19, *d* = -0.14,95% CI = [-0.36, 0.07]). We return to this result in a mini meta-analysis following Study 5 in the main text. Finally, though we had no specific prediction regarding the comparison between the control and fairness conditions, we found no significant difference in anticipated belonging across these conditions (*t*(506) = 1.72, *p* = 0.087, *d* = 0.19, 95% CI = [-0.03, 0.40]).

# Study 5

**Link to Anonymized Pre-Registration on the Open Science Framework (OSF)**

<https://osf.io/s2d3x/?view_only=aabe9dd6b53048288f4bb4b052d9f967>

**Link to Anonymized Study Materials, Data, and R Script (OSF)**

<https://osf.io/fqsya/?view_only=9b486e09c1ec471389f62602421cd64d>

**Full Scale for Feelings of Exploitation (*α* = 0.89)**

1. I feel the organization would use me to claim a good image regarding racial diversity.
2. I would feel treated as a racial diversity token in this organization.
3. I feel that the organization would try to exploit my race to enhance its reputation with regards to diversity.
4. I feel the organization would use my race to its advantage.

Scale: 1 “Strongly disagree” to 6 “Strongly agree”

Future work should fully validate this scale.

**Confirmatory Factor Analysis**

To confirm that the anticipated sense of belonging measure was multi-dimensional as in previous studies, we conducted a confirmatory factor analysis (CFA). A three-factor model fitted the data best (RMSEA = 0.07, CFI = 0.98, SRMR = 0.02, χ2(51) = 168.57). The factors were: sense of belonging – membership, sense of belonging – acceptance, and sense of belonging – rejection. The fit was significantly better than a one-factor model (RMSEA = 0.30, CFI = 0.60, SRMR = 0.20, χ2(54) = 2408.53), a two-factor model with anticipated rejection on one factor and anticipated membership and acceptance on a second factor (RMSEA = 0.10, CFI = 0.96, SRMR = 0.03, χ2(53) = 300.94), a two-factor model with anticipated acceptance on one factor and anticipated membership and rejection on a second factor (RMSEA = 0.30, CFI = 0.62, SRMR = 0.20, χ2(53) =2274.29), or a two-factor model with anticipated membership on one factor and anticipated acceptance and rejection on a second factor (RMSEA = 0.30, CFI = 0.62, SRMR = 0.20, χ2(53) =2265.20).

Thus, the sense of belonging measure is indeed a multi-dimensional, and its three facets assess distinct constructs. We thus followed our pre-registration plan of conducting analyses with the dependent variable – sense of belonging – split by subscale.

**Verification of Demographic Balance Across Conditions**

As per our lab’s policy, we verified the balance of the demographic variables across conditions to ensure that the core assumption of baseline comparability across cells in experimental designs is valid (Fives et al., 2013). We found a significant difference in the distribution of gender across conditions: χ2(2) = 23.54, *p* < 0.001. Specifically, women represented 79% of participants in the control condition, 67% in the fairness case condition, and 53% in the business case condition.

Consistent with the analysis plan we pre-registered on OSF, given the potential confounding role of gender in analyses without covariates, we controlled for gender in the analysis presented in the main text. Controlling for gender did not change the pattern or significance of the results.

**Additional Results Not Reported in the Main Text**

***Analyses of the Main Effect of Condition on Overall Sense of Belonging***

We conducted a one-way ANCOVA to investigate the effect of condition on the overall anticipated sense of belonging measure, controlling for gender (see above). As predicted and supporting H2, there was a significant effect of condition on overall anticipated sense of belonging (*F*(2, 476) = 6.82, *p* = 0.001, = 0.028, 90% CI [0.007, 0.054]). Pairwise comparisons revealed that African Americans randomly assigned to the business case condition anticipated significantly lower overall belonging (*MBusiness* = 4.36, *SE* = 0.08, 95% CI [4.20, 4.51])compared to their counterparts in the control case condition (*MControl* = 4.77, *SE* = 0.08, 95% CI [4.60, 4.93], *t*(476) = -3.64,  *p* < 0.001, *d* = -0.42, 95% CI = [-0.64, -0.20]) and in the fairness case condition (*MFairness* = 4.62, *SE* = 0.08,95% CI = [4.46, 4.78], *t*(476) = -2.39, *p* = 0.017, *d* = -0.27,95% CI = [-0.49, -0.05]). Finally, though we had no specific prediction regarding the comparison between the control and fairness conditions, there was no significant difference in overall anticipated belonging between these two conditions (*t*(476) = 1.32, *p* = 0.19, *d* = 0.15, 95% CI = [-0.07, 0.37]).

***Analyses of the Main Effect of Condition on Attraction to, and Desire to Join, the Organization***

We explored the effect of the manipulation on Attraction to the organization and Desire to join the organization using ANCOVAs, controlling for Gender. The results are reported in Table S9 below.

Table S9. *Study 5 Descriptive Statistics for, and Results of the One-Way (Condition: Business case vs. Control vs. Fairness case) ANCOVAs on Attraction to the Organization and Desire to Join the Organization.*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Outcome variables** | **Predictors** | **BC**  ***M***  **(*SE*)** | **Control**  ***M***  **(*SE*)** | **FC**  ***M***  **(*SE*)** | ***F*** | ***p*** | ***η2*** | **90% CI** | |
| **Lower bound** | **Upper bound** |
| 1. Attraction to the organization | Condition | 3.88  (0.07) | 4.00  (0.08) | 4.01  (0.07) | 1.01 | 0.36 | 0.004 | 0.000 | 0.016 |
|  | |  | | | ***t*** | ***p*** | ***d*** | **95% CI** | |
| ***Pairwise comparisons*** | | | **Lower bound** | **Upper bound** |
| FC-Control | | | 0.06 | 0.95 | 0.01 | -0.21 | 0.23 |
| BC-Control | | | -1.19 | 0.23 | -0.14 | -0.36 | 0.08 |
|  |  | BC-FC | | | -1.28 | 0.20 | -0.14 | -0.36 | 0.08 |
| 2. Desire to join the organization | Condition | 5.25  (0.10) | 5.67  (0.10) | 5.52  (0.10) | 4.38 | 0.013 | 0.018 | 0.002 | 0.040 |
|  | |  | | | ***t*** | ***p*** | ***d*** | **95% CI** | |
|  | | ***Pairwise comparisons*** | | | **Lower bound** | **Upper bound** |
| FC-Control | | | -1.05 | 0.29 | -0.12 | -0.34 | 0.10 |
| BC-Control | | | -2.92 | 0.004 | -0.33 | -0.55 | -0.11 |
|  |  | BC-FC | | | -1.92 | 0.056 | -0.22 | -0.44 | 0.004 |

*Note. N* = 480. For ANCOVAs, all between-groups degrees of freedom are equal to 2, and all within-groups degrees of freedom are equal to 476. For pairwise comparisons, all degrees of freedom are equal to 476.

**Secondary Measures**

***Feelings About the Case***

Participants were asked to write how the case for diversity made them feel.

***Expectations of Diversity in the Organization***

Participants were asked to indicate how diverse they expected the organization to be, on a scale ranging from 1 “Not at all” to 6 “Extremely”.

***Expectations of Confronting Behaviour***

Participants completed a 4-item scale measuring the extent to which they expected others in the organization to confront racial bias if they encountered it in the workplace, which was adapted from Rattan and Dweck (2010). A sample item was: “If someone made a racially biased comment, I expect people in this organization would not speak up to address it”, reverse scored (*α* = 0.71; scale: 1 “Strongly disagree” to 7 “Strongly agree”).

**Secondary Results**

***Indirect Effects Through Anticipated Sense of Belonging on Secondary Outcomes***

We explored the possibility that there would be indirect effects of type of organizational diversity case (X) on each of the additional measures in this appendix (Y), via the three facets of anticipated sense of belonging (membership (M1), acceptance (M2), and rejection (M3)), controlling for gender (Female = -1, Male = 1; Model 6 in Hayes 2013), which was identified as unbalanced across conditions (see above). We declared Type of organizational diversity case as a multicategorical variable, and effect-coded it to produce the relevant comparisons. The coefficients reported below are indirect effects and their bias-corrected, bootstrapped 95% CIs, computed with 10,000 resamples using the PROCESS macro in SPSS (Hayes, 2013).

The indirect effects through the membership and acceptance facets of anticipated sense of belonging were not supported for either outcome variable. Thus, we only describe in detail below the conditional indirect effects through the rejection facet of anticipated sense of belonging.

**Expectations of Diversity in the Organization.** The indirect effect of type of organizational diversity case on expectations of diversity in the organizations, through anticipated rejection was not supported when comparing the business case to the control and fairness case conditions (business vs. control: *b* = 0.02, *SE* = 0.03, 95% CI [-0.042; 0.092]; business vs. fairness: *b* = 0.01, *SE* = 0.02, 95% CI [-0.020; 0.052]). Neither was there a significant indirect effect when comparing the fairness case to the control case condition (fairness vs. control: *b* = 0.01, *SE* = 0.02, 95% CI [-0.023; 0.049]).

**Expectations of Confronting Behavior.** When comparing the business case to the control and fairness case conditions, there were significant indirect effects of type of diversity case on desire to join the organization, through anticipated rejection (business vs. control: *b* = -0.26, *SE* = 0.06, 95% CI [-0.377; -0.159]; business vs. fairness: *b* = -0.13, *SE* = 0.06, 95% CI [-0.239; -0.020]). Specifically, African Americans exposed to a business (vs. fairness, or control) case anticipated significantly greater rejection in the organization, which in turn predicted lower expectations that others in the organization would confront racial bias if they encountered it. The indirect effect was also significant when comparing the fairness condition to the control condition (*b* = -0.14, *SE* = 0.05, 95% CI [-0.234; -0.047]). However, as with the findings reported in the main text, the associated effect size was about half the size of the indirect effect obtained when comparing the business case to the control condition, suggesting that the fairness case decreases to a smaller extent African Americans’ expectations that others in the organization would confront racial bias if they encountered it.

***Serial Indirect Effects Through Social Identity Threat & Anticipated Sense of Belonging on Secondary Outcomes***

We also explored the possibility that there would be a serial indirect effect of type of organizational diversity case (X) through social identity threat (M1) to anticipated rejection (M2), on each of the additional measures in this appendix (Y), controlling for gender (Female = -1, Male = 1; Model 6 in Hayes 2013), which was identified as unbalanced across conditions (see above).

**Expectations of Diversity in the Organization.** The serial indirect effect of type of diversity case on expectations of diversity in the organizations, serially through social identity threat and anticipated rejection was not supported when comparing the business case to the control and fairness case conditions (business vs. control: *b* = -0.01, *SE* = 0.02, 95% CI [-0.066; 0.038]; business vs. fairness: *b* = -0.01, *SE* = 0.02, 95% CI [-0.043; 0.023]). Neither was there a significant serial indirect effect when comparing the fairness case to the control case condition (fairness vs. control: *b* = -0.01, *SE* = 0.01, 95% CI [-0.030; 0.017]).

**Expectations of Confronting Behavior.** When comparing the business case to the control and fairness case conditions, there was a significant serial indirect effect of type of diversity case on expectations of confronting behavior, through social identity threat and anticipated rejection (business vs. control: *b* = -0.13, *SE* = 0.03, 95% CI [-0.203; -0.070]; business vs. fairness: *b* = -0.07, *SE* = 0.03, 95% CI [-0.139; -0.026]). Specifically, African Americans randomly assigned to read to a business case reported significantly greater social identity threat compared to their counterparts in the control case or fairness case conditions, which in turn predicted greater anticipated rejection, and in turn, lower expectations that others in the organization would confront racial bias if they encountered it. The corresponding serial indirect effect was also significant when comparing the fairness case to the control case condition (*b* = -0.05, *SE* = 0.02, 95% CI [-0.105; -0.010]). However, as with the findings reported in the main text and above, the associated effect size was about half the size of the indirect effect obtained when comparing the business case to the control condition – suggesting that the fairness case decreases to a smaller extent African Americans’ expectations that others in the organization would confront racial bias if they encountered it.

# Study 6

**Link to Anonymized Pre-Registration on the Open Science Framework (OSF)**

<https://osf.io/unvfp/?view_only=07dc3c8e361e4f2b95b2444174455742>

**Link to Anonymized Study Materials, Data, and R Script (OSF)**

<https://osf.io/gbhn9/?view_only=dad4bcfbd12540ae985eb26f4420085c>

**Confirmatory Factor Analysis**

To confirm that the anticipated sense of belonging measure was multi-dimensional as in previous studies, we conducted a confirmatory factor analysis (CFA). A three-factor model fitted the data best (RMSEA = 0.07, CFI = 0.98, SRMR = 0.02, χ2(51) = 269.82). The factors were: sense of belonging – membership, sense of belonging – acceptance, and sense of belonging – rejection. The fit was significantly better than a one-factor model (RMSEA = 0.28, CFI = 0.69, SRMR = 0.18, χ2(54) = 4243.90), a two-factor model with anticipated rejection on one factor and anticipated membership and acceptance on a second factor (RMSEA = 0.11, CFI = 0.95, SRMR = 0.03, χ2(53) = 711.29), a two-factor model with anticipated acceptance on one factor and anticipated membership and rejection on a second factor (RMSEA = 0.27, CFI = 0.72, SRMR = 0.18, χ2(53) = 3866.62), or a two-factor model with anticipated membership on one factor and anticipated acceptance and rejection on a second factor (RMSEA = 0.26, CFI = 0.72, SRMR = 0.17, χ2(53) = 3770.58).

Thus, the sense of belonging measure is indeed a multi-dimensional, and its three facets assess distinct constructs. We thus followed our pre-registration plan of conducting analyses with the dependent variable – sense of belonging – split by subscale.

**Verification of Demographic Balance Across Conditions and Races**

Our lab’s policy is to verify the balance across conditions and races of all the demographics measured in experiments, in order to ensure that the core assumption of baseline comparability across cells in experimental designs is valid (Fives et al., 2013). In the course of these analyses, we found that three covariates significantly differed across conditions: Age (*F*(2, 1016) = 2.47, *p =* 0.085), Gender (χ2(2) = 7.03, *p* = 0.030), Student status (χ2(2) = 8.21, *p* = 0.016). Specifically, participants in the business case condition were marginally older than participants in the control condition (*MBusiness* = 26.0, *SE* = 0.37, 95% CI = [25.2, 26.7], *MControl* = 24.9, *SE* = 0.43, 95% CI = [24.1, 25.8], *t*(1016) = 1.85, *p* = 0.064, *d* = 0.15) and in the fairness case condition (*MFairness* = 24.9, *SE* = 0.36, 95% CI = [24.2, 25.7], *t*(1016) = 1.95, *p* = 0.051, *d* = 0.14). There was no significant age difference across the fairness and control conditions (*t*(1016) = 0.08, *p* = 0.94, *d* = 0.01). Moreover, women represented 56% of participants in the control condition, 52% in the fairness case condition, and 45% in the business case condition. Non-student participants (i.e., participants who had recently graduated) represented 14% in the fairness case condition, versus only 9% of participants in the control condition and 8% in the business case condition.

In addition, we found significant differences on five demographic variables across racial groups:

Education level (*t*(1017) = -4.22, *p* < 0.001), Student status (χ2(1) = 7.22, *p* = 0.007), Field of study (χ2(9) = 19.50, *p* = 0.021), Current employment status (χ2(2) = 5.49, *p* = 0.064), and Job seeker status (χ2(1) = 10.49, *p* = 0.001). Specifically, compared to Whites, African Americans had a significantly lower education level (*MAfrican Americans* = 3.82, *SE* = 0.04, 95% CI = [3.75, 3.88], *MWhites* = 4.02, *SE* = 0.03, 95% CI = [3.96, 4.09], *t*(1017) = -4.22, *p* < 0.001, *d* = -0.26). Africans Americans represented 51% of students among participants, but only 37% of recent graduates. In addition, among our participants, more African Americans than Whites were studying or had studied Law (60%), Medicine / Nursing (58%), Business / Economics (56%) and Computer Science (51%), but fewer than Whites were studying or had studied Engineering (45%), Other fields (45%), the Arts and Humanities (44%), Math (42%), the Physical Sciences (41%), and the Social Sciences (41%). In addition, African Americans represented 51% of participants not currently working and 53% of participants working part-time, versus only 43% of participants working full-time. Finally, African Americans accounted for 54% of participants looking for a job, and only 44% of participants who were not looking for one.

This imbalance across conditions (which appeared despite random assignment of participants to the two experimental conditions) and across races could thus have confounded the results of our Condition x Race analyses. In the main text, we thus controlled in all Condition x Race analyses for all six unbalanced covariates (Age, Gender, Student status, Education level, Field of study, Current employment status, Job seeker status)

**Additional Results Not Reported in the Main Text**

***Analyses of the Effect of Condition x Race on Overall Sense of Belonging***

We conducted two-way ANCOVAs to investigate the interaction effect of Condition x Race on the overall anticipated sense of belonging measure, controlling for demographic variables listed above. There was no significant Condition x Race interaction effect on overall belonging (*F*(2, 997) = 1.40, *p* = 0.25, = 0.003, 90% CI [0.000, 0.010]). However, in line with recommendations that hypothesis-driven comparisons be tested irrespective of the significance of the interaction (Castañeda et al., 1993; Ruxton & Beauchamp, 2008), we conducted simple effects analyses, which we report below.

**African American Participants.** Supporting H2, there was a significant effect of condition among African Americans, whereby those who had been randomly assigned to the business case condition anticipated significantly lower belonging to the organization (*MBusiness* = 4.17, *SE* = 0.09, 95% CI [3.99, 4.35]) relative to their counterparts in the control case condition (*MControl* = 4.58, *SE* = 0.11, 95% CI [4.37, 4.79],*t*(997) = -3.49,  *p* < 0.001, *d* = -0.40, 95% CI = [-0.63, -0.17]), and in the fairness case condition (*MFairness* = 4.38, *SE* = 0.09, 95% CI [4.21, 4.56],*t*(997) = -2.00, *p* = 0.045, *d* = -0.21, 95% CI = [-0.41, -0.01]). Finally, though we had no specific prediction regarding the control-fairness comparison, there was no significant difference in overall anticipated belonging between these two conditions (*t*(997) = 1.66, *p* = 0.098, *d* = 0.19, 95% CI = [-0.03, 0.42]).

**White American Participants.** Supporting H5, among Whites, there was no significant effect of condition on overall anticipated belonging across the business case and control conditions (*MBusiness* = 4.44, *SE* = 0.09, 95% CI [4.26, 4.62], *MControl* = 4.61, *SE* = 0.10, 95% CI [4.42, 4.80], *t*(997) = -1.52, *p* = 0.13, *d* = -0.17, 95% CI = [-0.40, 0.05]), across the business and fairness case (*MFairness* = 4.43, *SE* = 0.08, 95% CI [4.26, 4.59], *t*(997) = 0.08, *p* = 0.93, *d* = 0.01, 95% CI = [-0.20, 0.21]), or across the fairness and control conditions (*t*(997) = 1.67, *p* = 0.096, *d* = 0.18, 95% CI = [-0.03, 0.40]).

***Serial Indirect Effects Through Social Identity Threat & Anticipated Sense of Belonging***

Finally, we tested our pre-registered prediction that there would be a serial indirect effect of Type of organizational diversity case (X) through social identity threat (M1) to each facet of Anticipated sense of belonging (M2s), on each of the outcomes (Y) (Model 86 in Hayes, 2013; see Figure S4), and controlling for the same demographic variables as above. We note that none of the indirect effect analyses below had a significant index of moderated mediation, which we discuss in the Study 6 discussion.

X

Type of organizational diversity case

(Business case = 1,

Control case =0,

Fairness case = -1)

Y

* Attraction to the organization
* Desire to join the organization

M2s

Anticipated

sense of belonging

– Membership

– Acceptance

– Rejection

M1

Social identity threat

*Figure S4*. Serial indirect effects analyses conducted in Study 6 among African American and White participants, corresponding to Model 86 in Hayes (2013).

*Note.* The path represents the indirect effect of Type of organizational diversity case (X) on, respectively, Attraction to the organization and Desire to join the organization (Y), serially through Social identity threat (M1) and Anticipated rejection (M2s), controlling for unbalanced demographic variables across conditions.

W

Race

**African American Participants.**

***Attraction to the Organization.*** When comparing the business case to the control condition among African Americans, the predicted serial indirect effects of type of diversity case on desire to join the organization, through social identity threat and each facet of anticipated sense of belonging, were significant (via membership: *b* = -0.05, *SE* = 0.02, 95% CI [-0.090; -0.011]; via acceptance: *b* = -0.06, *SE* = 0.02, 95% CI [-0.110; -0.013]; via rejection: *b* = -0.04, *SE* = 0.02, 95% CI [-0.085; -0.011]). Specifically, African Americans randomly assigned to read a business case reported significantly greater social identity threat compared to their counterparts in the control condition, which in turn predicted significantly lower anticipated membership and acceptance, as well as greater anticipated rejection, which in turn predicted lower desire to join the organization. However, these serial indirect effects failed to reach significance when comparing the business and fairness case conditions (via membership: *b* = -0.03, *SE* = 0.02, 95% CI [-0.061; 0.006]; via acceptance: *b* = -0.03, *SE* = 0.02, 95% CI [-0.076; 0.007]; via rejection: *b* = -0.02, *SE* = 0.02, 95% CI [-0.058; 0.006]). Though we had no specific predictions regarding the control-fairness comparison, there were no significant indirect effects either when comparing the fairness case to the control case condition (via membership: *b* = -0.02, *SE* = 0.02, 95% CI [-0.058; 0.013]; via acceptance: *b* = -0.03, *SE* = 0.02, 95% CI [-0.072; 0.016]; via rejection: *b* = -0.02, *SE* = 0.02, 95% CI [-0.056; 0.013]).

***Desire to Join the Organization.*** The same pattern of significant indirect effects emerged with Desire to join the organization as the outcome variable. When comparing the business case to the control condition among African Americans, the predicted serial indirect effects of type of diversity case on desire to join the organization, through social identity threat and each facet of anticipated sense of belonging, were significant (via membership: *b* = -0.06, *SE* = 0.03, 95% CI [-0.112; -0.013]; via acceptance: *b* = -0.07, *SE* = 0.03, 95% CI [-0.132; -0.015]; via rejection: *b* = -0.08, *SE* = 0.03, 95% CI [-0.141; -0.018]). However, these serial indirect effects failed to reach significance when comparing the business and fairness case conditions (via membership: *b* = -0.03, *SE* = 0.02, 95% CI [-0.075; 0.007]; via acceptance: *b* = -0.04, *SE* = 0.03, 95% CI [-0.090; 0.009]; via rejection: *b* = -0.04, *SE* = 0.03, 95% CI [-0.097; 0.010]). Though we had no specific predictions regarding the control-fairness comparison, there were no significant indirect effects either when comparing these two conditions (via membership: *b* = -0.03, *SE* = 0.02, 95% CI [-0.072; 0.015]; via acceptance: *b* = -0.03, *SE* = 0.03, 95% CI [-0.086; 0.019]; via rejection: *b* = -0.03, *SE* = 0.03, 95% CI [-0.092; 0.020]).

**White American Participants.**

***Attraction to the Organization.*** Among White participants, when comparing the business case to the control condition, we found significant serial indirect effects of type of diversity case on desire to join the organization, through social identity threat and each facet of anticipated sense of belonging (via membership: *b* = -0.06, *SE* = 0.02, 95% CI [-0.098; -0.025]; via acceptance: *b* = -0.07, *SE* = 0.02, 95% CI [-0.121; -0.031]; via rejection: *b* = -0.06, *SE* = 0.02, 95% CI [-0.097; -0.024]). These effects were also significant when comparing the business case to the fairness case condition (via membership: *b* = -0.04, *SE* = 0.02, 95% CI [-0.075; -0.007]; via acceptance: *b* = -0.05, *SE* = 0.02, 95% CI [-0.091; -0.010]; via rejection: *b* = -0.04, *SE* = 0.02, 95% CI [-0.075; -0.006]), but not when comparing the fairness case to the control condition (via membership: *b* = -0.02, *SE* = 0.02, 95% CI [-0.053; 0.011]; via acceptance: *b* = -0.02, *SE* = 0.02, 95% CI [-0.064; 0.013]; via rejection: *b* = -0.02, *SE* = 0.02, 95% CI [-0.051; 0.010]).

***Desire to Join the Organization.*** Among White participants, when comparing the business case to the control condition, we found significant serial indirect effects of type of diversity case on desire to join the organization, through social identity threat and each facet of anticipated sense of belonging (via membership: *b* = -0.07, *SE* = 0.02, 95% CI [-0.121; -0.031]; via acceptance: *b* = -0.09, *SE* = 0.03, 95% CI [-0.145; -0.037]; via rejection: *b* = -0.10, *SE* = 0.03, 95% CI [-0.158; -0.041]). These effects were also significant when comparing the business case to the fairness case condition (via membership: *b* = -0.05, *SE* = 0.02, 95% CI [-0.092; -0.010]; via acceptance: *b* = -0.06, *SE* = 0.03, 95% CI [-0.111; -0.009]; via rejection: *b* = -0.07, *SE* = 0.03, 95% CI [-0.125; -0.011]), but not when comparing the fairness case to the control condition (via membership: *b* = -0.02, *SE* = 0.02, 95% CI [-0.065; 0.013]; via acceptance: *b* = -0.03, *SE* = 0.02, 95% CI [-0.079; 0.016]; via rejection: *b* = -0.03, *SE* = 0.03, 95% CI [-0.083; 0.018]).

***Analyses of the Effect of Condition x Race on Attraction to, and Desire to Join, the Organization***

We explored the effect of Condition x Race on Attraction to the organization and Desire to join the organization using ANCOVAs controlling for the unbalanced demographic variables listed above. The results are reported in Table S10 below.

Table S10. *Study 6 Descriptive Statistics and Results of the 2 (Condition: Business case vs. Fairness case vs. Control) x 2 (Race: African Americans vs. Whites) ANCOVAs on Attraction to the Organization and Desire to Join the Organization.*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Outcome variables** | **Predictors** | ***M***  **(*SE*)** | | ***M***  **(*SE*)** | | ***M***  **(*SE*)** | ***F*** | ***p*** | ***η2*** | **90% CI** | | |
| **Lower bound** | | **Upper bound** |
| 1. Attraction to the organization |  | **BC**  3.55  (0.07) | | **Control**  3.73  (0.08) | | **FC**  3.57  (0.07) |  |  |  |  | |  |
| Condition | 3.00 | 0.050 | 0.006 | 0.000 | | 0.015 |
|  |  | |  | |  |  |  |  |  | |  |
| Race | **African Americans**  3.58  (0.07) | | **Whites**  3.65  (0.06) | |  | 1.31 | 0.25 | 0.001 | 0.000 | | 0.008 |
| Interaction |  |  |  |  |  | 1.87 | 0.16 | 0.004 | 0.000 | | 0.011 |
| ***Pairwise comparisons*** | African Americans | **BC** | | **FC** | | **Control** | ***t*** | ***p*** | **Cohen’s d** | **95% CI** | | |
| 3.43  (0.09) | | 3.58  (0.08) | | 3.73  (0.10) | **Lower bound** | | **Upper bound** |
| FC-Control | | | | | -1.30 | 0.19 | -0.15 | -0.38 | | 0.08 |
| BC-Control | | | | | -2.68 | 0.007 | -0.31 | -0.53 | | -0.08 |
| BC-FC | | | | | -1.52 | 0.13 | -0.16 | -0.36 | 0.04 | |
| Whites | **BC** | | **FC** | | **Control** | ***t*** | ***p*** | **Cohen’s d** | **95% CI** | | |
| 3.67  (0.09) | | 3.56  (0.08) | | 3.73  (0.09) | **Lower bound** | **Upper bound** | |
| FC-Control | | | | | -1.57 | 0.12 | -0.17 | -0.39 | 0.04 | |
| BC-Control | | | | | -0.58 | 0.56 | -0.07 | -0.29 | 0.16 | |
| BC-FC | | | | | 1.00 | 0.32 | 0.11 | -0.10 | 0.31 | |
| **Outcome variables** | **Predictors** | ***M***  **(*SE*)** | | ***M***  **(*SE*)** | | ***M***  **(*SE*)** | ***F*** | ***p*** | ***η2*** | **90% CI** | | |
| **Lower bound** | | **Upper bound** |
| 2. Desire to join the organization |  | **BC**  5.01  (0.09) | | **Control**  5.27  (0.10) | | **FC**  5.13  (0.08) |  |  |  |  | | |
| Condition | 3.49 | 0.031 | 0.007 | 0.0003 | | 0.017 |
| Race | **African Americans**  5.08  (0.08) | | **Whites**  5.19  (0.08) | | 1.89 | 0.17 | 0.002 | 0.000 | | 0.009 |
|  | Interaction |  | |  | | 1.86 | 0.16 | 0.004 | 0.000 | | 0.011 |
| ***Pairwise comparisons*** | African Americans | **BC** | | **FC** | | **Control** | ***t*** | ***p*** | **Cohen’s d** | **95% CI** | | |
| 4.86  (0.11) | | 5.09  (0.11) | | 5.30  (0.13) | **Lower bound** | | **Upper bound** |
| FC-Control | | | | | -1.46 | 0.14 | -0.17 | -0.40 | | 0.06 |
| BC-Control | | | | | -3.17 | 0.002 | -0.36 | -0.59 | | -0.14 |
| BC-FC | | | | | -1.87 | 0.062 | -0.19 | -0.40 | | 0.01 |
| Whites | **BC** | | **FC** | | **Control** | ***t*** | ***p*** | **Cohen’s d** | **95% CI** | | |
| 5.17  (0.11) | | 5.16  (0.10) | | 5.25  (0.12) | **Lower bound** | | **Upper bound** |
| FC-Control | | | | | -0.64 | 0.53 | -0.07 | -0.28 | | 0.14 |
| BC-Control | | | | | -0.57 | 0.57 | -0.06 | -0.29 | | 0.16 |
| BC-FC | | | | | 0.05 | 0.96 | 0.005 | -0.20 | | 0.21 |

*Note.* *N* = 1,019. For ANCOVAs, between-groups degrees of freedom are equal to 2 for Condition, 1 for Race, 2 for their interaction, and all within-groups degrees of freedom are equal to 997. For pairwise comparisons, all degrees of freedom are equal to 997.

**Secondary Measures**

***Feelings About the Case***

Participants were asked to write how the case for diversity made them feel.

***Expectations of Diversity in the Organization***

Participants were asked to indicate how diverse they expected the organization to be, on a scale ranging from 1 “Not at all” to 6 “Extremely”.

**Secondary Results**

***Indirect Effects Through Anticipated Sense of Belonging on Secondary Outcomes***

We explored for indirect effects of Type of organizational diversity case (X) on each of the outcome variables measured (Y), via the three facets of Anticipated sense of belonging (membership (M1), acceptance (M2), and rejection (M3); Model 8 in Hayes, 2013), and controlling for the same demographic variable as above. we declared Type of organizational diversity case as a multicategorical variable, and effect-coded it to produce the relevant comparisons. We also effect-coded Race (White = -1; African American = 1). The coefficients reported below are indirect effects and their bias-corrected, bootstrapped 95% CIs, computed with 10,000 resamples using the PROCESS macro in SPSS (Hayes, 2013).

**Expectations of Diversity in the Organization.** We did not find any significant indirect effect of type of diversity case on expectations of diversity through any of the facets of anticipated sense of belonging, whether it be among African Americans or Whites.

***Serial Indirect Effects Through Social Identity Threat & Anticipated Sense of Belonging on Secondary Outcomes***

We also explored the possibility that there would be a serial indirect effect of Type of organizational diversity case (X) through social identity threat (M1) to each of the facets of Anticipated sense of belonging (M2s), on each of the outcomes (Y) (Model 86 in Hayes, 2013), and controlling for the same demographic variables as above.

**Expectations of Diversity in the Organization.** We did not find any significant serial indirect effect of type of diversity case on expectations of diversity, through any of the facets of anticipated sense of belonging as M2, whether it be among African Americans or Whites.

# Pilot Study for the Manipulation in Study 6

# *Method*

## Participants. We recruited 333 participants on MTurk, among which 34 participants failed to finish the survey and 13 were detected by the IPHub program as having suspicious IP addresses (non-US, or VPN/VPS). In keeping with our pre-registered exclusion criteria, these participants were excluded from analyses. Our final sample thus consisted of 286 participants (*Mage* = 39.91, *SD* = 11.51, range = [19; 70]; 143 men, 142 women, 1 nonbinary; 153 European Americans, 133 African Americans; 184 employed full-time, 53 employed part-time, 18 unemployed, 9 retired, 22 not currently working).

## Procedure. We used a between-subjects experimental design (Condition: Business case vs. Fairness case vs. Control). After providing informed consent, participants read: “You are going to see the diversity statement of a company called InvoCorp[[5]](#footnote-5). Please read this statement carefully, as you will subsequently be asked to evaluate it.” Participants were then randomly assigned to one of three conditions: business case, fairness case or control case.

## Organizational Diversity Case Manipulation. Participants were randomly assigned to see one of three versions of a purported corporate website, featuring one of the following diversity cases:

## Business case condition (first webpage)

Diversity and inclusion are part of our company's **commitment** **to** **performance**. Behind this focus is a simple but powerful idea: That diversity simply **makes good business sense**. In sum, we firmly believe that diversity and inclusion **can help our organization to achieve superior performance***.*

Fairnesss case condition (first webpage)

Diversity and inclusion are part of our company's **commitment** **to equality**. Behind this focus is a simple but powerful idea: That diversity simply **is the right thing to do**. In sum, we firmly believe that diversity and inclusion **can help our organization to achieve equal opportunity for all***.*

Control condition (first webpage)

Diversity and inclusion are part of our company's **commitment**. Behind this focus is a simple but powerful idea: That diversity simply **is all around us**. In sum, we firmly believe that diversity and inclusion **have a place in our organization**.

## Next, participants were presented with a second webpage, which repeated parts of the previous message that participants had read on the first webpage:

## Business case condition (second webpage)

“Ultimately, diversity simply **makes good business sense***.* We firmly believe that diversity and inclusion **can help our organization to achieve superior performance***.*”

## Fairness case condition (second webpage)

“Ultimately, diversity simply **is the right thing to do**. We firmly believe that diversity and inclusion **can help our organization to achieve equal opportunity for all***.*”

## Control condition (second webpage)

“Ultimately, diversity simply **is all around us**. We firmly believe that diversity and inclusion **have a place in our organization**.”

After reading the manipulation (see Figure S5), participants responded to the measures described below, as well as a manipulation check.

*Figure S5.* Manipulations in Study 6.

***Business case condition***

***Graphical user interface, application

Description automatically generated***

***Graphical user interface, application

Description automatically generated***

*Figure S5 (Continued).*

***Fairness case condition***

*Graphical user interface, application

Description automatically generated*

*Graphical user interface, application, website

Description automatically generated*

*Figure S5 (Continued).*

***Control condition***

**Graphical user interface, application

Description automatically generated**

**Graphical user interface, text, application

Description automatically generated**

## Measures.

***Manipulation Check - Perceived Instrumentality vs. Non-Instrumentality*.** Participants were asked to indicate the extent to which the diversity case they had read communicated an instrumental vs. non-instrumental approach to diversity by completing the following statement: “The company sees diversity as core to...” and using a bipolar scale ranging from 1 “Business performance” to 6 “Fairness and equality”.

***Affective Reactions.*** Participants were asked to complete a PANAS scale (Watson et al., 1988) to measure the extent to which they felt positive and negative affect following the experimental manipulation. Sample items to measure positive affect were: “Enthusiastic”, “Interested” (10 items, α = 0.93), and for negative affect were: “Distressed”, “Nervous” (10 items, α = 0.96). Participants responded using a Likert scale ranging from 1 “Very slightly or not at all” to 5 “Extremely”.

***Liking of the Case.*** Participants indicated the extent to which they liked the diversity case, using a Likert scale ranging from 1 “Strongly dislike” to 7 “Strongly like”.

***Interest in the Case.*** Participants indicated the extent to which they found the diversity case interesting, using a Likert scale ranging from 1 “Not at all interesting” to 5 “Very interesting”.

***Perceived Awkwardness of the Case.*** Participants indicated the extent to which they found the diversity case awkward, using a Likert scale ranging from 1 “Not at all awkward” to 5 “Very awkward”.

***Perceived Length.*** Participants indicated how long they perceived the diversity case to be, using a Likert scale ranging from 1 “Not at all long” to 5 “Very long”.

***Perceived Difficulty.*** Participants indicated how difficult they perceived the diversity case to be, using a Likert scale ranging from 1 “Not at all difficult” to 5 “Very difficult”.

***Perceptions of the Organization as Diverse.*** Participants responded to 3 items measuring the extent to which they agreed with the organization that had issued the diversity case was likely diverse. A sample item was: “InvoCorp likely has a high level of diversity” (α = 0.86). Participants responded using a Likert scale ranging from 1 “Strongly disagree” to 7 “Strongly agree”.

# *Results*

**Main Effects of Condition.**

Our lab’s policy is to verify that the demographics measured in experiments are balanced across conditions, to ensure that the core assumption of random assignment in experimental design is valid. We found no significant different across conditions for any of the demographic variables measured in this pilot study, and therefore did not include any controls in our analyses of the main effects of Condition.

***Manipulation Check - Perceived Instrumentality vs. Non-Instrumentality.*** We conducted a one-way ANOVA to investigate the effect of condition on perceived instrumentality vs. non-instrumentality. We found a significant main effect of condition (*F*(2, 283) = 110.21, *p* < 0.001, = 0.438, 90% CI [0.366, 0.495]). Pairwise comparisons revealed that participants in the business case condition perceived the diversity case as significantly more instrumental (or less non-instrumental; *MBusiness* = 2.19, *SD* = 1.66, 95% CI [1.90, 2.47]) compared to their counterparts in the control case condition (*MControl* = 4.43, *SD* = 1.33, 95% CI [4.15, 4.71], *t*(283) = -10.97,  *p* < 0.001 , *d* = -1.58, 95% CI = [-1.90, -1.21]) and in the fairness case conditions (*MFairness* = 5.12, *SD* = 1.24,95% CI = [4.83, 5.41], *t*(283) = -14.13, *p* < 0.001, *d* = -2.06, 95% CI = [-2.41, -1.71]). There was also a significant difference in perceived instrumentality between the control and fairness conditions (*t*(283) = -3.35, *p* = 0.001, *d* = -0.49, 95% CI = [-0.77, -0.20]).

***Affective Reactions.*** A one-way ANOVA showed no significant main effect of condition on positive affect (*F*(2, 283) = 0.91, *p* = 0.40, = 0.006, 90% CI [0.000, 0.025], *MBusiness* = 3.27, *SD* = 0.92, 95% CI [3.07, 3.46], *MControl* = 3.09, *SD =* 0.94, 95% CI [2.90, 3.28], *MFairness* = 3.12, *SD* = 1.04, 95% CI [2.92, 3.32]), or on negative affect (*F*(2, 283) = 0.25, *p* = 0.78, = 0.002, 90% CI [0.000, 0.011], *MBusiness* = 1.33, *SD* = 0.79, 95% CI [1.19, 1.47], *MControl* = 1.28, *SD* = 0.67, 95% CI [1.14, 1.42], *MFairness* = 1.26, *SD* = 0.63, 95% CI [1.11, 1.40]).

***Liking of the Case.*** We found no significant main effect of condition on the extent to which participants liked the diversity case (*F*(2, 283) = 0.77, *p* = 0.46, = 0.005, 90% CI [0.000, 0.023], *MBusiness* = 6.12, *SD* = 1.77, 95% CI [5.76, 6.49], *MControl* = 5.82, *SD* = 1.94, 95% CI [5.45, 6.18], *MFairness* = 5.88, *SD* = 1.73, 95% CI [5.51, 6.25]).

***Interest in the Case.*** There was no significant main effect of condition on the extent to which participants found the diversity case to be interesting (*F*(2, 283) = 2.13, *p* = 0.12, = 0.015, 90% CI [0.000, 0.042], *MBusiness* = 3.06, *SD* = 1.32, 95% CI [2.80, 3.32], *MControl* = 2.78, *SD* = 1.24, 95% CI [2.52, 3.03], *MFairness* = 2.70, *SD* = 1.31, 95% CI [2.43, 2.96]).

***Perceived Awkwardness of the Case.*** There was no significant main effect of condition on the extent to which participants found the diversity case to be awkward (*F*(2, 283) = 0.31, *p* = 0.73, = 0.002, 90% CI [0.000, 0.013], *MBusiness* = 1.78, *SD* = 1.16, 95% CI [1.54, 2.02], *MControl* = 1.91, *SD* = 1.29, 95% CI [1.67, 2.15], *MFairness* = 1.89, *SD* = 1.18, 95% CI [1.64, 2.14]).

***Perceived Length.*** There was no significant main effect of condition on the extent to which participants found the diversity case to be long (*F*(2, 283) = 1.21, *p* = 0.30, = 0.008, 90% CI [0.000, 0.030], *MBusiness* = 1.49, *SD* = 0.86, 95% CI [1.31, 1.67], *MControl* = 1.51, *SD* = 1.05, 95% CI [1.33, 1.69], *MFairness* = 1.33, *SD* = 0.71, 95% CI [1.14, 1.51]).

***Perceived Difficulty.*** We found no significant main effect of condition on the extent to which participants found the diversity case to be difficult (*F*(2, 283) = 2.09, *p* = 0.13, = 0.015, 90% CI [0.000, 0.041], *MBusiness* = 1.17, *SD* = 0.68, 95% CI [1.02, 1.32], *MControl* = 1.38, *SD* = 0.83, 95% CI [1.23, 1.52], *MFairness* = 1.23, *SD* = 0.70, 95% CI [1.08, 1.38]).

***Perceptions of the Organization as Diverse.*** We found no significant main effect of condition on the extent to which participants perceived the organization issuing the diversity case as diverse (*F*(2, 283) = 0.37, *p* = 0.69, = 0.003, 90% CI [0.000, 0.015], *MBusiness* = 5.32, *SD* = 1.38, 95% CI [5.03, 5.61], *MControl* = 5.15, *SD* = 1.56, 95% CI [4.86, 5.44], *MFairness* = 5.18, *SD* = 1.45, 95% CI [4.88, 5.48]).

**Condition x Race Interactions.**

To ensure that all results presented below are robust, we controlled for the demographic characteristic that varied across conditions despite random assignment to conditions (none), as well as for the demographic characteristics that varied across racial groups (Age), as per our lab policy. The results remain unchanged without controls.

***Manipulation Check - Perceived Instrumentality vs. Non-Instrumentality.*** We conducted a two-way ANCOVA to test the effect of Condition x Race on perceived instrumentality vs. non-instrumentality, controlling for Age. We found a significant interaction effect (*F*(2, 279) = 3.84, *p* = 0.023, = 0.027, 90% CI [0.002, 0.061]). Pairwise comparisons showed that as expected, White participants in the business case condition perceived the diversity case as significantly more instrumental (or less non-instrumental; *MBusiness* = 2.06, *SE* = 0.20, 95% CI [1.67, 2.45]) than their White counterparts in the control case condition (*MControl* = 4.51, *SE* = 0.20, 95% CI [4.12, 4.90], *t*(279) = -8.78,  *p* < 0.001 , *d* = -1.74, 95% CI = [-2.26, -1.22]) and in the fairness case condition (*MFairness* = 5.51, *SE* = 0.20,95% CI = [5.11, 5.90], *t*(279) = -12.37, *p* < 0.001, *d* = -2.45, 95% CI = [-3.06, -1.83]). Similarly, African American participants in the business case condition perceived the diversity case as significantly more instrumental (or less non-instrumental; *MBusiness* = 2.33, *SE* = 0.21, 95% CI [1.92, 2.75]) compared to their African American counterparts in the control case condition (*MControl* = 4.34, *SE* = 0.21, 95% CI [3.94, 4.75], *t*(279) = -6.84,  *p* < 0.001 , *d* = -1.43, 95% CI = [-1.93, -0.92]) and in the fairness case conditions (*MFairness* = 4.64, *SE* = 0.22,95% CI = [4.20, 5.08], *t*(279) = -7.55, *p* < 0.001, *d* = -1.64, 95% CI = [-2.18, -1.08]). However, while White participants in the fairness case condition perceived the diversity case as significantly more non-instrumental than their counterparts in the control case condition (*t*(279) = -3.55, *p* < 0.001, *d* = -0.71, 95% CI = [-1.12, -0.29]), African American participants did not perceive any significant difference in the diversity case's non-instrumentality across the control and fairness conditions (*t*(279) = -0.99, *p* = 0.32, *d* = -0.21, 95% CI = [-0.63, 0.21]).

***Affective Reactions.*** A two-way ANCOVA, controlling for Age, showed no significant effect of Condition x Gender on positive affect (*F*(2, 279) = 0.72, *p* = 0.49, = 0.005, 90% CI [0.000, 0.022]), or on negative affect (*F*(2, 279) = 0.17, *p* = 0.84, = 0.001, 90% CI [0.000, 0.009]).

***Liking of the Case.*** We found no significant effect of Condition x Race on the extent to which

participants liked the diversity case (*F*(2, 279) = 0.36, *p* = 0.70, = 0.003, 90% CI [0.000, 0.015]).

***Interest in the Case.*** There was no significant effect of Condition x Race on the extent to which participants found the diversity case to be interesting (*F*(2, 279) = 0.23, *p* = 0.79, = 0.002, 90% CI [0.000, 0.011]).

***Perceived Awkwardness of the Case.*** There was no significant effect of Condition x Race on the extent to which participants found the diversity case to be interesting (*F*(2, 279) = 0.01, *p* = 0.99, < 0.001, 90% CI [0.000, 1.000]).

***Perceived Length.*** There was no significant effect of Condition x Race on the extent to which participants found the diversity case to be long (*F*(2, 279) = 0.15, *p* = 0.86, = 0.001, 90% CI [0.000, 0.008]).

***Perceived Difficulty.*** We found no significant effect of Condition x Race on the extent to which participants found the diversity case to be difficult (*F*(2, 279) = 0.82, *p* = 0.44, = 0.006, 90% CI [0.000, 0.024]).

***Perceptions of the Organization as Diverse.*** We found no significant effect of Condition x Race on the extent to which participants perceived the organization issuing the diversity case as diverse (*F*(2, 279) = 1.03, *p* = 0.36, = 0.007, 90% CI [0.000, 0.028]).

***Discussion***

The manipulation did not have any significant effect on the variables measured, beyond the manipulation check. This suggests that the two cases did not trigger significantly different affective reactions, that participants equally liked the two cases, found them of comparable interest, awkwardness, length and difficulty, and did not perceive the organization issuing the diversity case as differentially diverse depending on the type of diversity case. There was, however, evidence that the fairness case was perceived differently across racial groups relative to the control case – Whites perceived it as significantly more non-instrumental than the control case, whereas African Americans perceived no significant difference.

**References**

Analytics Vidhya. (2020). *How to improve class imbalance using class weights in machine learning*. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2020/10/improve-class-imbalance-class-weights/

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46.

Eagly, A. H. (2016). When passionate advocates meet research on diversity, does the honest broker stand a chance? *Journal of Social Issues*, *72*(1), 199–222. https://doi.org/10.1111/josi.12163

Fernández, A., García, S., Galar, M., Prati, R. C., Krawczyk, B., & Herrera, F. (2018). Foundations on imbalanced classification. In *Learning from imbalanced data sets* (pp. 19–46). Springer International Publishing. https://doi.org/10.1007/978-3-319-98074-4\_2

Hayes, A. F. (2013). *An introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* Guilford Press.

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

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2017). *An introduction to statistical learning: With applications in R*. Springer.

Lever, J., Krzywinski, M., & Altman, N. (2016). Classification evaluation. *Nature Methods*, *13*(8), 603–604. https://doi.org/10.1038/nmeth.3945

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user’s guide.* (C. D. Creelman (ed.)). Cambridge University Press.

Pinel, E. C. (1999). Stigma consciousness: The psychological legacy of social stereotypes. *Journal of Personality and Social Psychology*, *76*(1), 114–128. https://doi.org/10.1037/0022-3514.76.1.114

Provost, F. (2000). Machine learning from imbalanced data sets 101. *Proceedings of the AAAI’2000 Workshop on Imbalanced Data Sets*, *68*(2000), 1–3.

Rattan, A., & Dweck, C. S. (2010). Who confronts prejudice?: The role of implicit theories in the motivation to confront prejudice. *Psychological Science*, *21*(7), 952–959. https://doi.org/10.1177/0956797610374740

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070.

1. As a reference point, an algorithm classifying arguments at random on our unbalanced test set would have a 30% F1-score, and a 30% error rate. [↑](#footnote-ref-1)
2. The characteristics of the algorithm on the testing set (greater precision than recall, greater “miss” rate than “false alarm” rate) are likely to extend to some extent to the classification of Fortune 500. The precise extent to which it does, however, is an open question, given that the MBA arguments in the testing set and the Fortune 500 diversity cases are two different bodies of text. [↑](#footnote-ref-2)
3. As suggested above, the characteristics of the algorithm on the testing set (greater recall than precision, greater “false alarm” rate than “miss” rate) are likely to extend to some extent to the classification of Fortune 500. The precise extent to which it does, however, is an open question, given that the MBA arguments in the testing set and the Fortune 500 diversity cases are two different bodies of text. [↑](#footnote-ref-3)
4. Three items from Pinel’s (1999) scale could not meaningfully be adapted to suit both women and men participants, and were thus not included in our survey. [↑](#footnote-ref-4)
5. The name and logo of this fictional organization were taken from Wilton et al. (2019). [↑](#footnote-ref-5)