

A. Methods

A.1. Cognitive Interviews

We recruited 9 participants from the DC Metropolitan area using Craigslist. We required participants to be over 18 years of age and fluent in English. Participants ranged between the ages of 20 and 66. These interviews took place on the University of Maryland campus and lasted about 1 hour. All participants signed a written consent form prior to the interview, and were paid \$30 for their time.

During these interviews, participants completed a preliminary version of the survey used in Study-1. After each survey question, we asked the participants several interview questions related to their comprehension of and feelings toward the survey. We found that some participants tended to use their own personal notions of fairness when answering comprehension questions rather than using the definition we provided. We were concerned that this would limit our ability to effectively measure comprehension. To address this problem, we rewrote several parts of our survey and added two new questions (Q14 and Q15).

A.2. Non-Expert Verification

We designed this study to assess *non-expert* understanding and opinions of ML fairness metrics. To this end, we asked respondents to self-rate their level of expertise in a variety of fields, including ML, at the end of the survey (see Appendix D.3). A number of participants did report having “expert” level experience in ML ($n = 2$ out of 147 in Study-1, and $n = 15$ out of 349 in Study-2). We considered removing these participants from the analyses, but ultimately did not because there was no relationship between self-reported ML expertise and comprehension score (Spearman’s rho, for both studies).

B. Study-1: Detailed Results

B.1. Our Survey Effectively Captures Rule Comprehension

We find that our survey is internally consistent, and effectively measures participant comprehension of demographic parity. The former we evaluated using Cronbach’s α and item-total correlation (discussed in §4.1.1), and the latter using two self-report measures and one free response question. See Fig. 7 for participant performance per question.

B.1.1. SELF-REPORTED RULE UNDERSTANDING AND USE ARE REFLECTED IN COMPREHENSION SCORE

First, we compared comprehension score to self-reported rule understanding (Q13). Higher comprehension scores were associated with greater confidence in understanding

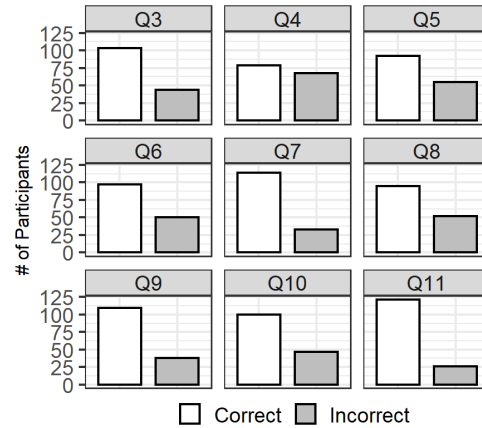


Figure 7. Number of participants answering each question correctly. Each panel contains all 147 participants.

(Spearman’s rho), suggesting that participants were accurately assessing their ability to apply the rule (see Fig. 8).

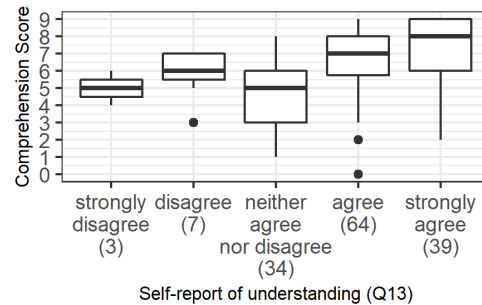


Figure 8. Comprehension score grouped by response to Q13. Self-reported understanding of the rule was associated with higher comprehension scores. X-axis is reversed for figure and correlation test.

Next, we compared comprehension score to a self-report question about the participant’s use of the rule (Q14) Participants who claimed to use only the rule tended to score higher than those who used their own notions of fairness or a combination thereof (K-W test, and post-hoc M-WU), suggesting that participants are answering somewhat honestly: when they try to apply the rule, comprehension scores improve (see Fig. 9).

B.1.2. PARTICIPANTS WITH HIGHER COMPREHENSION SCORES ARE BETTER ABLE TO EXPLAIN THE RULE

To further validate our comprehension score, we asked participants to explain the rule in their own words (Q12). Responses were qualitatively coded as one of five categories: **correct**, **partially correct**, **neither**, **incorrect**, or **none** (as discussed in §4.1.1). The results of this coding can be seen in Fig. 10. Participants providing correct ex-

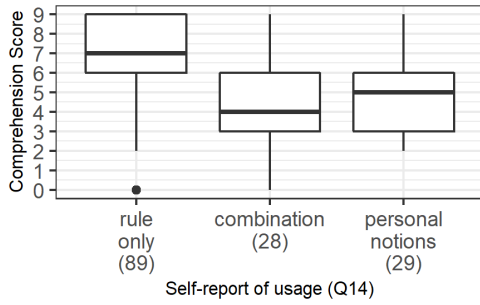


Figure 9. Comprehension score grouped by response to Q14. Rule compliance (leftmost on the x-axis) was associated with higher comprehension scores. One participant who did not provide a response was excluded from the figure and relevant analysis.

planations of the rule attained higher comprehension scores (k-W test, and post-hoc M-WU), further corroborating our claim that our comprehension score is a valid measure of fairness rule comprehension.

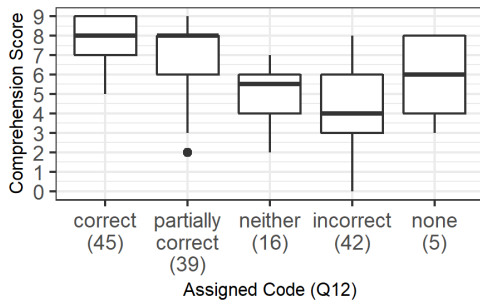


Figure 10. Comprehension score grouped by code assigned to Q12 response. Participants who provided either correct or partially correct responses tended to perform better.

B.2. Education Influences Comprehension

During the cognitive interview phase, we observed a possible trend of comprehension scores being lower for older participants and those with less educational attainment. If true, this would suggest that fairness explanations should be carefully validated to ensure they can be used with diverse populations. We investigated this hypothesis, in an exploratory fashion, using poisson regression models.

Three models were tested. The first regressed score against all four demographic categories as predictors (gender, age, ethnicity, and education), the second omitted education, and the third tested only education. Models were compared using Akaike information criterion (AIC), a standard method of evaluating model quality and performing model selection (Akaike, 1974). Comparison by AIC revealed that model 1 (all four categories) was a better predictor for comprehension score than models 2 or 3 (AIC = 643.3, 651.2, and 660.5, respectively; difference = 0.0, 7.9, and 17.1). In model 1, only education showed correlation with compre-

hension score (effect size = 1.40, $p < 0.05$). Further work is needed to confirm this exploratory result.

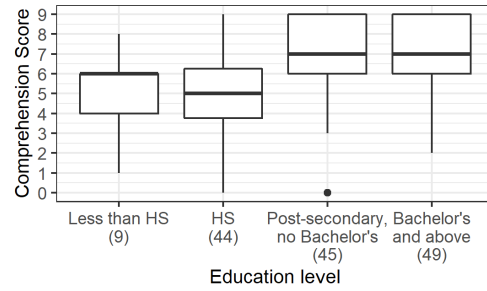


Figure 11. Comprehension score grouped by education level. Higher education level was associated with higher comprehension scores.

B.3. Disagreement with the Rule is Associated with Higher Comprehension Scores

Participants were asked for their opinion on the presented rule in another free response question (Q15). These responses were then qualitatively coded to capture participant sentiment towards the rule as one of five categories: **agree**, **depends**, **disagree**, **not understood**, or **none** (as discussed in §4.1.2).

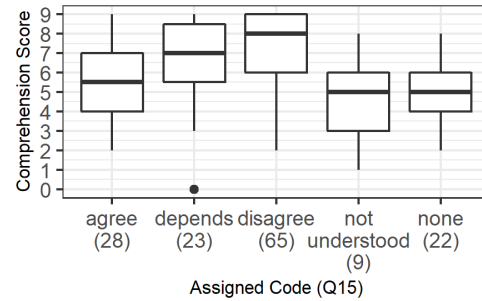


Figure 12. Comprehension score grouped by code assigned to Q15 response. Participants who exhibited negative sentiment toward the rule responses tended to perform better.

This question was added based on the cognitive interviews (see Appendix A.1), where perception seemed to influence compliance. The results of coding Q15 can be seen in Fig. 12. Participants who expressed disagreement with the rule performed better than those who expressed agreement, did not understand the rule, or provided no response to the question (K-W test, post-hoc M-WU). Note that this result should not be interpreted as an overall finding on the appropriateness of demographic parity. Instead we anticipate the perceptions of appropriateness of any fairness definition will be highly context-dependent.

B.4. Non-Compliance is Associated with Lack of Understanding

We were interested in understanding why some participants failed to adhere to the rule, as measured by their self-report of rule usage in Q14. After labeling participants as either “non-compliant” (NC, $n = 57$) or “compliant” (C, $n = 89$), we conducted a series of χ^2 tests to investigate this phenomenon.

Non-compliant participants were less likely to self-report high understanding of the rule in Q13 (see Fig. 13). Moreover, non-compliance also appears to be associated with a reduced ability to correctly explain the rule in Q12 (see Fig. 14). Further, negative participant sentiment towards the rule (Q15) also appears to be associated with greater compliance (see Fig. 15). Thus, non-compliant participants appear to behave this way because they do not *understand* the rule, rather than because they do not *like* it.

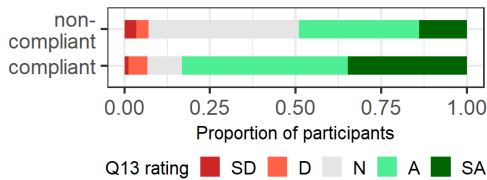


Figure 13. Self-report of understanding (Q13) split by compliance (Q14). NC participants tend to report less confidence in their ability to apply the rule. SD = strongly disagree, D = disagree, N = neither agree nor disagree, A = agree, SA = strongly agree.

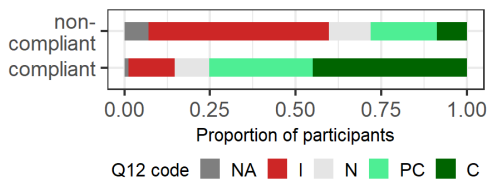


Figure 14. Correctness of rule explanation (Q12) split by compliance (Q14). NC participants tend to be less able to explain the presented rule in their own words. NA = none, I = incorrect, N = neither, PC = partially correct, C = correct.

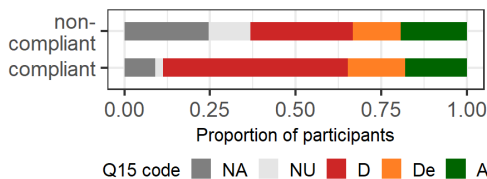


Figure 15. Participant agreement with rule (Q15) split by compliance (Q14). NC participants tend to harbor less negative sentiment towards the rule. NA = none, NU = not understood, D = disagree, De = depends, A = agree.

B.5. Decision Scenarios

For Study-1 we designed three decision-making scenarios to test whether the perceived importance or realism of a particular scenario influenced comprehension score. They are as follows:

- **Art Project (AP):** distributing awards for art projects to primary school students,
- **Employee Awards (EA):** distributing employee awards at a sales company, and
- **Hiring (HR):** distributing job offers to applicants.

In each scenario the students/employees/applicants are partitioned into two groups (parents’ occupation for the first scenario, and binary gender for the other two scenarios). We use a between-subjects design: participants are randomly partitioned into three conditions, one for each scenario (AP, EA, or HR). For each condition we define the *fairness rule* in the context of the decision-making scenario (see Appendix D for the full surveys).

Next we describe our main conclusion related to the different decision-making scenarios in Study-1: the scenario does not influence comprehension score.

B.5.1. SCENARIO DOES NOT INFLUENCE COMPREHENSION SCORES (RQ4)

We were concerned that less important and/or realistic scenarios would cause participants to take the survey less seriously, and therefore perform more poorly. To test this, participants were randomly assigned to a scenario, resulting in the following distribution: AP = 41, EA = 49, HR = 57.

A K-W test revealed no differences between scenarios in terms of comprehension score (mean comprehension scores: AP = 6.0, EA = 6.74, HR = 5.86). However, differences did exist between scenarios in terms of importance (assessed in Q2), measured in hours of effort deemed necessary to make the relevant decision (K-W, $p < 0.001$). Post-hoc M-WU revealed that participants believed making a decision in the AP scenario merited fewer hours of effort (mean = 3.15hrs) than in the EA (13.52hrs, $p < 0.001$) or HR (15.23hrs, $p < 0.001$) scenarios (corrected $\alpha = 0.05/3 = 0.017$). See Fig. 16 for distributions of responses.

Of note, it is possible that perceived realism, assessed in Q1 on a five-point Likert scale, was also influenced by scenario (K-W, $p = 0.051$), but we may need larger sample sizes to confirm this. Regardless, while the nature of a scenario does influence participant perception in terms of importance and (possibly) realism, it does not appear to influence comprehension (at least for the scenarios we chose). For this reason, we chose to test a single scenario (HR) in Study-2.

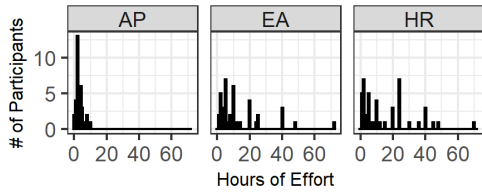


Figure 16. Importance of a scenario by proxy of hours of effort necessary to make a decision in each scenario. AP merited less hours of effort than both EA and HR.

C. Study-2: Detailed Results

C.1. Model Selection

In §4.2.2 we assessed eleven linear regression models for predicting comprehension scores. The best fit model, determined by model selection via AIC, included only education (edu) and fairness definition (def) as regressors. The results of model selection are below in Table 4.

Table 4. Models tested in §4.2.2, sorted by best to least fit. The first model in the table (edu + def) is the model of best fit. dAIC = difference from model with lowest AIC value.

Model regressors	AIC	dAIC
edu + def	-80.4	0
edu	-72.8	7.6
gender + edu	-70.3	10.1
age + edu	-63.7	16.7
gender + age + edu	-61.1	19.2
gender + age + eth + edu + def	-61.1	19.2
def	-60.8	19.6
gender + age + eth + edu	-55.5	24.9
gender + age + def	-46.4	34
gender + age + eth + def	-41.6	38.8
gender + age + eth	-37.2	43.2

C.2. Non-Compliance

In §4.2.4 we sought to further investigate the findings of Study-1 with regards to compliance (Q14). To do so, we labeled those who responded (in Study-2) with either having used their own personal notions of fairness ($n = 26$) or some combination of their personal notions and the rule ($n = 148$) as “non-compliant” (NC), with the remaining $n = 174$ labeled as “compliant” (C). One participant who did not provide a response was excluded from this analysis, conducted using KW and χ^2 tests.

Non-compliant participants were less likely to self-report high understanding of the rule in Q13 (KW test, $p < 0.001$, see Fig. 17). Moreover, non-compliance also appears to be associated with a reduced ability to correctly explain the rule in Q12 (χ^2 test, $p < 0.001$, see Fig. 18). This fits with the overall strong relationship we observed among comprehension scores, ability to explain the rule, and compliance.

Further, greater dislike towards the rule (Q15) also appears to be associated with greater compliance (KW test, $p < 0.05$, see Fig. 19). However, there was no relationship between disagreement towards the rule (Q16) and compliance (see Fig. 20).

These results largely corroborate the notion that non-compliant participants appear to behave this way because they do not *understand* the rule, rather than because they do not *like* it.

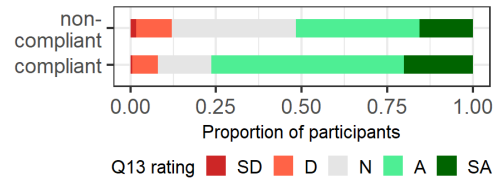


Figure 17. Self-report of understanding (Q13) split by compliance (Q14). NC participants tend to report less confidence in their ability to apply the rule. SD = strongly disagree, D = disagree, N = neither agree nor disagree, A = agree, SA = strongly agree.

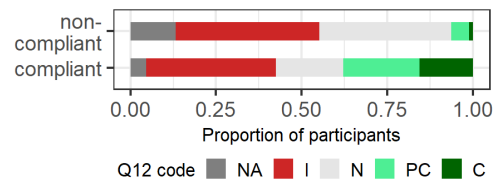


Figure 18. Correctness of rule explanation (Q12) split by compliance (Q14). NC participants tend to be less able to explain the presented rule in their own words. NA = none, I = incorrect, N = neither, PC = partially correct, C = correct.

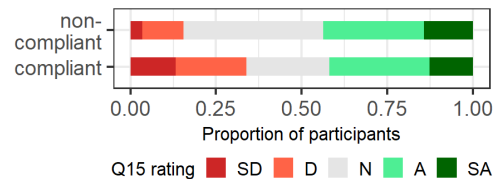


Figure 19. Participant liking for rule (Q15) split by compliance (Q14). NC participants tend to dislike the rule less than C participants. SD = strongly disagree, D = disagree, N = neither agree nor disagree, A = agree, SA = strongly agree.

D. Surveys

D.1. Study-1 Survey

Each of the surveys are split into four main sections. The first section is the consent form which can be found in Appendix E. The second section describes the scenario and asks questions about the given scenario (§D.1.1). The third section describes the fairness metric, defined as the rule, used (in this case it is demographic parity) and asks specific questions about the metric (§D.1.2). Finally the last section asks for demographic information (§D.3).

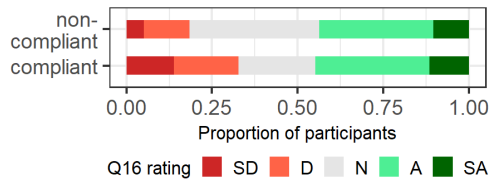


Figure 20. Participant agreement with rule (Q16) split by compliance (Q14). No differences were found between NC and C participants. SD = strongly disagree, D = disagree, N = neither agree nor disagree, A = agree, SA = strongly agree.

D.1.1.1. SCENARIO DESCRIPTIONS AND QUESTIONS

The following is shown to each participant:

It is very important that you read each question carefully and think about your answers. The success of our research relies on our respondents being thoughtful and taking this task seriously.

- I have read the above instructions carefully.

We then introduce one of three different decision making scenarios, described below, followed by two questions. Words that vary across scenario in the questions are shown as <art project, employee awards, hiring>.

Art project A fourth grade teacher is reviewing 20 student art projects. They will award lollipops to the top 4 students who put the most effort into their projects. The teacher knows that some of the students have artists as parents, who might have helped their children with their art project. The teacher’s goal is to give out lollipops only based on the amount of effort that the student *themselves* put into their projects.

The teacher uses the following criteria to decide who should get a lollipop:

- Elaborateness of each project.
- Creativity of each project.

About 50% of the students have artists as parents, and 50% do not.

In the past, students with artists as parents typically put more effort into their projects.

In this group of students there is a wide range of project quality (as measured by elaborateness and creativity). However, this range of quality is about the same between students with artists as parents and those without.

The teacher wants to make sure that they award lollipops in a fair way, no matter whether the students’ parents are artists or not.

Employee awards A manager at a sales company is deciding which of their 100 employees should receive each of 10 mid-year awards. The manager’s goal is to give awards to employees who *will* have high net sales at the end of the year.

The manager uses the following criteria to decide who should get an award:

- Recent performance reviews
- Mid-year net sales
- Number of years on the job

About 50% of the employees are men, and 50% are women.

In the past, men have achieved higher end-of-year net sales than women.

In this group of employees, there is a wide range of qualifications (as measured by performance reviews, mid-year net sales, and number of years on the job). However, this range of qualifications is about the same between male and female employees.

The manager wants to make sure that this awards process is fair to the employees, no matter their gender.

Hiring A hiring manager at a new sales company is reviewing 100 new job applications. Each applicant has submitted a resume, and has had an interview. The manager will send job offers to 10 out of the 100 applicants. Their goal is to make offers to applicants who will have high net sales after a year on the job.

The manager will use the following to decide which applicants should receive job offers:

- Interview scores
- Quality of recommendation letters
- Number of years of prior experience in the field

About 50% of the applicants are men, and 50% are women.

In the past, men have achieved higher net sales than women, after one year on the job.

In this applicant pool there is a wide range of applicant quality (as measured by interview scores, recommendation letters, and years of prior experience in the field). However, the range of quality is about the same for both male and female applicants.

The hiring manager wants to make sure that this hiring process is fair to applicants, no matter their gender.

Questions

1. To what extent do you agree with the following statement: a scenario similar to the one described above might occur in real life.
 - Strongly agree
 - Agree
 - Neither agree nor disagree
 - Disagree
 - Strongly Disagree
2. How much effort should the <teacher, manager, hiring manager> put in to make sure this decision is fair? [short answer - number of hours]

D.1.2. RULE DESCRIPTIONS AND QUESTIONS

Unless otherwise noted the rule description is shown above each of the questions for reference. Correct answers are noted in red.

Art project The teacher uses the following award rule to distribute lollipops: *The fraction of students who receive lollipops that have artist parents should equal the fraction of students in the class that have artist parents. Similarly, the fraction of students who receive lollipops that do not have artist parents should equal the fraction of students in the class that do not have artist parents.*

Example 1: If 10 out of the 20 students in the class have artist parents, then 2 out of the 4 lollipops would be awarded to students with artist parents (and the remaining 2 would be awarded to students without artist parents).

Example 2: If 5 out of the 20 students in the class have artist parents, then 1 out of the 4 lollipops would be awarded to students with artist parents (and the remaining 3 would be awarded to students without artist parents).

In the next section, we will ask you some questions about the information you have just read. Please note that this is not a test of your abilities. We want to measure the quality of the description you read, not your ability to take tests or answer questions.

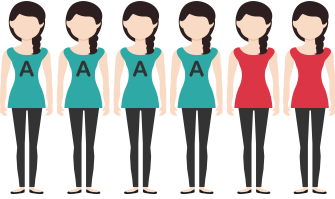
Please note that we ask you to apply and use ONLY the above award rule when answering the following questions. You will have an opportunity to state your opinions and feelings on the rule later in the survey.

3. Suppose a different teacher is considering awarding lollipops to the whole 4th grade. There are 100 students with artist parents, and 200 students without artist parents. The teacher decides to award 10 lollipops to students with artist parents. **Assuming the teacher is**

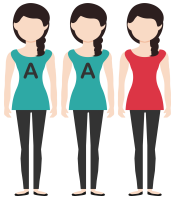
required to use the award rule above, how many students without artist parents need to receive lollipops?

- (a) 10
 - (b) 20
 - (c) 40
 - (d) 50
4. **Assuming the teacher is required to use the award rule above**, in which of these cases can a teacher award more lollipops to students without artist parents than to students with artist parents?
 - (a) When the students without artist parents have higher-quality projects (i.e., more elaborate and more creative) than those with artist parents.
 - (b) **When there are more students without artist parents than those with artist parents.**
 - (c) When students without artist parents have more creative projects than those with artist parents.
 - (d) This cannot happen under the award rule.
 5. **Assuming the teacher is required to use the award rule above**, is the following statement **TRUE OR FALSE**: Even if a student with artist parents has a project that is of the same quality (i.e., equally elaborate and equally creative) as another project by a student without artist parents, they can be treated differently (i.e., only one of the students might get a lollipop).
 6. **Assuming the teacher is required to use the award rule above**, is the following statement **TRUE OR FALSE**: If all students without artist parents have low-quality projects (i.e., low elaborateness and low creativity), but the teacher awards lollipops to some of them, then any lollipops awarded to students with artist parents must be awarded to those who have low-quality projects.
 7. **Assuming the teacher is required to use the award rule above**, is the following statement **TRUE OR FALSE**: Suppose the teacher is distributing 10 lollipops amongst a pool of students that includes students with and without artist parents. Even if all students with artist parents have low-quality (i.e., low elaborateness and low creativity) projects, some of them must still receive lollipops.
 8. **Assuming the teacher is required to use the award rule above**, is the following statement **TRUE OR FALSE**: This award rule always allows the teacher to award lollipops exclusively to the students who have the highest quality (i.e., most elaborate and most creative) projects.

In the two examples above there are 20 students. Consider a different scenario, with **6 students – 4 with artist parents and 2 without, as illustrated below**. The next three questions each give a potential outcome for all six students (i.e., which of the 6 students receive awards). Please indicate which of the outcomes follow **the award rule above**.

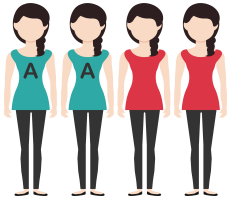


9. Alternative scenario 1:



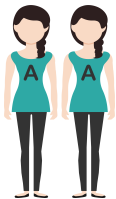
Does this distribution of awards obey the **award rule**?
Yes

10. Alternative scenario 2:



Does this distribution of awards obey the **award rule**?
No

11. Alternative scenario 3:



Does this distribution of awards obey the **award rule**?
No

12. In your own words, explain the **award rule**. [short answer] (The rule is not shown above this question)

13. To what extent do you agree with the following statement: I am confident I know how to **apply the award rule described above**?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly Disagree

14. Please select the choice that best describes your experience: When I answered the previous questions...

- (a) I applied the provided award rule only.
- (b) I used my own ideas of what the correct award decision should be rather than the provided award rule.
- (c) I used a combination of the provided award rule and my own ideas of what the correct award decision should be.

15. What is your opinion on the award rule? Please explain why. [short answer]

16. Suppose that you are the teacher whose job it is to distribute lollipops to students based on the criteria listed above (i.e., elaborateness of each project, creativity of each project). How would you ensure that this process is fair? [short answer]

17. Was there anything about this survey that was hard to understand or answer? [short answer]

Employee awards The manager uses the following award rule to distribute awards: *The fraction of employees who receive awards that are female should equal the fraction of employees that are female. Similarly, fraction of employees who receive awards that are male should equal the fraction of employees that are male.*

Example 1: If there are 50 female employees out of 100, then 5 out of the 10 awards should be awarded to female employees (and the remaining 5 would be made to male employees).

Example 2: If there are 30 female employees out of 100, then 3 out of the 10 awards should be awarded to female employees (and the remaining 7 would be made to male employees).

In the next section, we will ask you some questions about the information you have just read. Please note that this is not a test of your abilities. We want to measure the quality of the description you read, not your ability to take tests or answer questions.

Please note that we ask you to apply and use ONLY the above award rule when answering the following questions. You will have an opportunity to state your opinions and feelings on the rule later in the survey.

3. Suppose a different manager is considering employees for a different award. There are 100 male employees and 200 female employees, and they decide to give awards to 10 male employees. **Assuming the manager is required to use the award rule above**, how many female employees do they need to give awards to?

- (a) 10
- (b) 20
- (c) 40
- (d) 50

4. **Assuming the manager is required to use the award rule above**, in which of these cases can a manager give more awards to female employees than to male employees?

- (a) When there are more well-qualified female employees than well-qualified male employees (i.e., more women have better performance reviews, higher mid-year net sales, and more years on the job).
- (b) **When there are more female employees than male employees.**
- (c) When female employees receive higher performance reviews than male employees.
- (d) This cannot happen under the award rule.

5. **Assuming the manager is required to use the award rule above**, is the following statement **TRUE** OR **FALSE**: Even if a male employee's qualifications look similar to a female employee's (in terms of performance reviews, mid-year net sales, and years on the job), he can be treated differently (i.e., only one of the employees gets an award).

6. **Assuming the manager is required to use the award rule above**, is the following statement **TRUE** OR **FALSE**: If all female employees are unqualified (i.e., have low performance reviews, low mid-year net sales, and few years on the job), but you give awards to some of them, then awards given to male employees must be made to unqualified male employees.

7. **Assuming the manager is required to use the award rule above**, is the following statement **TRUE** OR **FALSE**: Suppose the manager is distributing 10 awards amongst a pool that includes both male and female employees. Even if all male employees are unqualified for an award (i.e., have low performance reviews, low mid-year net sales, and few years on the job), some of them must still receive awards.

8. **Assuming the manager is required to use the award rule above**, is the following statement **TRUE** OR

FALSE: This award rule always allows the manager to distribute awards exclusively to the most qualified employees (i.e., employees with better performance reviews, high mid-year net sales, and high number of years on the job).

In the two examples above there are 100 employees. Consider a different scenario, with **6 employees– 4 female and 2 male, as illustrated below**. The next three questions each give a potential outcome for all six employees (i.e., which of the 6 employees receive awards). Please indicate which of the outcomes follow **the award rule above**.



9. Alternative scenario 1:



Does this distribution of awards obey the **award rule**?
Yes

10. Alternative scenario 2:



Does this distribution of awards obey the **award rule**?
No

11. Alternative scenario 3:



Does this distribution of awards obey the **award rule**?
No

12. In your own words, explain the **award rule**. [short answer] (The rule is not shown above this question)
13. To what extent do you agree with the following statement: I am confident I know how to **apply the award rule described above**?
- Strongly agree
 - Agree
 - Neither agree nor disagree
 - Disagree
 - Strongly Disagree
14. Please select the choice that best describes your experience: When I answered the previous questions...
- (a) I applied the provided award rule only.
 - (b) I used my own ideas of what the correct award decision should be rather than the provided award rule.
 - (c) I used a combination of the provided award rule and my own ideas of what the correct award decision should be.
15. What is your opinion on the award rule? Please explain why. [short answer]
16. Suppose that you are the manager whose job it is to distribute mid-year awards to employees based on the criteria listed above (i.e., recent performance reviews, mid-year net sales, number of years on the job). How would you ensure that this process is fair? [short answer]
17. Was there anything about this survey that was hard to understand or answer? [short answer]

Hiring The hiring manager uses the following hiring rule to send out offers: *The fraction of applicants who receive job offers that are female should equal the fraction of applicants that are female. Similarly, fraction of applicants who receive job offers that are male should equal the fraction of applicants that are male.*

Example 1: If there are 50 female applicants out of the 100 applicants, then 5 out of the 10 offers would be made to female applicants (and the remaining 5 would be made to male applicants).

Example 2: If there are 30 female applicants out of the 100 applicants, then 3 out of the 10 offers would be made to female applicants (and the remaining 7 would be made to male applicants).

In the next section, we will ask you some questions about the information you have just read. Please note that this is not a test of your abilities. We want to measure the quality

of the description you read, not your ability to take tests or answer questions.

Please note that we ask you to apply and use ONLY the above hiring rule when answering the following questions. You will have an opportunity to state your opinions and feelings on the rule later in the survey.

3. Suppose a different hiring manager is considering applicants for a different job. There are 100 male applicants and 200 female applicants, and they decide to send offers to 10 male applicants. **Assuming the hiring manager is required to use the hiring rule above**, how many female applicants do they need to send offers to?
- (a) 10
 - (b) 20
 - (c) 40
 - (d) 50
4. **Assuming the hiring manager is required to use the hiring rule above**, in which of these cases can a hiring manager make more job offers to female applicants than to male applicants?
- (a) When there are more well-qualified female applicants than well-qualified male applicants (i.e., more women have higher interview scores, higher quality recommendation letters, and more years of prior experience in the field).
 - (b) **When there are more female applicants than male applicants.**
 - (c) When female applicants receive better interview scores than male applicants.
 - (d) This cannot happen under the hiring rule.
5. **Assuming the hiring manager is required to use the hiring rule above**, is the following statement **TRUE** OR **FALSE**: Even if a male applicant's qualifications look similar to a female applicant's (in terms of interview scores, recommendation letters, and years of prior experience in the field), he can be treated differently (i.e., only one of the applicants will receive a job offer).
6. **Assuming the hiring manager is required to use the hiring rule above**, is the following statement **TRUE** OR **FALSE**: If all female applicants are unqualified (i.e., have low interview scores, low-quality recommendation letters, and few years of prior experience in the field), but you send job offers to some of them, then any job offers made to male applicants must be made to unqualified male applicants.
7. **Assuming the hiring manager is required to use the hiring rule above**, is the following statement **TRUE**

OR FALSE: Suppose the hiring manager is sending out 10 job offers to a pool that includes male and female applicants. Even if all male applicants are unqualified (i.e., have low interview scores, low-quality recommendation letters, and few years of prior experience in the field), some of them must still receive job offers.



8. Assuming the hiring manager is required to use the hiring rule above, is the following statement TRUE OR FALSE: This hiring rule always allows the hiring manager to send offers exclusively to the most qualified applicants (i.e., applicants with high interview scores, high quality recommendation letters, and high number years of prior experience in the field).

Does this distribution of job offers obey the hiring rule? **No**

In the two examples above there are 100 applicants. Consider a different scenario, with **6 applicants – 4 female and 2 male, as illustrated below**. The next three questions each give a potential outcome for all 6 applicants (i.e., which of the 6 applicants receive job offers). Please indicate which of the outcomes follow **the hiring rule above**.

12. In your own words, explain the **hiring rule**. [short answer] (The rule is not shown above this question)
13. To what extent do you agree with the following statement: I am confident I know how to **apply the hiring rule described above**?
- Strongly agree
 - Agree
 - Neither agree nor disagree
 - Disagree
 - Strongly Disagree



9. Alternative scenario 1:



Does this distribution of job offers obey the hiring rule? **Yes**

10. Alternative scenario 2:



Does this distribution of job offers obey the hiring rule? **No**

11. Alternative scenario 3:

14. Please select the choice that best describes your experience: When I answered the previous questions...
- (a) I applied the provided hiring rule only.
 - (b) I used my own ideas of what the correct hiring decision should be rather than the provided hiring rule.
 - (c) I used a combination of the provided hiring rule and my own ideas of what the correct hiring decision should be.
15. What is your opinion on the hiring rule? Please explain why. [short answer]
16. Suppose that you are the hiring manager whose job it is to send job offers to applicants based on the criteria listed above (i.e., interview scores, quality of recommendation letters, number of years of prior experience in the field). How would you ensure that this process is fair? [short answer]
17. Was there anything about this survey that was hard to understand or answer? [short answer]

D.2. Study-2: Survey

Each of the surveys are split into four main sections. The first section is the consent form which can be found in Appendix E. The second section describes the hiring scenario and asks questions about it (§D.2.1). The third section describes the fairness metric, defined as the rule, used (in this case it is demographic parity) and asks specific questions about the metric (§D.2.2). Finally the last section asks for demographic information (§D.3).

D.2.1. SCENARIO DESCRIPTION AND QUESTIONS

The following is shown to each participant (note that Step 3 is not shown to participants with the DP definition):

It is very important that you read each question carefully and think about your answers. The success of our research relies on our respondents being thoughtful and taking this task seriously.

- I have read the above instructions carefully.

A company, Sales-a-lot, is reviewing their hiring process. They want to hire applicants who are high performing, and they also want to make sure that their hiring process is fair to their applicants, no matter their gender. To do this, Sales-a-lot employs an external firm, Recruit-a-matic, which keeps track of all applicants. This review will take place over one year.

For clarity at each stage of the hiring process we use images to represent the hiring pool.

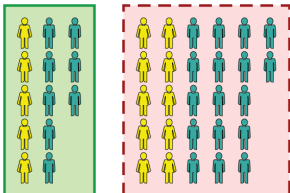
Step 1: Applicant Pool. At the beginning of the year, Sales-a-lot reviews all job applicants, and sends job offers to some of them. The initial applicant pool is shown with a gray background. For example, the following image shows an applicant pool with 15 female applicants and 25 male applicants:



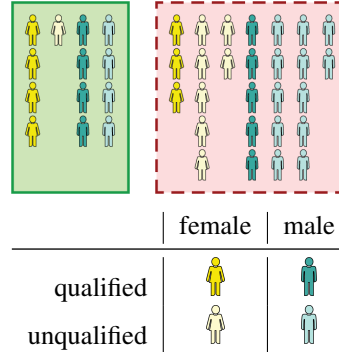
Step 2: Sending Job Offers. Next, Sales-a-lot sends job offers to some of these applicants, using the following criteria:

- Interview scores
- Quality of recommendation letters
- Number of years of prior experience in the field

Suppose that Sales-a-lot sends offers to 5 female applicants and 8 male applicants (so 10 female and 17 male applicants didn't receive offers). In the following image, applicants who received a job offer are shown on the left (with a green background) and applicants who didn't receive a job offer are shown on the right, with a red background):



Step 3: Applicant Evaluation. For the rest of the year, Recruit-a-matic (the external firm) keeps track of all applicants in the initial pool, whether they received job offers or not. At the end of the year, Recruit-a-matic finds out which applicants were high performers, i.e. qualified (shown in dark), and which applicants were low performers, i.e. unqualified (shown in light). For example, the following image shows that most of the high performers received job offers, but some did not.



Questions

- To what extent do you agree with the following statement: a scenario similar to the one described above might occur in real life.
 - Strongly agree
 - Agree
 - Neither agree nor disagree
 - Disagree
 - Strongly disagree
- How much effort, in hours, should Sales-a-lot put in to make sure these decisions were fair? [short answer - number of hours]

D.2.2. RULE DESCRIPTIONS AND QUESTIONS

The following sections provide fairness definitions (presented to participants as *rules*) for Demographic Parity, Equal Opportunity (FNR and FPR), and Equalized Odds. Unless otherwise noted the rule description is shown above each of the questions for reference. Correct answers are noted in red.

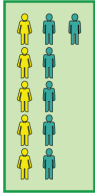
Demographic Parity. Recruit-a-matic uses the following rule to determine whether Sales-a-lot's hiring decisions were fair:

The fraction of male candidates who receive job offers should equal the fraction of female candidates who receive job offers.

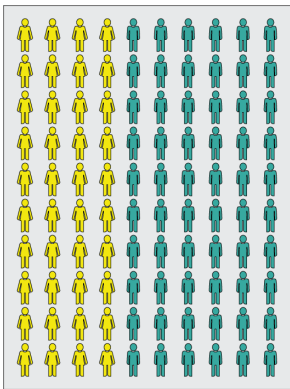
Example 1: Suppose that over the past year, Recruit-a-matic finds that Sales-a-lot received the following applicants (10 female and 12 male).



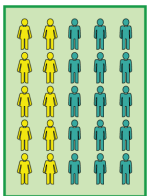
If Sales-a-lot sent job offers to the following number of applicants (5 female and 6 male), then this would be fair according to the hiring rule (note that there are other possible outcomes that are fair according to the hiring rule).



Example 2: Suppose that over the past year, Recruit-a-matic finds that Sales-a-lot reviewed a total of 100 applicants as follows (40 female and 60 male).



If Sales-a-lot sent job offers to the following number of applicants (10 female and 15 male), then this would be fair according to the hiring rule (note that there are other possible outcomes that are fair according to the hiring rule).

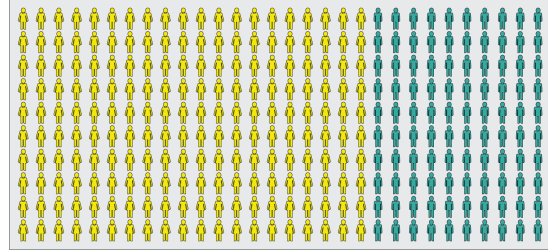


In the next section, we will ask you some questions about the information you have just read. Please note that this is not a test of your abilities. We want to measure the quality of the description you read, not your ability to take tests or answer questions.

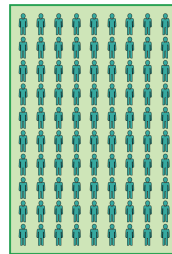
Please note that we ask you to apply and use ONLY the above hiring rule when answering the following questions. You will have an opportunity to state your opin-

ions and feelings on the rule later in the survey.

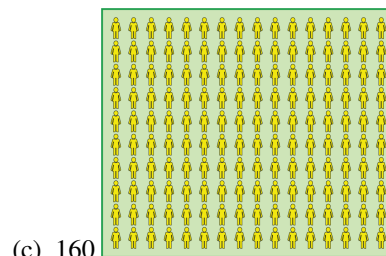
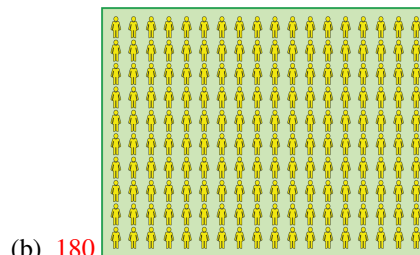
- Suppose a different company considered applicants for a different job. There were 200 female applicants and 100 male applicants,

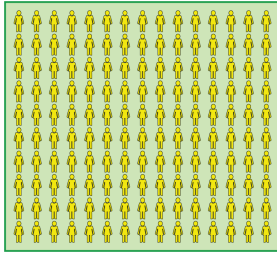


and they did send job offers to 90 male applicants.



Assuming that Recruit-a-matic reviews their decisions using the hiring rule above, how many female applicants should have received job offers?





(d) 150

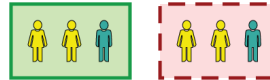
4. Assuming Recruit-a-matic reviews decisions using the hiring rule above, in which of these cases could Sales-a-lot have accepted more qualified female applicants than qualified male applicants?
 - (a) When there are more qualified female applicants than qualified male applicants (i.e., more women had low net sales at the end of the year).
 - (b) **When there are more female applicants than male applicants.**
 - (c) When female applicants receive worse interview scores than male applicants.
 - (d) This cannot happen under the hiring rule.
5. Consider one male applicant and one female applicant, both of whom are similarly qualified for the job (they achieve about the same net sales at the end of their first year). Is the following statement **TRUE OR FALSE**: The hiring rule above allows Sales-a-lot to make a job offer to one of these applicants and not the other.
6. Consider a situation where all female applicants were unqualified (they all achieve low net sales at the end of their first year), but some of them received job offers. Is the following statement **TRUE OR FALSE**: The hiring rule above requires that some job offers made to male applicants must have been made to unqualified male applicants.
7. Suppose Sales-a-lot received 100 male and 100 female applicants, and eventually made 10 job offers. Is the following statement **TRUE OR FALSE**: The hiring rule above requires that even if all male applicants were unqualified (they all achieve low net sales at the end of their first year), some of the unqualified males must have received job offers.
8. Is the following statement **TRUE OR FALSE**: The hiring rule above always allows Sales-a-lot to send job offers only to the most qualified applicants (those who achieve high net sales at the end of their first year).

Consider a different scenario than the two examples above, with 6 applicants – 4 female and 2 male, as illustrated below. The next three questions each give a different potential outcome for all 6 applicants (i.e., which of the 6 applicants do

receive job offers). Please indicate which of the outcomes follow the hiring rule above.



9. Sales-a-lot makes the following hiring decisions.



Do these decisions obey the hiring rule? **Yes**

10. Sales-a-lot makes the following hiring decisions.



Do these decisions obey the hiring rule? **No**

11. Sales-a-lot makes the following hiring decisions.



Do these decisions obey the hiring rule? **No**

12. In your own words, explain the hiring rule. [short answer] [The rule is not shown above this question]
13. To what extent do you agree with the following statement: I am confident I know how to apply the hiring rule described above?
 - Strongly agree
 - Agree
 - Neither agree nor disagree
 - Disagree
 - Strongly Disagree
14. Please select the choice that best describes your experience: When I answered the previous questions...
 - (a) I applied the provided hiring rule only.
 - (b) I used a combination of the provided hiring rule and my own ideas of what the correct hiring rule should be.
 - (c) I used only my own ideas of what the correct hiring decision should be rather than the provided hiring rule.
15. To what extent do you agree with the following statement: I like the hiring rule?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree

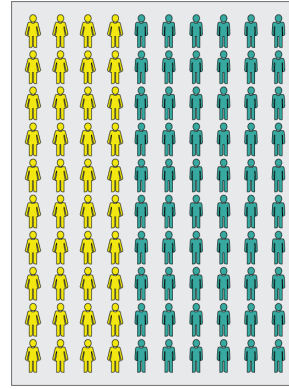
- Strongly Disagree

16. To what extent do you agree with the following statement: I agree with the hiring rule?

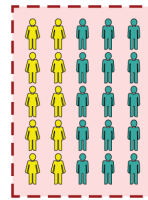
- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly Disagree

17. Please explain your opinion on the hiring rule. [short answer]

18. Was there anything about this survey that was hard to understand or answer? [short answer]



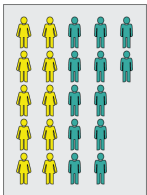
If Sales-a-lot did not send job offers to the following number of qualified applicants (10 female and 15 male), then this would be fair according to the hiring rule (note that there are other possible outcomes that are fair according to the hiring rule).



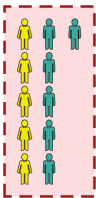
Equal Opportunity - FNR. Recruit-a-matic uses the following rule to determine whether Sales-a-lot’s hiring decisions were fair:

The fraction of qualified male candidates who do not receive job offers should equal the fraction of qualified female candidates who do not receive job offers.

Example 1: Suppose that over the past year, Recruit-a-matic finds that Sales-a-lot received the following qualified applicants (10 female and 12 male).



If Sales-a-lot did not send job offers to the following number of qualified applicants (5 female and 6 male), then this would be fair according to the hiring rule (note that there are other possible outcomes that are fair according to the hiring rule).



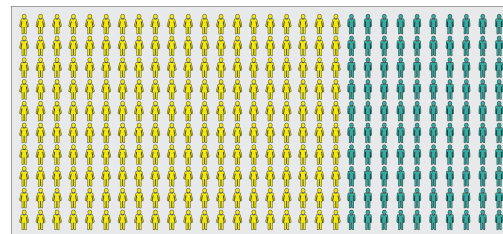
Example 2: Suppose that over the past year, Recruit-a-matic finds that Sales-a-lot reviewed a total of 100 qualified applicants as follows (40 female and 60 male).

Note that in the above examples the remaining qualified applicants received job offers, but are not displayed here.

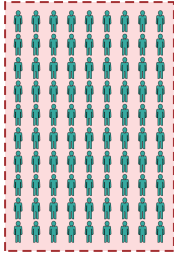
In the next section, we will ask you some questions about the information you have just read. Please note that this is not a test of your abilities. We want to measure the quality of the description you read, not your ability to take tests or answer questions.

Please note that we ask you to apply and use ONLY the above hiring rule when answering the following questions. You will have an opportunity to state your opinions and feelings on the rule later in the survey.

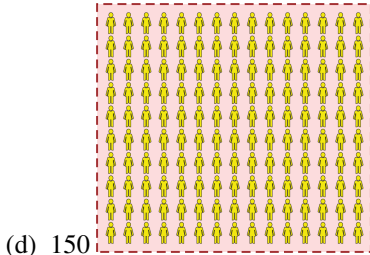
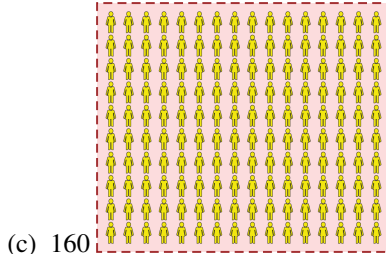
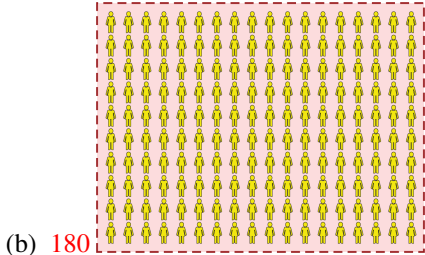
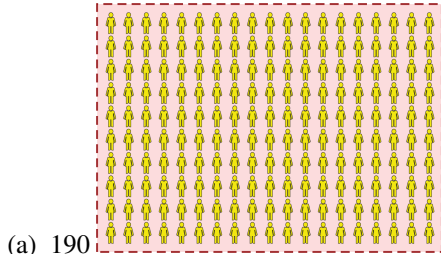
3. Suppose a different company considered applicants for a different job. There were 200 qualified female applicants and 100 qualified male applicants,



and they did not send job offers to 90 qualified male applicants.



Assuming that Recruit-a-matic reviews their decisions using the hiring rule above, how many qualified female applicants should not have received job offers?



4. Assuming Recruit-a-matic reviews decisions using the hiring rule above, in which of these cases could Sales-a-lot have rejected more qualified female applicants than qualified male applicants?

- (a) When there are more qualified female applicants than qualified male applicants (i.e., more women had low net sales at the end of the year).
- (b) When there are more female applicants than male applicants.

- (c) When female applicants receive worse interview scores than male applicants.
- (d) This cannot happen under the hiring rule.

- 5. Consider one male applicant and one female applicant, both of whom are similarly qualified for the job (they achieve about the same net sales at the end of their first year). Is the following statement **TRUE OR FALSE**: The hiring rule above allows Sales-a-lot to make a job offer to one of these applicants and not the other.
- 6. Consider a situation where all female applicants were unqualified (they all achieve low net sales at the end of their first year), but some of them received job offers. Is the following statement **TRUE OR FALSE**: The hiring rule above requires that some job offers made to male applicants must have been made to unqualified male applicants.
- 7. Suppose Sales-a-lot received 100 male and 100 female applicants, and eventually made 10 job offers. Is the following statement **TRUE OR FALSE**: The hiring rule above requires that even if all male applicants were unqualified (they all achieve low net sales at the end of their first year), some of the unqualified males must have received job offers.
- 8. Is the following statement **TRUE OR FALSE**: The hiring rule above always allows Sales-a-lot to send job offers only to the most qualified applicants (those who achieve high net sales at the end of their first year).

Consider a different scenario than the two examples above, with 6 qualified applicants – 4 female and 2 male, as illustrated below. The next three questions each give a different potential outcome for all 6 qualified applicants (i.e., which of the 6 applicants do not receive job offers). Please indicate which of the outcomes follow the hiring rule above.



9. Sales-a-lot makes the following hiring decisions.



Do these decisions obey the hiring rule? **Yes**

10. Sales-a-lot makes the following hiring decisions.



Do these decisions obey the hiring rule? **No**

11. Sales-a-lot makes the following hiring decisions.



Do these decisions obey the hiring rule? **No**

12. In your own words, explain the hiring rule. [short answer] [The rule is not shown above this question]

13. To what extent do you agree with the following statement: I am confident I know how to apply the hiring rule described above?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly Disagree

14. Please select the choice that best describes your experience: When I answered the previous questions...

- (a) I applied the provided hiring rule only.
- (b) I used a combination of the provided hiring rule and my own ideas of what the correct hiring rule should be.
- (c) I used only my own ideas of what the correct hiring decision should be rather than the provided hiring rule.

15. To what extent do you agree with the following statement: I like the hiring rule?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly Disagree

16. To what extent do you agree with the following statement: I agree with the hiring rule?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly Disagree

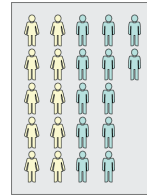
17. Please explain your opinion on the hiring rule. [short answer]

18. Was there anything about this survey that was hard to understand or answer? [short answer]

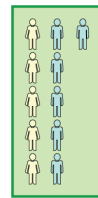
Equal Opportunity - FPR. Recruit-a-matic uses the following rule to determine whether Sales-a-lot's hiring decisions were fair:

The fraction of unqualified male candidates who receive job offers should equal the fraction of unqualified female candidates who receive job offers.

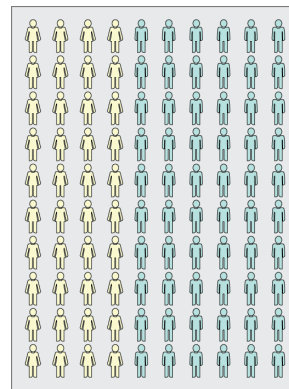
Example 1: Suppose that over the past year, Recruit-a-matic finds that Sales-a-lot received the following unqualified applicants (10 female and 12 male).



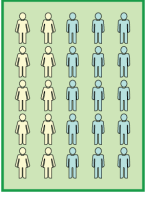
If Sales-a-lot sent job offers to the following number of unqualified applicants (5 female and 6 male), then this would be fair according to the hiring rule (note that there are other possible outcomes that are fair according to the hiring rule).



Example 2: Suppose that over the past year, Recruit-a-matic finds that Sales-a-lot reviewed a total of 100 unqualified applicants as follows (40 female and 60 male).



If Sales-a-lot sent job offers to the following number of unqualified applicants (10 female and 15 male), then this would be fair according to the hiring rule (note that there are other possible outcomes that are fair according to the hiring rule).

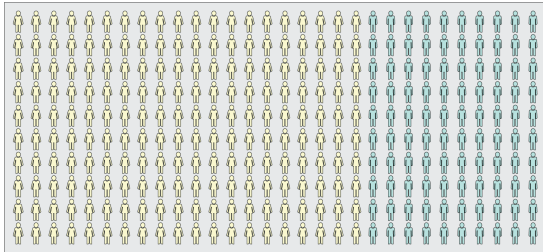


Note that in the above examples the remaining unqualified applicants did not receive job offers, but are not displayed here.

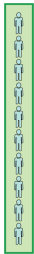
In the next section, we will ask you some questions about the information you have just read. Please note that this is not a test of your abilities. We want to measure the quality of the description you read, not your ability to take tests or answer questions.

Please note that we ask you to apply and use ONLY the above hiring rule when answering the following questions. You will have an opportunity to state your opinions and feelings on the rule later in the survey.

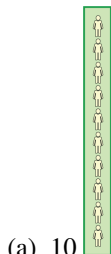
- Suppose a different company considered applicants for a different job. There were 200 unqualified female applicants and 100 unqualified male applicants,



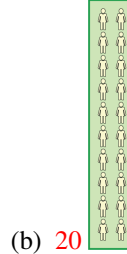
and they did send job offers to 10 unqualified male applicants.



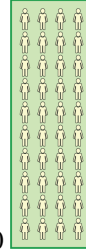
Assuming that Recruit-a-matic reviews their decisions using the hiring rule above, how many unqualified female applicants should have received job offers?



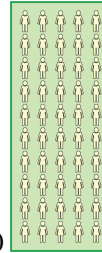
(a) 10



(b) 20



(c) 40



(d) 50

- Assuming Recruit-a-matic reviews decisions using the hiring rule above, in which of these cases could Sales-a-lot have accepted more unqualified female applicants than unqualified male applicants?
 - When there are more unqualified female applicants than unqualified male applicants (i.e., more women had low net sales at the end of the year).
 - When there are more female applicants than male applicants.
 - When female applicants receive worse interview scores than male applicants.
 - This cannot happen under the hiring rule.
- Consider one male applicant and one female applicant, both of whom are similarly qualified for the job (they achieve about the same net sales at the end of their first year). Is the following statement **TRUE** OR **FALSE**: The hiring rule above allows Sales-a-lot to make a job offer to one of these applicants and not the other.
- Consider a situation where all female applicants were unqualified (they all achieve low net sales at the end of their first year), but some of them received job offers. Is the following statement **TRUE** OR **FALSE**: The hiring rule above requires that some job offers made to male applicants must have been made to unqualified male applicants.
- Suppose Sales-a-lot received 100 male and 100 female applicants, and eventually made 10 job offers. Is the

following statement **TRUE OR FALSE**: The hiring rule above requires that even if all male applicants were unqualified (they all achieve low net sales at the end of their first year), some of the unqualified males must have received job offers.

8. Is the following statement **TRUE OR FALSE**: The hiring rule above always allows Sales-a-lot to send job offers only to the most qualified applicants (those who achieve high net sales at the end of their first year).

Consider a different scenario than the two examples above, with 6 unqualified applicants – 4 female and 2 male, as illustrated below. The next three questions each give a different potential outcome for all 6 applicants (i.e., which of the 6 applicants receive job offers). Please indicate which of the outcomes follow the hiring rule above.



9. Sales-a-lot makes the following hiring decisions.



Do these decisions obey the hiring rule? **Yes**

10. Sales-a-lot makes the following hiring decisions.



Do these decisions obey the hiring rule? **No**

11. Sales-a-lot makes the following hiring decisions.



Do these decisions obey the hiring rule? **No**

12. In your own words, explain the hiring rule. [short answer] [The rule is not shown above this question]

13. To what extent do you agree with the following statement: I am confident I know how to apply the hiring rule described above?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly Disagree

14. Please select the choice that best describes your experience: When I answered the previous questions...

- (a) I applied the provided hiring rule only.
- (b) I used a combination of the provided hiring rule and my own ideas of what the correct hiring rule should be.
- (c) I used only my own ideas of what the correct hiring decision should be rather than the provided hiring rule.

15. To what extent do you agree with the following statement: I like the hiring rule?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly Disagree

16. To what extent do you agree with the following statement: I agree with the hiring rule?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly Disagree

17. Please explain your opinion on the hiring rule. [short answer]

18. Was there anything about this survey that was hard to understand or answer? [short answer]

Equalized Odds. Recruit-a-matic uses the following rule to determine whether Sales-a-lot’s hiring decisions were fair:

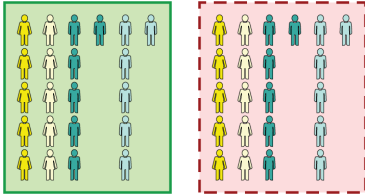
The fraction of qualified male candidates who do not receive job offers should equal the fraction of qualified female candidates who do not receive job offers. Similarly, the fraction of unqualified male candidates who receive job offers should equal the fraction of unqualified female candidates who receive job offers.

Example 1: Suppose that over the past year, Recruit-a-matic finds that Sales-a-lot received the following qualified applicants (10 female and 12 male) and unqualified applicants (10 female and 12 male).

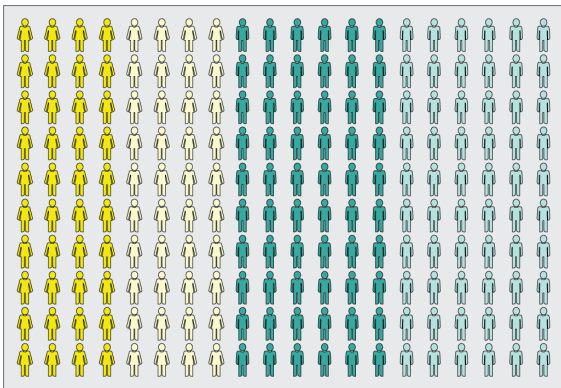


If Sales-a-lot did send offers to the following number of unqualified applicants (left, 5 female and 6 male), and did

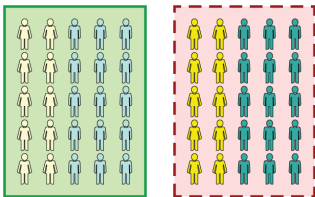
not send job offers to the following number of qualified applicants (right, 5 female and 6 male), then this would be fair according to the hiring rule (note that there are other possible outcomes that are fair according to the hiring rule).



Example 2: Suppose that over the past year, Recruit-a-lot finds that Sales-a-lot reviewed a total of 100 qualified applicants (40 female and 60 male) and 100 unqualified applicants (40 female and 60 male).



If Sales-a-lot did send offers to the following number of unqualified applicants (left, 10 female and 15 male), and did not send job offers to the following number of qualified applicants (right, 10 female and 15 male), then this would be fair according to the hiring rule (note that there are other possible outcomes that are fair according to the hiring rule).



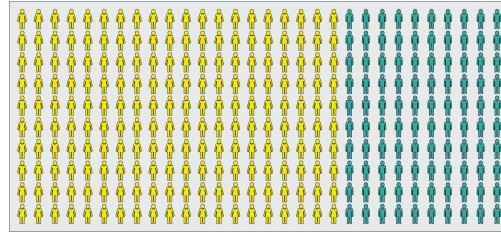
Note that in the above examples the remaining unqualified applicants did not receive job offers, but are not displayed here. Similarly, the remaining qualified applicants received job offers, but are not displayed here.

In the next section, we will ask you some questions about the information you have just read. Please note that this is not a test of your abilities. We want to measure the quality of the description you read, not your ability to take tests or answer questions.

Please note that we ask you to apply and use ONLY the above hiring rule when answering the following ques-

tions. You will have an opportunity to state your opinions and feelings on the rule later in the survey.

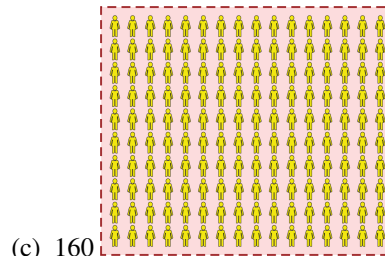
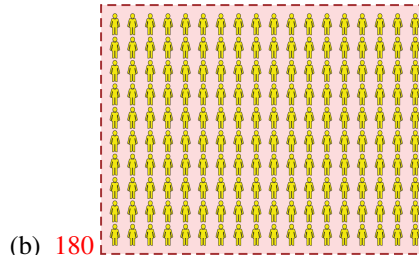
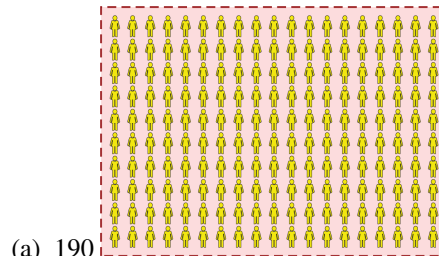
- Suppose a different company considered applicants for a different job. There were 200 qualified female applicants and 100 qualified male applicants,

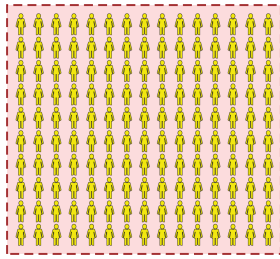


and they did not send job offers to 90 qualified male applicants.



Assuming that Recruit-a-matic reviews their decisions using the hiring rule above, how many qualified female applicants should not have received job offers?





(d) 150

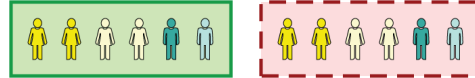
4. Assuming Recruit-a-matic reviews decisions using the hiring rule above, in which of these cases could Sales-a-lot have accepted more unqualified female applicants than unqualified male applicants?
- When there are more unqualified female applicants than unqualified male applicants (i.e., more women had low net sales at the end of the year).
 - When there are more female applicants than male applicants.
 - When female applicants receive worse interview scores than male applicants.
 - This cannot happen under the hiring rule.
5. Consider one male applicant and one female applicant, both of whom are similarly qualified for the job (they achieve about the same net sales at the end of their first year). Is the following statement **TRUE** OR **FALSE**: The hiring rule above allows Sales-a-lot to make a job offer to one of these applicants and not the other.
6. Consider a situation where all female applicants were unqualified (they all achieve low net sales at the end of their first year), but some of them received job offers. Is the following statement **TRUE** OR **FALSE**: The hiring rule above requires that some job offers made to male applicants must have been made to unqualified male applicants.
7. Suppose Sales-a-lot received 100 male and 100 female applicants, and eventually made 10 job offers. Is the following statement **TRUE** OR **FALSE**: The hiring rule above requires that even if all male applicants were unqualified (they all achieve low net sales at the end of their first year), some of the unqualified males must have received job offers.
8. Is the following statement **TRUE** OR **FALSE**: The hiring rule above always allows Sales-a-lot to send job offers only to the most qualified applicants (those who achieve high net sales at the end of their first year).

Consider a different scenario than the two examples above, with 6 qualified applicants – 4 female and 2 male; and 6 unqualified applicants – 4 female and 2 male. The next three questions each give a different potential outcome for the applicants (i.e., which of the applicants did or did not

receive job offers). Please indicate which of the outcomes follow the hiring rule above.

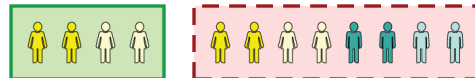


9. Sales-a-lot makes the following hiring decisions.



Do these decisions obey the hiring rule? **Yes**

10. Sales-a-lot makes the following hiring decisions.



Do these decisions obey the hiring rule? **No**

11. Sales-a-lot makes the following hiring decisions.



Do these decisions obey the hiring rule? **No**

12. In your own words, explain the hiring rule. [short answer] [The rule is not shown above this question]

13. To what extent do you agree with the following statement: I am confident I know how to apply the hiring rule described above?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly Disagree

14. Please select the choice that best describes your experience: When I answered the previous questions...

- I applied the provided hiring rule only.
- I used a combination of the provided hiring rule and my own ideas of what the correct hiring rule should be.
- I used only my own ideas of what the correct hiring decision should be rather than the provided hiring rule.

15. To what extent do you agree with the following statement: I like the hiring rule?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree

- Strongly Disagree
16. To what extent do you agree with the following statement: I agree with the hiring rule?
- Strongly agree
 - Agree
 - Neither agree nor disagree
 - Disagree
 - Strongly Disagree
17. Please explain your opinion on the hiring rule. [short answer]
18. Was there anything about this survey that was hard to understand or answer? [short answer]

D.3. Demographic Information

1. Please specify the gender with which you most closely identify:
- Male
 - Female
 - Other
 - Prefer not to answer
2. Please specify your year of birth
3. Please specify your ethnicity (you may select more than one):
- White
 - Hispanic or Latinx
 - Black or African American
 - American Indian or Alaska Native
 - Asian, Native Hawaiian, or Pacific Islander
 - Other
4. Please specify the highest degree or level of school you have completed:
- Some high school credit, no diploma or equivalent
 - High school graduate, diploma or the equivalent (for example: GED)
 - Some college credit, no degree
 - Trade/technical/vocational training
 - Associate's degree
 - Bachelor's degree
 - Master's degree
 - Professional or doctoral degree (JD, MD, PhD)
5. How much experience do you have in each of the following areas? (1 - no experience, 2 - limited experience, 3 - significant experience, 4 - expert)

- (a) Human resources (making hiring decisions)
- (b) Management (of employees)
- (c) Education (teaching)
- (d) IT infrastructure/systems administration
- (e) Computer science/programming
- (f) Machine learning/data science

We will maintain privacy of the information you have provided here. Your information will only be used for data analysis purposes.

E. Consent

E.1. Online Survey Consent Form

E.1.1. PROJECT TITLE

Fairness Evaluation and Comprehension

E.1.2. PURPOSE OF THE STUDY

This research is being conducted by Michelle Mazurek at the University of Maryland, College Park. We are inviting you to participate in this research project because you are above 18. The purpose of this research project is to understand lay comprehension of different fairness metrics.

E.1.3. PROCEDURES

The procedures will start with reading a brief description of a decision-making scenario. You will then be asked to answer some comprehension questions about the scenario. The questions will look like the following: What are the pros and cons of the notion of fairness described above?

Finally, you will be asked some demographics questions. The entire survey will take approximately 20 minutes or less.

E.1.4. POTENTIAL RISKS AND DISCOMFORTS

There are several questions to answer over the course of this study, so you may find yourself growing tired towards the end. Outside of this, there are minimal risks to participating in this research study. All data collected in this study will be maintained securely (see Confidentiality section) and will be deleted at the conclusion of the study.

However, if at any time you feel that you wish to terminate your participation for any reason, you are permitted to do so.

E.1.5. POTENTIAL BENEFITS

There are no direct benefits from participating in this research. We hope that, in the future, other people might

benefit from this study through improved understanding of fairness metrics and their applications.

E.1.6. CONFIDENTIALITY

Any potential loss of confidentiality will be minimized by storing all data (including information such as MTurk IDs and demographics) will be stored securely (a) in a password-protected computer located at the University of Maryland, College Park or (b) using a trusted third party (Qualtrics). Personally identifiable information that is collected (MTurk IDs, IP addresses, cookies) will be deleted upon study completion. All other data gathered will be stored for three years post study completion, after which it will be erased. The only persons that will have access to the data are the Principle Investigator and the Co-Investigators.

If we write a report or article about this research project, your identity will be protected to the maximum extent possible. Your information may be shared with representatives of the University of Maryland, College Park or governmental authorities if you or someone else is in danger or if we are required to do so by law.

E.1.7. COMPENSATION

You will receive \$3. You will be responsible for any taxes assessed on the compensation.

If you will earn \$100 or more as a research participant in this study, you must provide your name, address and SSN to receive compensation.

If you do not earn over \$100 only your name and address will be collected to receive compensation.

E.1.8. RIGHT TO WITHDRAW AND QUESTIONS

Your participation in this research is completely voluntary. You may choose not to take part at all. If you decide to participate in this research, you may stop participating at any time. If you decide not to participate in this study or if you stop participating at any time, you will not be penalized or lose any benefits to which you otherwise qualify.

If you decide to stop taking part in the study, if you have questions, concerns, or complaints, or if you need to report an injury related to the research, please contact the investigator:

Michelle Mazurek
5236 Irbe Center,
University of Maryland, College Park 20742
mmazurek@cs.umd.edu
(301) 405-6463

E.1.9. PARTICIPANT RIGHTS

If you have questions about your rights as a research participant or wish to report a research-related injury, please contact:

University of Maryland College Park
Institutional Review Board Office
1204 Marie Mount Hall
College Park, Maryland, 20742
E-mail: irb@umd.edu
Telephone: 301-405-0678

For more information regarding participant rights, please visit:

<https://research.umd.edu/irb-research-participants>

This research has been reviewed according to the University of Maryland, College Park IRB procedures for research involving human subjects.

E.1.10. STATEMENT OF CONSENT

By agreeing below you indicate that you are at least 18 years of age; you have read this consent form or have had it read to you; your questions have been answered to your satisfaction and you voluntarily agree to participate in this research study. Please ensure you have made a copy of the above consent form for your records.

Please ensure you have made a copy of the above consent form for your records. A copy of this consent form can be found here [link to digital copy].

- I am age 18 or older
- I have read this consent form
- I voluntarily agree to participate in this research study

E.2. Cognitive Interview Consent Form

E.2.1. PROJECT TITLE

Fairness Cognitive Interview

E.2.2. PURPOSE OF THE STUDY

This research is being conducted by Michelle Mazurek at the University of Maryland, College Park. We are inviting you to participate in this research project because you are above the age of 18, and fluent in English. The purpose of this research project is to understand lay comprehension of different fairness metrics.

E.2.3. PROCEDURES

The procedure involves completing an interview. The full procedure will be approximately 1 hour in duration.

During the interview you will be audio recorded, if you agree to be recorded. You will be asked to first read a brief description of a decision-making scenario. You will then be asked to fill out a survey about the scenario. While answering questions you will be asked verbal questions related to how you reached your answer in the survey.

Sample survey question: Is the following statement true or false? This hiring rule allows the hiring manager to send offers exclusively to the most qualified applicants.

Sample interview question: How did you reach your answer to that survey question?

E.2.4. POTENTIAL RISKS AND DISCOMFORTS

There are several questions to answer over the course of this study, so you may find yourself growing tired towards the end. Outside of this, there are minimal risks to participating in this research study. All data collected in this study will be maintained securely (see Confidentiality section) and will be deleted at the conclusion of the study.

However, if at any time you feel that you wish to terminate your participation for any reason, you are permitted to do so.

E.2.5. POTENTIAL BENEFITS

There are no direct benefits from participating in this research. We hope that, in the future, other people might benefit from this study through improved understanding of fairness metrics and their applications.

E.2.6. CONFIDENTIALITY

Any potential loss of confidentiality will be minimized by storing all data (including information such as demographics) securely (a) in a password protected computer located at the University of Maryland, College Park or (b) using a trusted third party (Qualtrics). Personally identifiable information that is collected will be deleted upon study completion. All other data gathered will be stored for three years post study completion, after which it will be erased. The only persons that will have access to the data are the principle Investigator and the Co-Investigators.

If we write a report or article about this research project, your identity will be protected to the maximum extent possible. Your information may be shared with representatives of the University of Maryland, College Park or governmental authorities if you or someone else is in danger or if we are required to do so by law.

E.2.7. COMPENSATION

You will receive \$30. You will be responsible for any taxes assessed on the compensation.

If you will earn \$100 or more as a research participant in this study, you must provide your name, address and SSN to receive compensation.

If you do not earn over \$100 only your name and address will be collected to receive compensation.

E.2.8. RIGHT TO WITHDRAW AND QUESTIONS

Your participation in this research is completely voluntary. You may choose not to take part at all. If you decide to participate in this research, you may stop participating at any time. If you decide not to participate in this study or if you stop participating at any time, you will not be penalized or lose any benefits to which you otherwise qualify.

If you decide to stop taking part in the study, if you have questions, concerns, or complaints, or if you need to report an injury related to the research, please contact the investigator:

Michelle Mazurek
5236 Iribe Center,
University of Maryland, College Park 20742
mmazurek@cs.umd.edu
(301) 405-6463

E.2.9. PARTICIPANT RIGHTS

If you have questions about your rights as a research participant or wish to report a research-related injury, please contact:

University of Maryland College Park
Institutional Review Board Office
1204 Marie Mount Hall
College Park, Maryland, 20742
E-mail: irb@umd.edu
Telephone: 301-405-0678

For more information regarding participant rights, please visit:

<https://research.umd.edu/irb-research-participants>

This research has been reviewed according to the University of Maryland, College Park IRB procedures for research involving human subjects.

E.2.10. STATEMENT OF CONSENT

Your signature indicates that you are at least 18 years of age; you have read this consent form or have had it read to you; your questions have been answered to your satisfaction and

Measuring Non-Expert Comprehension of Machine Learning Fairness Metrics

you voluntarily agree to participate in this research study.
You will receive a copy of this signed consent form.

Please initial all that apply (you may choose any number of these statements):

- I agree to be audio recorded
- I agree to allow researchers to use my audio recording in research publications and presentations.
- I do not agree to be audio recorded

If you agree to participate, please sign your name below.