

SUMMARY

Discrimination on the basis of race, gender, and other sociodemographic characteristics is common in hiring. Apart from **taste-based discrimination**, a second type of discrimination named as **statistical discrimination**, is based on inferences people draw about others depending on their membership in social groups. Such discrimination not necessarily be based on accurate beliefs, raising the question of when beliefs about group differences are systematically biased. Whereas psychologists have documented statistical fallacies that people fall victim, the implications for hiring decisions and discrimination remain largely unexplored.

In this project, we seek to examine an even more fundamental error that permeates high-stake decisions: *failures to adjust for selection* in the data generation process. More specifically, employers seek to identify the best candidates **based on performance signals**. The mechanisms generating these signals, however, may be systematically biased by selection. We expect that people fail to account for these biases, thus causing **statistical discrimination against the disadvantaged groups**. In addition, we expect that when **quota policy aiming at reducing discrimination** is conducted, people also fail to adjust to the fact that the majority group could potentially perform better, thus could potentially *facilitate* the implementation of quota policy aiming at reducing discrimination.

We conducted three studies to test to what extent people are able to account for selection. In general, we find evidence for (partial) selection neglect.

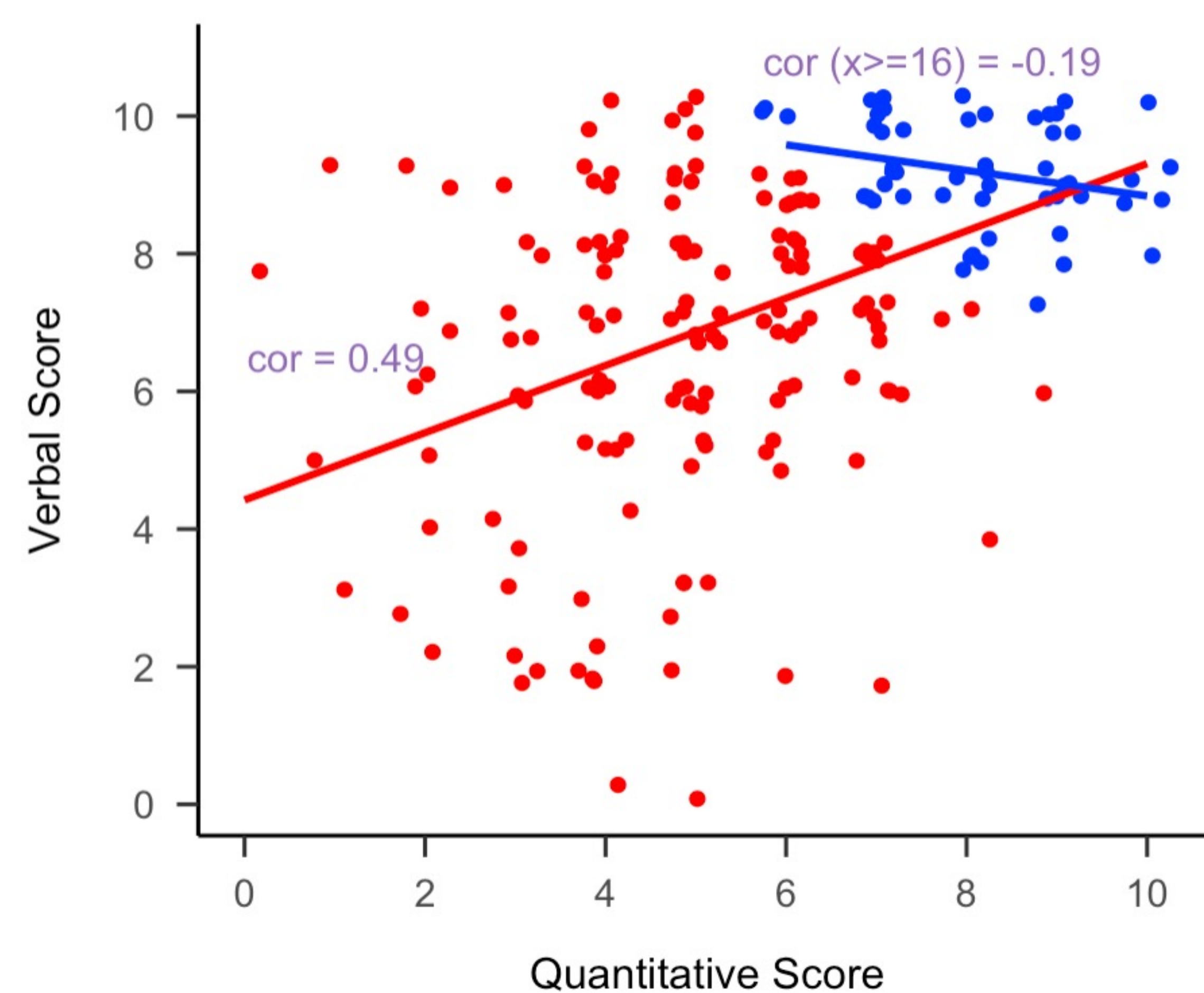
RESEARCH DESIGN (STUDY 1)



- 1000 participants act as employers to choose from 10 pairs of workers who previously completed math and verbal quizzes.

*Notice: The positive correlation between verbal and math score **reversed** when only top performers are considered. And we assume that people fail to account for such selection effect.

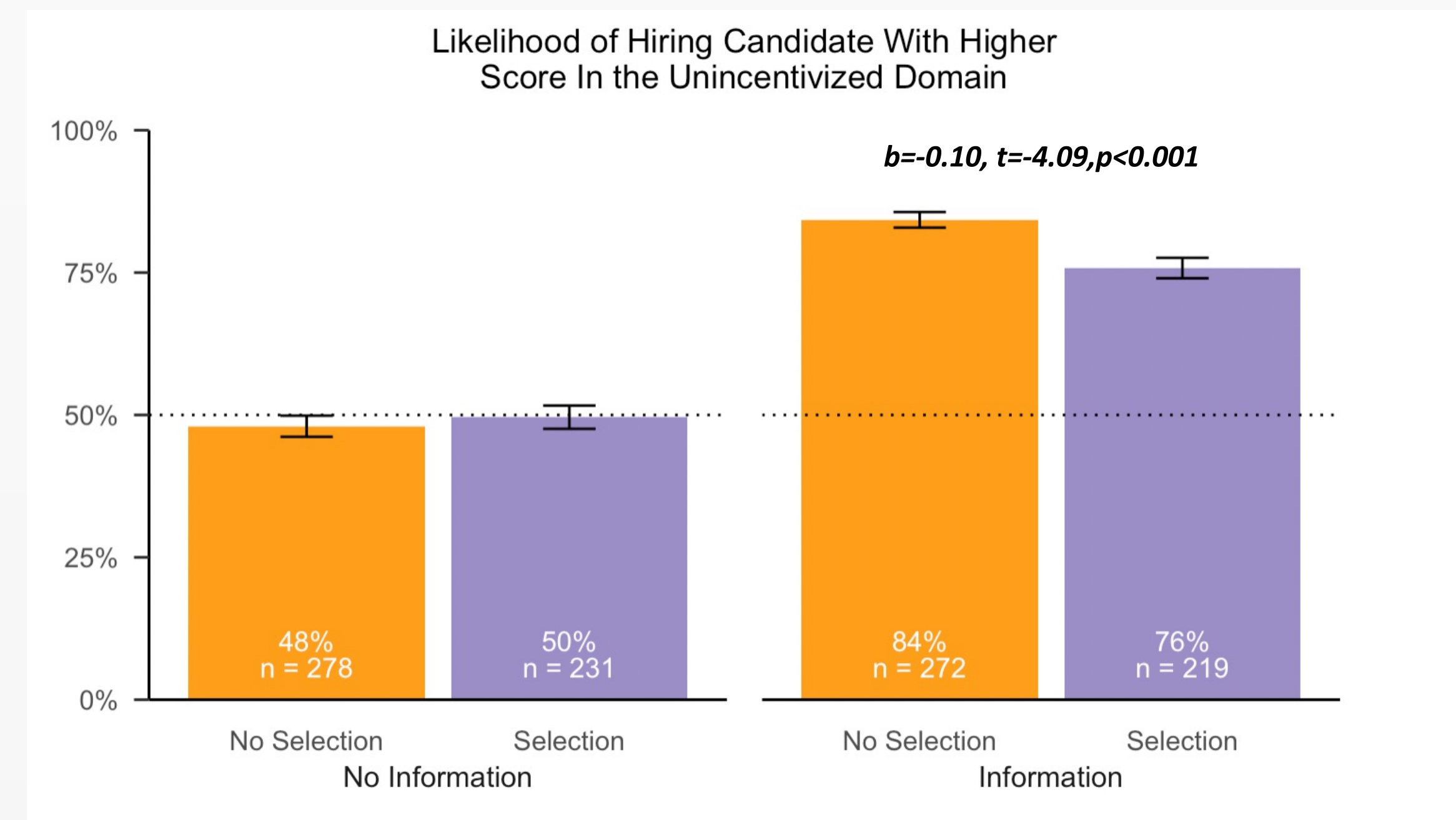
The relationship between verbal and quantitative score



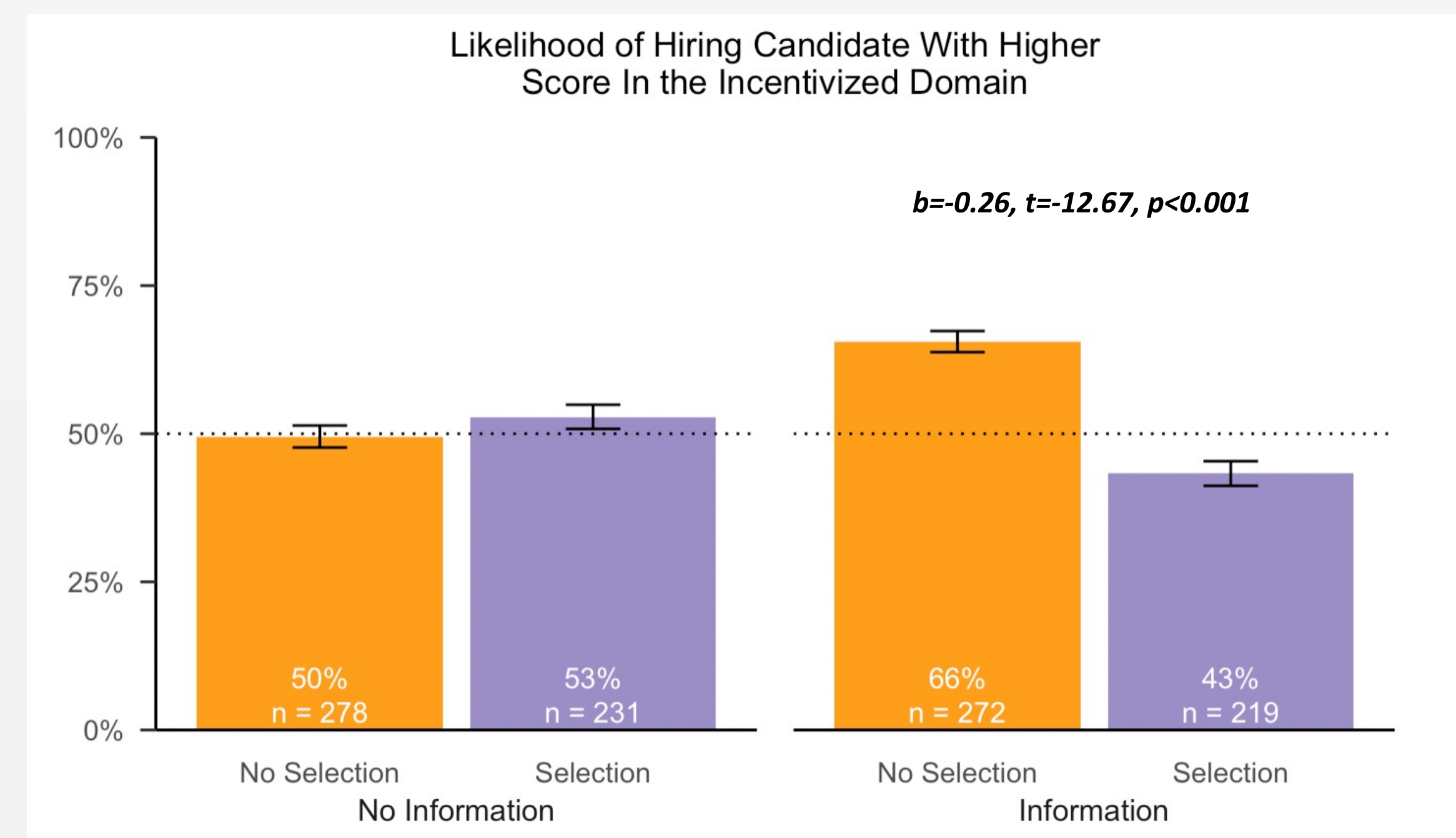
2X2 experimental design: Information X Selection

- Information 0:** Participants receive **no** information about the worker's score
- Information 1:** Participants receive information about the worker's score in the **unincented** domain.
- Selection 0:** 10 pairs of workers selected from 50 workers with the **highest** combined score across math and verbal tasks
- Selection 1:** 10 pairs of workers randomly selected from **all** 200 workers

RESULTS (STUDY 1)

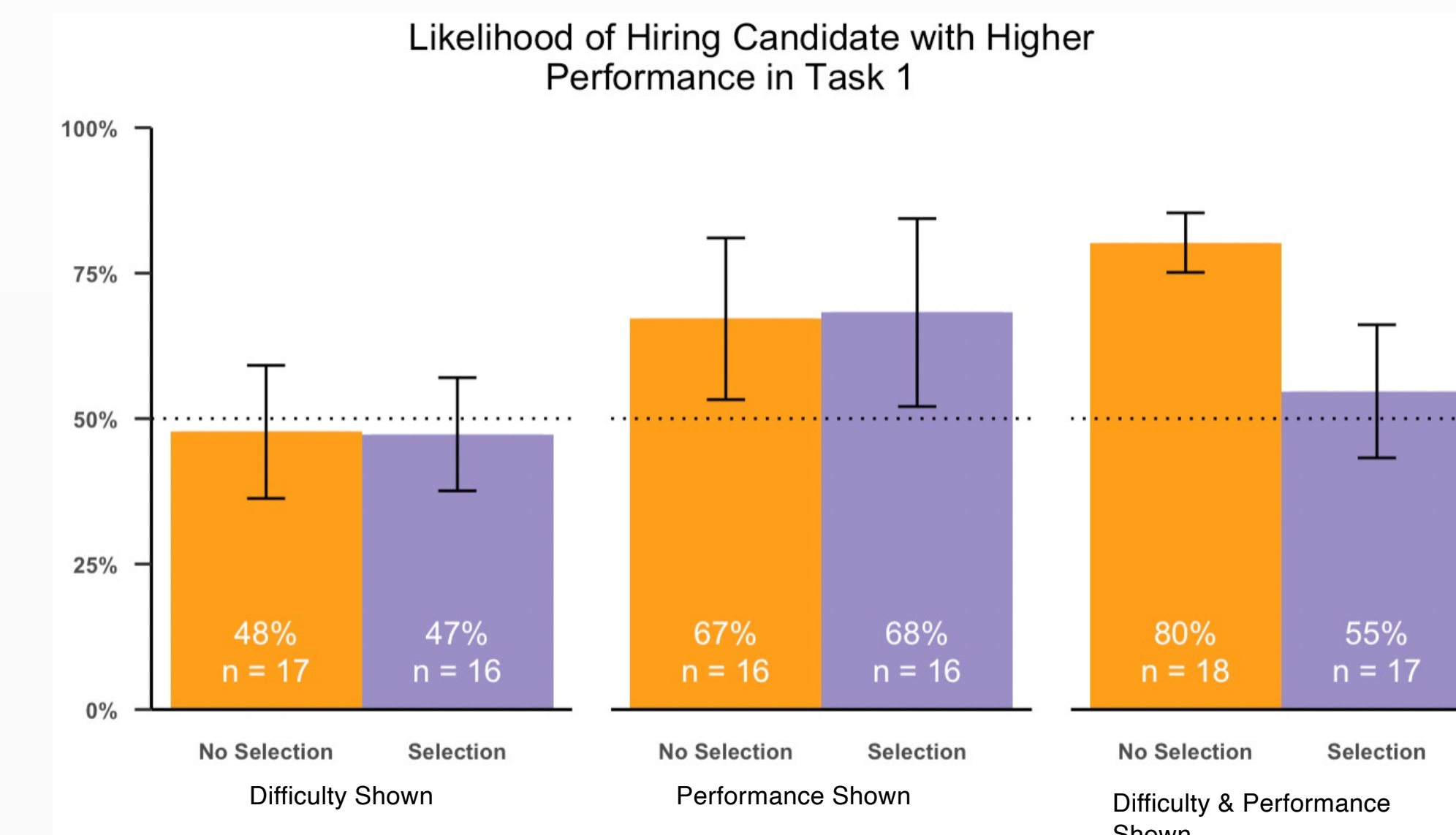


- Participants *always* prefer higher-score candidate when receiving score information on one of the quizzes, revealing that they *fail* to account for the effect of selection on the correlation of the two items in **Selection Scenario**.



- Participants make worse choices when receiving score information about one of the quizzes in **Selection Scenario**.

DESIGN & PILOT RESULTS (STUDY 2)



- 100 participants choose from 10 pairs of workers who previously completed string reversal tasks.
- Each worker completed two tasks, **task 1** is either **easy** (17 characters) or **hard** (23 characters). **Task 2** is a **medium** level task for everyone (20 characters).
- We show participants information **on task 1** and asked them to choose the worker in each pair that they believe performed better on task 2.

3X2 experimental design: Information X Selection

Information 1: Performance
Performance on task 1 is shown

Information 2: Difficulty
Difficulty on task 1 is shown

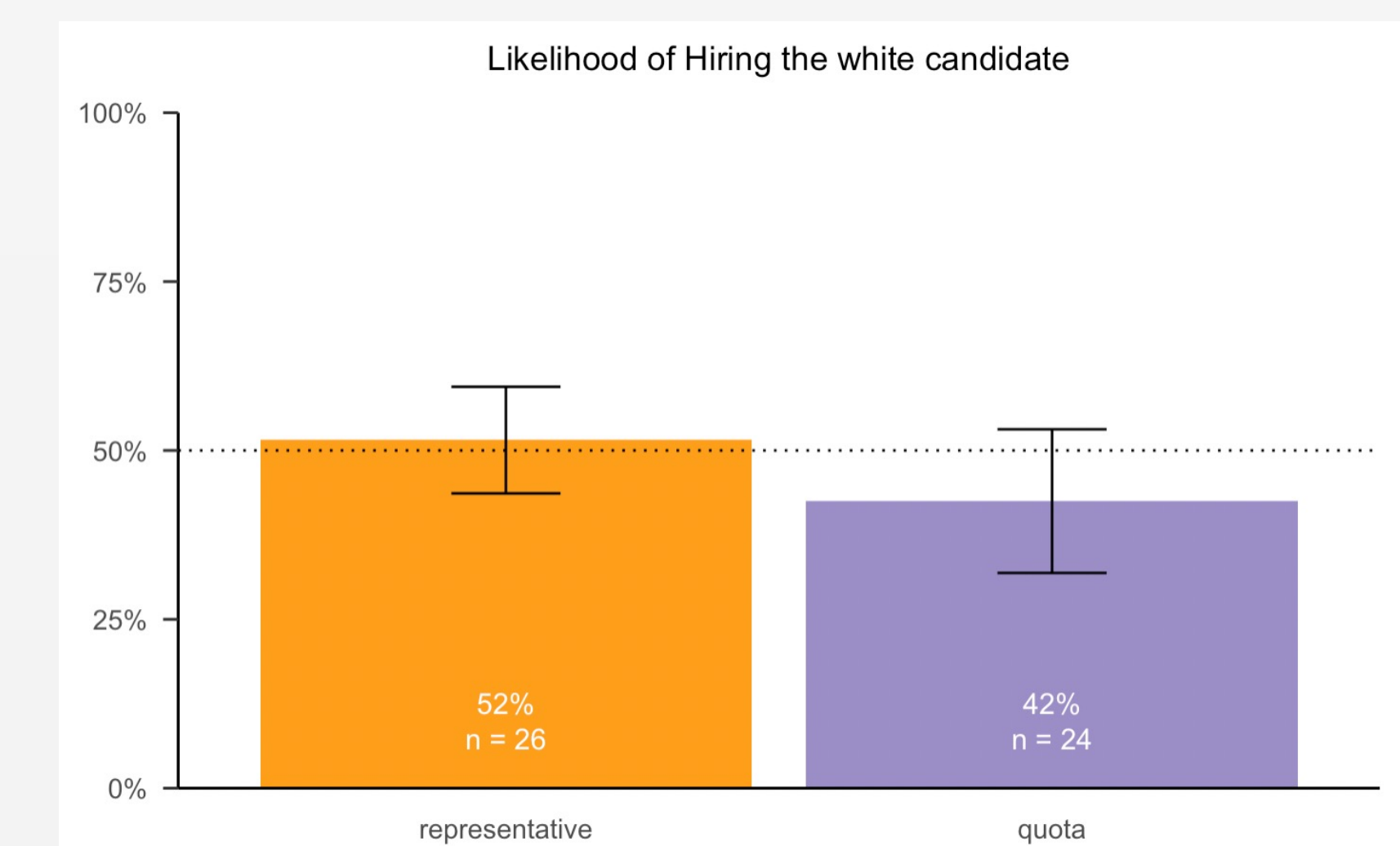
Information 3: Both
Difficulty and performance on task 1 is shown

Selection 0: Representative
10 pairs of workers randomly selected from all 400 participants

Selection 1: Selection
10 pairs of workers randomly selected from the 100 top performers from task 1 regardless of difficulty

- Participants account for selection somehow when **both** difficulty and performance information is shown, which makes their performance better than only **partial** information is given.

DESIGN & PILOT RESULTS (STUDY 3)



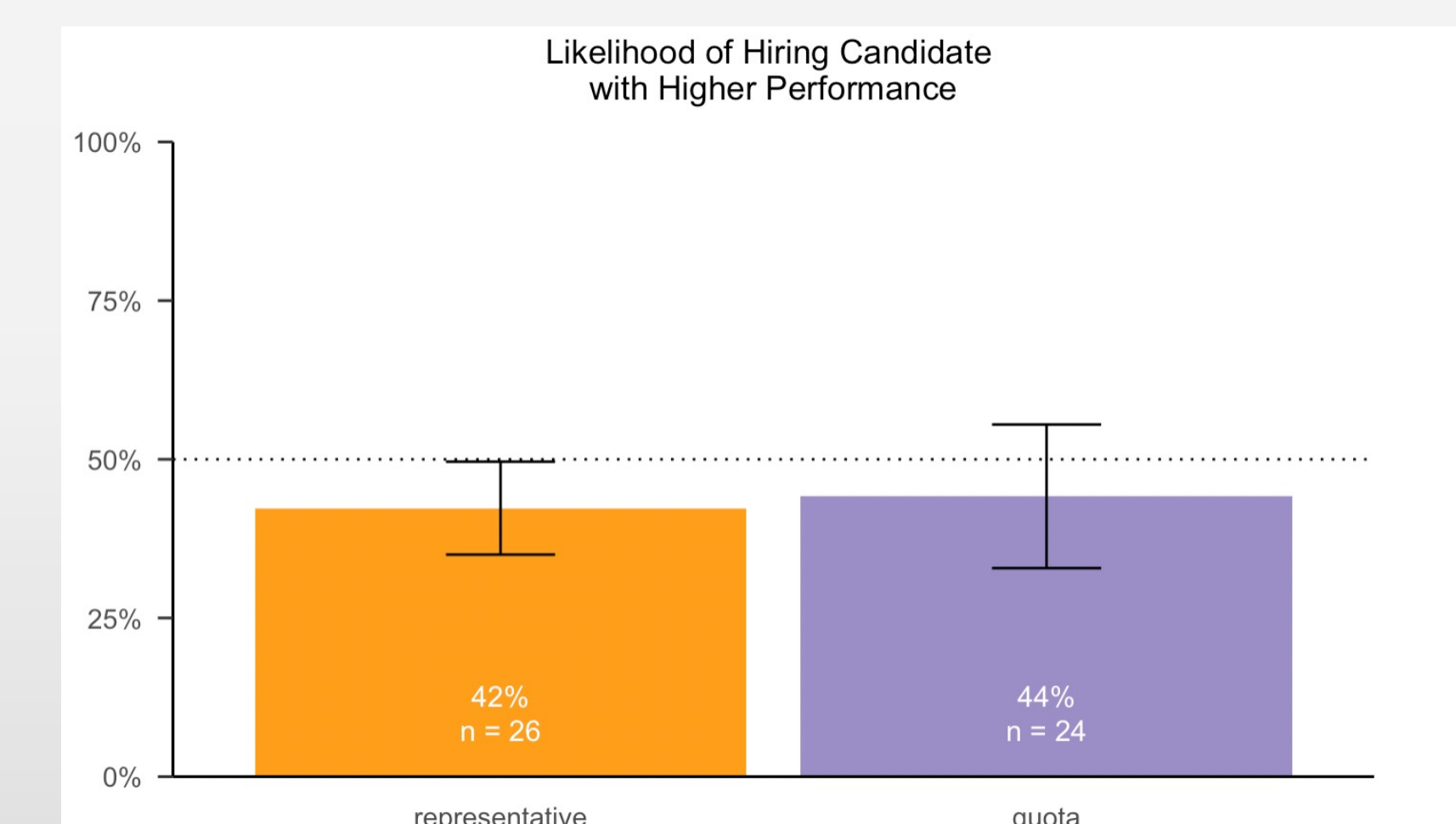
- 50 participants choose from 5 pairs of workers who previously completed a string reversal task.
- Only **demographic** information provided.
- Participants learn that the workers are randomly drawn from the 30 top performers of the overall pool of 200 workers.

Quota 0: Representative

Top performers are drawn to be ethnically **representative** of the US population (i.e., the 24 White top-performers, 4 Black top-performers, and 2 Asian top-performers)

Quota 1: Quota

Top performers are drawn based on a quota (i.e., an ethnically **balanced** sample of the 10 White top-performers, 10 Black top-performers, and 10 Asian top-performers)



- While participants choose white candidates less than 50% of chance in **Quota Condition** which reveals selection neglect, however, participants choose slightly even *better* candidates.