Gutenberg School of Management and Economics

& Research Unit “Interdisciplinary Public Policy”

Discussion Paper Series

Why do we Discriminate? The Role of Motivated Reasoning

Markus Eyting

August 26, 2022

Discussion paper number 2208

Johannes Gutenberg University Mainz
Gutenberg School of Management and Economics
Jakob-Welder-Weg 9
55128 Mainz
Germany
https://wiwi.uni-mainz.de/

Contact details

Markus Eyting
Chair of Digital Economics
Gutenberg School of Management and Economics
University of Mainz
Jakob-Welder-Weg 9
55128 Mainz
Germany
meyting@uni-mainz.de

Why do we Discriminate? The Role of Motivated Reasoning

Markus Eyting*†

August, 2022

Abstract

Identifying the cause of discrimination is crucial to design effective policies and to understand discrimination dynamics. Building on traditional models, this paper introduces a new explanation for discrimination: discrimination based on motivated reasoning. By systematically acquiring and processing information, individuals form motivated beliefs and consequentially discriminate based on these beliefs. Through a series of experiments, I show the existence of discrimination based on motivated reasoning and demonstrate important differences to statistical discrimination and taste-based discrimination. Finally, I demonstrate how this form of discrimination can be alleviated by limiting individuals' scope to interpret information.

*I thank Michael Kosfeld, Florian Hett, Guido Friebel, Alex Imas, Sendhil Mullainathan, Emir Kamenica, George Loewenstein, Vojtěch Bartoš, Simon Hess, Johannes Kasinger, Patrick Schmidt and seminar and conference audiences at the JGU Mainz, Goethe University Frankfurt, Discrimination & Diversity workshop at University of East Anglia, ESA Bologna, Behavioral Measurements conference Frankfurt, and EGPROC at University of Amsterdam for many valuable comments and discussions. I gratefully acknowledge the financial support from the SAFE Experiment Center Frankfurt and the Interdisciplinary Public Policy Mainz. The project received IRB approval at the joint ethics board of Goethe University Frankfurt and JGU Mainz. The experiments are pre-registered in the AEA RCT registry under AEARCTR-0008609.

†JGU Mainz, Heidelberg and Stanford University; email: meyting@stanford.edu

1 Introduction

Systematically disparate treatment of individuals from different social groups is widespread and has been documented in various contexts, such as the labor market (Blau and Kahn, 2017; Neumark, 2018), healthcare (Alsan et al., 2019), the justice system (Arnold et al., 2018; Arnold et al., 2021), policing (Goncalves and Mello, 2021), or education (Farkas, 2003). While discrimination by any cause can have fatal consequences for the discriminated, precise identification of the cause of discrimination has important implications for policy, welfare analyses, and discrimination dynamics (Bohren et al., 2019a).

Traditionally, economists have categorized discrimination as either taste-based discrimination (Becker, 1957) or statistical discrimination (Phelps, 1972; Arrow, 1973). Taste-based discrimination arises if an individual experiences an animus towards members of a particular group and therefore discriminates against them to not experience a disutility from interacting with them. Statistical discrimination is based on the notion that productivity of individuals is unobserved and therefore individuals rely on differences in group attributes to infer unobserved individual characteristics. In situations where there are no differences in group attributes, existing discrimination may still be statistical instead of taste-based as individuals may *incorrectly* perceive differences in group attributes and discriminate based on these wrong group-level beliefs (Bohren et al., 2019a).

This paper proposes and tests a new and different explanation of discrimination – *discrimination based on motivated reasoning*. I argue that (i) individuals hold systematically inaccurate beliefs about the productivity of individuals from different social groups, (ii) these incorrect beliefs are driven by motives, and (iii) resulting discrimination is a consequence of these motivated beliefs.[1]

While this explanation of discrimination links the existing concepts of taste-based discrimination and statistical discrimination, the distinction between these three concepts provides important implications for policy and discrimi-

[1]For now, I remain agnostic as to what 'motives' are based on. Section 4.1 discusses my view on 'motives' and its relation to 'taste' in more detail.

nation dynamics. While taste-based discrimination can only be decreased by affecting individuals' taste against members of particular groups, e.g. through direct confrontations, inaccurate statistical discrimination can effectively be alleviated by providing relevant group-level information to correct the wrongly perceived differences in group attributes. I show that providing this group-level information remains largely ineffective against discrimination based on motivated reasoning. Building on the concept of motivated reasoning as described by Epley and Gilovich (2016), I consider individuals' learning behavior and show that individuals systematically acquire and process information in line with their motives. In particular, I demonstrate that after receiving credible group-level and/or highly ambiguous individual-level information, individuals use the 'wiggle room' that this information provides for the formation of beliefs about unobserved individual-level characteristics. Second, I show that carefully designed information interventions that limit individuals' wiggle room to systematically interpret information may still counteract this form of discrimination, which contrasts the effect of information on taste-based discrimination. Taken together, this demonstrates why it is important to distinguish discrimination based on motivated reasoning from taste-based discrimination and inaccurate statistical discrimination.

In a series of online experiments among representative samples in the US[2], I provide a setting in which statistical discrimination is prevented by design. A critical assumption for statistical discrimination is that there are either differences in group-level statistics between individuals of two groups or differences in beliefs about group-level statistics. By carefully selecting the potentially discriminated and aligning beliefs among the potential discriminators, I make sure that neither statistics nor beliefs about these statistics differ in my experiments. Taste-based discrimination relies on the assumption that individuals experience disutility from interactions with members of a particular group. By design, my experiments do not include any interactions between the discriminators and the discriminated except for the mere selection of an indi-

[2]Representativeness is ensured in terms of age and gender and is established by Prolific (www.prolific.com).

vidual. This feature makes taste-based discrimination unlikely to be driving observed discrimination. Finally, my results show that individuals' behavior is inconsistent with any of these two forms of discrimination but consistent with discrimination based on motivated reasoning.

In the experiments, 'employers' repeatedly engage in binary hiring decisions between two experimental 'workers' from different races. Looking at decisions between Hispanic and Asian workers, I find that employers significantly discriminate against Hispanics after having been provided with actual group-level performance statistics that indicate equal productivity distributions between the two groups.

Second, I explore the information acquisition behavior of employers when they are provided with individual-level information. Discrimination based on motivated reasoning implies that employers 'fish for good news', which means they try to acquire information that supports their motive, e.g., by seeking additional information when a previous piece of information contradicts their motive and stopping to seek additional information once their information set is consistent with their motives. In line with this, I find that if the initial piece of provided information contradicts the employers' motive, they are significantly more likely to search for a second piece of information and acquire more pieces of information in total, as compared to when the initial piece of provided information confirms their motive.

Third, I study to what extent employers act consistently with the information, and how this depends on whether or not the information is in line with their motive. Further corroborating the mechanism of discrimination based on motivated reasoning, I find that employers are more likely to act consistently with the acquired information if the information confirms their motive and that this effect is larger if the wiggle room to interpret the information is larger.

Finally, I confirm that discrimination can be reduced by limiting the wiggle room of employers to interpret information. I find decreasing levels of discrimination when the wiggle room becomes smaller and provide evidence of how discrimination based on motivated reasoning can be avoided.

Taken together, this paper makes three main contributions. First, it proposes a new explanation of discrimination that links taste-based discrimination and inaccurate statistical discrimination: discrimination based on motivated reasoning. By systematically acquiring and processing information, individuals form motivated beliefs and consequentially discriminate based on these beliefs. Through a series of experiments, it provides first evidence of the existence and the underlying mechanisms of this explanation. Second, it demonstrates how discrimination based on motivated reasoning differs from traditional forms of discrimination and shows how it can arise in settings where taste-based discrimination and statistical discrimination is ruled out. Finally, it shows how these insights can be utilized to design an effective policy intervention to reduce discrimination. By varying the ambiguity of individual-level information, this paper demonstrates that limiting individuals' wiggle room to systematically engage with information may significantly reduce discrimination.

This paper adds to the existing literature in multiple ways. First, using an online experiment to study a particular cause of discrimination, it adds to the broad economic literature on theories of discrimination and empirical methods to measure it; see, e.g., Charles and Guryan (2011) for a review on challenges to measuring racial discrimination, Bohren et al. (2018) for a discussion about observational vs. experimental data, Bertrand and Duflo (2017), Heckman (1998), and Neumark et al. (2016) for discussions of field experiments, Onuchic (2022) for a recent review on theories of discrimination, and Bohren et al. (2022) for tools to model and measure systemic discrimination. In particular, it contributes to the recently growing economics literature, that aims to extend the traditional taxonomy of taste-based discrimination (Becker, 1957) versus statistical discrimination (Phelps, 1972; Arrow, 1973). So far, only a few studies have considered inaccurate beliefs as a source of discrimination in the economics literature.[3] In a closely related paper, Bohren et al. (2019a)

[3]In an in-depth literature review Bohren et al. (2019a) show that only 10.5% of 105 papers on discrimination that were published in 10 top economics journals between 1990 and 2018 differentiate between accurate and inaccurate beliefs. These and more recent studies include e.g. Fershtman and Gneezy (2001), Albrecht et al. (2013), Reuben et al. (2014), Bordalo et al. (2016), Bordalo et al. (2019), Bohren et al. (2019b), Bursztyn et al.

demonstrate that when economic agents have inaccurate beliefs about group attributes, resulting discrimination based on these beliefs can be mistaken for taste-based discrimination, statistical discrimination, or a combination thereof. They argue that providing credible information on relevant group-level distributions allows to separately identify inaccurate beliefs and animus as potential drivers of discrimination, as those with inaccurate beliefs should adjust their behavior upon receipt of credible group-level information while those who discriminate based on taste are unlikely to change their behavior in response to group-level information. This paper complements their idea by looking more closely at the behavior of those who do not change their behavior and would therefore traditionally be classified as taste-based discriminators. Other papers that aim to reveal the limits of this long-standing taxonomy by looking more closely at inaccurate beliefs as drivers for discrimination include Cornell and Welch (1996), who describe a form of ingroup vs. outgroup screening discrimination through better judgments of unknown qualities of candidates who belong to the same group as the decision maker, Coffman et al. (2021), who document that beliefs about average group differences drive gender discrimination in an artificial online hiring experiment, or Barron et al. (2022), who differentiate between explicit and implicit belief-based discrimination between genders.

On a more nuanced level, this paper adds to the economic literature on attending to information in discrimination contexts. Bartoš et al. (2016) pushed the research frontier in this context by modeling and documenting in three field experiments that rational (in)attention can amplify discrimination. They demonstrate that employers pay less attention to a priori less attractive applicants in cherry-picking markets, but more attention to a priori less attractive applicants in lemon-dropping markets. They reason that attention allocation is determined by the likelihood that costly information would change the status quo decision which is to not hire an applicant in the cherry picking market and to hire an applicant in the lemon dropping market. Focusing more on the extent to which information is considered in order to update beliefs, Mengel

(2020), and Esponda et al. (2022).

and Campos-Mercade (2021) attribute disparities in artificial hiring to signal neglect in the belief formation process. They show that employers conservatively update beliefs when confronted with new information and ultimately discriminate against disadvantaged workers. This paper complements these two studies by looking at both, the attention allocation decisions as well as the processing of seen information.

Finally, this paper builds upon a large literature on motivated reasoning (see e.g. Bénabou and Tirole, 2002; Bénabou, 2013; Brunnermeier and Parker, 2005; Di Tella et al., 2015; Epley and Gilovich, 2016; Grossman and Van der Weele, 2017; Köszegi, 2006). It translates findings of Chen and Heese (2021), who show that individuals are more (less) likely to continue acquiring information after they have received information that suggests that acting selfishly is harmful (harmless), into a discrimination context. I provide evidence that employers systematically search for information depending on whether previous information supports or contradicts their motive. In related work, Zimmermann (2020), finds that positive feedback has persistent effects on beliefs whereas negative feedback only affects beliefs in the short-run but fades over time. Thaler (2020) provides a novel experimental design to identify motivated reasoning and shows that motivated reasoning, even based on uninformative messages, can lead individuals' beliefs to become more polarized and less accurate. I apply a modified version of the proposed experimental design to identify motivated reasoning in the context of discrimination.

The remainder of this paper is structured as follows. The following section provides details on the experimental design. Section 3 demonstrates to what extent group-level beliefs of employers were aligned between treatments. Based on this, Section 4 explains the theoretical background of discrimination based on motivated reasoning and derives testable predictions. Main results are provided in Section 5. Finally, Section 6 concludes.

2 Experiment design

The data collection involves one survey and a series of pre-registered online experiments, programmed in Otree (Chen et al., 2016) and implemented on Prolific. The following subsections explain the experimental design. For more details and screenshots, see Appendix A.

2.1 The pool of workers

In the survey, I collect answers from 96 US participants on a logic quiz, a dictator game, and a real effort task. This is meant to mimic an 'assessment center' that provides proxies for workers' cognitive ability, social competence, and perseverance. Based on their answers in all three tasks, I calculate a score that defines their 'productivity' for the subsequent experiments. Additionally, I ask for race and other demographics, past school performance information, and psychological scales for resilience, cooperativeness, ambition, and diligence. This information is used to set up profiles of available experimental 'workers' for hire in the main experiments. To rule out statistical discrimination based on accurate beliefs as the cause for potential discrimination in the hiring experiment by design, I set up the final pool of workers for the hiring stage by selecting workers so that productivity distributions between race groups are equal. The final pool of workers that was used in the hiring experiments consists of 58 individuals from the US, equally balanced across gender, aged between 18 and 30 with a mean age of 22.83 years.[4]

2.2 The hiring experiments

In four different hiring experiments, subjects act as employers and are asked to hire workers from the constructed pool in a series of binary hiring decisions. In total, the aggregated pool of employers consists of 1338 subjects.[5]

[4]Of the 58 workers in the final pool, 17 identified as 'White', 16 as 'African American or Black', 13 as 'Asian', and as 12 'Hispanic or Latin' (hereinafter referred to as 'Hispanics').

[5]Appendix C.2 provides robustness checks with more limited employer pools, excluding subjects who correctly answered only 80% or less of training/test questions during the

2.2.1 Inducing motives

All experiments involve the same two treatment groups in a between-subjects design. In both groups, the employers make hiring decisions between two workers from different races. The two groups differ in the way the workers are labeled. In group 'Race', the employers could observe the races of the workers, whereas in group 'Neutral' the race labels are replaced by neutral shape labels, e.g. group 'Triangle', or group 'Diamond'. Importantly, the composition of workers in each hiring decision remains constant between treatments; only the labels are changed. This treatment variation induces a motive to hold particular productivity beliefs for employers of group Race, but not for those of group Neutral. Hence, results from employers in group Neutral serve as baseline levels in the subsequent analysis.

2.2.2 Experiment procedure

The experiments primarily consist of a belief stage followed by a hiring stage. In the belief stage, the employers are asked for their subjective belief distribution about the group productivities of the workers. These beliefs indicate the direction of their motives. Subsequently, they are provided with the true productivity distributions for each group of workers and are again asked for their beliefs.[6] All employers were informed that the pool of workers consisted of selected workers from the survey.

As in Coffman et al. (2021), the data on beliefs shows two things. First, it shows to what extent prior and posterior beliefs reflect the true productivity distributions. Second, and more importantly, it demonstrates the differences in beliefs between the two treatment groups. Updating beliefs between employers in group Race and employers in group Neutral is intended to generate identical and correct ex-post beliefs about the productivity distributions among employers in both groups. This renders the potential motives based on the workers' race the only difference between the two groups prior to the hiring

experiment. Results remain unaffected.

[6]Section A.2 of the Appendix shows screenshots of the belief stage.

decisions and rules out inaccurate statistical discrimination as the source for remaining discrimination.[7]

In the hiring stage, I ask the employers to make 20 incentivized hiring decisions between two workers from the constructed pool of workers. Each employer is repeatedly presented with a pair of workers and asked to hire one of them. After the experiment, one hiring decision is randomly chosen. If they hired the worker with the higher productivity score in this decision the employer gets a bonus payment of $2. The measurement of productivity has previously been explained to all employers. The workers did not receive an additional payment for being hired.

2.2.3 Varying wiggle room

Discrimination based on motivated reasoning implies that less wiggle room for the employers to form their beliefs about individual workers can decrease discrimination. For that reason, the experiments differ in the way individual-level information is provided for each hiring decision and thus in the extent of wiggle room that the employers have.

In Experiment 1 ('No Information'), employers have the most wiggle room as they do not receive any individual-level information. This means, without any further information about the two workers, in group Race employers are simply asked to hire e.g. the Asian worker or the Hispanic worker, whereas in group Neutral they are asked to hire e.g. the Triangle worker or the Diamond worker.[8] After all hiring decisions the employers answer a few demographic questions before the experiment ends. In total 605 subjects completed this experiment, 308 subjects in group Race and 297 subjects in group Neutral.

In Experiment 2 ('Ambiguous Information'), employers have a similar level of wiggle room. After the belief stage, they complete the same hiring task as

[7]Until here, the basic structure of the experiments is partly inspired by the design of Coffman et al. (2021) who study gender discrimination in binary hiring decisions, after aligning beliefs between employers in a gender treatment and a birth-month treatment.

[8]Importantly, the Triangle (Diamond) worker in treatment Neutral is the same worker as the Asian (Hispanic) worker in treatment Race. Moreover, I vary race-shape assignments between sessions.

the employers in the 'No Information' experiment. However, employers in this experiment receive an initial piece of individual-level information in each hiring decision and are able to request up to 9 additional signals. The signals consist of a binary message from one of two randomly-chosen information sources: True News or Fake News. The message from the True News source is always correct, the message from the Fake News source is never correct. The message reads "The better worker is X" where X is one of the two presented workers.[9] Note, that if the message comes from the Fake News source it implies that the worker who is not mentioned is indeed the better worker. Importantly, the employers do not know whether or not a message comes from the True news source or from the Fake News source. In fact, in this experiment, the employers are not informed about the likelihood with which a message comes from the True News source or the Fake News source.[10] Since the messages in this experiment are theoretically uninformative, they provide the employer with wiggle room to subjectively interpret each message according to their motive. Apart from this information structure during the hiring phase, this experiment is identical to the 'No Information' experiment. In total 499 subjects completed this experiment, 252 subjects in group Race and 247 subjects in group Neutral.

In Experiment 3 ('Uncertain Information') employers are provided with the exact same information structure as in the 'Ambiguous Information' experiment. However, in this experiment employers have previously been told that each message has a 60% likelihood to come from the True News source and a 40% likelihood to come from the Fake News source. This reduces their wiggle room compared to the 'Ambiguous Information' and 'No Information' experiments. Other than that, this experiment is identical to the 'Ambiguous Information' experiment. In total 120 subjects completed this experiment, 59 subjects in group Race and 61 subjects in group Neutral.

In Experiment 4 ('Tangible Information') employers do not receive mes-

[9] This feature of the experiment is inspired by the experimental tool to identify motivated reasoning in Thaler (2020).

[10] Subsequent elicitation of the perceived fraction of messages from the True News source reveals that the vast majority of employers guess that approximately 50% of all messages are true.

sages that directly displayed the (supposedly) better worker. Instead, employers are given individual-level information about past performances of the two workers.[11] Again, for each decision, employers receive one initial piece of information (e.g. their college GPA) and can request up to nine additional random pieces in each hiring decision. This environment still leaves wiggle room for participants but reduces it further as the signals are always true, but not necessarily conclusively predictive of the better of the two workers. In total 114 subjects completed this experiment, 62 subjects in group Race and 52 subjects in group Neutral.

Section A.3 of the Appendix shows screenshots of the hiring stages of each experiment. On average, the experiments took between 15 minutes ('No Information') to 24 minutes ('Tangible Information'). They were each conducted in December 2021 and with representative samples of the US population.

3 Beliefs about group productivities

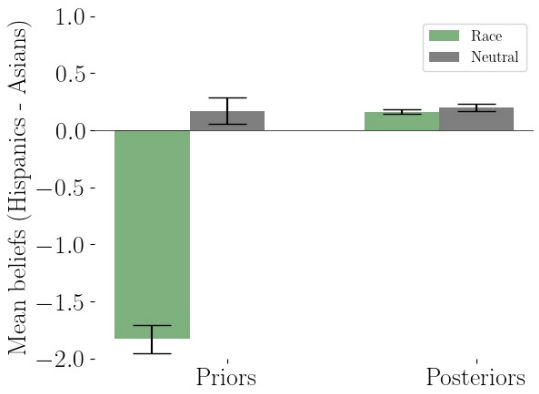
Before describing the predictions and hypotheses in more detail I first present results of the belief stage to identify employers' motives. I then demonstrate that employers in group Race and group Neutral hold identical group-level beliefs about the productivity of workers from different races, once they enter the hiring stage.

Looking at the prior beliefs about group productivity levels, the first bar of Figure 1 shows a large difference in mean beliefs about Asians and Hispanics in group Race. Employers of group Race believed Asian workers to be significantly more productive than Hispanic workers.[12] The second bar shows the differences in mean beliefs about Hispanic and Asian workers for employers in group Neutral and shows no significant differences in mean productivity levels

[11] The past performance information includes their college and high school GPA, SAT and ACT score, final high school math and English grade, and psychological measurements of their level of ambition, resilience, diligence, and agreeableness (based on Duckworth et al., 2007; Rammstedt and John, 2005; Sinclair and Wallston, 2004).

[12] For the entire prior belief distributions of employers in both groups, see Figure B1 in Appendix B.

Figure 1: Mean beliefs about group productivities



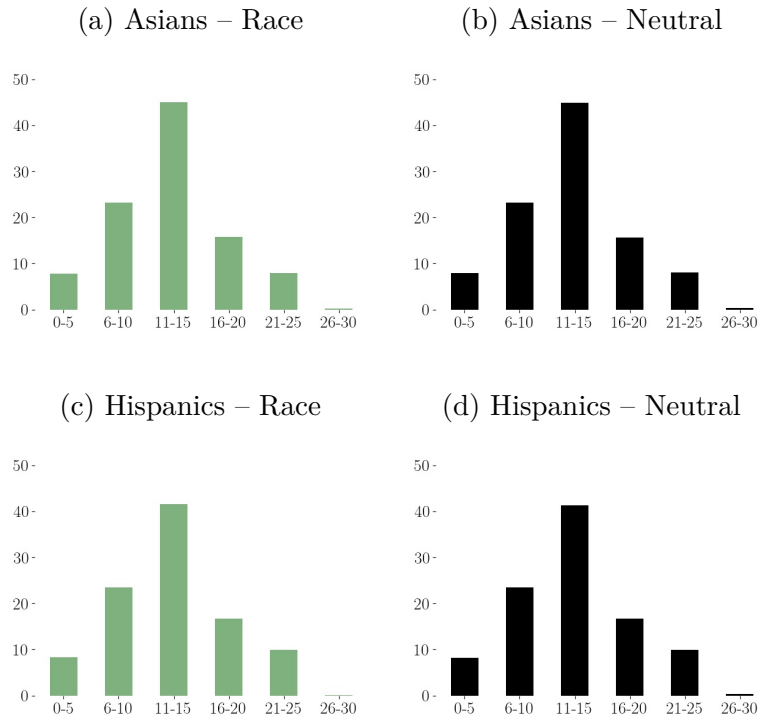
Notes: The first two bars show the mean productivity beliefs *before* the group-level information update of employers in group Race and group Neutral, respectively, the last two bars show the mean productivity beliefs *after* the group-level information update. Beliefs are plotted as the difference in mean beliefs about Hispanic workers and mean beliefs about Asian workers such that negative values indicate higher beliefs about the productivity of Asian workers than beliefs about the productivity of Hispanic workers, and vice versa. Error bars show the standard errors of the means.

between the two groups. Considering that the two groups were presented to them as group Triangle and group Diamond, this result seems almost generic and constitutes merely a verification. The difference in differences implies a potential for motivated beliefs that Asian workers are more productive than Hispanic workers.[13]

Next, I check whether the group-level information intervention was successful in aligning beliefs. The two bars on the right of Figure 1 show that the difference in differences in mean beliefs about group-level productivity of Asians and Hispanics between treatment groups disappeared after the information update. Importantly, this implies, that group-level beliefs about the productivity of Asians and Hispanics were equal across treatments when employers entered the hiring stage. In fact, Figure 2 demonstrates that employers in both treatment groups understand that the average, variance, and skewness

[13]I also elicited beliefs about black and white workers, but the analysis focuses on Asian and Hispanic workers. Hence, I consider decisions between members of two minority groups, which decreases the potential for social desirability bias, especially among white employers, e.g., due to recent Black Lives Matter movements. Mean beliefs about black and white workers are shown in Figure B2 in Appendix B.

Figure 2: Posterior belief distributions



Notes: Figures (a) and (b) show the employers' belief distributions about the productivities of Asian workers after the group-level information update in group Race and group Neutral, respectively, figures (c) and (d) show the employers' belief distributions about the productivities of Hispanic workers after the group-level information update in group Race and group Neutral, respectively.

of the productivity distributions of Asian and Hispanic workers are equal. This result holds for all four experiments and rules out statistical discrimination in the hiring decisions.[14] In the next section, I present testable predictions for discrimination based on motivated reasoning.

[14]Theoretically, it is still possible, that experimenter demand effects led employers to report similar belief distributions for all races, even though they did not actually believe in these distributions. However, employers were aware that the pool of workers was artificially constructed from a larger pool of Asian and Hispanic workers. This implies that there was no reason not to believe the provided group-level information.

4 Conceptual idea and hypotheses

This section describes the predictions and hypotheses in line with discrimination based on motivated reasoning. Taking the identical group-level belief distributions between employers of group Race and employers of group Neutral as a point of departure, this section presents various predictions for the hiring stages. The predictions depend on employers' wiggle room to systematically acquire and process information. Before, I start with a short discussion on the definition of a 'motive'.

4.1 What is a motive

Throughout the paper, I remain agnostic as to what defines a motive for a particular belief. A motive could be based on a taste against members of a particular group, but it could also be any other reason that leads an individual to wrongly believe that a member of one group is more/less productive than a member of another group. As an example, consider an employer who has always held the belief that members of one group are more productive than members of another group and who has based previous actions on this belief. If this employer would now come to realize that their long-standing belief was in fact untrue, they might experience a form of cognitive dissonance (Festinger, 1957). In response, to avoid experiencing cognitive dissonance the employer might prefer to stick with their initial belief. In this case, the employer might not hold any taste or animus against the member of a particular group but would still have a motive to hold a particular belief. Importantly, for discrimination based on motivated reasoning to explain disparate treatments of members from certain groups, it does not matter what the motive is based on, as long as it provides any reason to hold a particular belief.

4.2 Predictions under wiggle room

I predict that even though employers believe that there are no group-level productivity differences between Asians and Hispanics, there can still be dis-

crimination against Hispanics. The reasoning is that without any individual-level information, employers have enough wiggle room to reinstate their initial beliefs of a productive Asians vs. a less productive Hispanic when confronted with a binary decision between two individuals. While they know that on a group-level there are no differences, they will reason that between particular two workers from either group, the Asian worker will still be the more productive worker.

Hypothesis 1a *There is significant discrimination against Hispanics in group Race, but not in group Neutral, among employers in experiments 'No Information' and 'Ambiguous Information'.*

Providing employers with ambiguous individual-level information allows having a first look at the actual mechanism through which beliefs are reinstated, by yielding insights into the information acquisition and processing behavior of employers. Discrimination based on motivated reasoning predicts that individuals will 'fish for good news', meaning that they will request additional information signals if the previous signal contradicts their motive, but will not request additional signals if the previous signal confirms their motive. Since all individuals received one initial signal by default, I test for each hiring decision whether or not the content of the initial information signal affects (i) the likelihood to request a second information signal, and (ii) how many signals employers request per decision.

Hypothesis 1b *If the initial information signal suggests to hire the Hispanic worker in the 'Ambiguous Information' experiment, employers in group Race are more likely to request a 2nd signal and request more signals than employers in group Neutral.*

Moreover, discrimination based on motivated beliefs predicts that individuals overweight information signals that confirm their motive and underweight signals that contradict their motive, if their wiggle room allows. To test this, I analyze whether or not the overall direction of requested signals in a decision affects the likelihood to which employers act according to the signals. More

specifically, I test whether individuals are more likely to follow the signals, if the majority of requested signals confirms their motive, and less likely if the majority of requested signals contradict their motive.

Hypothesis 1c *If the majority of all considered signals in a decision suggests to hire the Hispanic worker in the 'Ambiguous Information' experiment, employers in group Race are less likely to act according to the suggestion than employers in group Neutral.*

4.3 Predictions under reduced wiggle room

When the provided individual-level information becomes less ambiguous, employers have less wiggle room to interpret information according to their motives. In response, motivated belief-based discriminators might still systematically search for news that supports their motive. However, knowing that this kind of individual-level information provides meaningful suggestions, motivated belief-based discriminators will reduce the systematic overweighting of information signals that are in line with their motives (and vice versa).

Hypothesis 2a *If the overall direction of all considered signals suggests to hire the Hispanic worker, the difference in the likelihood to follow the suggestion between group Race and group Neutral is larger in the 'Ambiguous Information' experiment than in the 'Uncertain Information' and the 'Tangible Information' experiments.*

Consequently, limiting employers' wiggle room should be an effective way to reduce discrimination.

Hypothesis 2b *There is less discrimination against Hispanics in decisions of employers in the 'Uncertain Information' and the 'Tangible Information' experiments, than in decisions of employers in the 'Ambiguous Information' or the 'No Information' experiments.*

By design of the experiments, accurate statistical discrimination is ruled out as there are no group-level differences between Asians and Hispanics in my

constructed pool of workers. Following Bohren et al. (2019a), the successful group-level information intervention rules out inaccurate statistical discrimination as the cause for potential subsequent discrimination as this form of discrimination is based on differences in group-level beliefs. Finally, taste-based discrimination is unlikely, as it assumes that economic agents experience a disutility from interactions with certain individuals, but my experimental setting does not include any employer-worker interaction and employers are aware of this right from the start.[15]

5 Results

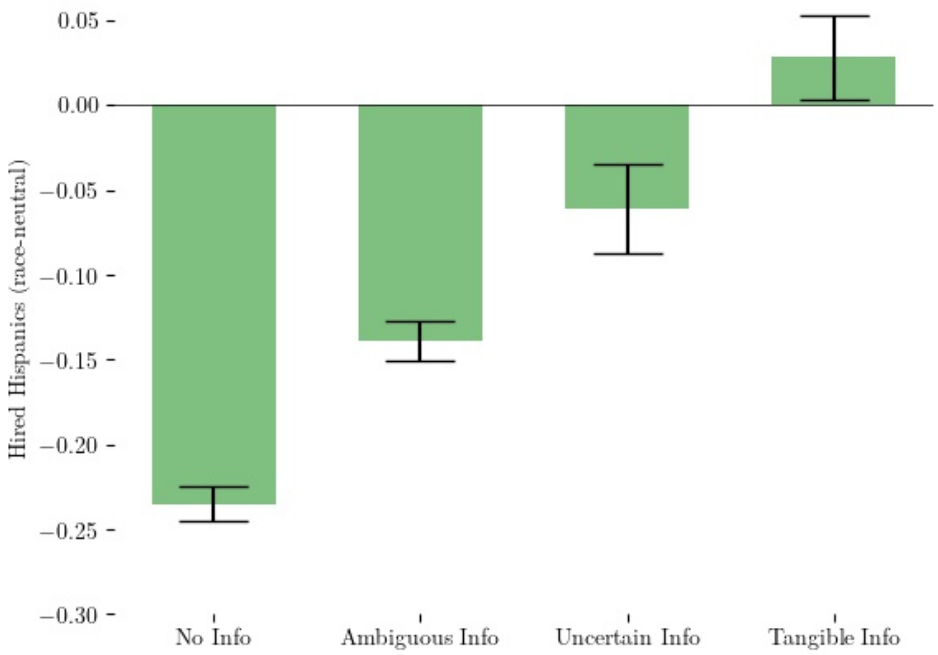
I will first consider to what extent employers discriminate in the experiments that leave the most wiggle room. The subsequent analysis of the information acquisition and processing behavior in the 'Ambiguous Information' provides first insights into the extent of potential motivated reasoning.

5.1 Wiggle room - No Information & Ambiguous Information

In line with hypothesis 1a, the first two bars in Figure 3 show that discrimination against Hispanics is substantial in the two experiments that leave the most wiggle room for employers. Providing employers with the race labels of workers in the 'No Information' experiment significantly decreases the hiring rate of Hispanics by 23.71pp. In the 'Ambiguous Information' experiment, there is also a significant 13.92pp difference in hiring rates of Hispanics between the two treatments. Albeit this difference is lower than in the case with no additional information, both experiments show significant discrimination

[15]In other studies, this design feature is used to rule out taste-based discrimination (e.g. Barron et al., 2022). One could argue, that there is still the theoretical possibility that individuals consider the mere selection of a worker from the less preferred group as interaction with that worker. In response to this concern, I present further evidence of behavior that is inconsistent with taste-based discrimination in the analyses of the data from experiments 2 to 4.

Figure 3: Hiring rates of Hispanic workers across experiments



Notes: The vertical axis displays the fraction of hired Hispanics in group Race minus the fraction of hired Hispanics in group Neutral. The horizontal axis groups the decisions by experiments ('No Information', 'Ambiguous Information', 'Uncertain Information', and 'Tangible Information'). Error bars indicate standard errors of the means.

against Hispanics, even though employers are aware that group-level productivity scores between the two groups are equal. Usually, this form of discrimination would be described as taste-based discrimination even though there is no interaction between employers and workers. To further investigate the potential cause of this discrimination, I now turn to the information acquisition and processing behavior of the employers in the 'Ambiguous Information' experiment.

Table 1 provides first evidence that discrimination is based on motivated reasoning. Column 1 shows that an initial signal that suggests to hire the Hispanic worker has a significantly larger positive effect on the likelihood to

acquire a second signal in group Race than in the group Neutral (coeff=0.1046, p-value=0.004). Column 2 illustrates that an initial signal that suggests to hire the Hispanic worker has a significantly larger positive effect on the total number of signals requested per decision in group Race than in group Neutral (coeff=0.5161, p-value=0.028). These results are consistent with hypothesis 1b and demonstrate that employers were 'fishing for good news', where 'good' refers to an information signal that confirms their motive to hire the Asian worker instead of the Hispanic worker. Finally, column 3 indicates that employers of group Race are significantly less likely to follow the signals if the majority of requested signals suggest to hire the Hispanic worker than employers of group Neutral (coeff=-0.1881, p-value< 0.001). This lends support to hypothesis 1c. Taken together, the results imply that employers discriminate after having systematically acquired and processed ambiguous information about the workers. Usually, under the assumption that employers derive disutility from the mere selection of a Hispanic worker, the results so far would still allow classifying observed discrimination as purely taste-based discrimination, in which the discriminators use the available information to justify their taste. In the next section, I provide evidence against this interpretation.

5.2 Reduced wiggle room - Uncertain Information & Tangible Information

To deepen our understanding of discrimination based on motivated reasoning I now consider the results from the two 'debiasing' experiments. In experiments 'Uncertain Information' and 'Tangible Information', I limit employers' wiggle room to systematically use the information to form motivated beliefs by providing more meaningful individual-level information. If discrimination in the previous two experiments was based on taste and employers simply used the information to justify their taste, we should not observe any differences between behavior of employers in the 'Ambiguous Information' experiment and behavior of employers in the experiments with reduced wiggle room. This is because the hiring behavior of taste-based discriminators should not respond

Table 1: Information acquisition and processing behavior under ambiguous information

	(1) 2nd signal	(2) number of signals	(3) follow signals
race * hispanic	0.1046***	0.5161**	−0.1881***
	(*0.0361*)	(*0.2345*)	(*0.0403*)
race	−0.0667	−0.3422	0.0596***
	(*0.0427*)	(*0.2804*)	(*0.0201*)
hispanic	−0.0390	−0.2801	−0.0695***
	(*0.0255*)	(*0.1705*)	(*0.0268*)
Observations	3290	3290	3290
Baseline mean dep. var.	0.5483	3.6246	0.8676

Notes: This table shows OLS results of the 'Ambiguous Information' experiment. *2nd signal*, is a dummy equal to 1 if an employer requested a second signal. *number of signals*, counts the number of requested signals. *follow signals*, is a dummy equal to 1 if an employer's hiring decision followed the majority of considered signals. *difference* refers to the differences in coefficients for *follow signals* and tests hypothesis 2a. *race* is the treatment dummy equal to 1 if the decision was made in treatment group Race. *hispanic* is a dummy equal to 1 if the initial signal (columns 1 and 2) or the direction of all considered signals (column 3) suggests to hire the Hispanic worker. Note that this holds for both treatment groups even though employers in group Neutral did not observe the race of the workers. Units of observation are decision specific. Standard errors are clustered at the individual level and displayed in parentheses. *p< 0.1, **p< 0.05, ***p< 0.01.

to variations in information. While one could argue that less wiggle room might complicate the ex-post rationalization of animus towards Hispanics, it does not affect the animus itself. The driver of discrimination is therefore unaffected by changes in information and the hiring behavior should not change in experiments 'Uncertain Information' and 'Tangible Information'.

First, I consider the case where employers were given similar information as in the 'Ambiguous Information' experiment, except that employers are told that signals have a 60% likelihood to come from the True News source and a 40% likelihood to come from the Fake News source. The information acquisition and processing behavior of employers in this experiment is shown in Table 2. Column 1 indicates that employers in group Race are insignificantly more likely to acquire a second information signal if the initial signal suggests

to hire the Hispanic worker than employers in group Neutral and column 2 shows that employers acquire insignificantly more information signals in group Race than in group Neutral. Note that coefficient sizes are similar to the results from the 'Ambiguous Information' experiment and the insignificance is likely to be an artefact of the much smaller sample size in this experiment. Looking at whether or not employers follow the overall direction of signals systematically differently, columns 3 and 4 of Table 2, illustrate that reducing ambiguity and hence the potential wiggle room to interpret the information substantially decreases the treatment difference in the likelihood with which employers act in line with the acquired information. A majority of requested signals that suggests to hire the Hispanic worker, reduces the likelihood to follow the suggestion by 6.08pp more in group Race than in group Neutral. This is a 12.74pp reduction in the absolute effect size compared to the results from the 'Ambiguous Information' experiment. Column 4 shows the results from the triple interaction analysis and provides evidence that this reduction is significant.

Next, I consider the case where employers were given individual-level information about past performances of the two workers in each decision. Across all three outcomes (probability to acquire a second signal, number of signals considered in each decision, likelihood to follow the majority of considered signals) Table 3 displays further reduced effect sizes compared to the 'Uncertain Information' experiment. An initial signal that suggests to hire the Hispanic worker has no significant treatment effect on the likelihood to acquire a second signal and on the number of signals requested. Finally, column 4 shows that the effect on the likelihood to follow the signals of (-2.06pp) has also significantly decreased by 16.75pp, compared to the effect in the 'Ambiguous Information' experiment. Together with the results from the 'Uncertain Information' experiment above, this supports hypothesis 2a.

Summarizing these results, I find that reducing employers' wiggle room to interpret individual-level information affects the extent to which they follow the signals. While in the 'No Information' and the 'Ambiguous Information' case, employers 'fish for good news' and refuse to follow the signals if this search

Table 2: Information acquisition and processing behavior under uncertain information

	Uncertain Information			Uncertain - Ambiguous
	(1) 2nd signals	(2) number of signals	(3) follow signals	(4) follow signals
race * hispanic	0.0898	0.6035	−0.0608	0.1274**
	(*0.0614*)	(*0.5416*)	(*0.0494*)	(*0.0638*)
race	0.0083	0.2919	0.0456*	
	(*0.0800*)	(*0.7498*)	(*0.0261*)	
hispanic	−0.0141	0.1543	−0.0292	
	(*0.0462*)	(*0.3632*)	(*0.0320*)	
Observations	756	756	756	
Baseline mean dep. var.	0.7313	5.4478	0.9292	

Notes: This table shows OLS results of the 'Uncertain Information' experiment. *2nd signal*, is a dummy equal to 1 if an employer requested a second signal. *number of signals*, counts the number of requested signals. *follow signals*, is a dummy equal to 1 if an employer's hiring decision followed the majority of considered signals. *difference* refers to the differences in coefficients for *follow signals* and tests hypothesis 2a. *race* is the treatment dummy equal to 1 if the decision was made in treatment group Race. *hispanic* is a dummy equal to 1 if the initial signal (columns 1 and 2) or the direction of all considered signals (column 3) suggests to hire the Hispanic worker. Note that this holds for both treatment groups even though employers in group Neutral did not observe the race of the workers. Units of observation are decision specific. Standard errors are clustered at the individual level and displayed in parentheses. *p< 0.1, **p< 0.05, ***p< 0.01.

was unsuccessful, a reduction of the wiggle room to interpret the information reduces the reluctance to follow unwanted signals. Employers may still search for information that confirms their motive, but if they cannot find this kind of information, they are now less likely to still act against the information than in the experiments with more wiggle room.

Finally, the last two bars in Figure 3 show how the adapted information processing behavior translates into less discrimination. While there was substantial discrimination in the experiments 'No Information' and 'Ambiguous Information', absolute treatment differences in hiring rates of Hispanics in the experiments with less wiggle room for the employers decrease significantly to −6.14pp in the Uncertain Information experiment and to an even positive (but insignificant) rate of 2.81pp in the Tangible Information experiment (see Table C1 in Appendix C.1). This supports hypothesis 2b and illustrates that lim-

Table 3: Information acquisition and processing behavior under tangible information

	Tangible Information			Tangible - Ambiguous
	(1) 2nd signals	(2) number of signals	(3) follow signals	(4) follow signals
race * hispanic	0.0260	0.5709	−0.0206	0.1675***
	(*0.0318*)	(*0.3937*)	(*0.0492*)	(*0.0620*)
race	0.0381	0.8460	−0.0328	
	(*0.0374*)	(*0.5347*)	(*0.0287*)	
hispanic	0.0051	0.1056	−0.0860**	
	(*0.0235*)	(*0.3137*)	(*0.0365*)	
Observations	742	742	742	
Baseline mean dep. var.	0.9135	5.5240	0.8894	

Notes: This table shows OLS results of the 'Tangible Information' experiment. *2nd signal*, is a dummy equal to 1 if an employer requested a second signal. *number of signals*, counts the number of requested signals. *follow signals*, is a dummy equal to 1 if an employer's hiring decision followed the majority of considered signals. *difference* refers to the differences in coefficients for *follow signals* and tests hypothesis 2a. *race* is the treatment dummy equal to 1 if the decision was made in treatment group Race. *hispanic* is a dummy equal to 1 if the initial signal (columns 1 and 2) or the direction of all considered signals (column 3) suggests to hire the Hispanic worker. Note that this holds for both treatment groups even though employers in group Neutral did not observe the race of the workers. Units of observation are decision specific. Standard errors are clustered at the individual level and displayed in parentheses. *p< 0.1, **p< 0.05, ***p< 0.01.

iting the wiggle room to interpret individual-level information can effectively decrease discrimination based on motivated reasoning.

As mentioned above, if discrimination was based on taste, we should not have observed differences in employers' behavior in these two experiments compared to employers' behavior in the first two experiments as the content of information should not affect taste-based discriminators. Since we do observe significant differences in (i) the way employers engage with the information and (ii) hiring rates of Hispanic vs. Asian workers, the results cannot be explained by taste-based discrimination. By design, (inaccurate) statistical discrimination was ruled out before employers entered the hiring stage.

5.3 Further evidence

I pre-registered and conducted two further experiments using the past performance information of the workers. These experiments were identical to the previous experiments, except for the provided individual-level information.

In the first additional experiment ('One Information'), employers are provided with individual-level information about past performances of the two workers (as in the Tangible Information experiment). However, for each decision, employers receive one piece of information and can not request any additional piece.

In the second additional experiment ('All Information'), employers are immediately provided with all ten available individual-level pieces of information about past performances of the two workers.

Both experiments reduce employers' wiggle room at least as much as the Tangible Information experiment, as employers cannot engage in systematic information search and provided information is true and closely related to the respective workers. Hiring rates between treatment groups in these experiments are almost identical (see D1 of Appendix D). This provides further evidence consistent with discrimination based on motivated reasoning.

6 Conclusion

While the existence of discrimination in many contexts has been documented extensively, we still know much less about particular mechanisms of discrimination. This study contributes to filling this gap by studying a new explanation that links taste-based discrimination and statistical discrimination. By taking a closer look at how individuals deal with individual-level information about other individuals, I identify a form of discrimination that differs from purely taste-based discrimination as well as from documented forms of statistical discrimination. In a series of experiments, participants act as employers and repeatedly engage in binary hiring decisions. When confronted with information that leaves a lot of wiggle room for interpretation, individuals make use of

the inconclusiveness of the information and engage with the information systematically differently, depending on the informational content. They search for information that confirms their motive and put less weight on information that contradicts their motive, and ultimately discriminate. Only when the wiggle room of individuals to interpret information decreases through more conclusive information structures, discrimination decreases. This implies that employers respond to information, but that this response is influenced by their motive.

Distinguishing between various forms of discrimination has important implications for designing targeted policy interventions. This paper complements the idea by Bohren et al. (2019a) who state that when group statistics are equal between two groups, remaining discrimination looks like taste-based discrimination but might actually be belief-based discrimination, as discriminators are not aware of the equal group statistics. I extend their argument by showing that even if employers are aware of equal group statistics, remaining discrimination still does not need to be taste-based. Discriminators might believe that particular individuals of one group outperform individuals of another group, and these beliefs might be driven by motives. In light of discrimination based on motivated reasoning, it seems important that policy interventions take wiggle room of information into account to effectively fight discrimination. Since credible individual-level information is often difficult to provide to decision-makers, the potential for discrimination is large in many contexts. As such, contexts in which individual-level information is sometimes not existent (e.g. some delivery services) or is mainly provided by the potentially discriminated (e.g. labor market) are prone to this form of discrimination. Decision-makers might reason that individual-level information from and about a particular individual is only true if it confirms their motive but false if it contradicts their motive. It is therefore important, that information interventions provide means to signal the validity of information, which reduces decision-makers' wiggle room to engage in motivated reasoning.

Even though this study provides first evidence for the existence of discrimination based on motivated reasoning the exact mechanisms of how information

is processed such that decision-makers ultimately discriminate remain yet to be studied. Future research could employ process tracing techniques to look more closely at the information processing behavior (Chen and Fischbacher, 2016; Lahey and Oxley, 2016). Future studies could also study the interplay of individual-level and group-level information in discrimination contexts in more detail. It remains unclear to what extent different information acquisition and processing behavior can be linked to concepts like groupiness (Kranton and Sanders, 2017), or other individual characteristics and attitudes. Finally, studying discrimination based on motivated reasoning in the field could yield first evidence of this form of discrimination in real-world settings.

References

Albrecht, K., Von Essen, E., Parys, J., and Szech, N. (2013). "Updating, self-confidence, and discrimination". *European Economic Review* 60, pp. 144–169.

Alsan, M., Garrick, O., and Graziani, G. (2019). "Does diversity matter for health? Experimental evidence from Oakland". *American Economic Review* 109 (12), pp. 4071–4111.

Arnold, D., Dobbie, W., and Hull, P. (2021). "Measuring racial discrimination in algorithms". *AEA Papers and Proceedings*. Vol. 111, pp. 49–54.

Arnold, D., Dobbie, W., and Yang, C. S. (2018). "Racial bias in bail decisions". *The Quarterly Journal of Economics* 133 (4), pp. 1885–1932.

Arrow, K. J. (1973). "Innovation and Intellectual Property Rights". *Discrimination in Labor Markets*. Ed. by O. Ashenfelter and A. Rees. Princeton NJ: Princeton University Press.

Barron, K., Ditlmann, R., Gehrig, S., and Schweighofer-Kodritsch, S. (2022). "Explicit and implicit belief-based gender discrimination: A hiring experiment".

Bartoš, V., Bauer, M., Chytilová, J., and Matějka, F. (2016). "Attention discrimination: Theory and field experiments with monitoring information acquisition". *American Economic Review* 106 (6), pp. 1437–75.

Becker, G. S. (1957). *1971:" The economics of discrimination", (1971)*.

Bénabou, R. (2013). "Groupthink: Collective delusions in organizations and markets". *Review of Economic Studies* 80 (2), pp. 429–462.

Bénabou, R. and Tirole, J. (2002). "Self-confidence and personal motivation". *The quarterly journal of economics* 117 (3), pp. 871–915.

Bertrand, M. and Duflo, E. (2017). "Field experiments on discrimination". *Handbook of economic field experiments* 1, pp. 309–393.

Blau, F. D. and Kahn, L. M. (2017). "The gender wage gap: Extent, trends, and explanations". *Journal of economic literature* 55 (3), pp. 789–865.

Bohren, A., Imas, A., and Rosenberg, M. (2018). "The Language of Discrimination: Using Experimental versus Observational Data". *AEA Papers and Proceedings*. Vol. 108, pp. 169–74.

Bohren, J. A., Haggag, K., Imas, A., and Pope, D. G. (2019a). *Inaccurate statistical discrimination*. Tech. rep. National Bureau of Economic Research.

Bohren, J. A., Hull, P., and Imas, A. (2022). *Systemic Discrimination: Theory and Measurement*. Tech. rep. National Bureau of Economic Research.

Bohren, J. A., Imas, A., and Rosenberg, M. (2019b). "The dynamics of discrimination: Theory and evidence". *American economic review* 109 (10), pp. 3395–3436.

Bordalo, P., Coffman, K., Gennaioli, N., and Shleifer, A. (2016). "Stereotypes". *The Quarterly Journal of Economics* 131 (4), pp. 1753–1794.

— (2019). "Beliefs about gender". *American Economic Review* 109 (3), pp. 739–73.

Brunnermeier, M. K. and Parker, J. A. (2005). "Optimal expectations". *American Economic Review* 95 (4), pp. 1092–1118.

Bursztyn, L., González, A. L., and Yanagizawa-Drott, D. (2020). "Misperceived social norms: Women working outside the home in saudi arabia". *American economic review* 110 (10), pp. 2997–3029.

Charles, K. K. and Guryan, J. (2011). "Studying discrimination: Fundamental challenges and recent progress". *Annu. Rev. Econ.* 3 (1), pp. 479–511.

Chen, D. L., Schonger, M., and Wickens, C. (2016). "oTree – An open-source platform for laboratory, online, and field experiments". *Journal of Behavioral and Experimental Finance* 9, pp. 88–97.

Chen, F. and Fischbacher, U. (2016). "Response time and click position: Cheap indicators of preferences". *Journal of the Economic Science Association* 2 (2), pp. 109–126.

Chen, S. and Heese, C. (2021). *Fishing for good news: Motivated information acquisition*. Tech. rep. University of Bonn and University of Mannheim, Germany.

Coffman, K. B., Exley, C. L., and Niederle, M. (2021). "The role of beliefs in driving gender discrimination". *Management Science* 67 (6), pp. 3551–3569.

Cornell, B. and Welch, I. (1996). "Culture, information, and screening discrimination". *Journal of political Economy* 104 (3), pp. 542–571.

Di Tella, R., Perez-Truglia, R., Babino, A., and Sigman, M. (2015). "Conveniently upset: avoiding altruism by distorting beliefs about others' altruism". *American Economic Review* 105 (11), pp. 3416–42.

Duckworth, A. L., Peterson, C., Matthews, M. D., and Kelly, D. R. (2007). "Grit: perseverance and passion for long-term goals." *Journal of personality and social psychology* 92 (6), p. 1087.

Epley, N. and Gilovich, T. (2016). "The mechanics of motivated reasoning". *Journal of Economic perspectives* 30 (3), pp. 133–40.

Esponda, I., Oprea, R., and Yuksel, S. (2022). "Discrimination Without Reason: Biases in Statistical Discrimination".

Farkas, G. (2003). "Racial disparities and discrimination in education: What do we know, how do we know it, and what do we need to know?" *Teachers College Record* 105 (6), pp. 1119–1146.

Fershtman, C. and Gneezy, U. (2001). "Discrimination in a segmented society: An experimental approach". *The Quarterly Journal of Economics* 116 (1), pp. 351–377.

Festinger, L. (1957). *A theory of cognitive dissonance*. Vol. 2. Stanford university press.

Goncalves, F. and Mello, S. (2021). "A few bad apples? Racial bias in policing". *American Economic Review* 111 (5), pp. 1406–41.

Grossman, Z. and Van der Weele, J. J. (2017). "Self-image and willful ignorance in social decisions". *Journal of the European Economic Association* 15 (1), pp. 173–217.

Heckman, J. J. (1998). "Detecting discrimination". *Journal of economic perspectives* 12 (2), pp. 101–116.

Köszegi, B. (2006). "Ego utility, overconfidence, and task choice". *Journal of the European Economic Association* 4 (4), pp. 673–707.

Kranton, R. E. and Sanders, S. G. (2017). "Groupy versus non-groupy social preferences: Personality, region, and political party". *American Economic Review* 107 (5), pp. 65–69.

Lahey, J. N. and Oxley, D. (2016). "The power of eye tracking in economics experiments". *American Economic Review* 106 (5), pp. 309–13.

Mengel, F. and Campos-Mercade, P. (2021). "Irrational Statistical Discrimination". *Available at SSRN 3843579*.

Neumark, D. (2018). "Experimental research on labor market discrimination". *Journal of Economic Literature* 56 (3), pp. 799–866.

Neumark, D., Burn, I., and Button, P. (2016). "Experimental age discrimination evidence and the Heckman critique". *American Economic Review* 106 (5), pp. 303–08.

Onuchic, P. (2022). "Recent Contributions to Theories of Discrimination". *arXiv preprint arXiv:2205.05994*.

Phelps, E. S. (1972). "The statistical theory of racism and sexism". *American economic review* 62 (4), pp. 659–661.

Rammstedt, B. and John, O. P. (2005). "Kurzversion des big five inventory (BFI-K)". *Diagnostica* 51 (4), pp. 195–206.

Reuben, E., Sapienza, P., and Zingales, L. (2014). "How stereotypes impair women's careers in science". *Proceedings of the National Academy of Sciences* 111 (12), pp. 4403–4408.

Sinclair, V. G. and Wallston, K. A. (2004). "The development and psychometric evaluation of the Brief Resilient Coping Scale". *Assessment* 11 (1), pp. 94–101.

Thaler, M. (2020). "The Fake News Effect: Experimentally Identifying Motivated Reasoning Using Trust in News". *Available at SSRN 3717381*.

Zimmermann, F. (2020). "The dynamics of motivated beliefs". *American Economic Review* 110 (2), pp. 337–61.

For Online Publication

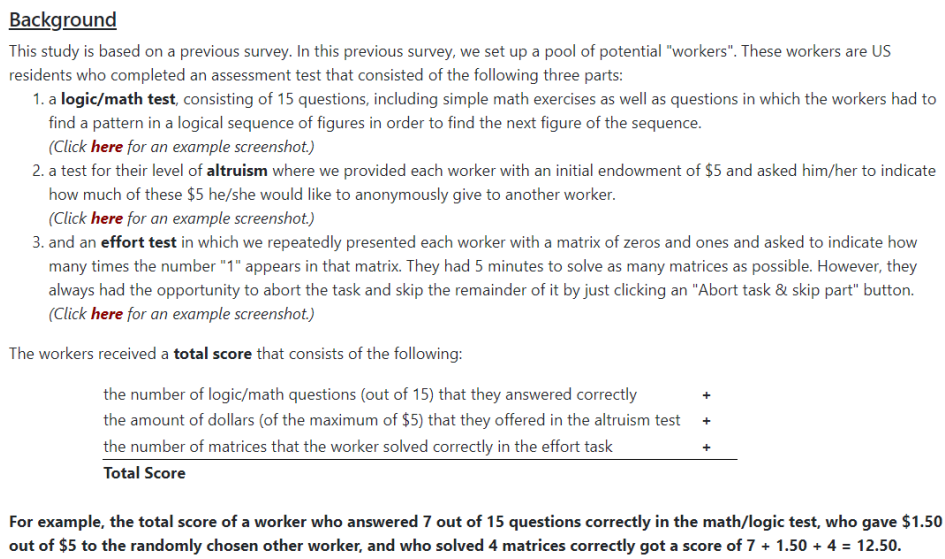
Appendix

A Experimental Design

This section provides example screenshots of all relevant screens for both treatments and across experiments.

A.1 Instructions

Figure A1: Background



Background

This study is based on a previous survey. In this previous survey, we set up a pool of potential "workers". These workers are US residents who completed an assessment test that consisted of the following three parts:

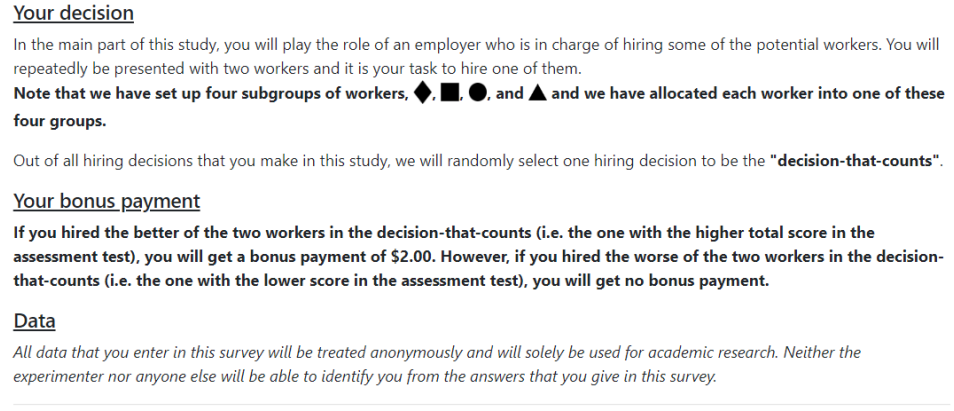
1. a **logic/math test**, consisting of 15 questions, including simple math exercises as well as questions in which the workers had to find a pattern in a logical sequence of figures in order to find the next figure of the sequence.
   *(Click **here** for an example screenshot.)*
2. a test for their level of **altruism** where we provided each worker with an initial endowment of $5 and asked him/her to indicate how much of these $5 he/she would like to anonymously give to another worker.
   *(Click **here** for an example screenshot.)*
3. and an **effort test** in which we repeatedly presented each worker with a matrix of zeros and ones and asked to indicate how many times the number "1" appears in that matrix. They had 5 minutes to solve as many matrices as possible. However, they always had the opportunity to abort the task and skip the remainder of it by just clicking an "Abort task & skip part" button.
   *(Click **here** for an example screenshot.)*

The workers received a **total score** that consists of the following:

the number of logic/math questions (out of 15) that they answered correctly	+
the amount of dollars (of the maximum of $5) that they offered in the altruism test	+
the number of matrices that the worker solved correctly in the effort task	+
Total Score	

For example, the total score of a worker who answered 7 out of 15 questions correctly in the math/logic test, who gave $1.50 out of $5 to the randomly chosen other worker, and who solved 4 matrices correctly got a score of 7 + 1.50 + 4 = 12.50.

Figure A2: Decision (group Neutral)



Your decision

In the main part of this study, you will play the role of an employer who is in charge of hiring some of the potential workers. You will repeatedly be presented with two workers and it is your task to hire one of them.
Note that we have set up four subgroups of workers, ♦, ■, ●, and ▲ and we have allocated each worker into one of these four groups.

Out of all hiring decisions that you make in this study, we will randomly select one hiring decision to be the **"decision-that-counts"**.

Your bonus payment

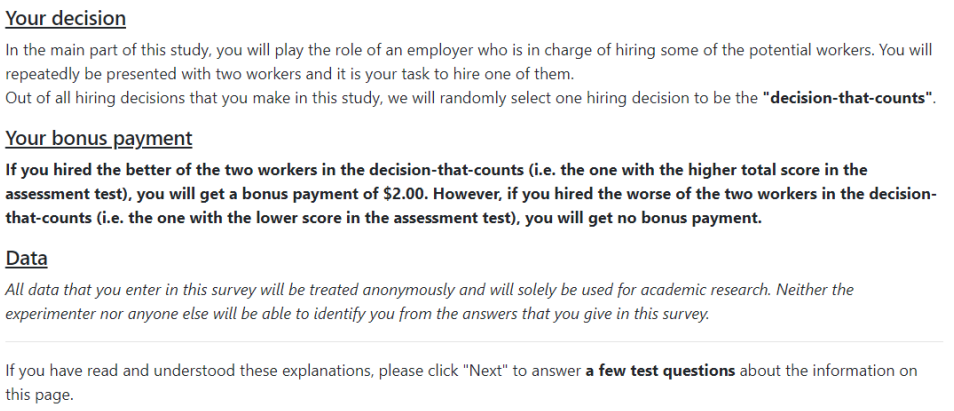
If you hired the better of the two workers in the decision-that-counts (i.e. the one with the higher total score in the assessment test), you will get a bonus payment of $2.00. However, if you hired the worse of the two workers in the decision-that-counts (i.e. the one with the lower score in the assessment test), you will get no bonus payment.

Data

All data that you enter in this survey will be treated anonymously and will solely be used for academic research. Neither the experimenter nor anyone else will be able to identify you from the answers that you give in this survey.

If you have read and understood these explanations, please click "Next" to answer **a few test questions** about the information on this page.

Figure A3: Decision (group Race)



Your decision

In the main part of this study, you will play the role of an employer who is in charge of hiring some of the potential workers. You will repeatedly be presented with two workers and it is your task to hire one of them.
Out of all hiring decisions that you make in this study, we will randomly select one hiring decision to be the **"decision-that-counts"**.

Your bonus payment

If you hired the better of the two workers in the decision-that-counts (i.e. the one with the higher total score in the assessment test), you will get a bonus payment of $2.00. However, if you hired the worse of the two workers in the decision-that-counts (i.e. the one with the lower score in the assessment test), you will get no bonus payment.

Data

All data that you enter in this survey will be treated anonymously and will solely be used for academic research. Neither the experimenter nor anyone else will be able to identify you from the answers that you give in this survey.

If you have read and understood these explanations, please click "Next" to answer **a few test questions** about the information on this page.

Figure A4: Test Questions

True or False: The workers completed a logic/math test, an altruism test and an effort test.

- ○ True
- ○ False

True or False: A very good worker to hire is one that has very good logic/math skills, is very altruistic and puts a lot of effort into tasks.

- ○ True
- ○ False

What is the total score of a worker who solved 5 math/logic questions correctly, who gave $2.44 to the randomly chosen other worker and who solved 3 matrices correctly in the effort task?

How high would your bonus payoff be, if you hired a worker with a total score of 15.3 instead of a worker with a total score of 10.5 in the decision-that-counts?

$

How high would your bonus payoff be, if you instead hired a worker with a total score of 2.60 instead of a worker with a total score of 5 in the decision-that-counts?

$

A.2 Belief Stage

This section shows screenshots of the belief stage. The screenshots show the elicitation of beliefs about Asian workers. The elicitation of beliefs about workers from the other groups was conducted identically.

Figure A5: Elicitation of Priors (group Neutral)

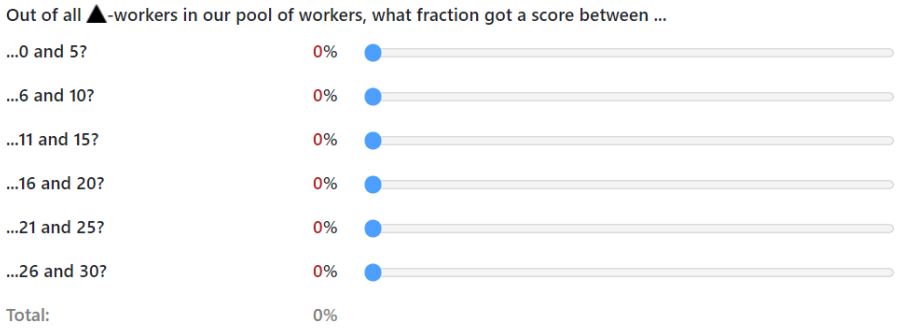


Figure A6: Elicitation of Priors (group Race)

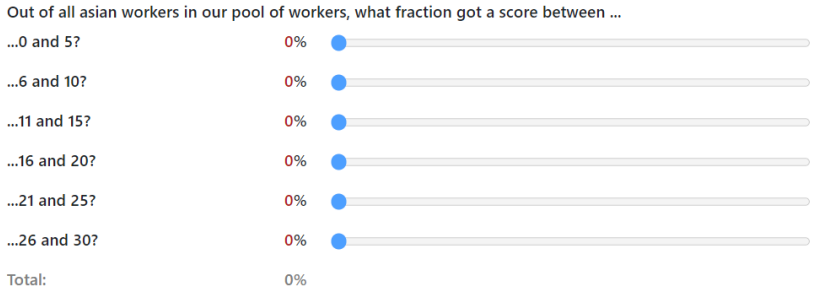


Figure A7: Elicitation of Posteriors (group Neutral)

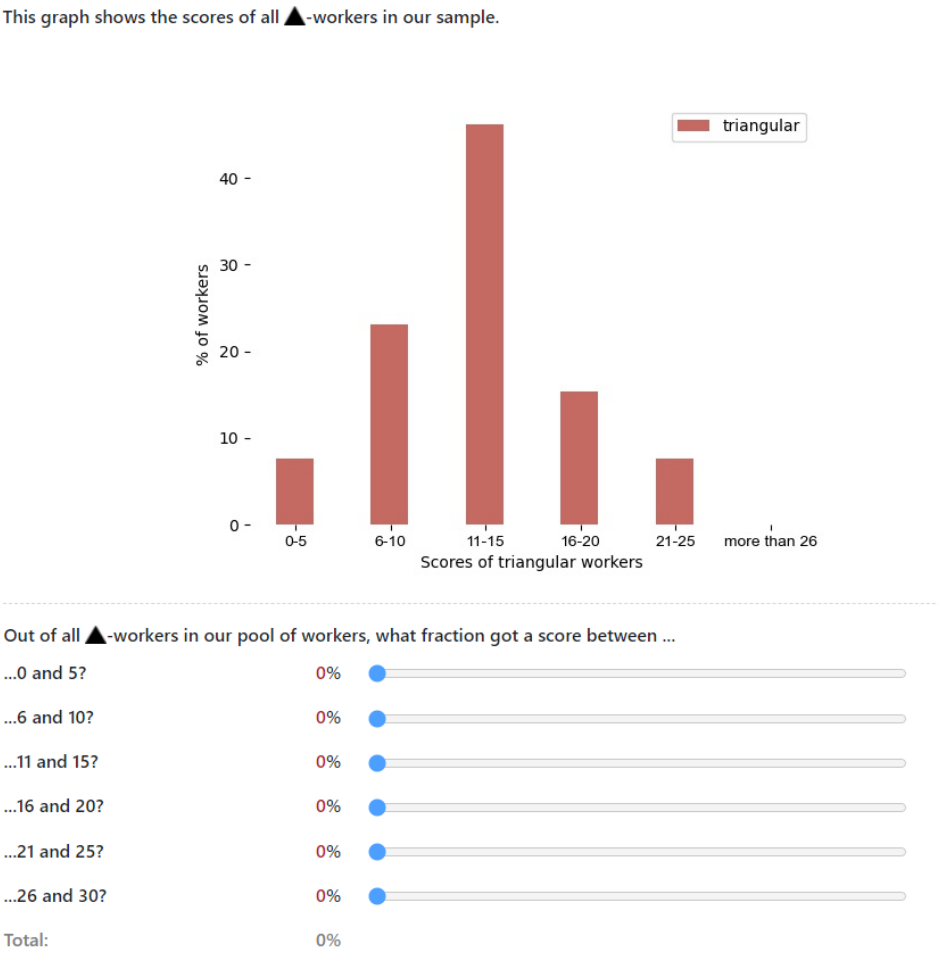
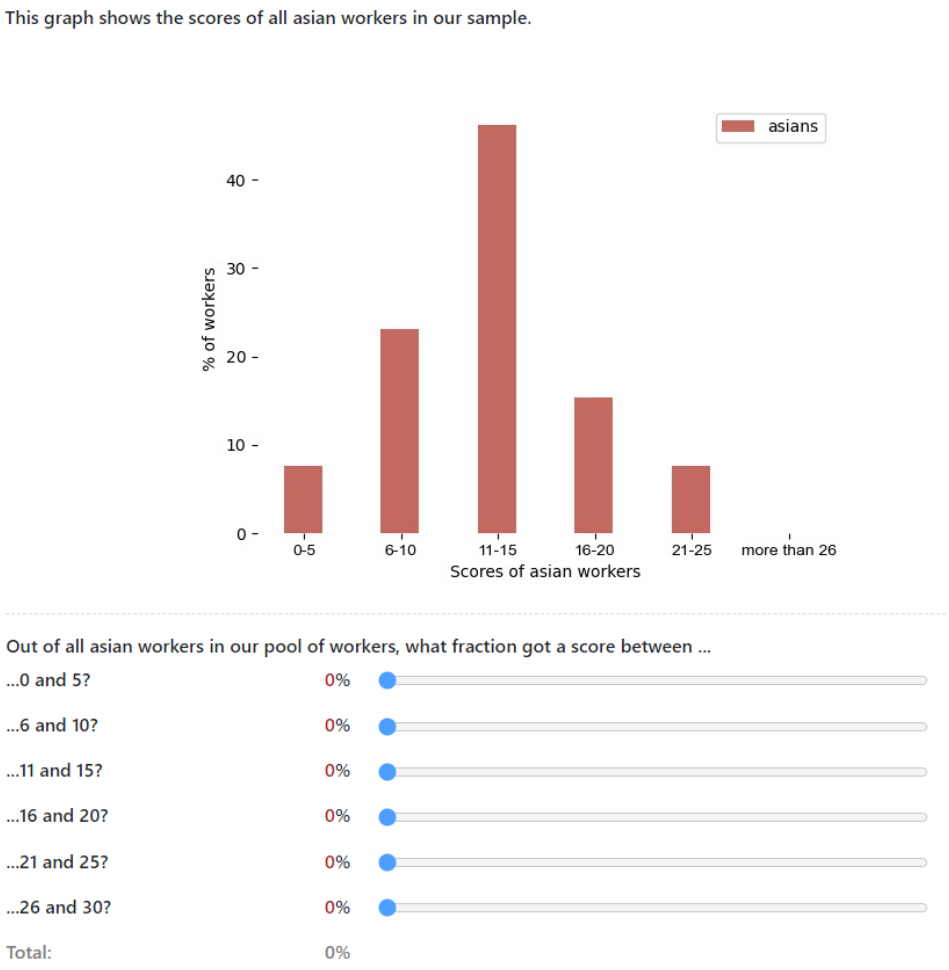
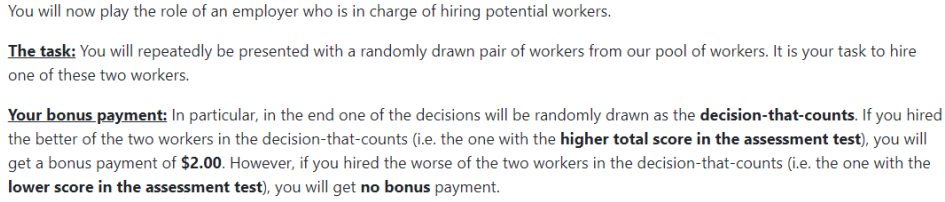


Figure A8: Elicitation of Posteriors (group Race)



A.3 Hiring Stage

Figure A9: Hiring Instructions (Experiment 'No Information')

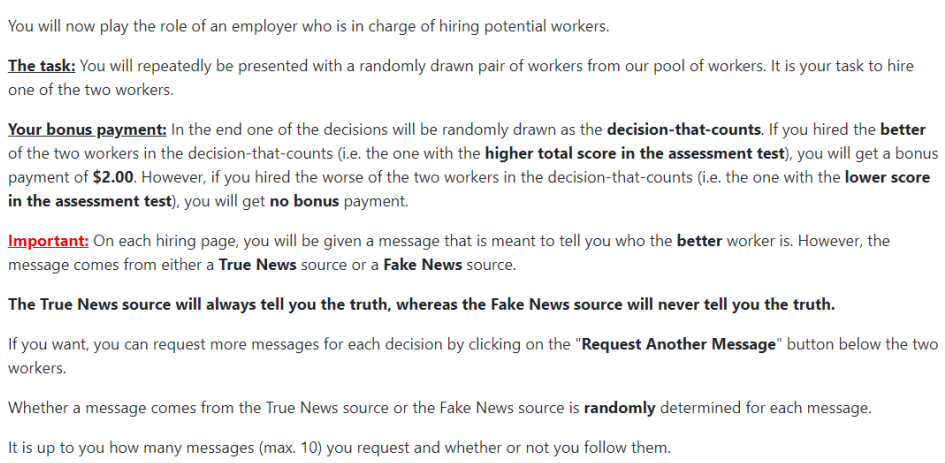


You will now play the role of an employer who is in charge of hiring potential workers.

The task: You will repeatedly be presented with a randomly drawn pair of workers from our pool of workers. It is your task to hire one of these two workers.

Your bonus payment: In particular, in the end one of the decisions will be randomly drawn as the **decision-that-counts**. If you hired the better of the two workers in the decision-that-counts (i.e. the one with the **higher total score in the assessment test**), you will get a bonus payment of **$2.00**. However, if you hired the worse of the two workers in the decision-that-counts (i.e. the one with the **lower score in the assessment test**), you will get **no bonus** payment.

Figure A10: Hiring Instructions (Experiment 'Ambiguous Information')



You will now play the role of an employer who is in charge of hiring potential workers.

The task: You will repeatedly be presented with a randomly drawn pair of workers from our pool of workers. It is your task to hire one of the two workers.

Your bonus payment: In the end one of the decisions will be randomly drawn as the **decision-that-counts**. If you hired the **better** of the two workers in the decision-that-counts (i.e. the one with the **higher total score in the assessment test**), you will get a bonus payment of **$2.00**. However, if you hired the worse of the two workers in the decision-that-counts (i.e. the one with the **lower score in the assessment test**), you will get **no bonus** payment.

Important: On each hiring page, you will be given a message that is meant to tell you who the **better** worker is. However, the message comes from either a **True News** source or a **Fake News** source.

The True News source will always tell you the truth, whereas the Fake News source will never tell you the truth.

If you want, you can request more messages for each decision by clicking on the "**Request Another Message**" button below the two workers.

Whether a message comes from the True News source or the Fake News source is **randomly** determined for each message.

It is up to you how many messages (max. 10) you request and whether or not you follow them.

Figure A11: Hiring Instructions (Experiment 'Uncertain Information')

You will now play the role of an employer who is in charge of hiring potential workers.

The task: You will repeatedly be presented with a randomly drawn pair of workers from our pool of workers. It is your task to hire one of the two workers.

Your bonus payment: In the end one of the decisions will be randomly drawn as the **decision-that-counts**. If you hired the **better** of the two workers in the decision-that-counts (i.e. the one with the **higher total score in the assessment test**), you will get a bonus payment of **$2.00**. However, if you hired the worse of the two workers in the decision-that-counts (i.e. the one with the **lower score in the assessment test**), you will get **no bonus** payment.

Important: On each hiring page, you will be given a message that is meant to tell you who the **better** worker is. However, the message comes from either a **True News** source or a **Fake News** source.

The True News source will always tell you the truth, whereas the Fake News source will never tell you the truth.

If you want, you can request more messages for each decision by clicking on the "**Request Another Message**" button below the two workers.

The likelihood that a message comes from the **True News source is 60%**, the likelihood that a message comes from the **Fake News source is 40%**.

It is up to you how many messages (max. 10) you request and whether or not you follow them.

Figure A12: Hiring Instructions (Experiment 'Tangible Information')

You will now play the role of an employer who is in charge of hiring potential workers.

The task: You will repeatedly be presented with a randomly drawn pair of workers from our pool of workers. It is your task to hire one of these two workers.

Your bonus payment: In particular, in the end one of the decisions will be randomly drawn as the **decision-that-counts**. If you hired the better of the two workers in the decision-that-counts (i.e. the one with the **higher total score in the assessment test**), you will get a bonus payment of **$2.00**. However, if you hired the worse of the two workers in the decision-that-counts (i.e. the one with the **lower score in the assessment test**), you will get **no bonus** payment.

Important: On each hiring page, you will be given a random piece of information about the two workers.

If you want you can request more information for each decision by clicking on the **"Get more information"** button below the two workers.

It is up to you how many pieces of information (max. 10) you request and whether or not you follow them.

Figure A13: Hiring Decision (Experiment 'No Information', group Neutral)

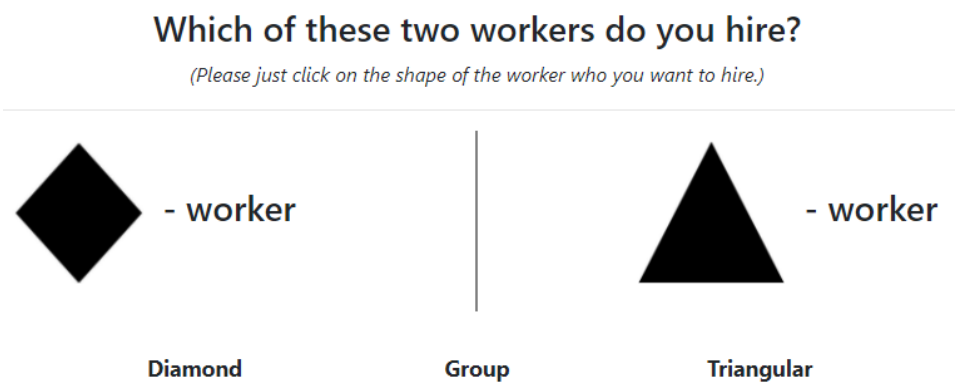


Figure A14: Hiring Decision (Experiment 'No Information', group Race)

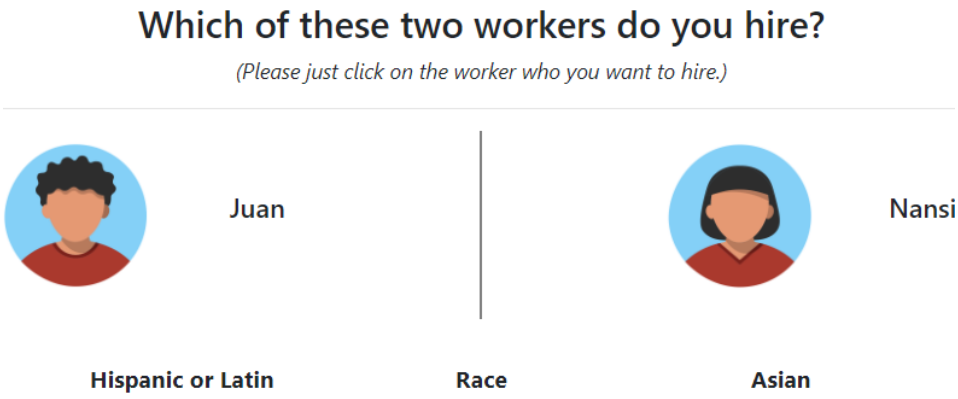


Figure A15: Hiring Decision (Experiment 'Ambiguous Information' & 'Uncertain Information', group Neutral)

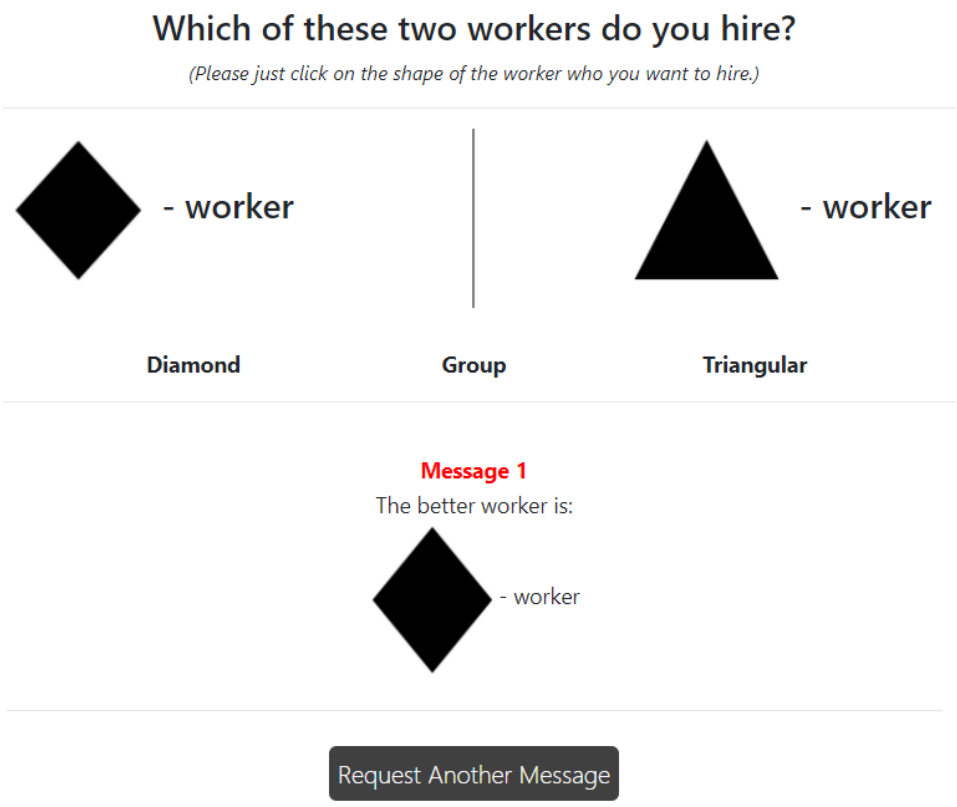


Figure A16: Hiring Decision (Experiment 'Ambiguous Information' & 'Uncertain Information', group Race)

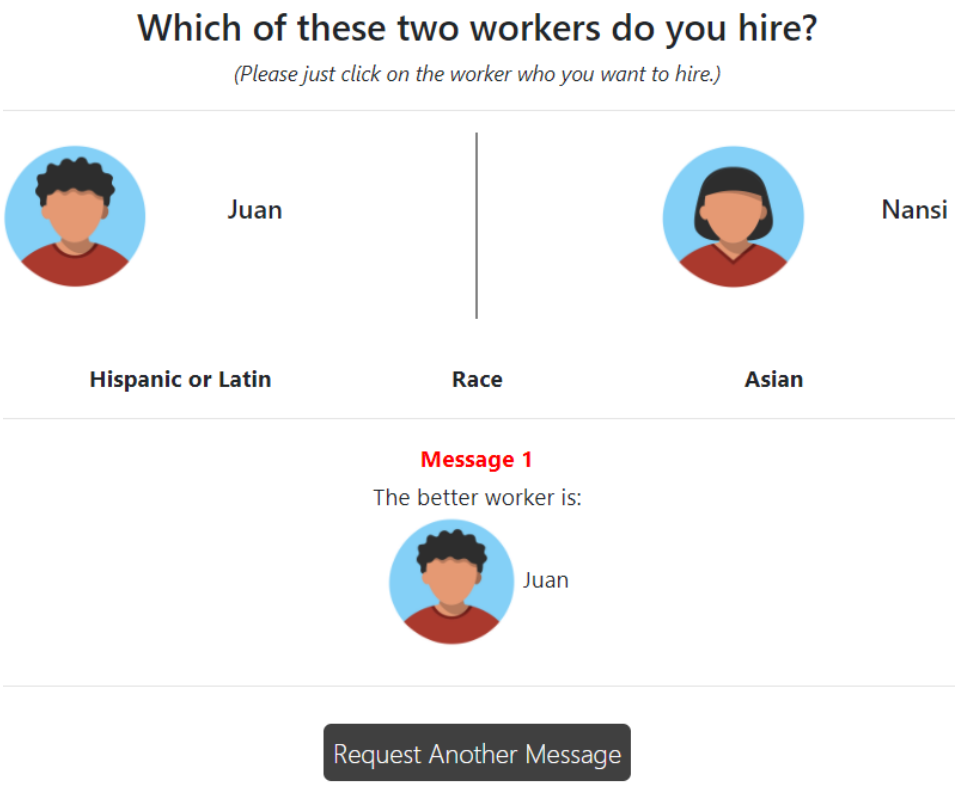


Figure A17: Hiring Decision (Experiment 'Tangible Information', group Neutral)

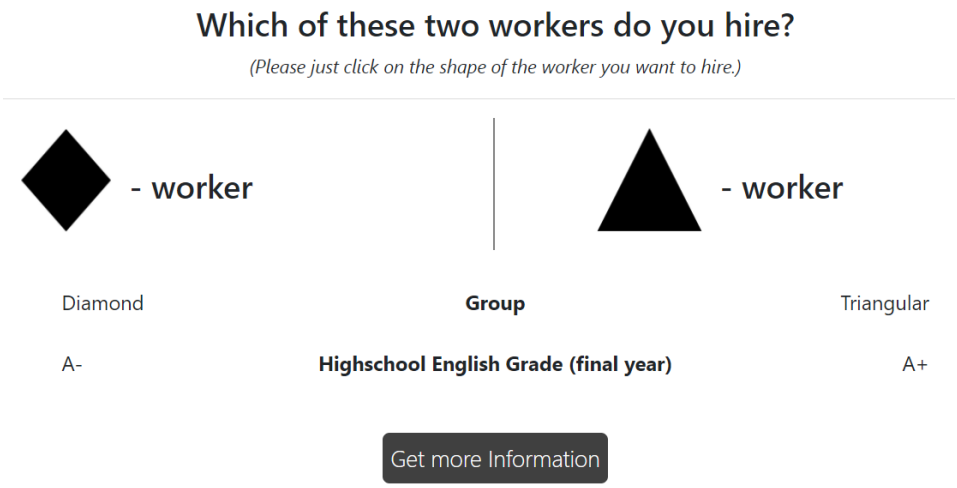
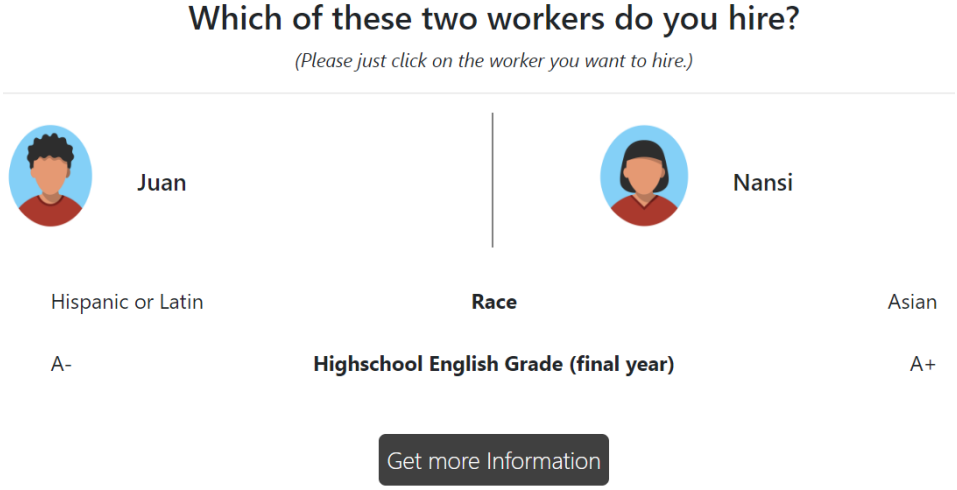


Figure A18: Hiring Decision (Experiment 'Tangible Information', group Race)



A.4 Deviation from Preregistration

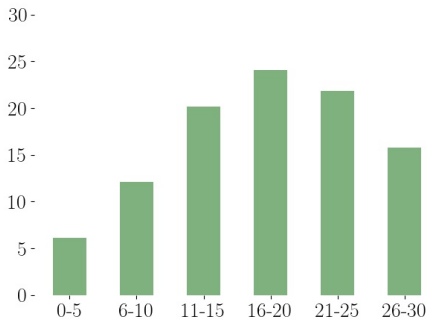
The sample sizes of the different experiments deviate from what has been specified in the preregistration file. While the preregistration aimed at 500 to 600 employers per experiment, especially the experiments with decreased wiggle room ('Uncertain Information') and ('Tangible Information') include less participants than preregistered. While this is due to resource constraints, the analyses of the hypotheses are well-powered with current sample sizes.

B Group-level Beliefs

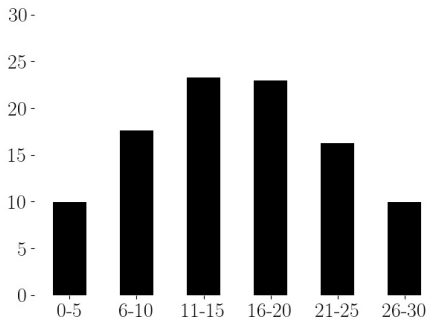
This section first shows the separate prior belief distributions for Asian and Hispanic workers. For this, I pool data on beliefs from all four experiments, as distributions were elicited before the hiring stage. Distributions for particular experiments are available on request, however, they do not differ from the pooled distributions.

Figure B1: Prior belief distributions

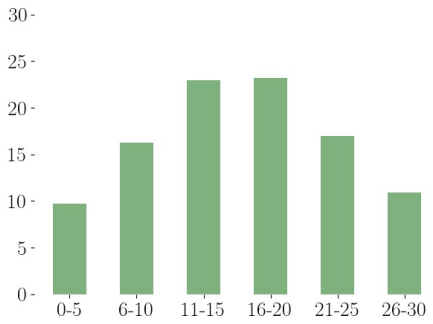
(a) Priors about Asians – Race



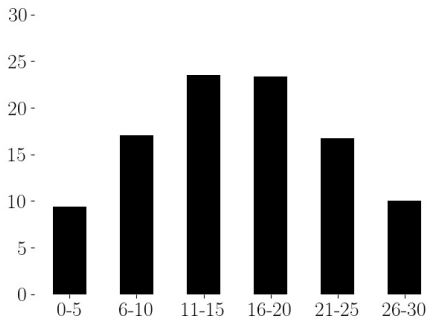
(b) Priors about Asians – Neutral



(c) Priors about Hispanics – Race



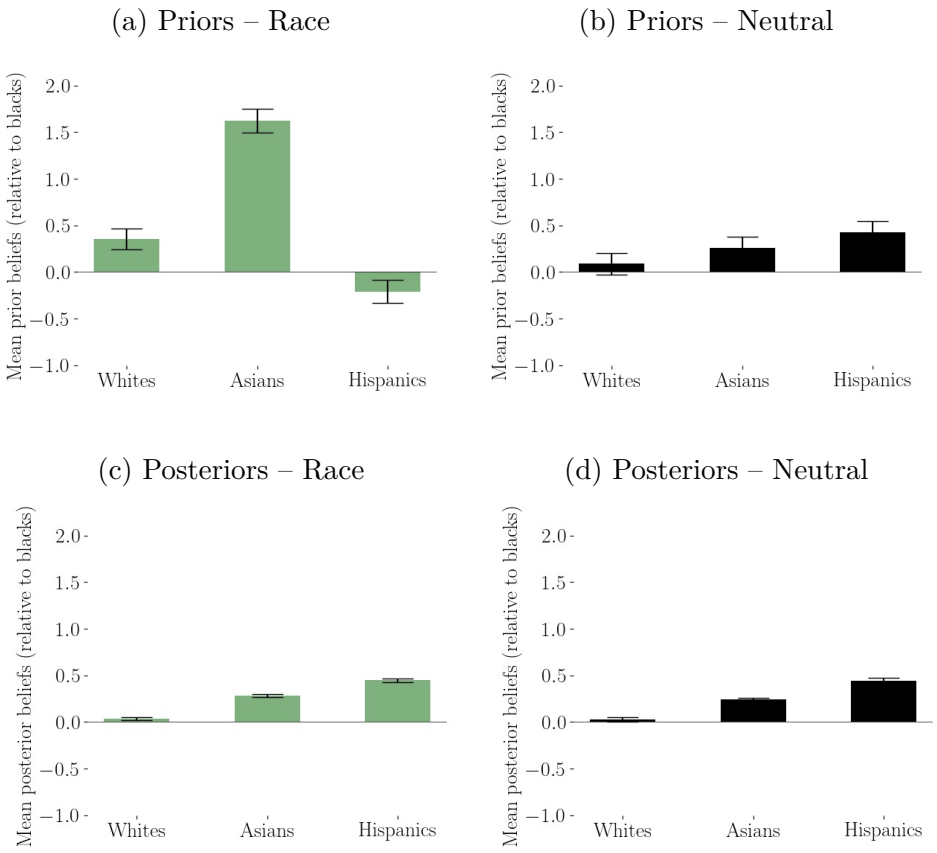
(d) Priors about Hispanics – Neutral



Notes: Figures (a) and (b) show the employers' belief distributions about the productivities of Asian workers before the group-level information update in group Race and group Neutral, respectively, figures (c) and (d) show the employers' belief distributions about the productivities of Hispanic workers before the group-level information update in group Race and group Neutral, respectively.

Figure B2 shows the relation between mean beliefs about black, white, Asian and Hispanic workers and provides evidence, that the largest difference in prior beliefs, and hence in potential for motivated beliefs, lies in the difference between Asian and Hispanic workers.

Figure B2: Mean beliefs about group productivities



Notes: Figures (a) and (b) show the mean beliefs before the group-level information update of employers in group Race and group Neutral, respectively, figures (c) and (d) show the mean beliefs after the group-level information update of employers in group Race and group Neutral, respectively. Beliefs are plotted as the difference in means compared to beliefs about blacks. Error bars show the standard errors of the means.

C Results

This section provides additional results. Section C.1 adds the numbers to Figure 3 of the main text. Subsequently, Section C.2 provides results of some robustness exercises.

C.1 Discrimination

Table C1 shows discrimination rates across experiments and thereby provides the numbers for Figure 3 of the main text.

Table C1: Discrimination rates across experiments

	Dep. var: hired hispanic			
	(1) no info	(2) ambiguous info	(3) uncertain info	(4) tangible info
race	−0.2353***	−0.1392***	−0.0614	0.0281
	(*0.0285*)	(*0.0229*)	(*0.0408*)	(*0.0302*)
Observations	3633	3290	756	742
Baseline mean dep. var.	0.5064	0.4994	0.5230	0.4606

Notes: In column (1) employers did not receive any individual-level information regarding the two workers, in column (2) they received ambiguous information, in column (3) uncertain information, and in column (4) tangible information. The dependent variable in both models is a dummy equal to 1 if the Hispanic worker was hired and 0 if the Asian worker was hired. 'race' is the treatment dummy and equal to 1 if the decision was made in treatment group race (showing the respective races) and 0 otherwise. Standard errors are clustered at the individual level and displayed in parentheses. *p< 0.1, **p< 0.05, ***p< 0.01.

C.2 Robustness Checks

During the Instructions, participants in all four experiments had to answer a few training questions in order to make sure that instructions were understood. While the main text includes all participants that answered at least 3 out of 5 of these questions correctly, this section shows results for stricter limits. Section C.2.1 shows results for participants who answered at least 4 out of 5 questions correctly, Section C.2.2 for those who answered all test questions correctly. The results of the main text are robust to these variations.

C.2.1 At least 4 test questions correct

Table C2: Hiring Rates of Hispanic workers across experiments

	(1) No info	(2) Ambiguous info	(3) Uncertain info	(4) Tangible info
race	-0.2295^{***}	-0.1328^{***}	-0.0674	0.0102
	(*0.0316*)	(*0.0243*)	(*0.0448*)	(*0.0319*)
Observations	3071	2917	663	586
Baseline mean dep. var.	0.5000	0.4960	0.5228	0.4728

Notes: This table includes decisions of participants who completed at least 4 test questions correctly. In column (1) employers did not receive any individual-level information regarding the two workers, in column (2) they received ambiguous information, in column (3) uncertain information, and in column (4) tangible information. The dependent variable in both models is a dummy equal to 1 if the Hispanic worker was hired and 0 if the Asian worker was hired. 'race' is the treatment dummy and equal to 1 if the decision was made in treatment group race (showing the respective races) and 0 otherwise. Standard errors are clustered at the individual level and displayed in parentheses. *p< 0.1, **p< 0.05, ***p< 0.01.

Table C3: Information Behavior across Experiments

(a)	Ambiguous Information		
	(1) 2nd signals	(2) number of signals	(3) follow signals
race * hispanic	0.0945**	0.4630*	−0.1736***
	(*0.0389*)	(*0.2526*)	(*0.0430*)
race	−0.0572	−0.3650	0.0550***
	(*0.0458*)	(*0.2988*)	(*0.0210*)
hispanic	−0.0268	−0.2019	−0.0752***
	(*0.0271*)	(*0.1792*)	(*0.0280*)
Observations	2919	2919	2919
Baseline mean dep. var.	0.5496	3.6684	0.8740

(b)	Uncertain Information			Uncertain - Ambiguous
	(1) 2nd signals	(2) number of signals	(3) follow signals	(4) follow signals
race * hispanic	0.1437**	0.6262	−0.0501	0.1235*
	(*0.0634*)	(*0.5870*)	(*0.0682*)	(*0.0620*)
race	−0.0400	−0.0779	0.0296	
	(*0.0836*)	(*0.8118*)	(*0.0258*)	
hispanic	−0.0533	0.1056	−0.0478	
	(*0.0468*)	(*0.3979*)	(*0.0329*)	
Observations	665	665	665	
Baseline mean dep. var.	0.7738	5.8571	0.9432	

(c)	Tangible Information			Tangible - Ambiguous
	(1) 2nd signals	(2) number of signals	(3) follow signals	(4) follow signals
race * hispanic	0.0260	0.5709	−0.0206	0.1675***
	(*0.0318*)	(*0.3937*)	(*0.0492*)	(*0.0620*)
race	0.0381	0.8460	−0.0328	
	(*0.0374*)	(*0.5347*)	(*0.0287*)	
hispanic	0.0051	0.1056	−0.0860**	
	(*0.0235*)	(*0.3137*)	(*0.0365*)	
Observations	742	742	742	
Baseline mean dep. var.	0.9135	5.5240	0.8894	

Notes: This table includes decisions of participants who completed at least 4 test questions correctly. In panel (a) employers received ambiguous information, in panel (b) uncertain information, and in panel (c) tangible information. *2nd signal*, is a dummy equal to 1 if an employer requested a second signal. *number of signals*, counts the number of requested signals. *follow signals*, is a dummy equal to 1 if an employer's hiring decision followed the majority of considered signals. *race* is the treatment dummy equal to 1 if the decision was made in treatment group race. *hispanic* is a dummy equal to 1 if the initial signal (columns 1 and 2) or the direction of all considered signals (column 3) suggests to hire the Hispanic worker. Units of observation are decision specific. Standard errors are clustered at the individual level and displayed in parentheses. *p< 0.1, **p< 0.05, ***p< 0.01.

C.2.2 All 5 test questions correct

Table C4: Hiring Rates of Hispanic workers across experiments

	Dep. var: hired hispanic			
	(1) no info	(2) ambiguous info	(3) uncertain info	(4) tangible info
race	-0.2324^{***}	-0.1665^{***}	-0.0625	0.0131
	(*0.0396*)	(*0.0303*)	(*0.0599*)	(*0.0388*)
Observations	1911	1869	448	364
Baseline mean dep. var.	0.4936	0.5143	0.4955	0.4845

Notes: This table includes decisions of participants who completed all 5 test questions correctly. In column (1) employers did not receive any individual-level information regarding the two workers, in column (2) they received ambiguous information, in column (3) uncertain information, and in column (4) tangible information. The dependent variable in both models is a dummy equal to 1 if the Hispanic worker was hired and 0 if the Asian worker was hired. 'race' is the treatment dummy and equal to 1 if the decision was made in treatment group race (showing the respective races) and 0 otherwise. Standard errors are clustered at the individual level and displayed in parentheses. *p< 0.1, **p< 0.05, ***p< 0.01.

Table C5: Information Behavior across Experiments

(a)	Ambiguous Information		
	(1) 2nd signals	(2) number of signals	(3) follow signals
race * hispanic	0.0873*	0.4469	−0.1984***
	(*0.0485*)	(*0.3259*)	(*0.0549*)
race	−0.0111	−0.0960	0.0846***
	(*0.0577*)	(*0.3964*)	(*0.0247*)
hispanic	0.0043	−0.1242	−0.0608*
	(*0.0321*)	(*0.2189*)	(*0.0343*)
Observations	1869	1869	1869
Baseline mean dep. var.	0.5210	3.6806	0.8669

(b)	Uncertain Information			Uncertain - Ambiguous
	(1) 2nd signals	(2) number of signals	(3) follow signals	(4) follow signals
race * hispanic	0.0860	0.0222	−0.0385	0.1600*
	(*0.0752*)	(*0.7070*)	(*0.0731*)	(*0.0917*)
race	0.0154	0.6590	−0.0072	
	(*0.0958*)	(*0.9525*)	(*0.0202*)	
hispanic	−0.0120	0.5375	−0.0931**	
	(*0.0552*)	(*0.4779*)	(*0.0440*)	
Observations	448	448	448	
Baseline mean dep. var.	0.7826	5.7652	0.9754	

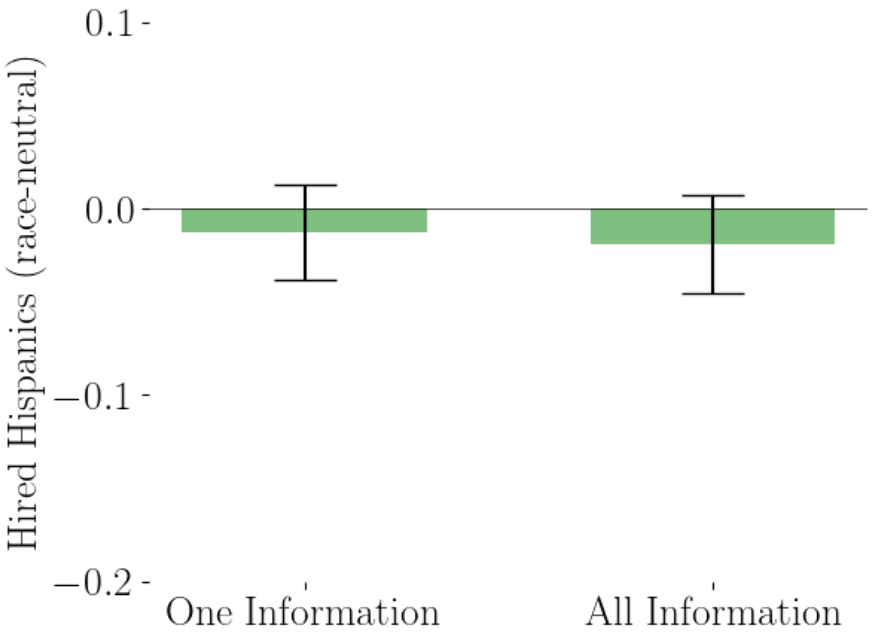
(c)	Tangible Information			Tangible - Ambiguous
	(1) 2nd signals	(2) number of signals	(3) follow signals	(4) follow signals
race * hispanic	0.0568	0.9549*	−0.1000	0.0984
	(*0.0503*)	(*0.5274*)	(*0.0670*)	(*0.0834*)
race	0.0486	0.7249	0.0134	
	(*0.0664*)	(*0.7616*)	(*0.0402*)	
hispanic	0.0021	−0.1436	−0.0124**	
	(*0.0291*)	(*0.3887*)	(*0.0496*)	
Observations	364	364	364	
Baseline mean dep. var.	0.8812	5.6436	0.8426	

Notes: This table includes decisions of participants who completed all 5 test questions correctly. In panel (a) employers received ambiguous information, in panel (b) uncertain information, and in panel (c) tangible information. *2nd signal*, is a dummy equal to 1 if an employer requested a second signal. *number of signals*, counts the number of requested signals. *follow signals*, is a dummy equal to 1 if an employer's hiring decision followed the majority of considered signals. *race* is the treatment dummy equal to 1 if the decision was made in treatment group race. *hispanic* is a dummy equal to 1 if the initial signal (columns 1 and 2) or the direction of all considered signals (column 3) suggests to hire the Hispanic worker. Units of observation are decision specific. Standard errors are clustered at the individual level and displayed in parentheses. *p< 0.1, **p< 0.05, ***p< 0.01.

D Additional Experiments

Since neither the One Information experiment nor the All Information experiment allow to study the information acquisition and processing behavior without any process tracing techniques, results of the 'One Information' and 'All Information' experiments are not included in the main text but mainly serve as additional evidence and are briefly summarized in Figure D1 below. Neither of the experiments shows rates of discrimination, which is consistent with the limited wiggle room in both experiments.

Figure D1: Hiring rates of Hispanic workers under wiggle room



Notes: The vertical axis displays the fraction of hired Hispanics in group Race minus the fraction of hired Hispanics in group Neutral. The horizontal axis groups the decisions by experiments, either the 'One Information' experiment or the 'All Information' experiment. Error bars indicate standard errors of the means.