

Why do we Discriminate? The Role of Motivated Reasoning

Markus Eyting*

JGU Mainz, Stanford University, University Heidelberg

This version: April 28, 2022

Please do not cite or distribute

Abstract

This study contributes to our understanding of the underlying causes of discrimination by providing a link between taste and belief-based discrimination through the mechanism of motivated reasoning. It provides first evidence on how this insight can be utilized to design policy interventions that effectively alleviates this form of discrimination. In a series of online experiments, ‘experimental employers’ engage in repeated hiring decisions for which they are provided with various forms of individual-level information. It can be shown, that the employers systematically search for information that supports their motive while they disregard information that contradicts their motive. Finally, I show that reducing employers’ wiggle room to interpret information signals alleviates discrimination by reducing systematic information processing behavior.

1 Introduction

Understanding the drivers of disparate treatments of individuals from different social groups has been a goal of economic research for many years. To date, economists typically categorize direct discrimination as either taste-based (Becker, 1957) or belief-based discrimination,

**Address for correspondance:* Markus Eyting, JGU Mainz, Jakob-Welder-Weg 4, 55128 Mainz, Germany. E-mail: meyting@uni-mainz.de. Phone: +49-6131-39-22166.

based on either accurate (Phelps, 1972; Arrow, 1973) or inaccurate beliefs (Bohren et al., 2019). While taste-based discriminators experience a disutility from interactions with members of a particular group, belief-based discriminators (in)correctly rely on differences in group attributes to infer unobserved individual characteristics. While any kind of discrimination can have fatal consequences for the discriminated, a precise identification of the cause of discrimination has important implications for policy, welfare analyses and discrimination dynamics.

This study makes two contributions to the literature on discrimination: First, it adds to our understanding of its underlying causes by providing a link between taste and belief-based discrimination through the mechanism of motivated reasoning. Second, it provides first evidence on how this insight can be utilized to design a specific policy intervention that effectively alleviates this form of discrimination.

I build on the work of Bohren et al. (2019), who argue that providing credible information on relevant group-level distributions allows to separately identify inaccurate beliefs and animus as potential drivers of discrimination. In particular, they argue that those with inaccurate beliefs should adjust their behavior upon receipt of credible group-level information while those who hold inaccurate beliefs merely in order to mask an underlying animus are unlikely to change their behavior in response to information. I extend their argument by looking more closely at those who do not change their behavior in order to understand a potential channel through which animus drives discrimination. In particular, I provide evidence that after receiving credible group-level information, these individuals use the ‘wiggle room’ that group-level information provides for the formation of beliefs about unobserved individual-level characteristics. I show that these individuals apply motivated reasoning to form inaccurate beliefs about individual characteristics, by systematically acquiring and processing individual-level information. Finally, by varying the reliability of individual-level information, I show that carefully designed information interventions can still change behavior of these agents, which demonstrates that it is important to identify this particular channel of motivated reasoning.

In order to understand more precisely which specific channel I study, consider the following example: An agent makes a series of binary hiring decisions between an applicant from group A and an applicant from group B. The agent is not necessarily aware, that individuals from group A and B are equally productive but forms their own belief about respective group productivity levels. If the agent discriminates against individuals from one group, we would

usually conclude that discrimination could either be based on taste or on inaccurate beliefs. Now, the agent is provided with credible information on the relevant distributions, indicating that individuals of group A and B are equally productive (in mean and variance). Following the argument by Bohren et al. (2019), if the agent still discriminates, we would conclude, that discrimination is based on taste. I take a step further and identify one potential channel through which taste affects discrimination – motivated reasoning. Even though the agent now knows that individuals from group A and B are equally productive, the agent might still always decide to hire from the one group in repeated binary decisions by repeatedly believing that “in this particular case individual from group A is better than individual from group B”. This implies, that the agent uses the ‘wiggle room’ that group-level (or ambiguous individual-level) information provides in order to form individual-level beliefs in line with their motives. I study to what extent and under which conditions these motives drive the formation of individual-level beliefs to rationalize discrimination.

In order to answer this question, I conduct a series of online experiments to analyze the information acquisition and processing behavior of individuals in potentially discriminating decisions. Similar to Chen and Heese (2021), I find that individuals selectively acquire and interpret information in line with their motives if they have the necessary ‘wiggle room’ to do so. They systematically search for information signals, update their beliefs in direction of their motives and ultimately discriminate based on these beliefs. I also show to what extent limiting this ‘wiggle room’ can be an effective measure to fight this form of discrimination.

2 Experiment Design

The data collection involves a series of pre-registered online experiments and one preliminary online survey, all conducted in the US and implemented on Prolific¹. In the preliminary survey, I collect answers from participants on a logic quiz, a dictator game and a real effort task. Based on their answers in all three tasks, I calculate a score that defines their ‘productivity’ in the subsequent experiments. Additionally, I ask for demographics and past school performance information. The participants are then used as available ‘experimental workers’ for hire in the main experiments. The final pool of workers consists of 58 individuals.

All subsequent experiments involve the same two treatment groups, that differ in the way the experimental workers are labelled. In group ‘race’, ‘experimental employers’ make hiring

¹The experiments are pre-registered under AEARCTR-0008609.

decisions between two experimental workers from different races. In group ‘neutral’ the race labels are replaced by neutral shape labels so that employers make decisions between two workers from different neutral groups, e.g. group ‘circle’, or group ‘square’. This treatment variation is meant to induce a motive for productivity beliefs for experimental employers in group ‘race’, but not for those of group ‘neutral’. Hence, results from employers in group ‘neutral’ serve as baseline levels in the subsequent analysis.

Previous to the hiring decisions, all experimental employers are asked for their subjective expectations about the group productivities of workers. These beliefs indicate the direction of their motives. Subsequently, they are given the true group statistics (in form of a distribution) and are again asked for their beliefs. This is intended to update beliefs about groups of workers and align them between the two treatment groups. In the hiring phase, I ask the experimental employers to make incentivized hiring decisions. Each experimental employer is repeatedly presented with a pair of experimental workers and asked to hire the one with the higher productivity. The measurement of productivity was previously explained to the employers. After the experiment, one hiring decision is randomly chosen and the experimental employer gets a bonus payment of \$2 if they hired the better worker in this decision.

The experiments differ in the way individual-level information is provided for each hiring decision and thus in the extent of wiggle room the employers have to form their beliefs about individual workers.

- In the ‘no info’ experiment, experimental employers have the most wiggle room as they do not receive any additional information.
- In the ‘ambiguous info’ experiment, experimental employers’ wiggle room is equally large in theory as they receive practically uninformative signals. Specifically, they receive the name of the better of the two workers for each hiring decision, but this information comes from either a True News or a Fake News source. The True News source always tells the truth, whereas the Fake News source never tells the truth². For each decisions, employers receive one initial information signal per default and can request up to nine additional signals, each from either the True News or the Fake News source. Employers do not know from which source a message is delivered.
- The ‘uncertain info’ is similar to the ‘ambiguous info’ experiment, except that now the

²This design is inspired by the experimental tool to identify motivated reasoning in Thaler, 2020.

employers have previously been 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. This reduces their wiggle room compared to the ‘ambiguous info’ and ‘no info’ experiments.

- In the ‘tangible info’ experiment, instead of signals directly displaying the (supposedly) better of the two workers, employers are given individual-level information about past performances of the two workers, such as e.g. their college GPA, their final high school math grade or a self-report of their level of resilience. Again, for each decision, employers receive one initial piece of information and can request up to nine additional pieces. This environment still leaves wiggle room for participants, but reduces it further as the signals are always true, but not necessarily point towards the better of the two workers.

3 Beliefs about Group Productivities

In order to identify motives of the experimental employers, I first look at initial beliefs about group productivity levels. I find that initially Asians were considered to be the superior workers compared to Hispanics³. Next, I check whether the group-level information intervention was successful in aligning beliefs. Indeed, I find that the difference in beliefs about group-level productivity of Asians and Hispanics disappeared after the information update. In fact, experimental employers understand that average and variance of the productivity distributions are equal between the two groups. This means when entering the hiring phase, average group level beliefs about the productivity of groups of workers were homogeneous across treatments. This results holds for all four experiments and, following Bohren et al. (2019), rules out discrimination based on inaccurate beliefs as the cause for any subsequent discrimination.

³I initially asked for beliefs about the productivity of Whites, Blacks, Asians, and Hispanics, respectively. Since there was no difference between Whites, Blacks, and Hispanics, I only focus on the decisions between an Asian and an Hispanic worker in the hiring stage. This implies that I consider discrimination between two minority groups, and thereby decrease the potential impact of social desirability concerns e.g. due to recent Black Lives Matter movements.

4 Predictions and Hypotheses

This section describes the predictions and hypotheses in line with discrimination based on motivated reasoning.

4.1 Baseline Discrimination

When motivated belief based discriminators are not given any or only highly ambiguous individual-level information, their wiggle room in binary hiring decisions is large. I therefore expect them to discriminate in the direction of their motives. This would also be consistent with taste-based discrimination.

Hypothesis 1 *There is significant discrimination against Hispanics in the ‘race’ group, but not in the ‘neutral’ group, among experimental employers in experiments ‘no info’ and ‘ambiguous info’.*

$$\text{HiredHispanic}_d = \alpha + \beta_1 \text{Race}_d + \epsilon_d$$

$$\beta_1 < 0$$

4.2 Baseline Information Acquisition and Processing

When motivated belief based discriminators are given ambiguous individual-level information, they have the necessary wiggle room to interpret information according to their motives and therefore they continue to search for more information if initial information contradicts their motives (and vice versa). This would be inconsistent with taste-based discrimination.

Hypothesis 2 *In the ‘ambiguous info’ experiment, employers in the ‘race’ group are more likely to request a 2nd signal and request more signals when the initial signal suggests to hire the Hispanic worker. These differences is larger in the ‘race’ group than in the ‘neutral’ group.*

$$\text{Get2ndSignal}_d = \alpha + \beta_1 \text{Race}_d * 1stHispanic_d + \beta_2 \text{Race}_d + \beta_3 1stHispanic_d + \epsilon_d$$

$$\beta_1 > 0$$

$$\#Signals_d = \alpha + \beta_1 \text{Race}_d * 1stHispanic_d + \beta_2 \text{Race}_d + \beta_3 1stHispanic_d + \epsilon_d$$

$$\beta_1 > 0$$

When motivated belief based discriminators are given ambiguous individual level information, they overweigh information signals that are in line with their motives (and vice versa).

Hypothesis 3 *In the ‘ambiguous info’ experiment, employers in the ‘race’ group are less likely to follow a signal when the overall direction of all considered signals suggests to hire the Hispanic worker. This difference is larger in the ‘race’ group than in the ‘neutral’ group.*

$$FollowSignals_d = \alpha + \beta_1 Race_d * AllHispanics_d + \beta_2 Race_d + \beta_3 AllHispanics_d + \epsilon_d$$

$$\beta_1 < 0$$

4.3 Debiasing: What happens under less wiggle room?

When motivated belief based discriminators are given less ambiguous individual level information they have less wiggle room to interpret information according to their motives and therefore reduce (i) discrimination, (ii) the systematic search for more information if the initial information contradicts their motives (and vice versa), and (iii) the systematic overweighing of information signals that are in line with their motives (and vice versa). This would be inconsistent with taste-based discrimination.

5 Results

In order to analyze the role of discrimination based on motivated reasoning I will first consider to what extent experimental employers discriminate in the experiments that leave the most wiggle room. The analysis of the information acquisition and processing behavior in the ‘ambiguous info’ allows first insights about the extent of potential motivated reasoning.

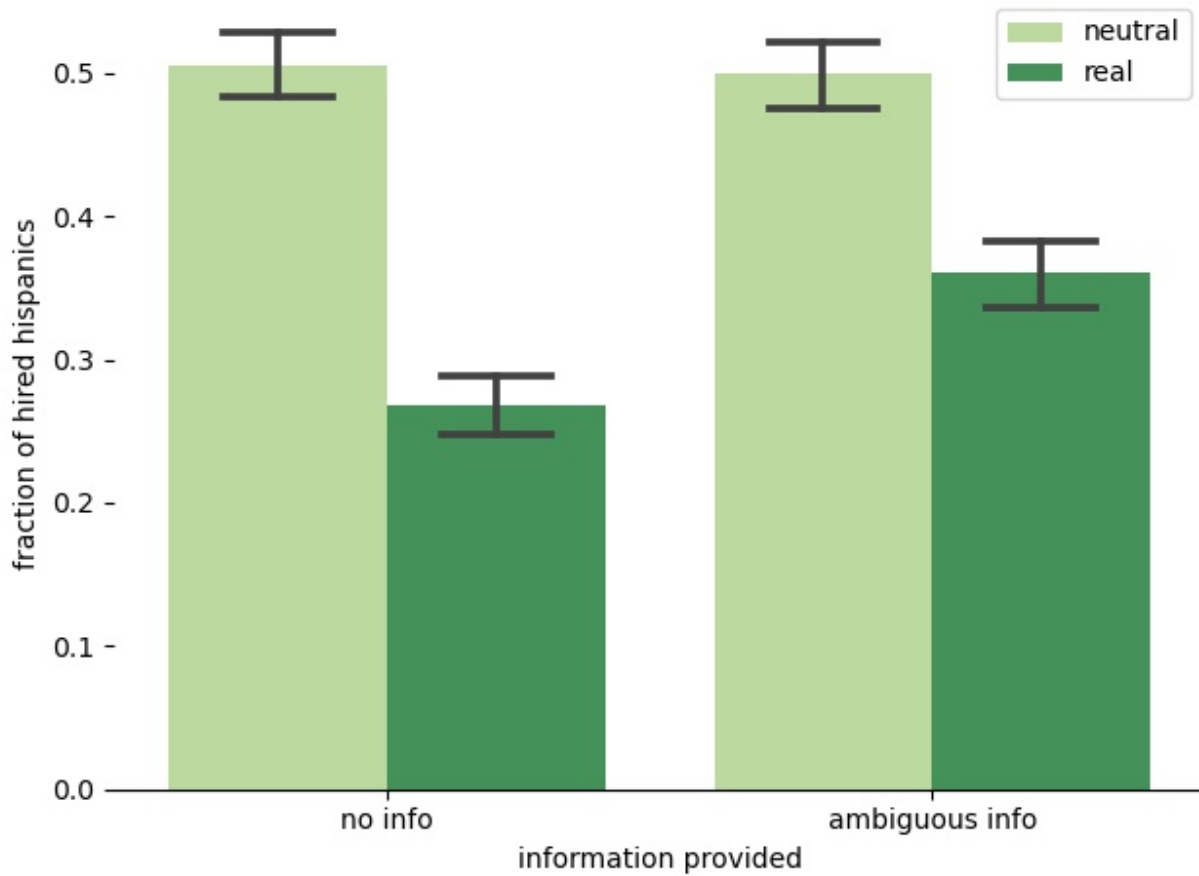
To further deepen our understanding of discrimination based on motivated reasoning I will subsequently consider experiments ‘uncertain info’ and ‘tangible info’. These experiments reduce the wiggle room for the experimental employers as the individual-level information is meaningful and suggestive about who the better worker is. Reduced information acquisition and processing behavior in these experiments compared to the first two experiments would further corroborate the existence of discrimination based on motivated reasoning and provide a first indication for a potential pathway to alleviate this form of discrimination.

5.1 Providing Wiggle Room - No Info & Ambiguous Info

In line with hypothesis 1, figure 1 shows that discrimination against Hispanics is substantial in the two experiments that leave the most wiggle room for employers. In the ‘no info’ experiment, Hispanics are hired in about 50.57% of all decisions whenever employers were not provided with the race of the respective employees. Providing employers with the race labels of workers significantly decreases the hiring rate of Hispanics by 23.71pp. In the ‘ambiguous info’ experiment, there is also 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 against Hispanics, even though employers are aware that group-level scores between the two groups are equal. Usually, this form of discrimination would therefore be described as taste-based discrimination. In order to further investigate the potential cause of this discrimination, I now turn to the information acquisition and processing behavior of the experimental employers in the ‘ambiguous info’ experiment.

Table 1a 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 the ‘race’ group than in the ‘neutral’ group (coeff=0.1046, p-value=0.004). In line with this, columns 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 the ‘race’ group than in the ‘neutral’ group (coeff=0.5161, p-value=0.028). This is consistent with hypothesis 2 and provides first evidence that employers were ‘fishing for good news’, with ‘good’ meaning that an information signal confirms their motive to hire the Asian worker instead of the Hispanic worker. Finally, column 3 indicates that an initial signal that suggests to hire the Hispanic worker has a significantly larger negative effect on the likelihood to follow the signal in the ‘race’ group than in the ‘neutral’ group (coeff=-0.1881, p-value< 0.001). This lends support for hypotheses 3. Taken together, this implies that the experimental employers systematically acquire and process the ambiguous information – a practice that is inconsistent with taste-based discrimination but consistent with the formation of motivated beliefs to rationalize discrimination.

Figure 1: Hiring rates of Hispanic workers under wiggle room



Notes: The vertical axis displays the fraction of hired hispanics. The horizontal axis groups the decisions by experiments, either the 'no information' experiment or the 'ambiguous information' experiment. Yellow bars show decisions of the 'neutral' group, blue bars show decisions of the 'race' group. Error bars indicate 95% confidence intervals.

Table 1: Information acquisition and processing behavior across experiments

(a) ambiguous information			
Dependent variable	2nd signal	number of signals	follow signals
real * hispanic	0.1046 (0.004)	0.5161 (0.028)	-0.1881 (< 0.001)
real	-0.0667 (0.118)	-0.3422 (0.222)	0.0596 (0.003)
hispanic	-0.0390 (0.126)	-0.2801 (0.100)	-0.0695 (0.009)
Observations	3290	3290	3290
Baseline mean dep. var.	0.5483	3.6246	0.8676
(b) uncertain information			
Dependent variable	2nd signal	number of signals	follow signals
race * hispanic	0.0898 (0.143)	0.6035 (0.2652)	-0.0608 (0.219)
race	0.0083 (0.917)	0.2919 (0.6971)	0.0456 (0.080)
hispanic	-0.0141 (0.761)	0.1543 (0.6709)	-0.0292 (0.361)
Observations	756	756	756
Baseline mean dep. var.	0.7313	5.4478	0.9292
(c) tangible information			
Dependent variable	2nd signal	number of signals	follow signals
race * hispanic	0.0801 (0.094)	0.5858 (0.0547)	-0.0836 (0.008)
race	0.0150 (0.704)	0.4316 (0.1135)	-0.0012 (0.932)
hispanic	-0.0058 (0.870)	0.5380 (0.0123)	-0.0600 (0.003)
Observations	2246	2246	2246
Baseline mean dep. var.	0.6027	2.9432	0.9181

Notes: This table shows results of (a) the ‘ambiguous information’ experiment, (b) the ‘uncertain information’ experiment, and (c) 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. *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. P-values are in parentheses, standard errors are clustered at the individual-level.

5.2 Reducing Wiggle Room - Uncertain Info & Tangible Info

Finally, I consider the results from the two ‘debiasing’ attempts. In experiments ‘uncertain info’ and ‘tangible info’, I limit employers’ wiggle room to systematically treat the information in order to form motivated beliefs by providing more meaningful individual-level information.

First, I consider the case where employers were given similar information as in the ‘ambiguous information’ experiment, except that now employers have previously been 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 Fake experiment is shown in table 1b. Column 1 in table 1b indicates that employers still systematically acquire a second information signal depending on the initial information signal, but compared to the ambiguous information setting, the effect size decreased by 2pp and is now insignificant. However, the insignificance might also be an artefact of the smaller sample size in this model. In line with this, column 2 shows that employers acquire more information signals in the ‘race’ group than in the ‘neutral’ group. Looking at whether or not employers also systematically follow the overall direction of signals, column 3 of table 1b, illustrates that reducing ambiguity and hence the potential wiggle room to interpret information substantially decreased the extent to which employers systematically follow the direction of the signals. If the majority of signals considered in a decision suggests to hire the Hispanic worker, has a significantly larger negative effect (-6.08pp) to the follow this suggestion in the ‘race’ group than in the ‘neutral’ group. This is a 12.73pp . reduction in the absolute effect size compared to the ‘ambiguous information’ experiment.⁴

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 1c displays similar results as in the ‘uncertain information’ experiment. Again, an initial signal that suggests to hire the Hispanic worker has a significantly larger positive effect on the likelihood to acquire a second signal and on the number of signals requested in the ‘race’ group than in the ‘neutral’ group. Moreover, the effect on the likelihood to follow the signals of (-8.36pp) is similar to the effect in the ‘uncertain information’ experiment .

⁴A preliminary analysis (not yet included in this version) including a triple interaction term shows that this decrease is statistically significant at the 5% level.

Summarizing these findings, we can see 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 ‘fished for good news’ and even refused to follow the signals if this search was unsuccessful, a reduction of the wiggle room to interpret the information reduced the reluctance to follow unwanted signals. Employers still fish for news that confirm 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, figure 2, shows how the adapted information processing behavior translates into less discrimination. Figures 2a and 2b show that discrimination in the experiments with less wiggle room for the employers is insignificant and between 6.14pp in the uncertain info experiment (figure 2c) and 1.98pp in the tangible information experiment (figure 2d). Taken together, these results illustrate the discrimination alleviating effect of decreased wiggle room to interpret individual-level information and thereby corroborate the existence of discrimination based on motivated reasoning.

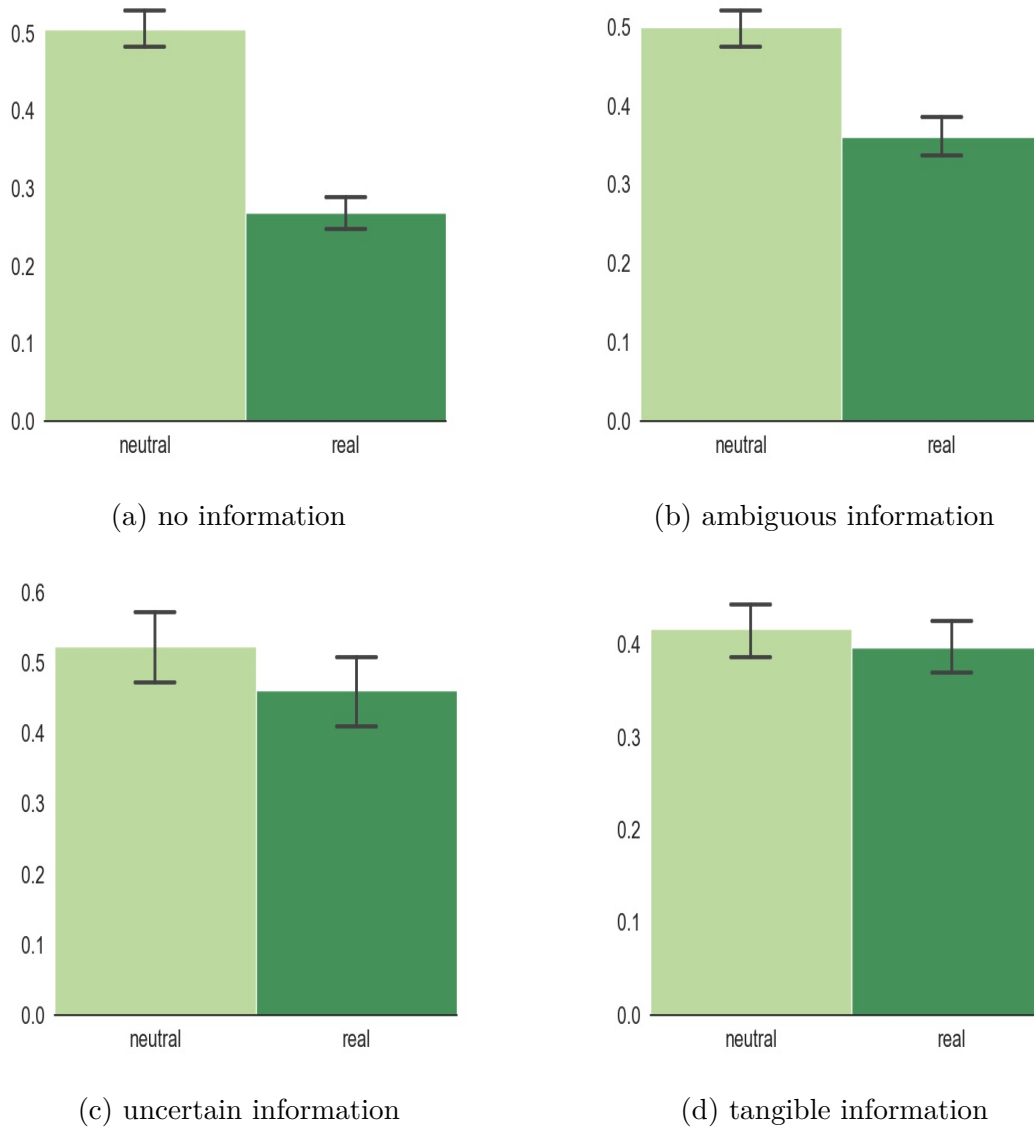
6 Conclusion

This study contributes to the literature on the causes of discrimination by studying a link between taste-based and belief-based discrimination. By taking a closer look at the way in which 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 purely belief-based discrimination. Instead, individuals use wiggle room provided by the inconclusiveness of information in order to discriminate in line with their motives. This form of motivated discrimination requires that policy interventions take wiggle room of information into account in order to effectively fight discrimination.

Acknowledgements

I gratefully acknowledge the financial support from the SAFE Experiment Center Frankfurt and the Interdisciplinary Public Policy Mainz. I also wish to express my sincere gratitude to Florian Hett, Michael Kosfeld, Alex Imas, Simon Hess, Johannes Kasinger, and participants of the Grueneburgseminar Frankfurt for many valuable comments and discussions.

Figure 2: Hiring rates of Hispanic workers across experiments



Notes: This figure shows the hiring rates of Hispanic workers for (a) the 'no information' experiment, (b) the 'ambiguous information' experiment, (c) the 'uncertain information' experiment, and (d) the 'tangible information' experiment. The vertical axes display the fraction of hired Hispanics. The horizontal axis groups the decisions by treatment groups. Yellow bars show decisions of the 'neutral' group, blue bars show decisions of the 'race' group. Error bars indicate 95% confidence intervals. (a) and (b) repeat results of figure 1 and are shown here again for comparisons.

References

- 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.
- Becker, G. S. (1957). 1971:” *The economics of discrimination*”, (1971).
- Bohren, J. A., Haggag, K., Imas, A., and Pope, D. G. (2019). *Inaccurate statistical discrimination*. Tech. rep. National Bureau of Economic Research.
- Chen, S. and Heese, C. (2021). *Fishing for good news: Motivated information acquisition*. Tech. rep. University of Bonn and University of Mannheim, Germany.
- Phelps, E. S. (1972). “The statistical theory of racism and sexism”. *American economic review* 62 (4), pp. 659–661.
- Thaler, M. (2020). “The Fake News Effect: Experimentally Identifying Motivated Reasoning Using Trust in News”. *Available at SSRN 3717381*.