Proceedings of the Ninth AAAI Conference on Human Computation and Crowdsourcing (HCOMP 2021)

Modeling Simultaneous Preferences for Age, Gender, Race, and Professional Profiles in Government-Expense Spending: a Conjoint Analysis

Lujain Ibrahim,[1] Mohammad M. Ghassemi,[2] Tuka Alhanai[1]
[1]New York University Abu Dhabi
[2]Michigan State University
li431@nyu.edu, ghassem3@msu.edu, twa2013@nyu.edu

Abstract

Bias can have devastating outcomes on everyday life, and may manifest in subtle preferences for particular attributes (age, gender, race, profession). Understanding bias is complex, but first requires identifying the variety and interplay of individual preferences. In this study, we deployed a sociotechnical web-based human-subject experiment to quantify individual preferences in the context of selecting an advisor to successfully pitch a government-expense. We utilized conjoint analysis to rank the preferences of 722 U.S. based subjects, and observed that their ideal advisor was White, middle-aged, and of either a government or STEM-related profession (0.68 AUROC, $p < 0.05$). The results motivate the simultaneous measurement of preferences as a strategy to offset preferences that may yield negative consequences (e.g. prejudice, disenfranchisement) in contexts where social interests are being represented.

Introduction

Governments, foundations, and private equity firms alike provide funds to advance innovative ideas and causes. Funding decisions are impacted by the characteristics of the proposed cause as well as the human *promoters* of the cause (De Martino et al. 2006). The human factors that may play a role in a funding decision include the demographic background (age, gender, and race) and personal experiences (educational history, profession) of the promoter (Pedersen and Nielsen 2020). Given that funding decisions can have long lasting impacts on communities, quantifying the influence of human factors on funding decision-making patterns may provide insight into the degree and polarity of biases in these decisions (Mavrodiev, Tessone, and Schweitzer 2013; Schöbel, Rieskamp, and Huber 2016; Kinsey et al. 2019).

Several recent studies provide evidence that implicit bias (ethnic, gender, and political among others) impacts both human and machine decision-making in many settings (finance, investment, hiring) (Larson et al. 2016; Lindner, Graser, and Nosek 2014; Parikh, Teeple, and Navathe 2019). Despite the accumulation of scientific evidence, members of the general population have divergent opinions on the impact of implicit bias. Furthermore, surveys of U.S. American adults indicate that these divergent opinions are associated with demographic factors (generation, race, and political) (Doherty 2020). Given the increasing prevalence of sociotechnical decision-making systems, studying biases and preferences using a sociotechnical setup, where humans process information and make decisions via human-computer interfaces, has increased in importance.

Objective & Research Hypotheses

Objective In this paper, we developed a sociotechnical system through which we studied the *relative* influence of and preference for demographic factors (age, gender, race, and profession), on decision-making of U.S. based crowdworkers. More specifically, we collected data from 722 subjects who participated in a web-based hiring task. The task required subjects to choose the optimal demographic properties of an advisor to secure competitive government funding for projects in seven categories (e.g. health services, defense). While giving subjects the ability to create an advisor profile with any combination of age, gender, race, and profession of their liking, we also presented subjects with a suggested advisor profile. This study expands on previous national polls, by outlets like Gallup & Pew Research Center, that gauge the public's preferences for these attributes, by utilizing more recent developments in survey and analysis techniques to investigate the *relative* and *simultaneous* influence of preferences (Reinhart 2020; Doherty 2020; Shafranek 2019; Carey et al. 2020). We analyze preferences at the national level due to easier access to a larger number of people at that level, but this interface and statistical approach are primarily of great utility at smaller organizational levels, from government to companies, where opportunities for targeted interventions are possible. We present a visual and conjoint analysis of the collected data herein, and make the codebase publicly available to facilitate extensibilty[1].

Terminology Throughout this paper, we refrain from labelling what we observed as *biases* because we can not know for certain what drove subjects to make the decisions they made. Instead, we will refer to subjects' decisions as *stated preferences* where we define a *preference* as a greater liking for one option over another (or others) resulting from a combination of implicit bias, strategic considerations, and

[1]https://github.com/x-labs-xyz/aaai-hcomp21-preferences

other behavioral factors (Elsesser and Lever 2011; Dictionary n.d.; Hainmueller, Hopkins, and Yamamoto 2014). In the context of our experiment, strategic considerations include the selection of an advisor's profession based on relevant authority that profession brings to the funding request (e.g. medical doctor requesting funds for medical research), and the selection of advisor demographic attributes that improve the odds of success because of discrimination and/or systemic bias that are not held by the subject, but are nevertheless perceived to be relevant for success. For example, a subject may not personally feel that men are more qualified advisors, but may select a male nonetheless because they perceive that society has a positive bias towards men, and thus the selection of a male advisor will provide a strategic advantage in the acquisition of funds (Gong, Xu, and Takeuchi 2017; Santana 2018).

Hypotheses Based on previous U.S. based surveys of individual preferences for a set of age, gender, race, and profession options, and taking into consideration current debates of cultural and political significance, we investigated whether our experiment supports the following research hypotheses:

- **H1**: There is a preference for younger advisors over older advisors.

Over 40 years ago, the U.S. congress passed the Age Discrimination in Employment Act, which forbade employment discrimination against workers over the age of 40 (of Labor n.d.). However, in an increasingly aging U.S. population, there remains to be subtle and explicit manifestations of discrimination against older workers, including a few relatively recent and high profile age discrimination lawsuits. Studies on age discrimination include *Lössbroek et al*'s analyses which showed that the valuation of applicants' skills and age are largely independent (Lössbroek et al. 2020), and *Carlsson and Eriksson*'s study which showed that women experience more age discrimination than men, highlighting the need for an intersectional examination of age discrimination. (Carlsson and Eriksson 2019; Neumark, Burn, and Button 2017).

- **H2**: There is a preference for STEM-related professions over non-STEM-related professions.

A recent 2019 Gallup poll on the trustworthiness of professions showed that nurses, engineers, doctors, and pharmacists are viewed as the most trustworthy professions, despite debates over the public's increasing distrust in science and scientists (Fink 2019; Krause et al. 2019; Reinhart 2020). Studies on trust in professions have largely focused on independent examinations of select professions/industries (e.g. journalism), and less on relative perceptions (Lewis 2020; Willnat, Weaver, and Wilhoit 2019; Webster 2018). Our study attempts to paint a picture of *relative* preferences to better reflect the reality of forming opinions and making choices in the real world.

- **H3**: There is variation in gender and racial preferences along political party lines.

A recent report by the Pew Research Center showed that U.S. voters' attitudes towards gender and race are even more divided along party lines in 2020 than they were in 2016, indicating an increasingly critical relationship between political party self-identification and preferences (Doherty 2020; Enns 2018). Studies also point to the influence of political party self-identification on nonpolitical decisions, making it an important factor to examine. For instance, *Gift and Gift* show how political signals in resumes affect hiring, and *Shafranek* shows, through conjoint analysis, how political considerations factor into roommate choice (Gift and Gift 2014; Shafranek 2019).

Related Work

Implicit Bias There have been numerous studies that investigate implicit biases and preferences, particularly through web-based crowdsourced experiments. These studies have focused on specific human attributes such as gender, race, and age. Studies have focused on identifying gender biases in language; *Tang et al.* studied the effects of gender-biased terminology in job listings and its impact on who applies to what jobs (Tang et al. 2017). Other research expanded from studying gender biases to studying both gender and racial biases; *Thebault-Spieker et al.* used an Amazon Mechanical Turk (AMT) based study and a Bayesian approach to investigate racial, gender, and reputation bias in the evaluation of gig workers (Thebault-Spieker et al. 2017). Further, age biases have also been investigated; *Kaufmann et al.* and *Lindner et al.* focused their research on age-based biases in hiring decisions, where *Kaufmann et al.* used an AMT experiment to study hiring decisions using facial age appearance and *Lindner et al.* surveyed more than 1,000 U.S. based adults to investigate the relationship between age-based hiring discrimination and self-perceived objectivity (Lindner, Graser, and Nosek 2014; Kaufmann et al. 2017). In our own work, we are interested in utilizing a similar web-based, crowdsourced experimental framework applied to study individual preferences across multiple attributes.

Conjoint Analysis for Choice-based Data Studies that examined multidimensional preferences in choice selections, ranging from roommate choices to immigrant choices, typically utilized sociotechnical setups and conjoint analysis to generate estimations of the causal effects of multiple treatment components simultaneously (Meyerding and Merz 2018; Verma and Chandra 2018; Anand, Bansal, and Aggrawal 2018; Hainmueller, Hopkins, and Yamamoto 2014). In *Awad et al.*'s The Moral Machine experiment, which was designed to explore people's decisions in the moral dilemmas of autonomous vehicles (i.e. trolley problem), conjoint analysis was used to compute the average marginal component effect of nine preferences simultaneously using a sample size of over 30 million (Awad et al. 2018). Also, in the field of electoral politics, *Gutiérrez-Romero and LeBas* utilized conjoint analysis to inspect the role of electoral violence with respect to a candidate's profile and performance, on vote choice and willingness to vote in Kenya (Gutiérrez-Romero and LeBas 2020).

While the scope and scale of the analysis of previous work have had a strong impact in the field, less work has focused on studying the multidimensional nature of biases and preferences in a context where a decision made by one individual would have a more direct, tangible, and medium-to-long-term social impact (as is the case of allocating funds for government-expenses which is the focus of this paper); a presidential candidate may be thought to manage a system indirectly (Meier and O'Toole Jr 2006), immigrants may seem like a distant cause (Markowitz and Slovic 2020), while the trolley problem results in an immediate life-or-death situation (Edmonds 2013).

Experiment Setup & Data Collection

This study was approved by the Human Subjects Research Program Institute Review Board (IRB) at New York University Abu Dhabi.

Dataset Creation

Real-world Government Expenses A total of 49 expenses were curated for this experiment in the following seven categories: (1) health services, (2) defense, (3) public assistance, (4) environmental protections, (5) education, (6) scientific and medical research, and (7) debt. The categories, as well as the specific expenses that fall under them, were derived from available online data on federal income tax expenditures (Center 2020). Individual experimental trials included a specific expense drawn from a given expense category.

Profiles from Census Data and Research Surveys Four advisor attributes were selected to study subject preferences: (1) age, (2) gender, (3) race, and (4) profession. In this paper, we refer to a specific value within an attribute as an attribute-level (e.g. *male* is an attribute-level of the gender attribute). For the age attribute, there were three attribute-levels: *25*, *45*, and *75* years of age; these attribute-levels were meant to represent *young*, *middle-aged*, and *senior* individuals respectively (Richardson et al. 2013; Derous and Decoster 2017). For the gender attribute, there were two attribute-levels: *male* and *female*; these attribute levels were selected to capture stereotypical gender-related responses and preferences (Wijenayake et al. 2019). For the race/ethnicity attribute, there were four attribute-levels: *White*, *Black*, *Latino*, and *Asian*; these attribute-levels constitute the four largest racial/ethnic groups in the U.S. (QuickFacts 2019). For the profession attribute, 16 professions were derived from a recent Gallup poll that explored the perceptions of individuals towards professions (Reinhart 2020). For the purposes of our analysis, these professions were represented as four attribute-levels that reflect professional category: (1) *governance-related professions* (state governor, congress member, police officer, union leader), (2) *science, technology, engineering, and math (STEM) - related professions* (engineer, medical doctor, nurse, pharmacist), (3) *business-related professions* (insurance salesperson, business executive, stockbroker, advertising practitioner), and (4) *other professions* (lawyer, university professor, journalist, clergy). Along with these four attributes, advisors were given stereotypical racial and gendered representations with names as well as avatars designed using Apple's Memoji application (Support 2020; Herring et al. 2020; Stark 2018). A dataset consisting of all the possible combinations of expenses, ages, genders, races, and professions was then synthesized and randomly sampled during the experiment.

Experiment Task

Task Web Layout Our web-based experiment presented subjects with an expense and instructed them to create an advisor that they believe would successfully pitch that expense for government funding. To this end, subjects could customize the advisor's attributes (age, gender, race, and profession) from a pre-specified set (e.g. *25*, *45*, or *75* years old for age). The specific expense fell under one of the seven categories outlined in the previous section, where both the specific expense and the expense category were displayed on the screen in each task. Additionally, subjects were informed that an advisor had been suggested to them in each task and were rendered with that advisor's specific attributes (e.g. *Age*: 25 [years old], *Gender*: female, *Race*: White, and *Profession*: journalist). The attributes of the suggested advisor were selected at random, remained fixed, and served to establish a reference point from which the subject may specify their custom advisor (Boncinelli et al. 2020). Subjects were able to either re-select the same attributes as that displayed for the suggested advisor, or select alternative attributes. This approach was adopted to avoid confounding factors that could arise from an incongruent user interface design (i.e. one-click to select the suggested advisor versus eight clicks to customize an advisor profile). As subjects customized their advisor, they were able to either select the same attribute-levels as that of the suggested advisor, or choose entirely new attribute-levels. As attributes were specified, the custom advisor was rendered with an avatar and name conditioned on the selected age, gender, and race. Once the subject completed their selection of all four attributes, a submit button was enabled, allowing subjects to submit their selections and move on to the next task. The layout of the task web page can be seen in Appendix Figure 6.

Task Information Order & Presentation Each experimental session consisted of 15 tasks and a survey on subjects' demographic backgrounds, socioeconomic conditions, and political outlooks. Within each of these 15-task sessions, one task consisted of an expense from a randomly selected expense category, while the other 14 tasks consisted of two tasks under each of the seven expense categories. As subjects customized their advisor, they were presented with all the age attribute-levels (three options), all the gender attribute-levels (two options), all the race attribute-levels (four options), but only four options from the 16 profession attribute-levels. The number of options presented under professions was reduced from the full number to avoid overwhelming subjects and to ensure they consider each option as they choose their preferred profession. Out of these four profession options, one of them was always the suggested advisor's profession to ensure that option was always available for re-selection. As for the other three options, one profession was randomly selected from each of the other three unrepresented categories of professions. The order the options were presented in the drop-down menus was random and changed in each task in a session to reduce the effect of order on results (Arechar, Kraft-Todd, and Rand 2017).

End-of-Experiment Demographic Survey The study concluded with a survey collecting information on subjects' demographic backgrounds, socioeconomic conditions, and political outlooks. The survey consisted of eight multiple choice questions on subjects' gender, age, educational degree level, ethnicity, income, political views, voting behavior in the 2020 U.S. national elections, and religion.

Data Collection, Crowd-sourcing, and Processing

The study was conducted using AMT human intelligence tasks (HITs). This was motivated by AMT's quick and easy access to a large number of people located in the U.S., our target population (Buhrmester, Talaifar, and Gosling 2018; Stewart et al. 2015), and by the success of previous studies on age, gender, and race biases using AMT (Berinsky, Huber, and Lenz 2012; Gardner, Brown, and Boice 2012; Nadler and Kufahl 2014).

AMT HIT Parameters To ensure the quality and relevance of the collected data, a number of qualifications were placed which AMT workers had to meet to complete our HITs: (1) workers completing our HITs should have previously completed at least 1000 HITs, (2) workers completing our HITs should have an approval rating of at least 95%, and (3) workers should be located in the U.S. (Hunt and Scheetz 2019). Our study also included two attention checks to flag subjects who may have not been paying attention to the tasks (Aruguete et al. 2019). These two checks involved replacing the field that usually includes information about the expense with a direct instruction to choose the same attribute-levels as the suggested advisor (in the first check) and entirely different attribute-levels than the suggested advisor (in the second check) (Abbey and Meloy 2017). A single HIT consisted of a single complete study in addition to the attention checks (Mellis and Bickel 2020).

Data Collection Whenever a task in a HIT was submitted, the task identification number and the customized advisor attributes were collected and sent to a remote server and stored in a secure PostgreSQL (PSQL) database. Each subject was then redirected to the next task until they reached the first attention check, second attention check, and finally the survey. The subject's survey answers were collected and stored in the database along with the HIT ID, assignment ID, anonymized subject ID, IP address, submission timestamp, and the two attention check binary variables (whether or not they passed the attention check).

Following collection, the data were then processed to exclude submissions from subjects who: (1) failed either one of the attention checks, (2) were located outside the U.S. (filtered by the collected IP addresses), and/or (3) had incomplete results. This processing excluded data from 698 subjects, reducing the number of subjects from 1420 to 722 subjects. In total, this resulted in 10,830 examples available for use in the analyses (722 subjects $*$ 15 tasks-per-subject).

Binary Encoding of Changes in Choice (4 features) We captured in four binary features ({*age*, *gender*, *race*, *profession*} *changed*) whether subjects selected attributes for their customized advisor that were the *same* as the suggested advisor, or *different* from the suggested advisor. A difference in value was indicated by a '1' and no difference in value was indicated by a '0'. This was used to generate the conjoint analysis dataset.

Methods

Conjoint analysis is a methodological approach used to breakdown factors involved in decision-making through evaluating multiple hypothetical choices of profiles. In discrete choice-based conjoint analysis, subjects are presented with two or more profiles with varying attribute-levels and are asked to choose between these profiles (Green and Srinivasan 1990). Compared to traditional survey techniques, conjoint analysis presents users with multiple attributes simultaneously thus reducing social desirability response bias and allowing us to measure the different trade-offs subjects make (Hainmueller, Hopkins, and Yamamoto 2014). In this experiment, using the customization exercise, we cast our tasks as choice-based conjoint tasks involving choosing between two profiles: (1) the *suggested* advisor profile or (2) the *customized* advisor profile.

Outcome Target: 'selected' Profile The dependant feature for this model was a binary feature that represented whether the advisor profile was or was not selected by a subject; '1' indicated a selected profile, and '0' otherwise.

Advisor Attributes as Input (9 features) The four advisor attribute features (*age*, *gender*, *race*, and *profession category*) were dummy-coded where *25/young* was the reference group for age, *female* was the reference group for gender, *White* was the reference group for race, and *governance-related professions* was the reference group for profession category. This encoding yielded nine features as model inputs.

Modeling Approach There are several approaches to analyzing conjoint tasks from Bayesian hierarchical modelling to ordinary least squares (OLS) regression (Green, Krieger, and Wind 2001). In our analysis, we employed *Hainmuller et al.*'s approach using logistic regression instead of an OLS regression (Hainmueller, Hopkins, and Yamamoto 2014). Logistic regression was performed on our data where the outcome target 'selected' was regressed on the nine dummy input features of the attribute-levels. This was performed in two ways: (1) on the entire dataset for the global analysis, and (2) on the subsets of the dataset with subjects whose self-identified political affiliation was either 'Democrat' or 'Republican'. All the models were trained and tested on both the entire dataset/subsets of the dataset and using leave-one-subject-out (LOSO) cross validation (Koul, Becchio, and Cavallo 2018). The performance metric utilized was the area under the receiver operating characteristic curve (AUROC);

it provides a single value robust to problems where data is not necessarily balanced, and evaluates across various levels of classification thresholds (Hosmer Jr, Lemeshow, and Sturdivant 2013).

Given the advisor attribute-level features X as model inputs, and the model output as probability Pr that an advisor profile outcome target Y would be 'selected' by subjects ($Y = 1$), the logistic regression model is defined as:

$$Pr(Y = 1) = \frac{exp(\beta_0 + \beta_1 X_1 + ... + \beta_9 X_9)}{1 + exp(\beta_0 + \beta_1 X_1 + ... + \beta_9 X_9)} \quad (1)$$

where $X_1, ..., X_9$ are the nine dummy-coded input features of the advisor attribute-levels, $\beta_1, ..., \beta_9$ are the estimated regression coefficients for each of the input features, and β_0 is the regression intercept (Hosmer Jr, Lemeshow, and Sturdivant 2013).

Average Marginal Component Effect The average marginal component effect (AMCE) is the marginal effect of an attribute-level averaged over the joint distribution of the remaining attribute-levels. Since our experiment utilized completely independent randomization, making attributes mutually independent, the estimation of the AMCE and interpretation of the results was significantly simplified. Thus, the estimated AMCE of each attribute-level was interpreted as the average change in the probability that a profile will be selected when it includes that attribute-level instead of the reference attribute-level.

The results of each regression were used to estimate the AMCE π of each attribute-*level* j (e.g. *middle-aged*, *senior*) in a given attribute i (e.g. age). The AMCE was then used to calculate the importance Imp for a given attribute (i.e. Imp_i), which is the difference between the *maximum* and *minimum* AMCE values across all attribute-levels. It was calculated as follows:

$$Imp_i = max(\pi_i^j) - min(\pi_i^k) \quad (2)$$

where k is the k_{th} attribute-level of the i_{th} attribute. Given the value of attribute importance, we then calculated the *relative* importance $Rel - Imp$ of each attribute (i.e. $Rimp_i$), which is the normalized importance. It was calculated as follows:

$$Rimp_i = \frac{Imp_i}{\sum_{i=1}^{m} Imp_i} \quad (3)$$

where m is the total number of attributes.

Results & Discussion

Demographics of AMT Subjects

Each HIT in the experiment was prized at 0.5 USD and took an average of 3 to 5 minutes to complete. The most represented states subjects completed the experiment from were California (71 subjects), Texas (45 subjects), New York (42 subjects), Florida (39 subjects), and New Jersey (35 subjects), which was generally proportional to the population across states in the U.S[2].

[2]https://www.census.gov/

In this study, slightly more women (55%) than men (44%) participated. Subjects had a mean age of 35-44 years with 26% of subjects falling in that age range. More than half of the subjects had reported an education level of a bachelor's degree or higher (60.3%), and a majority self-identified as White (72.3%). Approximately half of the subjects' self-identified political affiliation was as a Democrat (46.8%), about one-quarter self-identified as a Republican (23.9%), and the rest self-identifying as either an Independent or Other. In terms of religiosity, the most prevalent response of subjects was No Religion, Agnostic, or Atheist (44.3%), while the next most prevalent response was Protestant (27.7%).

Subjects' Decisions on Age, Gender, Race, and Profession

Visualizing Choices To understand subjects' selection of an appropriate advisor for the presented expense, we visually represented their decisions. Figures 1, 2, and 3 are choice flow diagrams that show the frequency and direction of the change from different attribute-levels of the suggested advisors to different selected attribute-levels of the customized advisors. The widths of the bands are proportional to the frequency of the choices made in that direction.

Younger Is Not Always Better: Middle-aged Advisors Preferred Over Both Younger & Older Advisors While we hypothesized (**H1**) that subjects would have a stronger preference toward younger advisors (age = 25), it may be inferred from Figure 1 that when subjects chose an age that differed from that of the suggested advisor, subjects' strongest preference was for choosing the middle-aged option (age = 45), both when suggested a younger option (age = 25) and an older option (age = 75), thus nullifying **H1**. The

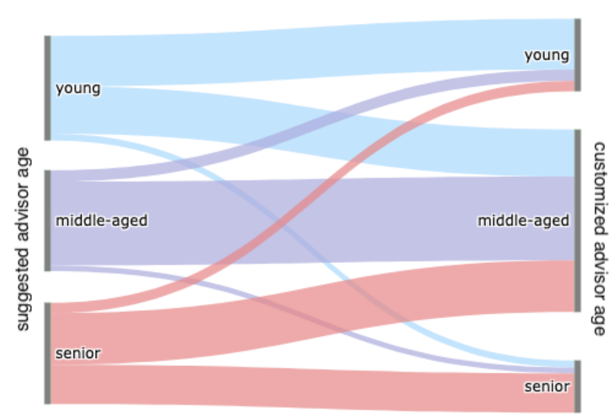


Figure 1: Age choice flow diagram. Diagram visualizes the frequency and direction of transitions between the age attribute-levels of the suggested advisors and the selected age attribute-levels of the customized advisors. The diagram shows that the middle-aged age option, 45, was the most preferred option (*25/young* decreased by 30.8%, *45/middle-aged* increased by 80.5%, and *75/senior* decreased by 48.5%).

number of times *45/middle-aged* was selected as the age attribute of the customized advisors increased by 80.5% relative to the number of times it was suggested as the age attribute of the suggested advisors, whereas the *25/young* option decreased by 30.8% and the *75/senior* option decreased by 48.5%.

These results suggest that subjects viewed age as a proxy for experience and previous success that could predict future success of getting an expense approved for government funding. However, since the oldest option (age = 75) was still the least preferred option, it seems that preferences were not entirely driven by perceptions of work experience and competency, but may have accounted for an advisor holding a more contemporary outlook since the mean age of subjects was closer to the middle-aged group than the other two age groups (Carlsson and Eriksson 2019).

No Clear Preference for an Advisor's Gender For the gender choice, Figure 2 seems to indicate that no clear preference initially appeared in the global analysis. 71% of tasks involved subjects choosing the same gender as the suggested advisor for their customized advisor where 51.7% of these cases involved choosing *female* again and 48.3% involved choosing *male* again. As for the remaining 29% of tasks where the gender other than the suggested advisor gender was selected, 49.3% involved choosing *female* and 50.7% involved choosing *male*. The number of times *male* was selected as the gender attribute of the customized advisors increased by only 0.74% relative to the number of times it was suggested as the gender attribute of the suggested advisors, whereas *female* decreased only by 0.70%. The gender attribute was further explored in the conjoint analysis.

Majority White Subjects Seem to Prefer White As Advisor's Race As for the race choice, Figure 3 indicates the strongest preference that appeared when subjects chose a race for their customized advisor that differed from that of the suggested advisor was choosing *White* when the suggested advisor race was any of the other races. The number of times *White* was selected as the race attribute of the customized advisors increased by 27.2% relative to the number of times it was suggested as the race attribute of the suggested advisors. The *White* attribute-level was the only race attribute-level to experience an increase. The second strongest preference was for choosing *Black*, which experienced a decrease of 1.51%, followed by choosing *Asian* with a decrease of 14.2%, finally followed by choosing *Latino* with a decrease of 19.4%.

The strong preference towards choosing *White* advisors may be explained in a number of ways. Given that 72.3% of our subjects identified as White, this preference may be a result of similarity or affinity bias which has been studied and demonstrated to be influential in decision-making (Roth et al. 2020; Ye et al. 2017; de Kock and Hauptfleisch 2018). Secondly, since subjects were asked to create a profile which they believe will successfully obtain funding from the government, the strong preference towards White advisors can be a result of people's belief that there is existing bias in high-stakes decision-making that favors White individuals and existing discrimination towards people from other races which would make a White advisor more likely to succeed in obtaining funding (Dehon et al. 2017; Capers IV et al. 2017; Einstein and Glick 2017). Finally, this could be attributed to implicit biases favoring White individuals among White participants, as well as participants of all ethnicities. However, our sample size of non-White subjects was small and thus this interpretation may not be mildly supported by the data. Preferences for different ethnicites and genders varied across different political party self-identifications where Republican subjects showed a stronger preference for White and male advisors compared to Democrat subjects.

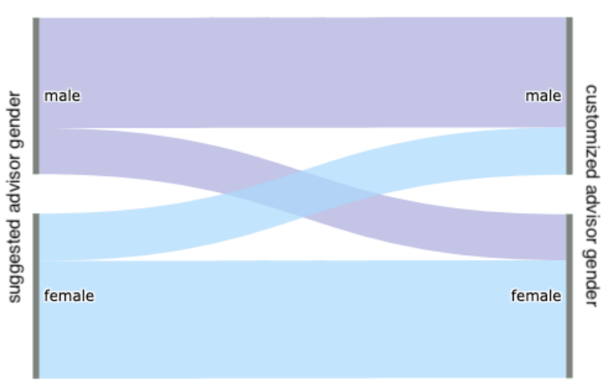


Figure 2: Gender choice flow diagram. Diagram visualizes the frequency and direction of transitions between the gender attribute-levels of the suggested advisors and the selected gender attribute-levels of the customized advisors. The diagram shows that there was no clear preference between male and female at this level of analysis (*male* increased by 0.74% and *female* decreased by 0.70%).

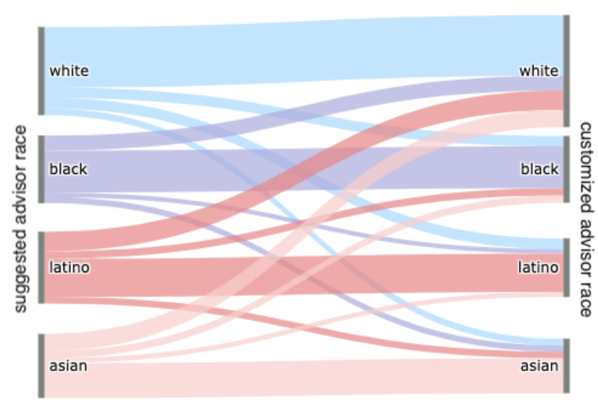


Figure 3: Race choice flow diagram. Diagram visualizes the frequency and direction of transitions between the race attribute-levels of the suggested advisors and the selected race attribute-levels of the customized advisors. The diagram shows that White was the most preferred option (*White* increased by 27.2%, *Black* decreased by 1.51%, *Latino* decreased by 19.4%, and *Asian* decreased by 14.2%).

Academics, Government Officials, Business Executives, & Scientists Are Among Most Preferred Professions In Appendix Figure 7, we plotted a transition matrix which visualizes the frequency and direction of the selection for professions between the suggested advisor attribute to that of the customized advisor. There are a number of key patterns that we can observe from the figure. First, we observe that the most popular selection and relatively consistent change from all of the suggested professions was to the *university professor* profession. This may be because university professors are viewed as both knowledgeable and relatively impartial, and thus being a more appropriate choice for an advisor both overall and relative to other professions.

Next, at the level of profession categories, switching to a *governance-related* or a *STEM-related* profession occurred more frequently than switching out of these professions. The popularity of *governance-related* professions may be attributed to our experiment involving government funds which may have naturally led to subjects selecting *governance-related* professions (who are already in government) as being the most likely to achieve success. The popularity of *STEM-related* professions may be due to a perception of subject expertise that would be appropriate for several healthcare / science / education related expenses.

On the other hand, *business-related* professions showed the opposite pattern with the exception of the *business executive* profession. The *business executive* profession appeared notably more preferable than the other *business-related* professions and even more preferable than some of the professions in other categories (e.g. *police officer*, *journalist*, or *lawyer*). This trend may be due to a perception that executives have leadership and managerial skills and so can be convincing, and thus have a more relevant set of skills than a *stockbroker*, *insurance salesperson* or *advertising practitioner* who may be perceived to have a tangential set of skills.

Finally, the *journalist* profession was one of the least popular professions in the profession categories other than the *business-related professions*. This may be due to a negative perception that journalists are not truly impartial, and reduced public trust in the media as evidenced by extensive on-going discussions around media and trust (Tandoc Jr, Jenkins, and Craft 2019; Willnat, Weaver, and Wilhoit 2019; Lewis 2020; Newman and Fletcher 2017).

Overall, these results only partially support our hypothesis **H2**, as even though *STEM-related* professions were quite preferable, *governance-related* professions and the *business executive* profession were also quite popular, at times as popular as or more popular than some of the *STEM-related* professions.

Global Conjoint Analysis: Confirms Choice-flow and Transition Trends for Advisor Attributes

To examine the relative importance of advisor attributes simultaneously, we computed the AMCE of each attribute-level and the relative/normalized importance of each attribute. The AUROC of the logistic regression model used in this analysis was 0.68 when trained and tested on the entire dataset and also 0.68 when using LOSO cross validation. All attribute-levels were found to be statistically significant ($p < 0.05$) except for *other professions*. Table 1 shows the AMCE of all the possible advisor attribute-levels relative to the unplotted reference levels of each attribute (age: *25/young*, gender: *male*, race: *White*, and profession: *governance-related professions*).

The attribute-level that had the greatest effect not only for the age attribute but for all the attributes was *45/middle-aged*. Additionally, as previously observed, the *White* attribute-level had the highest positive change effect for the race attribute, and the difference between the effect of the *male* and *female* attribute-levels is very small. The results from this analysis also confirmed the trends seen in Appendix Figure 7 at the level of profession categories showing a stronger preference for *STEM*, *governance-related*, and *other professions* (mostly driven by the popularity of the *university professor* profession) compared to *business-related professions*. The greatest variation in preferences within one of age, gender, race, or profession was seen in the age attribute with the difference between the preference for *45/middle-aged* and *75/senior*, followed by the profession attribute with the difference between the preference for *STEM-related professions* and *business-related professions*, followed by the race attribute with the difference between the preference for *White* and *Latino*, and finally followed by the gender attribute with the difference between the preference for *female* and *male*.

The relative/normalized importance was also calculated showing that age was the most important attribute in this decision-making task, followed by profession, race, and finally gender which was 56.4% less important than age.

Local Conjoint Analysis: Political Outlook is Associated with Gender & Race Preferences

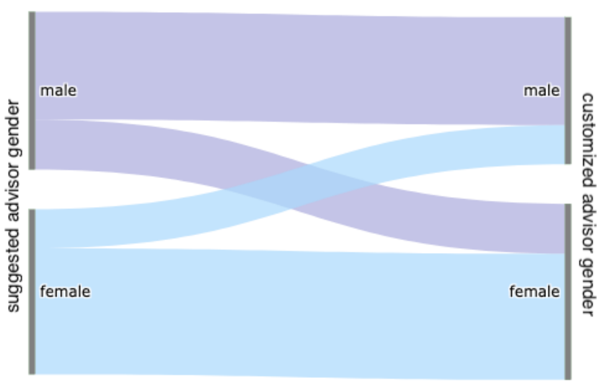
We were interested in understanding if cross-sections of subjects demonstrated differences in preferences from the aggregate set of subjects; we term this as *local* conjoint analysis. For the local conjoint analysis, we were interested in

	dy/dx	std err	p-val
mid-aged	0.251	0.006	< 0.001
senior	-0.052	0.008	< 0.001
female	0.014	0.006	< 0.001
black	-0.032	0.009	< 0.001
latino	-0.088	0.009	< 0.001
asian	-0.067	0.009	< 0.001
stem prof.	0.023	0.008	0.005
business prof.	-0.084	0.009	< 0.001
other prof.	0.004	0.008	0.631

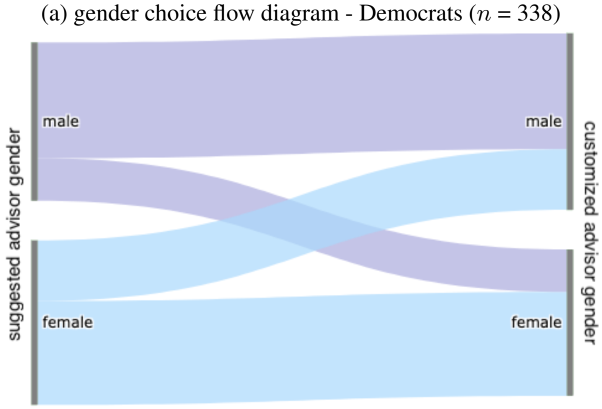
Table 1: AMCE values (dy/dx) of the various attribute-levels in age, gender, race, and profession on subjects' selection of an advisor. All attribute-levels were found to be statistically significant ($p < 0.05$) except for *other professions*. The attribute-level with the highest AMCE was an age of *45/middle-aged*. Within gender, race, and profession, the attribute-levels with the largest AMCE were *female*, *White*, and *STEM-related professions* respectively.

the following three properties: (1) a cross-section of subjects with a self-selected outlook (i.e. not physiological or a function of birth), (2) which in the specific context of government-expense funding would be an identity in which decision-making values are derived from, and (3) contain a sample size greater than 20% of the total number of subjects. Therefore, satisfying these three conditions, a cross-section of subjects' self-identified political outlook was studied.

Two subsets of the dataset were created, one with responses only from subjects who self-identified as Democrat (n = 338) and one with responses only from subjects who self-identified as Republican (n = 173). Similar analyses to the ones previously described were then performed on these two subsets and the results were compared. The AUROC values of the logistic regression models used in this analysis were 0.68 / 0.68 (full data / LOSO cross-validation) and 0.70 / 0.69 (full data / LOSO cross-validation) for the Democrat and Republican subsets respectively. All attribute-levels were found to be statistically significant ($p < 0.05$) except *Black*, *STEM-related professions*, and *other professions* for Democrat subjects, and *business-related professions* and *other professions* for Republican subjects.



(a) gender choice flow diagram - Democrats (n = 338)



(b) gender choice flow diagram - Republicans (n = 173)

Figure 4: Gender choice flow diagrams for Democrat subjects (n = 338) and Republican subjects (n = 173). Democrat subjects made more transitions from *male* to *female* advisors, whereas Republican subjects made more transitions from *female* to *male* advisors.

Gender Observations Along Political Party Lines Figure 4 shows the gender choice flow diagrams for Democrat subjects (Figure 4a) and Republican subjects (Figure 4b). The diagrams show that Democrat subjects made more transitions from *male* to *female* advisors than from *female* to *male* advisors, whereas Republican subjects made more transitions from *female* to *male* advisors than *male* to *female* advisors. For Democrat subjects, the number of times *female* was selected as the gender attribute of the customized advisors increased by 6.68% relative to the number of times it was suggested as the gender attribute of the suggested advisors, whereas *male* decreased by 6.97%. Republican subjects showed the opposite trend; the number of times *female* was selected as the gender attribute of the customized advisors decreased by 11.0% relative to the number of times it was suggested as the gender attribute of the suggested advisors, whereas *male* increased by 11.5%.

The difference in gender preferences along party lines was also seen in the local conjoint analysis AMCE results in Tables 2 and 3. Conjoint analysis was completed on the two subsets of the dataset, one with responses from Democrat subjects the other with responses from Republican subjects. Table 2 shows that for Democrat subjects, the *female* attribute-level had a larger positive effect on the choice than the *male* attribute-level indicating a greater preference for female advisors among Democrat subjects. On the other hand, the table shows that for Republican subjects, the *male* attribute-level had a larger positive effect than the *female* attribute-level indicating a greater preference for male advisors among Republican subjects. The difference in the effect between *male* and *female* among Democrat subjects was observed to be slightly greater than the difference among Republican subjects, indicating that the preference for female advisors among Democrat subjects was slightly stronger than the preference for male advisors among Republican subjects. These differences may be due to differences in outlook, and/or a more nuanced view on differences in effectiveness of advisor profiles as perceived by a third-party; for example, subjects from either political identity may acknowledge that government is a male-dominated environment, so Republican subjects may feel that a male would be better received as an advisor, whereas Democrats may feel strongly for female representation regardless of the perceived efficacy of such a candidate as perceived by a third-party.

Race Observations Along Political Party Lines Figure 5 shows the race choice flow diagrams for Democrat subjects (Figure 5a) and Republican subjects (Figure 5b). The diagrams for both show that most transitions to a single race were transitions to *White*. However, Democrat subjects made more transitions from *White* to other races and less transitions out of non-White races compared to Republican subjects, whose difference between transitioning from other races to *White* and transitioning from *White* to other races was more profound. For Democrat subjects, both the *White* and *Black* attribute-levels experienced an increase, by 15.8% and 7.90% respectively, in the number of times they were selected as the race attributes of the customized advisors rela-

	dy/dx	std err	p-val
mid-aged	0.2296	0.009	< 0.001
senior	-0.0802	0.012	< 0.001
female	0.0425	0.009	< 0.001
black	0.0052	0.013	0.680
latino	-0.0633	0.013	< 0.001
asian	-0.0471	0.013	< 0.001
stem prof.	0.0112	0.012	0.361
business prof.	-0.1032	0.013	< 0.001
other prof.	0.0010	0.012	0.937

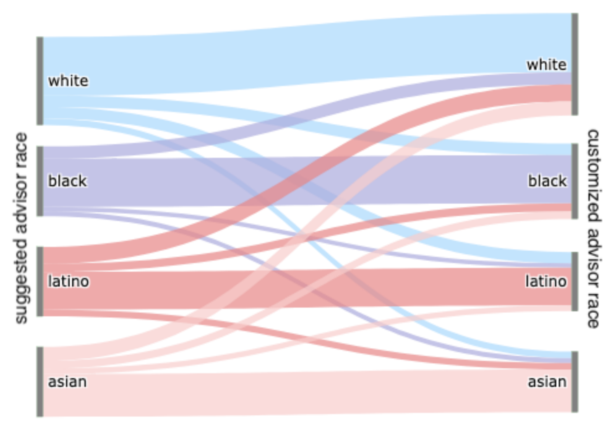
Table 2: AMCE values (dy/dx) of advisor attribute levels for Democrat subjects. The preference for White and male advisors compared to advisors of other races was lower for Democrat subjects than for Republican subjects. All attribute-levels were found to be statistically significant ($p < 0.05$) except for *Black*, *STEM-related professions*, and *other professions*.

	dy/dx	std err	p-val
mid-aged	0.273	0.012	< 0.001
senior	-0.039	0.016	0.019
female	-0.035	0.013	0.008
black	-0.088	0.017	< 0.001
latino	-0.123	0.018	< 0.001
asian	-0.100	0.017	< 0.001
stem prof.	0.045	0.017	0.007
business prof.	-0.024	0.018	0.168
other prof.	0.016	0.017	0.332

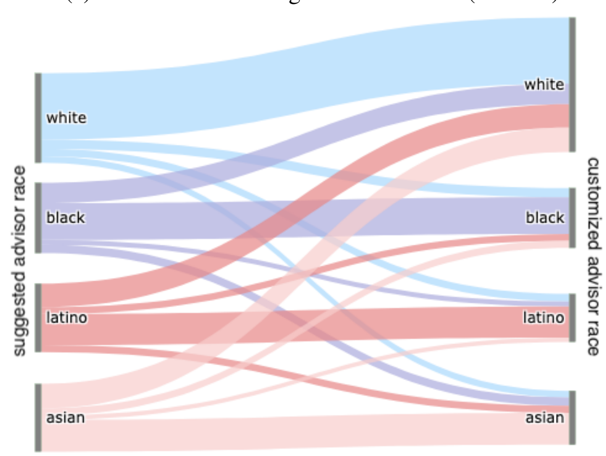
Table 3: AMCE values (dy/dx) of advisor attribute levels for Republican subjects. The preference for White and male advisors compared to advisors of other races was lower for Democrat subjects than for Republican subjects. All attribute-levels were found to be statistically significant ($p < 0.05$) except for *business-related professions* and *other professions*.

tive to the number of times they were suggested as the race attributes of the suggested advisors. *Latino* and *Asian*, however, both experienced a decrease by 15.1% and 12.8% respectively. The favorability of the *White* attribute-level was lower and the favorability of the other race attribute-levels was higher in the Democrat subset of the dataset compared to the entire dataset. However, the opposite trend is seen with Republican subjects, where the favorability of the *White* attribute-level was much higher and the favorability of the other race attribute-levels was lower compared to the entire dataset. For Republican subjects, only the *White* attribute-level experienced an increase, and it was a significant one of 50.1%, in the number of times it was selected as the race attribute of the customized advisors relative to the number of times it was suggested as the race attributes of the suggested advisors, whereas the *Black*, *Latino*, and *Asian* attribute-levels all experienced a decrease of 14.9%, 29.7%, and 20.7% respectively.

The difference in race preferences along party lines can also be seen in the conjoint analysis Tables 2 and 3. For both Democrat subjects and Republican subjects, the *Latino* and *Asian* attribute-levels had a negative effect on the probability that an advisor profile was selected. However, the negative effect was greater for Republican subjects than for Democrat subjects. Additionally, while all three *Black*, *Latino*, and *Asian* race attribute-levels decreased the probability that an advisor was selected by Republican subjects, the *Black* attribute-level slightly increased the probability that an advisor is selected by Democrat subjects, although this increase was not statistically significant. This indicates a stronger preference for White advisors specifically among Republican subjects compared to Democrat subjects. Overall, these results support our hypothesis **H3** that gender and race preferences varied along party lines.



(a) race choice flow diagram - Democrats (n = 338)



(b) race choice flow diagram - Republicans (n = 173)

Figure 5: Race choice flow diagrams for Democrat subjects (n = 338) and Republicans subjects (n = 173). The diagrams for both show that most transitions to a single race were transitions to *White*. However, in the data collected from Democrat subjects, more transitions were made from White to other races and less transitions were made out of non-White races.

Limitations & Future Directions

Studies have found AMT workers to be younger, female, more liberal, and with lower incomes than the general U.S. population (Shank 2015; Berinsky, Huber, and Lenz 2012). This was the case in our AMT worker population which consisted of more females, more self-identified Democrats, and a majority that held bachelor's degrees or higher levels of education. These trends, along with the observation that the majority of our subjects self-identified as White, posed a limitation on the analyses, results, and conclusions that were possible in this study. With a more diverse and representative sample, deeper inquiries would have been pursued. Thus, additional research is needed to gauge whether the results of this study are generalizable at the national level.

Our study is also limited when it comes to identifying the underlying sources of preferences; it does not disentangle what choices result from implicit biases towards/against certain demographics versus what choices result from the perception of bias and discrimination in the present day U.S.. Given the current polarized cultural/political climate in the U.S., a subject's preferences may be strongly driven by either one or both of a subjects own implicit biases and/or their perception of external bias. These hidden factors motivate an experimental design which builds on our existing observations, and would help answer these nuanced questions in future studies.

Conclusion

Understanding people's preferences is necessary to consolidate a better understanding of political and cultural tensions and their influence on decisions that impact society. A traditional approach is to ask survey questions directly on perceived competence and electing representation. However, an alternative approach is to conduct a discrete-choice exercise to select between two candidate profiles. This second approach has the advantage of revealing the underlying rank of preferences through statistical modeling. We applied such an approach through a web-based experiment on AMT tasking 722 U.S. based individuals to select attributes of an advisor that could successfully pitch government-expenses to fund. Utilizing conjoint analysis, we observed strong preferences for advisors that were White, middle-aged, and held government/STEM-related professions (0.68 AUROC, $p < 0.05$). We also observed shifts in gender preferences across self-identified political affiliations (0.70 AUROC, $p < 0.05$). This work motivates further studies in understanding the underlying reasoning for preferences that individuals may hold to help distinguish between (sub)conscious prejudices, lack of cross-boundary experiences, conscious adoption of a utilitarian outlook, and/or other possible reasons.

Appendix

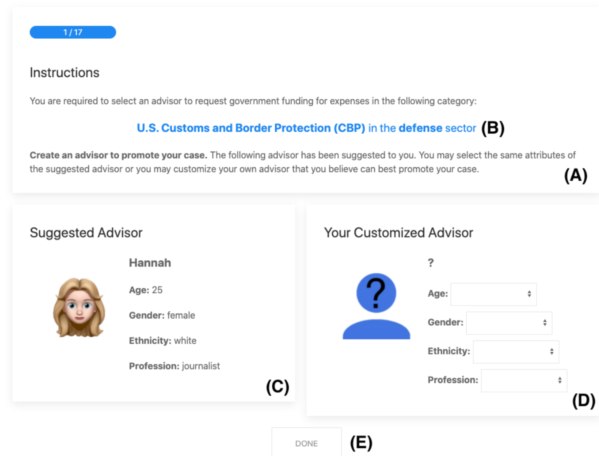


Figure 6: Screengrab of a single task. Tasks consisted of an instructions section (A) with the expense information (B), a suggested advisor section (C), a customized advisor section (D) with drop-down menus for customization, and a submit button (E) which was only enabled after subjects customized all attributes.

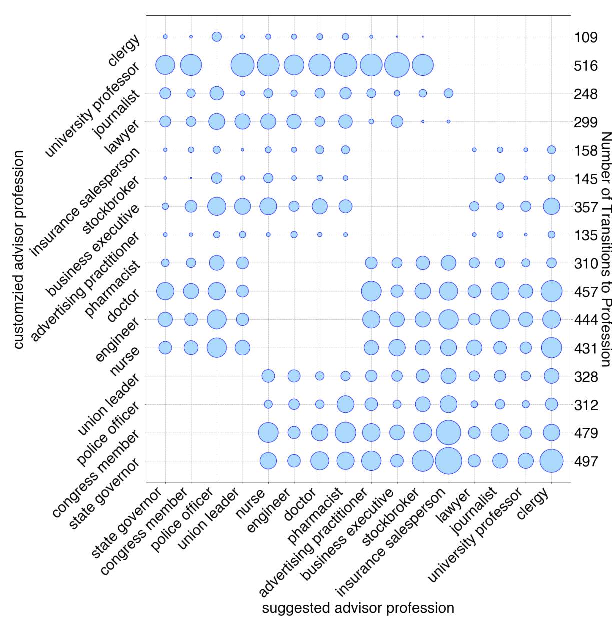


Figure 7: Professions choices bubble plot visualizing the frequency and direction of transitions between the profession attribute-levels of the suggested advisors and the selected profession attribute-levels of the customized advisors. Transitions between professions in the same profession category (which were not possible), and results from those who re-selected the suggested advisor profession are not plotted. The plot shows a preference for *university professor*, *STEM-related*, and *governance-related* professions.

References

Abbey, J. D.; and Meloy, M. G. 2017. Attention by design: Using attention checks to detect inattentive respondents and improve data quality. *Journal of Operations Management* 53: 63–70.

Anand, A.; Bansal, G.; and Aggrawal, D. 2018. Choice based diffusion model for predicting sales of mobile phones using conjoint analysis. *The Journal of High Technology Management Research* 29(2): 216–226.

Arechar, A. A.; Kraft-Todd, G. T.; and Rand, D. G. 2017. Turking overtime: how participant characteristics and behavior vary over time and day on Amazon Mechanical Turk. *Journal of the Economic Science Association* 3(1): 1–11.

Aruguete, M. S.; Huynh, H.; Browne, B. L.; Jurs, B.; Flint, E.; and McCutcheon, L. E. 2019. How serious is the 'carelessness' problem on Mechanical Turk? *International Journal of Social Research Methodology* 22(5): 441–449.

Awad, E.; Dsouza, S.; Kim, R.; Schulz, J.; Henrich, J.; Shariff, A.; Bonnefon, J.-F.; and Rahwan, I. 2018. The Moral Machine Experiment. *Nature* 563. doi:10.1038/s41586-018-0637-6.

Berinsky, A. J.; Huber, G. A.; and Lenz, G. S. 2012. Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk. *Political Analysis* 20(3): 351–368. doi:10.1093/pan/mpr057.

Boncinelli, F.; Contini, C.; Gerini, F.; Romano, C.; Scozzafava, G.; and Casini, L. 2020. The Role of Context Definition in Choice Experiments: a Methodological Proposal Based on Customized Scenarios. *Wine Economics and Policy* 9(2): 49–62.

Buhrmester, M. D.; Talaifar, S.; and Gosling, S. D. 2018. An evaluation of Amazon's Mechanical Turk, its rapid rise, and its effective use. *Perspectives on Psychological Science* 13(2): 149–154.

Capers IV, Q.; Clinchot, D.; McDougle, L.; and Greenwald, A. G. 2017. Implicit racial bias in medical school admissions. *Academic Medicine* 92(3): 365–369.

Carey, J. M.; Carman, K. R.; Clayton, K. P.; Horiuchi, Y.; Htun, M.; and Ortiz, B. 2020. Who wants to hire a more diverse faculty? A conjoint analysis of faculty and student preferences for gender and racial/ethnic diversity. *Politics, Groups, and Identities* 8(3): 535–553.

Carlsson, M.; and Eriksson, S. 2019. Age discrimination in hiring decisions: Evidence from a field experiment in the labor market. *Labour Economics* 59: 173–183.

Center, T. P. 2020. What are the largest tax expenditures? https://www.taxpolicycenter.org/briefing-book/what-are-largest-tax-expenditures. Accessed: 2021-03-01.

de Kock, F. S.; and Hauptfleisch, D. B. 2018. Reducing racial similarity bias in interviews by increasing structure: A quasi-experiment using multilevel analysis. *International Perspectives in Psychology* 7(3): 137–154.

De Martino, B.; Kumaran, D.; Seymour, B.; and Dolan, R. J. 2006. Frames, Biases, and Rational Decision-Making in the Human Brain. *Science* 313(5787): 684–687. ISSN 0036-8075. doi:10.1126/science.1128356. URL https://science.sciencemag.org/content/313/5787/684.

Dehon, E.; Weiss, N.; Jones, J.; Faulconer, W.; Hinton, E.; and Sterling, S. 2017. A systematic review of the impact of physician implicit racial bias on clinical decision making. *Academic Emergency Medicine* 24(8): 895–904.

Derous, E.; and Decoster, J. 2017. Implicit Age Cues in Resumes: Subtle Effects on Hiring Discrimination. *Frontiers in Psychology* 8: 1321. ISSN 1664-1078. doi:10.3389/fpsyg.2017.01321. URL https://www.frontiersin.org/article/10.3389/fpsyg.2017.01321.

Dictionary, O. E. n.d. Preference (noun) Definition - Oxford English Dictionary. https://www.oed.com/search?searchType=dictionary&q=preference&_searchBtn=Search. Accessed: 2021-03-01.

Doherty, C. 2020. Voters' Attitudes About Race and Gender Are Even More Divided Than in 2016. https://www.pewresearch.org/politics/2020/09/10/voters-attitudes-about-race-and-gender-are-even-more-divided-than-in-2016/. Accessed: 2021-03-01.

Edmonds, D. 2013. *Would you kill the fat man?: The trolley problem and what your answer tells us about right and wrong*. Princeton University Press.

Einstein, K. L.; and Glick, D. M. 2017. Does race affect access to government services? An experiment exploring street-level bureaucrats and access to public housing. *American Journal of Political Science* 61(1): 100–116.

Elsesser, K. M.; and Lever, J. 2011. Does gender bias against female leaders persist? Quantitative and qualitative data from a large-scale survey. *Human Relations* 64(12): 1555–1578.

Enns, P. K. 2018. Clarifying the Role of Racism in the 2016 US Presidential Election: Opinion Change, Anti–Immigrant Sentiment, and Vote Choice. In *presentation at the 2018 meeting of the American Political Science Association*.

Fink, K. 2019. The biggest challenge facing journalism: A lack of trust. *Journalism* 20(1): 40–43.

Gardner, R. M.; Brown, D. L.; and Boice, R. 2012. Using Amazon's Mechanical Turk website to measure accuracy of body size estimation and body dissatisfaction. *Body Image* 9(4): 532 – 534. ISSN 1740-1445. doi:https://doi.org/10.1016/j.bodyim.2012.06.006. URL http://www.sciencedirect.com/science/article/pii/S174014451200085X.

Gift, K.; and Gift, T. 2014. Does Politics Influence Hiring? Evidence from a Randomized Experiment. *Political Behavior* 37(3): 653–675. doi:10.1007/s11109-014-9286-0.

Gong, F.; Xu, J.; and Takeuchi, D. T. 2017. Racial and ethnic differences in perceptions of everyday discrimination. *Sociology of Race and Ethnicity* 3(4): 506–521.

Green, P. E.; Krieger, A. M.; and Wind, Y. 2001. Thirty years of conjoint analysis: Reflections and prospects. *Interfaces* 31(3_supplement): S56–S73.

Green, P. E.; and Srinivasan, V. 1990. Conjoint analysis in marketing: new developments with implications for research and practice. *Journal of marketing* 54(4): 3–19.

Gutiérrez-Romero, R.; and LeBas, A. 2020. Does electoral violence affect vote choice and willingness to vote? Conjoint analysis of a vignette experiment. *Journal of peace research* 57(1): 77–92.

Hainmueller, J.; Hopkins, D. J.; and Yamamoto, T. 2014. Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments. *Political Analysis* 22(1): 1–30. doi:10.1093/pan/mpt024.

Herring, S. C.; Dainas, A. R.; Lopez Long, H.; and Tang, Y. 2020. Animoji performances. *Language@ Internet* 18(1).

Hosmer Jr, D. W.; Lemeshow, S.; and Sturdivant, R. X. 2013. *Applied logistic regression*, volume 398. John Wiley & Sons.

Hunt, N. C.; and Scheetz, A. M. 2019. Using MTurk to distribute a survey or experiment: Methodological considerations. *Journal of Information Systems* 33(1): 43–65.

Kaufmann, M. C.; Krings, F.; Zebrowitz, L. A.; and Sczesny, S. 2017. Age Bias in Selection Decisions: The Role of Facial Appearance and Fitness Impressions. *Frontiers in Psychology* 8: 2065. ISSN 1664-1078. doi:10.3389/fpsyg.2017.02065. URL https://www.frontiersin.org/article/10.3389/fpsyg.2017.02065.

Kinsey, M.; Gwynne, S.; Kuligowski, E. D.; and Kinateder, M. 2019. Cognitive biases within decision making during fire evacuations. *Fire technology* 55(2): 465–485.

Koul, A.; Becchio, C.; and Cavallo, A. 2018. Cross-validation approaches for replicability in psychology. *Frontiers in psychology* 9: 1117.

Krause, N. M.; Brossard, D.; Scheufele, D. A.; Xenos, M. A.; and Franke, K. 2019. Trends—Americans' Trust in Science and Scientists. *Public Opinion Quarterly* 83(4): 817–836. ISSN 0033-362X. doi:10.1093/poq/nfz041. URL https://doi.org/10.1093/poq/nfz041.

Larson, J.; Angwin, J.; Mattu, S.; and Kirchner, L. 2016. Machine Bias. http://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Accessed: 2021-03-01.

Lewis, S. C. 2020. Lack of trust in the news media, institutional weakness, and relational journalism as a potential way forward. *Journalism* 21(3): 345–348.

Lindner, N. M.; Graser, A.; and Nosek, B. A. 2014. Age-Based Hiring Discrimination as a Function of Equity Norms and Self-Perceived Objectivity. *PLOS ONE* 9(1): 1–6. doi:10.1371/journal.pone.0084752. URL https://doi.org/10.1371/journal.pone.0084752.

Lössbroek, J.; Lancee, B.; van der Lippe, T.; and Schippers, J. 2020. Age Discrimination in Hiring Decisions: A Factorial Survey among Managers in Nine European Countries. *European Sociological Review* 37(1): 49–66. ISSN 0266-7215. doi:10.1093/esr/jcaa030. URL https://doi.org/10.1093/esr/jcaa030.

Markowitz, D. M.; and Slovic, P. 2020. Social, psychological, and demographic characteristics of dehumanization toward immigrants. *Proceedings of the National Academy of Sciences* 117(17): 9260–9269.

Mavrodiev, P.; Tessone, C. J.; and Schweitzer, F. 2013. Quantifying the effects of social influence. *Scientific reports* 3: 1360.

Meier, K. J.; and O'Toole Jr, L. J. 2006. Political control versus bureaucratic values: Reframing the debate. *Public administration review* 66(2): 177–192.

Mellis, A. M.; and Bickel, W. K. 2020. Mechanical Turk data collection in addiction research: Utility, concerns and best practices. *Addiction* 115(10): 1960–1968.

Meyerding, S. G.; and Merz, N. 2018. Consumer preferences for organic labels in Germany using the example of apples–Combining choice-based conjoint analysis and eye-tracking measurements. *Journal of cleaner production* 181: 772–783.

Nadler, J. T.; and Kufahl, K. M. 2014. Marital status, gender, and sexual orientation: Implications for employment hiring decisions. *Psychology of Sexual Orientation and Gender Diversity* 1(3): 270–278. doi:10.1037/sgd0000050.

Neumark, D.; Burn, I.; and Button, P. 2017. Age discrimination and hiring of older workers. *Age* 6.

Newman, N.; and Fletcher, R. 2017. Bias, bullshit and lies: Audience perspectives on low trust in the media. *Available at SSRN 3173579* .

of Labor, U. D. n.d. Age Discrimination. https://www.dol.gov/general/topic/discrimination/agedisc. Accessed: 2021-03-01.

Parikh, R. B.; Teeple, S.; and Navathe, A. S. 2019. Addressing bias in artificial intelligence in health care. *Jama* 322(24): 2377–2378.

Pedersen, M. J.; and Nielsen, V. L. 2020. Bureaucratic decision-making: A multi-method study of gender similarity bias and gender stereotype beliefs. *Public Administration* 98(2): 424–440.

QuickFacts, C. B. 2019. U.S. Census Bureau QuickFacts: United States. URL https://www.census.gov/quickfacts/fact/table/US/RHI125219#RHI125218.

Reinhart, R. 2020. Nurses Continue to Rate Highest in Honesty, Ethics. https://news.gallup.com/poll/274673/nurses-continue-rate-highest-honesty-ethics.aspx. Accessed: 2021-03-01.

Richardson, B.; Webb, J.; Webber, L.; and Smith, K. 2013. Age discrimination in the evaluation of job applicants. *Journal of Applied Social Psychology* 43(1): 35–44.

Roth, P. L.; Thatcher, J. B.; Bobko, P.; Matthews, K. D.; Ellingson, J. E.; and Goldberg, C. B. 2020. Political affiliation and employment screening decisions: The role of similarity and identification processes. *Journal of Applied Psychology* 105(5): 472.

Santana, E. 2018. Situating perceived discrimination: How do skin color and acculturation shape perceptions of discrimination among Latinos? *The Sociological Quarterly* 59(4): 655–677.

Schöbel, M.; Rieskamp, J.; and Huber, R. 2016. Social Influences in Sequential Decision Making. *PLOS ONE* 11(1): 1–23. doi:10.1371/journal.pone.0146536. URL https://doi.org/10.1371/journal.pone.0146536.

Shafranek, R. M. 2019. Political considerations in nonpolitical decisions: a conjoint analysis of roommate choice. *Political Behavior* 1–30.

Shank, D. B. 2015. Using Crowdsourcing Websites for Sociological Research: The Case of Amazon Mechanical Turk. *The American Sociologist* 47(1): 47–55. doi:10.1007/s12108-015-9266-9.

Stark, L. 2018. Facial recognition, emotion and race in animated social media. *First Monday* .

Stewart, N.; Ungemach, C.; Harris, A. J.; Bartels, D. M.; Newell, B. R.; Paolacci, G.; Chandler, J.; et al. 2015. The average laboratory samples a population of 7,300 Amazon Mechanical Turk workers. *Judgment and Decision making* 10(5): 479–491.

Support, A. 2020. Use Memoji on your iPhone or iPad Pro. https://support.apple.com/en-us/HT208986. Accessed: 2021-03-01.

Tandoc Jr, E. C.; Jenkins, J.; and Craft, S. 2019. Fake news as a critical incident in journalism. *Journalism Practice* 13(6): 673–689.

Tang, S.; Zhang, X.; Cryan, J.; Metzger, M.; Zheng, H.; and Zhao, B. 2017. Gender Bias in the Job Market: A Longitudinal Analysis. *Proceedings of the ACM on Human-Computer Interaction* 1: 1–19. doi:10.1145/3134734.

Thebault-Spieker, J.; Kluver, D.; Klein, M.; Halfaker, A.; Hecht, B.; Terveen, L.; and Konstan, J. 2017. Simulation Experiments on (the Absence of) Ratings Bias in Reputation Systems. *Proceedings of the ACM on Human-Computer Interaction* 1: 1–25. doi:10.1145/3134736.

Verma, V. K.; and Chandra, B. 2018. Sustainability and customers' hotel choice behaviour: a choice-based conjoint analysis approach. *Environment, development and sustainability* 20(3): 1347–1363.

Webster, S. W. 2018. Anger and declining trust in government in the American electorate. *Political Behavior* 40(4): 933–964.

Wijenayake, S.; van Berkel, N.; Kostakos, V.; and Goncalves, J. 2019. Measuring the effects of gender on online social conformity. *Proceedings of the ACM on Human-Computer Interaction* 3(CSCW): 1–24.

Willnat, L.; Weaver, D. H.; and Wilhoit, G. C. 2019. The American journalist in the digital age: How journalists and the public think about journalism in the United States. *Journalism Studies* 20(3): 423–441.

Ye, T.; Alahmad, R.; Pierce, C.; Robert, L.; et al. 2017. Race and rating on sharing economy platforms: The effect of race similarity and reputation on trust and booking intention in Airbnb. AIS.