I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18239

Foreign Accents and Employer Beliefs: Experimental Evidence on Hiring Discrimination

Elisa Taveras
Ozlem Tonguc
Maria Zhu
Nicola Miller

OCTOBER 2025

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18239

Foreign Accents and Employer Beliefs: Experimental Evidence on Hiring Discrimination

Elisa Taveras
University of Texas Rio Grande Valley

Ozlem Tonguc
Binghamton University

Maria Zhu
Syracuse University and IZA

Nicola Miller
Binghamton University

OCTOBER 2025

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Foreign Accents and Employer Beliefs: Experimental Evidence on Hiring Discrimination*

This study investigates whether employers in an online hiring experiment exhibit discrimination based on workers' accents that indicate English is not their primary language. To assess accent bias, we implement a randomized treatment design in which participants acting as employers are assigned to one of two conditions: a treatment where the worker's accent is revealed ("Accent Revealed") or a control where it is not ("Accent Blind"). Using incentive-compatible methods, we elicit employers' beliefs about the productivity of randomly assigned workers, providing brief demographic information and audio clips that either reveal or mask accent characteristics. We evaluate worker productivity in two skills: Mathematics and Verbal reasoning. We find that employers rate accented workers as less capable than their non-accented counterparts in both skills, and this gap persists after providing employers with a signal of a worker's test score. Employers also display lower willingness to pay, particularly in Verbal skills tasks, even when provided with performance signals.

JEL Classification: C91, D90, J01, J71

Keywords: foreign accent, labor market discrimination, laboratory experiments

Corresponding author:
Maria Zhu
Syracuse University
900 S Crouse Ave
13244 Syracuse NY
USA

E-mail: mzhu33@syr.edu

* This project was reviewed by Binghamton University's Institutional Review Board and determined exempt under protocol STUDY00005035. All mistakes are our own.

1 Introduction

"My mom didn't speak English as a first language. Her English was pretty good, but she still had a strong accent. And it was so infuriating for me as a little kid to see the way people would treat her–basically as if she was dumber."– Alexis Ohanian on the WIRED's Uncanny Valley podcast; Aug 26, 2025

Many developed economies, including the United States, rely heavily on foreign-born workers. In the US, the proportion of foreign-born individuals in the civilian labor force reached a record high of 18.1 percent in 2022 (Bureau of Labor Statistics, 2023). A substantial share of these workers are non-native English speakers: 80 percent of foreign-born individuals in the US speak a language other than English at home and 41 percent report limited English proficiency (Ruggles *et al.*, 2024).[1] As the share of foreign-born workers in the US labor force continues to grow, it is important to understand their experiences in the labor market. This is particularly relevant from a policy perspective, as attitudes towards immigrants in the US have become increasingly politically polarized. In 2017, 84 percent of Democrats believed that immigrants do more to strengthen than burden the country, compared to only 42 percent of Republicans (Pew Research Center, 2017).

From a labor market perspective, speech is a critical determinant of perceived competence and employability. Even when highly proficient in English, non-native English speakers may exhibit pronunciation patterns that differ from native norms. Non-U.S. English accent-related speech characteristics–hereafter referred to as an "accent"–can make workers for whom English is not a primary language especially vulnerable to potential bias. Indeed, research finds that people evaluate those with non-native accents differently than their counterparts with native accents when it comes to trustworthiness (Munro *et al.*, 2006; Burns & Keswell, 2015; Caballero & Pell, 2020; Spence *et al.*, 2024). Accent can also affect listeners' comprehension, leading to communication delays (van Wijngaarden *et al.*, 2001; Floccia *et al.*, 2009).[2] The presence of non-native accents has also been linked to perceptions of lower trust and competence, even when there are no measurable communication costs (Rubin, 1992; Adank *et al.*, 2009; Shuck, 2004, 2006). These perceptions potentially contribute to wage disparities, as Isphording & Sinning (2012) find that higher English speaking ability is associated with increased wages in the US. Moreover, language skills are widely regarded

[1] These values were calculated using the 2022 American Community Survey(ACS) data. The ACS measures English proficiency on a five-point scale: 1: Does not speak English, 2: Yes, but not well, 3: Yes, speaks well, 4: Yes, speaks very well, 5: Yes, speaks only English. We define 1–3 as having limited English proficiency.

[2] This holds true even though strong accents are often intelligible and listeners adapt to them, mitigating such delays over time (Munro & Derwing, 1995; Derwing & Munro, 2009; Bradlow & Bent, 2008; Baese-Berk *et al.*, 2013).

as a form of human capital, yielding significant labor market returns (Chiswick & Miller, 1995).

This paper examines whether employers are biased against workers who speak with an accent but are otherwise comparable to workers without an accent.[3] Identifying accent-based discrimination in the labor market requires isolating accent from other correlated characteristics, such as race and ethnicity, education, language proficiency, and communication ability. To overcome the challenge of disentangling the effect of accent from the effect of confounding factors, we conduct two hiring experiments using participants residing in the US via an online labor platform.[4]

We conduct a lab experiment to address some of the limitations of correspondence studies, which typically infer discrimination based on callback rates in response to fictitious resumes submitted to real job postings. In contrast, a lab experiment allows us to construct a richer dataset that includes both hiring decisions and employer beliefs about worker productivity. This allows for a more nuanced analysis of the sources of accent-based discrimination across multiple dimensions. Also, in real-world labor markets, employers typically speak with candidates only after the initial callback stage. Although audit studies may replicate this interaction, such as by phone calls in early stages or interviews with fictitious applicants (e.g., Schmaus & Kristen, 2022), they limit researchers' control over the accent and content of speech. In contrast, in a lab experiment, we are able to present *standardized* audio clips of actual participants, ensuring consistency and precision in the manipulation of accent.

Our study follows a two-step design.[5] First, we created a pool of workers that belong to one of the three major foreign-born racial-ethnic groups that participate in the U.S. labor market (i.e., Asian, Hispanic and White)[6] for whom we collect information on demographics and preferences, and a voice recording. The recording consists of the worker reading a paragraph with standardized content, allowing us to hold constant English proficiency (i.e., vocabulary, grammar, sentence structure). In addition, we collected data on performance in two skill areas: Mathematics and Verbal reasoning.

In the second step, we conducted two experiments that both elicit employers' incentivized beliefs about worker performance and their willingness to pay to hire workers for their test performance. In the spirit of a correspondence study, we presented employers with a brief and standardized worker profile (henceforth referred to as "CV") that included demographic information from the worker and educational attainment. As part of the pro-

[3]In this paper, we refer to "accents" to mean accents of whom English is not a primary language and not having an accent means having a speech pattern of a native speaker of U.S. English.

[4]Prolific, www.prolific.com

[5]This design closely follows Bohren *et al.* (2023).

[6]Bureau of Labor Statistics, 2024: https://www.bls.gov/news.release/pdf/forbrn.pdf

file, the employer listened to an audio clip related to the worker's preferences. To minimize the influence of speech characteristics other than those related to accent on evaluations, all audio clips were generated using artificial intelligence (AI) software using a standardized (gender-specific) AI voice. We use two types of audio clips. First, using AI, we voiceover each worker's original recording to create an *Accent Revealed* audio clip that preserves the worker's accent-related speech characteristics (e.g., pronunciation, intonation) while standardizing non-accent-related characteristics (e.g., timbre). Second, we created an *Accent Blind* audio clip of the written transcript of a worker's recording, using the same AI voice with a neutral U.S. English accent.

In the first experiment ("No Signal"), we implemented a 2×2 between-subject treatment design in which employers are randomized to one skill, Mathematics or Verbal, and to one type of audio clips: *Accent Revealed* or *Accent Blind*. We collected beliefs and hiring decisions when only the worker's CV and audio clip is available. Although the CV can aid employers in assessing worker ability, employers may still hold inaccurate beliefs about worker performance. To investigate whether inaccurate beliefs could be corrected, we conducted a second 2×2 experiment ("Signal Intervention") where employers received a signal on test performance in some of their evaluations.

Having access to worker CVs and audio clips, employers in both experiments were asked to predict different workers' performance on skills tests. We find that when the worker accent is revealed, employers believe that accented workers perform worse in both the Mathematics and Verbal skill tests compared to non-accented workers. Specifically, the accent gap in beliefs over worker performance is -0.33 points for Mathematics and -0.66 for Verbal skills, corresponding to 4.6 percent and 6.7 percent of the average score in these skills, respectively. Strikingly, this bias arises in the beliefs about Mathematics skills, which, employers are informed of being measured in a test that requires minimal English proficiency. Furthermore, we find that these accent gaps in beliefs do not accurately reflect the actual accent gaps in workers' abilities. In looking at heterogeneity across racial/ethnic groups, we find that accent bias is present in beliefs about all three groups (Asian, Hispanic, White) in the Verbal skills test. In the Mathematics test, we find accent gaps in beliefs about Hispanic and White workers, while the estimated gap for Asians is much smaller and not statistically significant. Notably, while employers adjust their overall beliefs about performance in response to being provided a signal on the worker's test performance, the signal does not largely reduce the accent gap in beliefs about worker performance, particularly for Verbal skills.

We find accent gaps in employers' willingness to pay to hire a given worker for their skill test performance as well. Precisely, in the absence of the performance signal, when the worker accent is revealed, employers are willing to pay 3.0 and 7.2 cents less, for an accented worker for their Mathematics and Verbal skills, respectively. These imply 3.5% and 7.9% lower earnings for accented workers. When an informative performance signal is available, this gap disappears for Mathematics but persists for Verbal skills, with employers choosing to pay 10.2 cents less (16%) to accented workers. Moreover, we demonstrate a tight relationship between the actual accent gaps in willingness to pay to those implied by the employers' beliefs, which favors belief-based theories of discrimination against accented workers in our context.

Finally, we demonstrate the robustness of our results to using an alternative measure of accent and controlling for additional factors such as audio clip characteristics, worker evaluation order effects and sequence effects, employer characteristics and sample selection.

Our results demonstrate that accent-related speech characteristics affect employer beliefs and hiring decisions, even in tasks that do not require English proficiency, such as in the Mathematics skills test. Part of this accent gap stems from employers inaccurately using accent as a proxy for worker ability. A signal about the worker productivity can help reduce the accent penalty, but only when little English knowledge is needed (Mathematics skills test). The persistence of the gap in Verbal skills test when employers receive an informative performance signal suggests that accent may dominate an informative performance signal in forming employer beliefs about a worker's verbal skills.

Related Literature. Our work builds on existing experimental research on labor market discrimination, which has examined factors such as race, immigration status, physical disability, customer preferences and the cost of discrimination (Bertrand & Mullainathan, 2004; Oreopoulos, 2011; Bartoš *et al.*, 2016; Hedegaard & Tyran, 2018; Acquisti & Fong, 2020; Chan, 2022; Bellemare *et al.*, 2023; Ge & Wu, 2024; Tomlin, 2024; Eames, 2024).[7] While most of the literature relies on field-based audit or correspondence studies, a smaller number of lab-based correspondence studies have investigated labor market discrimination based on race (Park, 2024; Rackstraw, 2022) and country of origin (Bohren *et al.*, 2023). Our paper contributes to this literature by focusing on accent-based discrimination–a salient but underexplored dimension of identity–using a novel lab-based design that isolates accent from other correlated characteristics such as race, language proficiency, and communication ability.

Our paper also contributes to understanding the effectiveness of informative signals in diminishing biases in candidate evaluations.[8] Informative signals, group- or individual-level, are a central design element in discrimination experiments (Castillo & Petrie, 2010; Albrecht

[7]See Bertrand & Duflo (2017) and Neumark (2018) for a review of the experimental literature on discrimination in economics.

[8]See Haaland *et al.* (2023) for a review of the literature on information provision experiments.

et al., 2013; Ewens *et al.*, 2014; Reuben *et al.*, 2014; Bohren *et al.*, 2019; Wozniak & MacNeill, 2020; Coffman *et al.*, 2021; Laouénan & Rathelot, 2022; Bohren *et al.*, 2023; Coffman *et al.*, 2025). Our results on the effects of an individual-level signal are in line with that of Coffman *et al.* (2025) in that individual-level signals are under-utilized by evaluators for Verbal skills, but not for Mathematics skills.

This paper also contributes to research on how voice and speech characteristics influence perception and treatment in social and labor market contexts. Within this literature, research finds that language skills (Chiswick & Miller, 1995; Isphording & Sinning, 2012) and voices perceived as attractive (Zuckerman & Driver, 1989) are associated with more positive assessments and labor market outcomes. At the same time, other observational or qualitative studies have documented negative effects of accents (Rubin, 1992; Li & Campbell, 2009; Gluszek & Dovidio, 2010; Tenzer *et al.*, 2014; Schroeder & Epley, 2015).[9] Fuertes *et al.* (2012) approaches the subject from a social psychology perspective, providing qualitative insights into how accents influence interpersonal evaluations. Moreover, lab-based studies have explored the influence of non-native accents in labor market perceptions and hiring decisions using recorded interviews with actors or entrepreneurs' fund pitches (Deprez-Sims & Morris, 2010; Huang *et al.*, 2013). Our study extends this literature by using a novel lab-based design to provide new insights on accent discrimination. We investigate employer beliefs about worker ability by accent status and the interplay between these beliefs and pay decisions. Therefore, our paper contributes to the existing literature by using real participants in a setting where, instead of asking hypothetical decisions, worker productivity is relevant and impacts the employer's payoff. Notably, we also explore the role of information in shaping accent gaps in employer evaluations. Furthermore, our experiment evaluates various accent types and skills, offering nuanced insights into how accent discrimination manifests.

Lastly, our paper contributes to the small but emerging body of literature utilizing AI tools to complement experiment design (such as by improving the perceived authenticity of fictional job candidates in Evsyukova *et al.*, 2025). Our utilization of AI tools standardizes audio recordings from real humans, which are complex and multi-attributed stimuli.

The remainder of the paper is structured as follows: Section 2 outlines the experiment and describes the worker pool and employer samples. Section 3 describes our empirical specifications, and Section 4 presents our results. Finally, Section 5 concludes the paper.

[9] See Spence *et al.* (2024) for a meta-analysis on the role of accent in hiring, and Spence *et al.* (2021) for a meta-analysis that investigates studies on accent discrimination among children.

2 Experimental Design

We implemented a two-stage online experiment with participants recruited through the Prolific platform.[10] In the first stage, hereafter "Worker Study," we recruited participants residing in the United States to play the role of a "worker."[11] We collected information on each worker's demographics and preferences to construct a profile and measured their performance in different knowledge tests focusing on two separate skills: Mathematics or Verbal. In the second stage, hereafter "Employer Study," we recruited new participants residing in the US ("employers") to evaluate the workers in two separate experiments, **No Signal** and **Signal Intervention.** After observing the profiles of the workers, employers reported their incentivized beliefs on workers' test performances and chose an amount they were willing to pay (WTP) to hire the workers for their test performance. In the following subsections, we provide details on each part.

2.1 The Worker Study

A total of 522 participants completed the worker study. For each worker, we have demographic information, an audio recording, and their performance on the skill tests. The worker study consisted of three parts.[12]

In Part 1, workers self-reported their demographic information. In Part 2, following recent online experiments on discrimination (e.g., Bohren *et al.*, 2023), we collected workers' preferences on "filler items": characteristics that are unlikely to be related to worker performance, but help in making worker profiles distinct from each other. In our design, these items consisted of information on workers' preferences for beverages, desserts, and colors. Once workers completed a multiple choice survey on the preferences questions, they were presented with a paragraph summarizing their answers and tasked to record themselves reciting the paragraph. Thus, the recording of the filler items in our design played the additional role of showcasing workers' accents within the following standardized context:[13]

> *"I prefer* ***[coffee/tea]*** *over* ***[coffee/tea]*** ***[and/but do not]*** *like to have dessert with it. As an ice cream flavor, I prefer* ***[chocolate/vanilla]*** *over* ***[chocolate/vanilla]****. I* ***[/do not]*** *like the color yellow* ***[and/but]*** *prefer* ***[yellow gold/bright yellow]*** *over* ***[yellow gold/bright yellow]****."*

[10] https://prolific.com/

[11] To ensure a sufficient number of accented speakers, we utilized additional Prolific pre-screening conditions such as the participants' native language being Chinese, Spanish, etc.

[12] Information was also collected on workers' willingness to accept for compensation. Although we do not include that information here, it is available upon request.

[13] Brackets reflect text that varies based on each individual worker's reported preferences.

In Part 3, workers completed multiple 15-question tests of interest that measure a worker's knowledge across the following topics. In this paper, we focus on the following tests:[14]

- The **Mathematics Skills** test consisted of 15 open-ended questions, measuring a worker's arithmetic and algebraic knowledge. Questions were presented with minimal verbal expressions to avoid differences in performance that could be attributed to knowledge of English. The workers had a 20-second time limit to answer each question. See Figure A.1 for an example.

- The **Verbal Skills** test consisted of 15 multiple choice questions, measuring knowledge of vocabulary and logic. As in the math test, the workers had a time limit of 20 seconds per question. Figure A.2 shows a sample question.

The median completion time of the worker study was 17 minutes. Participants were paid a fixed amount of \$2.75 for completion and a bonus of up to \$1.50. The bonus payment was randomly determined based on one of the knowledge tests (10 cents for each correct answer) or the worker's payoff from Part 4 of the experiment. The average total payment was \$3.48. In Section 2.3, we describe attributes of the "Worker Pool," which is the subsample of workers who were ultimately evaluated by employers.

2.2 The Employer Study

The employer study consisted of two separate experiments: the **No Signal** experiment and the **Signal Intervention** experiment. A total of 1,472 Prolific participants completed the employer study, from which approximately 60% (883) participated in the **No Signal** experiment. In both experiments, each participant was randomly assigned to a single treatment cell in a 2×2 design (skill × accent information):

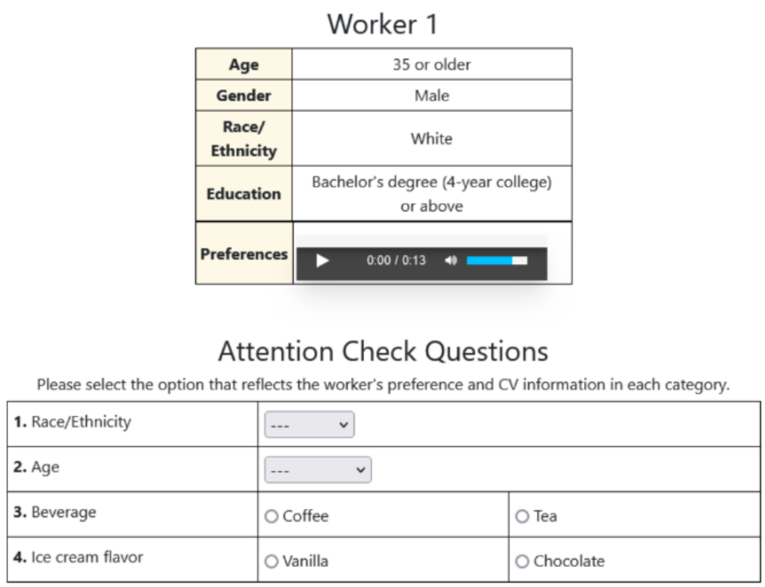
- **Skills:** Employers were randomized into evaluating workers based on only one of the tests ("skills"): *Mathematics* or *Verbal*. To facilitate their decision-making, employers were provided with a sample test question (see Figure A.1 and Figure A.2) as well as the average test score of the worker pool for the skill they were assigned to.

- **Accent Information:** Employers were also randomized into listening to a version of *audio clips*: either the workers' speech characteristics related to accent (e.g., intonation) were masked (*Accent Blind*) or revealed (*Accent Revealed*).

[14] Although we also collected a third skill test on Social Norms knowledge, the employers' analysis of this paper focuses on the Mathematics and Verbal skills test. Thus, we exclude the details of this third test here for brevity.

Each employer was randomly matched to a set of workers and were then shown profiles one at a time. For each worker, the employer was presented with a simple profile ("CV") containing information on the worker's age, gender, race/ethnicity, and education level. Employers also listened to an audio clip about the worker's preferences on filler information (beverages, desserts and colors). To ensure that employers paid attention to both the worker's demographic information and the audio clip, they were asked attention check questions regarding two random CV items and two statements from the audio clip for each worker. Figure 2.1 shows an example worker CV and the random attention check questions attached to it. Note that the version of the audio clip depends on the information treatment.

Figure 2.1: Example Worker CV and Attention Check Questions

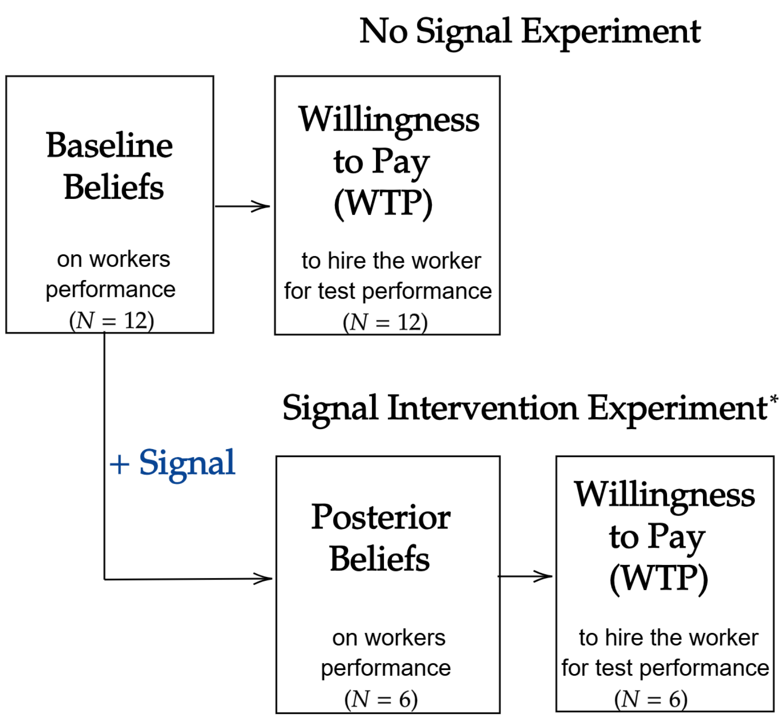


In the *Accent Revealed* treatment arm, to provide information on the worker's accent, we used an Artificial Intelligence (AI) voice-transformed version of each worker's own recording collected as part of the Worker Study. The transformed voice preserved the pronunciation and intonation of the original recordings. Thus, in the *Accent Revealed* treatment, the audio clip conveys information on whether a worker has a foreign accent or not to the employer. On the other hand, in the *Accent Blind* treatment arm, the audio clip, while still accurately stating the worker's preferences, is generated by a commercial text-to-speech AI tool using a US-native-accented voice. Thus, in this treatment, we *mask* accents of workers. Importantly, because there might be slight variations in speech characteristics even within native English speakers, the *Accent Blind* treatment allows us to standardize the recordings of all workers. See Section 2.3 for more details on the creation of the audio recordings.

In both experiments, we collected demographic information in addition to incentivized employers' decisions. Figure 2.2 depicts the sequence of decision tasks in each experiment. Employers perform two tasks: (1) reporting beliefs about workers' test performance and subsequently, (2) reporting WTP to hire different workers based on their test performance.

In both experiments, we measure employers' baseline beliefs about workers' test performance by asking employers, *"What is the number of questions this worker answered correctly in the **[Math/Verbal]** skills test?"* To incentivize truthful reporting, payoffs were based on the Binarized Scoring Rule (BSR) (Hossain & Okui, 2013) with a maximum allowed error of 3 points.[15] We collected baseline beliefs for 12 separate workers one by one. Workers differed from each other by their gender, race/ethnicity and accent characteristics.[16]

Figure 2.2: Timeline of Employers' Incentivized Tasks



Note. The figure shows the sequence of decision tasks in each experiment. In each experiment, employers are assigned to an information treatment, *Accent Revealed* or *Accent Blind*, and a skill type: *Mathematics* or *Verbal*. *In the Signal Intervention experiment, employers had a fourth task that is omitted in this study.

Because CVs may contain information beyond demographic characteristics–and because employers might hold inaccurate beliefs–we also investigate the effectiveness of a performance

[15]With this scoring rule, the employer's probability of winning a fixed payoff of $4.20 increases quadratically from zero to one as the difference between employer's prediction the worker's true score diminishes.

[16]Section 2.3 provides details on how workers were classified into different categories.

signal in the **Signal Intervention** experiment. After the initial 12 workers, employers were provided a signal for the test score of each of the remaining workers they made decisions. A worker's signal was their score in a random subset of 5 questions in the 15-question-test ("Random-5"; see Figure A.3), which is positively correlated with the overall performance of the workers in the Worker Pool. The employers were informed that the test questions selected for the "Random-5" were the same across all workers. We refer to the beliefs elicited after the introduction of the signal as employers' *posterior beliefs*, and we collected these beliefs for six separate workers. Importantly, because we collected both baseline and posterior beliefs for a given employer in the **Signal Intervention** experiment, we can measure the within-subject role of a signal intervention.

Following Bohren *et al.* (2023), in both of our experiments we included a hiring task in addition to eliciting beliefs on performance in order to differentiate statistical discrimination from taste-based discrimination in the aggregate. In this task employers were assigned an additional set of workers, and decided whether to hire each worker, one by one, by reporting the highest wage they were willing to pay to each worker. By hiring a worker, employers earned a revenue that increased linearly with the worker's test score.[17] The hired worker's wage was subtracted from this revenue.

In the **No Signal** experiment, employers submitted their WTP after observing the worker's CV and the audio clip for 12 workers. We refer to these decisions as the "Baseline WTP" given that no further information is given about the workers' performance. On the other hand, employers in the **Signal Intervention** experiment made hiring decisions for 6 workers. When making these decisions, the employers also received the "Random-5" signal. The reported WTPs are, therefore, the "WTP with signal." To incentivize truthful responses, the employer's payoff in this part increased in the worker's test score, and the wage for hiring was based on the Becker-DeGroot-Marschak mechanism (Becker *et al.*, 1964; Healy, 2017).[18] Note that the employers were informed that the hiring decisions were not hypothetical: if hired, the wage could be an additional bonus to the worker.

Employers in the **No Signal** experiment made decisions regarding a total of 24 workers, while the employers in the **Signal Intervention** performed an additional task for which

[17]Employers earned a revenue of 14 cents for each correct answer the worker answered correctly on the 15-question test.

[18]The payoff was calculated as follows: a wage between \$0 and \$2.10 was randomly drawn and compared to the employer's WTP. If the employer's WTP was at least as large as the randomly drawn wage, the employer hired the worker, earning a payoff of $\$2.10 + \$0.14 \times WorkerTestScore - RandomWage$. Otherwise, both the employer's and the worker's payoffs were \$0. The employer received the payoff as a bonus if the worker was chosen as the worker-that-counts for the employer's bonus, and the worker-that-counts got a bonus equal to their own payoff from the decision.

they evaluated 6 additional workers, giving them a total of 30 evaluated workers.[19] Each participant was paid a fixed amount of $4.50 for completing the study and received a bonus payment between $0 and $4.20. The bonus payment amount was determined based on the payoff of a randomly chosen decision.[20]

In the next section, we discuss the workers and the characteristics of employers.

2.3 Workers

There are 72 workers in the Worker Pool, with each worker belonging to one of 12 possible categories based on demographic and accent characteristics. The demographic characteristics are based on self-reported gender (male or female) and race/ethnicity (White, Asian, or Hispanic). A worker's status as accented or not is based on whether the speech pattern in their audio recording matches that of a native U.S. English speaker or not, as rated by an AI tool described in detail below.

Audio Clips

We used a state-of-the-art voice generator and transformation software that used AI technology to create the audio clips presented to employers.[21] The technology served two purposes. First, the software performed **speech-to-speech transformation** to voiceover the workers' original recordings using a pre-selected AI voice.[22] This allowed us to both anonymize worker recordings and also standardize voice characteristics, such as pitch and timbre, as much as possible. By doing so, we reduced potential confounding from unobservable vocal traits (e.g., vocal attractiveness and stereotypes; Zuckerman & Driver, 1989). Importantly, the speech-to-speech transformation preserved the non-voice characteristics of the worker's speech, namely pronunciation, intonation, and pace. Hence, the voice-transformed recording still conveyed the worker's accent. Second, used the tool to generate **text-to-speech** audio clips based on the written transcript of workers' original recordings and used the same AI voice by gender as in the voice-transformed clips. These text-to-speech audio clips preserved the content of the worker's recording and used the same AI voice, but masked the worker's accent with a neutral U.S. English accent. In our experiment, the speech-to-speech transformed recordings were used in the *Accent Revealed* treatment arm while the text-to-speech clips were used on the *Accent Blind* treatment arm.

[19]The additional task is related to the workers' willingness to accept.

[20]Data collection for the **No Signal** experiment took place in April 2025, while the **Signal Intervention** experiment included two waves (November 2024 and March 2025).

[21]https://elevenlabs.io/

[22]There were two separate voices, one for men and one for women.

We measured a worker's accent in their voice-transformed recording using English Language Speech Assistant (ELSA), software designed to help non-native speakers of English improve their accents.[23] ELSA evaluates audio clips by comparing their speech patterns to the speech patterns of native English speakers, scoring them between 0 (complete beginner) and 100 percent (completely native) based on how close they are. The overall score from 0 to 100 is based on sub-scores across various components of English speech: intonation, pronunciation, fluency, grammar, and vocabulary.[24,25]

To measure a worker's *Accent*, we collected the intonation and pronunciation scores of the speech-to-speech recordings in ELSA.[26] We then classified workers as accented if their ELSA score on intonation or pronunciation was less than 90%, and non-accented otherwise. See an example of a recording and classification in Section A.2.

Worker Pool

To create the Worker Pool sample from the 522 participants in the Worker Study, we first excluded workers who 1) did not self-report their gender as male or female, 2) were over 65 years of age, or 3) did not belong to one of our target race/ethnicity groups: White, Asian, or Hispanic. We further eliminated workers whose recordings were silent or low-quality.[27] Finally, for each of the 12 worker categories (e.g., Female-White-Not Accented), we selected 6 workers for the Worker Pool. These workers were chosen for each category based on their representative demographic characteristics such as education level and overall performance. Additionally, we selected workers keeping in mind the quality of the recordings in terms of perceived accent and overall sound quality (e.g., no background noise).

We selected 6 workers per worker category based on various trade-offs. To begin, a smaller number of workers per category allowed us to collect as many employer evaluations per worker as possible. Since one of the employer experiments had four incentivized decision parts (the three used in this study depicted in Figure 2.2), the lower limit on the number of workers per category was four. Next, workers might be perceived differently from each other for reasons that might be difficult to quantify. For example, while we made our best effort to homogenize the workers' recordings, voice-transformed recordings still have slight

[23] https://elsaspeak.com/en/

[24] These are common components utilized in measuring English proficiency. For example, see the rubric for the TOEFL-Speaking Section: https://www.ets.org/pdfs/toefl/toefl-ibt-speaking-rubrics.pdf

[25] ELSA also gives some approximated grades on various international tests.

[26] We excluded other categories because they are controlled factors in our experimental design: fluency was controlled through minimal variation in the length of the audio clips, while vocabulary and grammar were standardized due to workers reading a paragraph we provide them. Thus, the recordings do not reflect the workers' language ability in these categories.

[27] When possible, the workers' voice recordings were edited for quality (e.g., remove empty spaces).

variations (e.g., due to the speaker's pace) across workers. By holding workers constant, for which having fewer workers is paramount, we were able to reduce the effect of underlying characteristics. Lastly, to have more variation in the workers presented to the employers, we preferred a number larger than 4 workers per category.

Table 2.1 presents the characteristics of the Worker Pool by race/ethnicity. Column (1) presents the characteristics of the complete sample, followed by the self-identified characteristics of Asians (column (2)), Hispanics (column (3)), and Whites (column (4)). On average, our sample of workers is 35 years old and more than half have a college degree or higher. Asians are 15 percentage points (hereafter, *pp*) more likely to have a college degree than the average worker, while the least likely to have college or higher educational attainment are Hispanics (-0.10 *pp*) followed by Whites (-0.06 *pp*).

Table 2.1: Worker Pool Descriptive Statistics

	(1) All	(2) Asian	(3) Hispanic	(4) White
Age	35.14	30.92	34.38	40.13
	(10.35)	(9.83)	(9.82)	(9.61)
College or More	0.60	0.75	0.50	0.54
	(0.49)	(0.44)	(0.51)	(0.51)
Math Score	7.07	9.33	5.75	6.13
	(3.53)	(3.27)	(3.33)	(2.94)
Verbal Score	9.90	10.50	9.79	9.42
	(1.84)	(2.00)	(1.79)	(1.61)
N	72	24	24	24

Note. This table presents means and standard deviations (in parentheses) of characteristics for the sample of workers. The columns refer to the worker's self-identified race/ethnicity. "College or More" is an indicator variable that takes a value of 1 if the worker reported educational attainment is a Bachelor's degree in college (4-year) or higher, and 0 otherwise. Math score and Verbal score represent the number of questions a worker answered correctly out of 15 in the Math and Verbal tests, respectively.

The average scores of the Worker Pool workers are 7.1 out of 15 (or 47% of the total score) for Mathematics and 9.9 (66%) for Verbal tests. Relative to the Worker Pool, Asians performed better on the Mathematics test, 32% higher than the average, while Hispanics and Whites performed worse (-19% and -13%). We observe a similar pattern in the Verbal test, where Asians' scores are, on average, 6% higher than the average, while the scores of Hispanics and Whites are lower by -1% and -5%, respectively. Table A.1 in the appendix further breaks down these summary statistics with respect to accent.

Table 2.2: Relationship Between Accent and Test Scores

	(1) All	(2) Asian	(3) Hispanic	(4) White
DV: Math Score				
Accent	-0.917	-0.167	-1.833	-0.750
	(0.831)	(1.363)	(1.332)	(1.216)
DV: Verbal Score				
Accent	-0.806*	-0.833	-0.250	-1.333**
	(0.426)	(0.816)	(0.747)	(0.610)
N	72	24	24	24

Note. This table shows the correlation between accent and test scores in the Worker Pool. * p<0.10, ** p<0.05, *** p<0.01.

We further explore whether there is an accent score gap by estimating the relationship between accent status and skill test score in Table 2.2. Column (1) presents the accent gap for the complete sample, followed by the results for race/ethnicity groups (columns (2)–(4)). Overall, there is qualitative evidence of an accent gap in performance, with the strongest evidence in Verbal skills (-0.806, significant at the 10% level). These gaps are also present across race/ethnicity categories, but are not statistically significant, possibly due to the small number of workers per category.

Across skill tests and race/ethnicity, Asians have the smallest accent gap in Mathematics skills, while Hispanics have the smallest accent gap in Verbal skills. Hispanics have the largest accent gap in Mathematics, while Whites have the largest for Verbal skills.

2.4 Employers

Table 2.3 presents the descriptive statistics of the employers. Columns (1) and (2) show summary statistics for the **No Signal** experiment, while columns (3) and (4) are the summary statistics for the **Signal Intervention** experiment. For each experiment, columns (1) and (3) present statistics for the employers assigned to the *AccentBlind* treatment arms, while columns (2) and (4) present summary statistics of those assigned to the *AccentRevealed* treatments.

Table 2.3: Employer Pool Description

	(1)	(2)	(3)	(4)
	No Signal Experiment		**Signal Intervention Experiment**	
	AccentBlind	*AccentRevealed*	*AccentBlind*	*AccentRevealed*
Age	40.07	38.79	37.75	37.94
	(13.21)	(12.10)	(12.34)	(12.06)
Female	0.49	0.50	0.46	0.46
	(0.50)	(0.50)	(0.50)	(0.50)
Asian	0.05	0.06	0.12	0.10
	(0.21)	(0.24)	(0.33)	(0.30)
Hispanic	0.05	0.05	0.07	0.09
	(0.23)	(0.21)	(0.25)	(0.28)
Black	0.24	0.22	0.19	0.17
	(0.43)	(0.41)	(0.39)	(0.38)
White	0.67	0.69	0.67	0.69
	(0.47)	(0.46)	(0.47)	(0.46)
College or More	0.72	0.68	0.63	0.57
	(0.45)	(0.47)	(0.48)	(0.50)
Reside in Large City	0.35	0.36	0.36	0.34
	(0.48)	(0.48)	(0.48)	(0.47)
Multilingual	0.21	0.21	0.18	0.16
	(0.41)	(0.41)	(0.39)	(0.37)
Employed Full-Time	0.58	0.57	0.59	0.53
	(0.49)	(0.50)	(0.49)	(0.50)
N	445	438	292	297

Note. This table presents means and standard deviations (in parentheses) of characteristics of the sample of employers. Race/ethnicity categories are not mutually exclusive. College or More equals 1 if the employer educational attainment is a Bachelor's degree in college (4-year) or higher and 0 otherwise. Reside in Large City is equal 1 if the employer resides in a city with a population above 250,000, and 0 otherwise.

The average age of employers in the sample is 39 years old, and slightly less than half are female. About 70% of the employers are White, and the majority of employers have a college degree or higher education. Reassuringly, the summary statistics show that employers are comparable across the information treatment groups within each experiment. Therefore, in terms of observable characteristics, we have balanced samples across the accent information treatment arms. Employer characteristics across experiments are similar too, enhancing comparability across the two experiments.

3 Empirical Strategy

The central objective of this paper is to examine whether employers' beliefs and decisions are influenced by a worker's accent. The data consist of employer evaluations at the employer–worker pair level. As outlined in Section 2.2, the study elicited two primary types of employer decisions: beliefs about workers' performance and employers' WTP. These outcomes provide the basis for addressing the research questions detailed below.

3.1 Does a worker's foreign accent affect employer assessment of that worker's performance?

The first question we address is whether employers display accent discrimination in assessing worker ability. We answer this question by first establishing employers' baseline beliefs, as defined below. Then, we explore how these beliefs respond to receiving information on an individual worker's performance.

Baseline beliefs. To establish the baseline beliefs of the employers, we use the first 12 decisions of each employer in both the **No Signal** and **Signal Intervention** experiments and estimate:

$$y_{ie} = \beta_0 + \beta_1 (Accent \times AccRev)_{ie} + \beta_2 Accent_i + \beta_3 AccRev_e + \mathbf{X}'_{ie}\Gamma + \epsilon_{ie}, \tag{1}$$

where the outcome variable y_{ie} refers to the employer e's belief about the of number of questions that worker i answered correctly out of 15. $\mathbf{X}'_{ie}$ is a vector of control variables that includes the worker's gender, age, education, and race and ethnicity: Asian and Hispanic. $Accent_i$ is an indicator variable for whether the worker has an accent of a non-native U.S. English speaker, and $AccRev_e$ is an indicator variable for whether the employer is in the *Accent Revealed* treatment group. Our treatment effect coefficient of interest is β_1, which captures the effect of revealing a worker's accent status to employers on the gap in evaluation between workers who have an accent and those who do not. Hereafter, we refer to this coefficient as the accent gap for brevity.

A separate question of interest is whether employer accent bias in evaluating worker ability is based on accurate or inaccurate perceptions. By controlling for actual test performance, we test whether differences in the perception of worker ability are attenuated. In practice, this means estimating an augmented version of Equation 1 that includes a control for the worker's actual test score, $Score_i$:

$$y_{ie} = \beta_0 + \beta_1(Accent \times AccRev)_{ie} + \beta_2 Accent_i + \beta_3 AccRev_e + \mathbf{X}'_{ie}\Gamma + \sigma Score_i + \epsilon_{ie}. \quad (2)$$

If the magnitude of β_1 is significantly attenuated in Equation 2 in comparison to Equation 1, this would indicate that when the worker's accent status is revealed to employers, employers' beliefs in differences in performance between accented and non-accented workers are accurate, largely reflecting actual differences in performance between the two groups. On the other hand, if the magnitude of β_1 is similar or larger in Equation 2 compared to Equation 1, this indicates employers' beliefs do not reflect actual differences in performance between the two groups.

Response of beliefs to a performance signal. To disentangle the various factors contributing to baseline accent gaps in assessments estimated in Equation 1, we assess the extent to which employers update their beliefs about worker ability when given a signal about a worker's performance (the "Random-5" signal, see Section 2.2). Additionally, we investigate whether belief updating varies depending on whether an accent is present or not. To answer these questions, we utilize the observations from the **Signal Intervention** experiment, comparing beliefs reported in the first 12 decisions ($Signal_{ie} = 0$) to those reported in decisions 13 - 18 ($Signal_{ie} = 1$), where $Signal_{ie}$ is an indicator variable equal to 1 if the employer was shown the performance signal ("Random-5"):

$$\begin{aligned} y_{ie} = {} & \delta_0 + \delta_1(Accent \times AccRev \times Signal)_{ie} + \delta_2(Accent \times AccRev)_{ie} + \delta_3(Accent \times Signal)_{ie} + \\ & + \delta_4(AccRev \times Signal)_{ie} + \delta_5 Accent_i + \delta_6 AccRev_e + \delta_7 Signal_{ie} + \mathbf{X}'_{ie}\Gamma + \epsilon_{ie}, \quad (3) \end{aligned}$$

The outcome variable and other right-hand-side variables are as in Equation 1. The coefficient of interest in this specification, δ_1, captures the effect of providing a performance signal on employers' assessments of workers with accents relative to their non-accented counterparts, when the accent is revealed. δ_7 also provide insights on how employers' overall workers' assessments respond to the informative signal.

3.2 Do employer beliefs related to foreign accent result in wage differences?

From a policy perspective, a critical question is whether employer beliefs translate into differences in worker compensation. Using employer decisions in the hiring task, we estimate:

$$WTP_{ie} = \phi_0 + \phi_1(Accent \times AccRev)_{ie} + \phi_2 Accent_i + \phi_3 AccRev_e + \mathbf{X}'_{ie}\Gamma + \epsilon_{ie}, \tag{4}$$

where the outcome variable, WTP_{ie}, is employer e's WTP to hire worker i. The rest of the variables are the same as in Equation 1. Our coefficient of interest, ϕ_1, measures the effect a worker having a foreign accent on the employer's WTP to hire the worker, relative to a non-accented counterpart with similar characteristics. We estimate this equation separately for the **No Signal** experiment (decisions 13-24) and the **Signal Intervention** experiment (decisions 19-24) samples because in the latter, employers had access to the Random-5 signal while making their decisions.

Results from the **No Signal** experiment gives us differences in the baseline WTP, which might be due to statistical (i.e., belief-based) discrimination by employers, as well as taste-based discrimination (Becker, 1957).[28] To investigate the role of taste-based discrimination in accent-based WTP gaps, we use employer decisions in the belief elicitation task (decisions 1-12) to generate a new outcome variable on implied willigness to pay, $IWTP_{ie}$. This variable is calculated as 14 cents times employer e's belief about the of number of questions that worker i answered correctly out of 15,[29] and estimates the WTP that is implied by the employer beliefs:

$$IWTP_{ie} = \zeta_0 + \zeta_1(Accent \times AccRev)_{ie} + \zeta_2 Accent_i + \zeta_3 AccRev_e + \mathbf{X}'_{ie}\Gamma + \epsilon_{ie}, \tag{5}$$

A comparison between ζ_1 and ϕ_1 in Equation 4 overall as well as across different worker types allows us to interpret the extent that actual WTP gaps can be explained with taste-based theories of discrimination against accented workers. Specifically, ζ_1 being smaller than ϕ_1 in absolute value would be consistent with taste-based discrimination against accented workers.

Since hiring decisions are made after observing the worker's performance signal in the **Signal Intervention** experiment, we construct $IWTP_{ie}$ in Equation 5 utilizing correspond-

[28] This theory predicts the decision maker to give up on profits due to animus towards a group.

[29] In this task, the employer's revenue was 14 cents times the worker's test score.

ing decisions in the belief elicitation task (decisions 13-18). The results from this specific experiment test to what extent WTP gaps are related to belief-based differences about worker productivity, this time in the presence of a signal aims to correct inaccurate beliefs.

4 Empirical Analysis

4.1 Main Results

4.1.1 Does a worker's foreign accent affect employer assessment of that worker's performance?

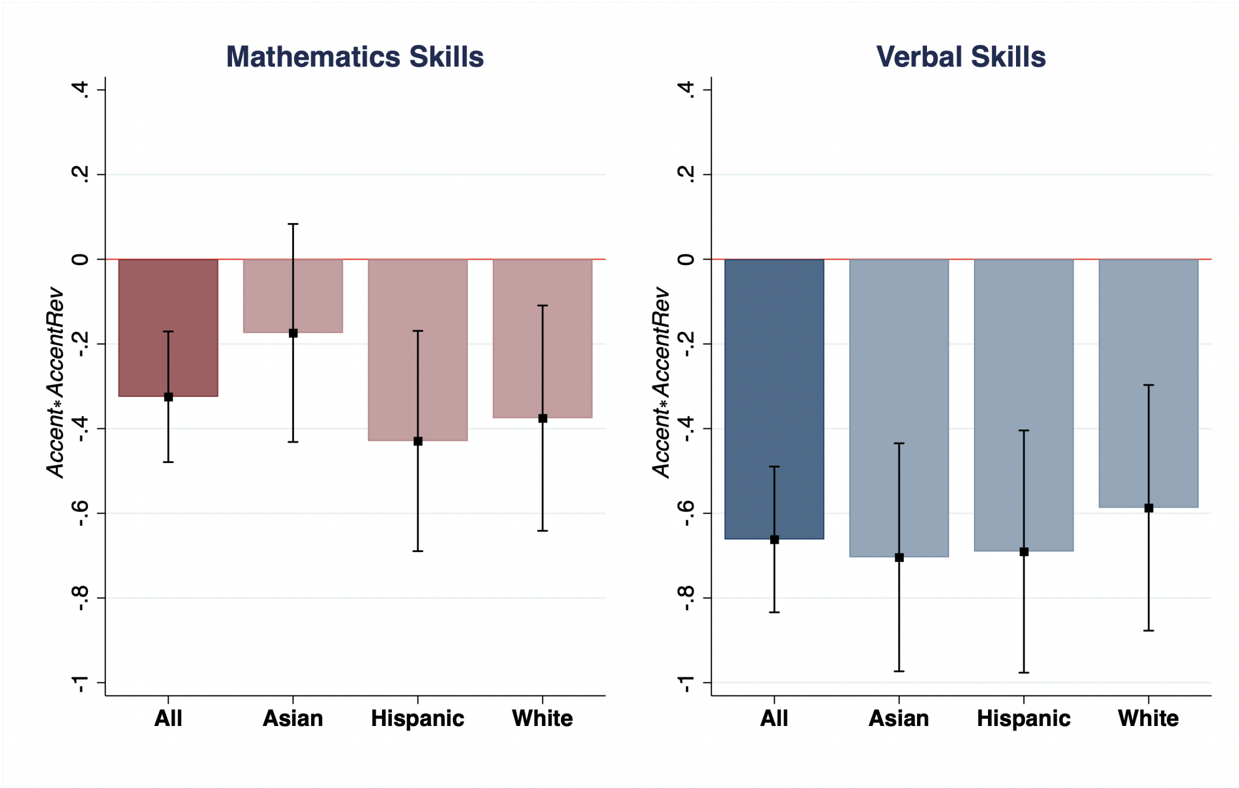
Baseline beliefs. Figure 4.1 displays the results for our variable of interest from estimating Equation 1, the estimated number of correct answers. The first subfigure details the estimation results for Mathematics skills, while the second displays the same for Verbal skills. The bars are organized to sequentially present the results for the entire sample, followed by a breakdown of results for Asian, Hispanic, and White workers.

We first note that employers assigned to the Mathematics skills test are overly optimistic about workers' performance. Their average belief is 8.40 points, exceeding the Worker Pool average score of 7.1 points. By contrast, employers assigned to the Verbal skills test hold slightly pessimistic beliefs, with an average of 9.43 points compared to the Worker Pool average of 9.9 points.

Despite the miscalibrated average beliefs of employers, we find that when a worker's accent is revealed, accented workers are believed to perform worse on both skills. Specifically, employers evaluate these workers' test scores to be 0.33 points less on the Mathematics test and 0.66 points less on the Verbal test. These effects represent 4.6% of the average score for Mathematics and 6.7% of the average score for Verbal tests. Comparing across skills, we note that the estimated accent gap for the Verbal skills is twice that of the Mathematical skills. As shown in Figure 4.1, these coefficients are qualitatively different from each other. Notably, these differences in baseline beliefs exist even when the gaps in our Worker Pool are very similar across skills (see Table 2.2).

Figure 4.1 shows that, relative to their other workers of the same race/ethnicity, when the worker's accent is revealed, employers believe that accented Asians perform worse in the Verbal test by -0.70 points. Accented White and Hispanic workers are perceived as performing worse in both tests, with the gaps being slightly larger for Hispanics. In both cases, the accent gaps for the Verbal skills are about 1.5 to 2 times that of the Mathematics skills test.

Figure 4.1: Baseline Beliefs over Worker Performance



Notes: The figure displays the estimated coefficients and 95% confidence intervals of $Accent \times AccRev$ in equation (1). The estimations include evaluations 1-12 from both experiments. The following covariates are included in the estimations but are excluded from the figure for brevity: *Accent*, *AccRev*, age, education, race and ethnicity: Asian and Hispanic (the reference group is White), and a year and an experiment dummy. Bars: All include the sample of the employers' evaluations 1-12 for all workers; Asian, Hispanic and White include the subsample of the employer evaluations 1-12 for which workers were Asian, Hispanic and White, respectively. For each skill, the information that follows is presented in the following order (All, Asian, Hispanic, White). Average beliefs: B_{Math} =(8.40; 9.45; 7.6; 8.07), B_{Verbal} =(9.43; 9.87; 8.95; 9.45). Observations: N_{Math} =(9,336; 3,112; 3,112; 3,112), N_{Verbal} =(8,328; 2,776; 2,776; 2,776).

To explore the accuracy of employer beliefs, we estimate prior beliefs conditional on test scores (Equation 2). If the accent gap stems from accurate beliefs, controlling for the worker test score should then reduce the estimated accent gap. Estimates of Equation 2 in Table A.2–Table A.3 in the appendix indicate that controlling for worker test score does not meaningfully reduce the accent gap in employer assessments in either skill.

Our results suggest that employers might inaccurately use worker accent as a signal of worker ability. Interestingly, this accent gap is present even when very little English knowledge would be needed, such as in the Mathematics skills test. That said, overall, we find the largest effect for Verbal skills, indicating that employers use the presence of a foreign accent as a signal of English proficiency, even when the *Accent Revealed* audio clip is based on a standardized paragraph.

Response of beliefs to a performance signal. Figure 4.2 shows the results of estimating Equation 3, which explores the role of a performance signal on employers' beliefs about worker performance. Therefore, the figure includes only the beliefs reported at the **Signal Intervention** experiment. We find that the presence of a performance signal updates employers' overall beliefs about worker performance. Specifically, when the "Random-5" signal is available, employers lower the number of questions they believe the worker answered correctly by -1.39 and -1.63 for Mathematics and Verbal skills tests, respectively. We observe qualitatively similar effects across all race/ethnicity groups. These results suggest that without a signal, employers overestimate the performance of workers but adjust their priors after receiving the performance signal.[30]

Moreover, in the Accent Revealed treatment arm, we do not find a statistically significant estimated effect of the signal on the differences in beliefs over performance due to accent. This suggests qualitatively that the accent gap estimated in Figure 4.1 still persists after controlling for the presence of the signal. However, it is worth first noting that the signs that most of the $Accent \times AccRev \times Signal$ coefficients are positive.[31] This sign hints that employers incorporate the signal when updating their beliefs by accent. Secondly, while none of the $Accent \times AccRev \times Signal$ coefficients are significant, the role of the signal in the accent gap seems to vary with skill. Put precisely, Figure 4.2 shows that the coefficient is about 40% of the accent gap coefficient ($Accent \times AccRev$) for Mathematics, while the same relationship is only 4% for Verbal skills. Thus, while not significant, the results for the Mathematics skills hint that an informative signal about worker performance may reduce the accent gap in employer beliefs about worker performance.

Furthermore, we can interpret our results from Figure 4.1 as employers inaccurately using worker accent as a signal of worker ability, since employer beliefs are largely uncorrelated with actual worker performance (Table A.2–Table A.3). Because the performance signal aims to reduce inaccurate beliefs, and because the same signal is correlated with a worker's performance, employer posterior beliefs should be positively correlated with actual worker performance. We show that this is true in Tables A.4–A.5, the estimation results for the augmented Equation 3, to include test score controls. The results are qualitatively robust to including test score controls.[32]

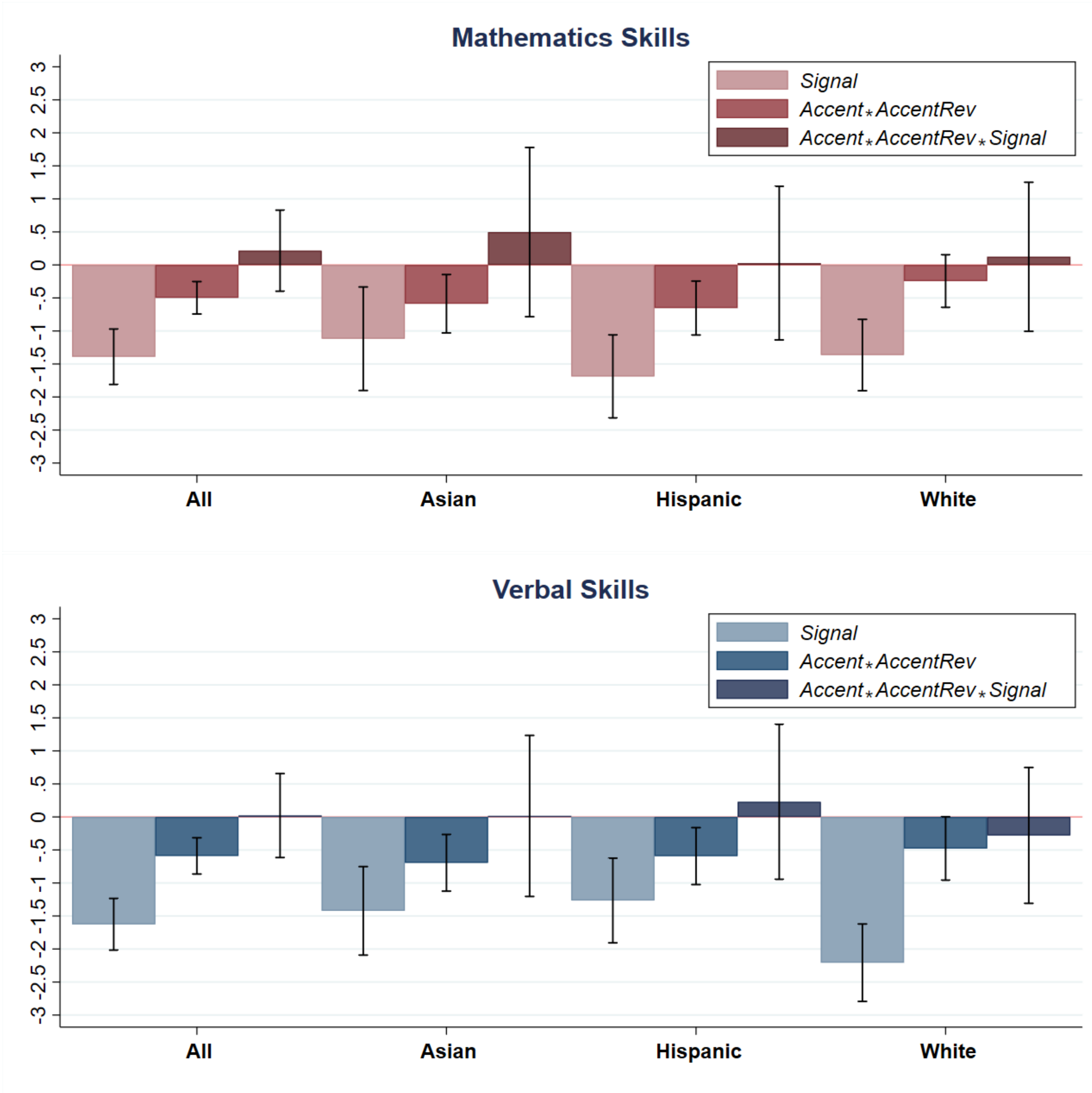
The results thus far indicate that when worker accent is revealed, employers believe ac-

[30] Notably, there is no differential effect of the signal by treatment arm. Results are excluded for brevity.

[31] The coefficient is only negative for the subsample of White workers in the Verbal skill test.

[32] On average, the employers are correctly using the signals. Precisely, the signal represents a third (0.30) of the test score, and the 6 out of 18 beliefs evaluations have the signal available, or 0.33 of the sample. Therefore, the relationship between $TestScore$ and y_{ie} should be $0.3 \times 0.33 \approx 0.10$. The estimated coefficients for $TestScore$ in Tables A.4–A.5 qualitatively align with this approximate effect.

Figure 4.2: Beliefs over Worker Performance: Response to Informative Signal



Notes: The figure displays the estimated coefficients and 95% confidence intervals in equation (3). The estimations include evaluations 1-18 from the Signal Intervention experiment. The following covariates are included in the estimations but are excluded from the figure for brevity: $Accent$, $AccRev$, $AccRev \times Signal$, $Accent \times Signal$, age, education, race and ethnicity: Asian and Hispanic (the reference group is White), and a year dummy. For each skill, the information that follows is presented in the following order (All, Asian, Hispanic, White). Average beliefs: B_{Math} =(8.15; 9.27; 7.49; 7.72), B_{Verbal} =(9.33; 9.81; 8.91; 9.28). Observations: N_{Math} =(5,310; 1,733; 1,778; 1,799), N_{Verbal} =(5,292; 1,739; 1,783; 1,770).

cented workers perform worse than their non-accented counterparts across all skills. These differences are largely driven by inaccurate beliefs, particularly in Verbal skills, as the accent gap remains even after controlling for workers' test scores. Importantly, the gap persists even when employers are provided with a signal of workers' performance. While not statistically significant, the evidence suggests that the signal modestly adjusts employer beliefs in Mathematics, implying that informative signals may mitigate accent penalties in skills where accent might be less relevant to the task. In contrast, for Verbal skills the signal is less effective. While accents convey salient information about ability, employers interpret information from accents inaccurately, believing that accents are a more negative signal of ability than they actually are. These differences across skills are striking, given the strength of the performance signal. In the next section, we examine whether this bias extends to differential wage payments by worker accent.

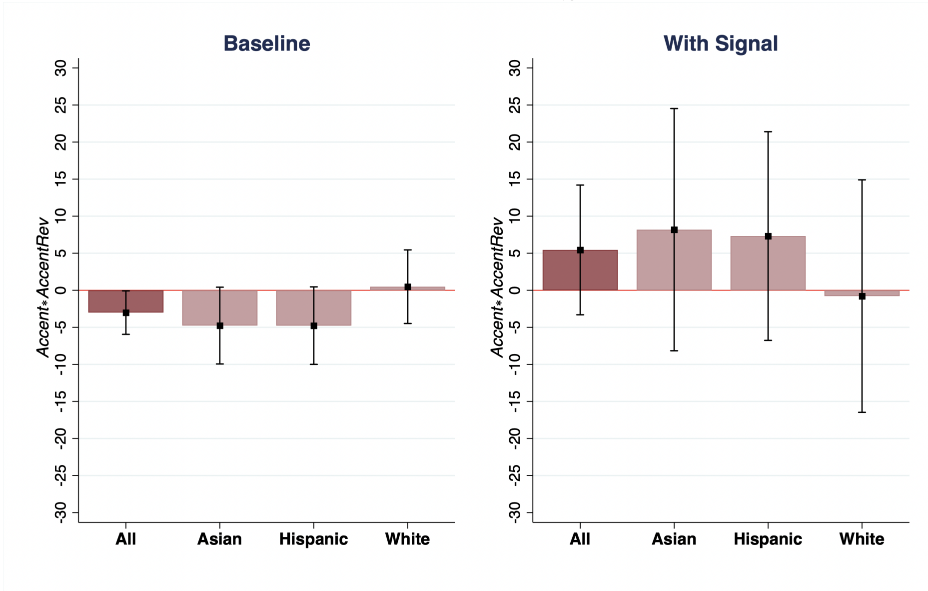
4.1.2 Do employer beliefs related to foreign accent result in wage differences?

Figure 4.3 presents the results for the WTP of the employer to hire accented workers compared to non-accented workers, when accent is revealed, estimated by Equation 4. Panel A presents the results for the Mathematics skills and Panel B presents the results for the Verbal skills. Within each panel, the first subfigure shows the results in the absence of a performance signal (**No Signal** experiment), and the second subfigure shows WTP in the presence of a performance signal (**Signal Intervention** experiment). The notes include the average WTP across the different samples.

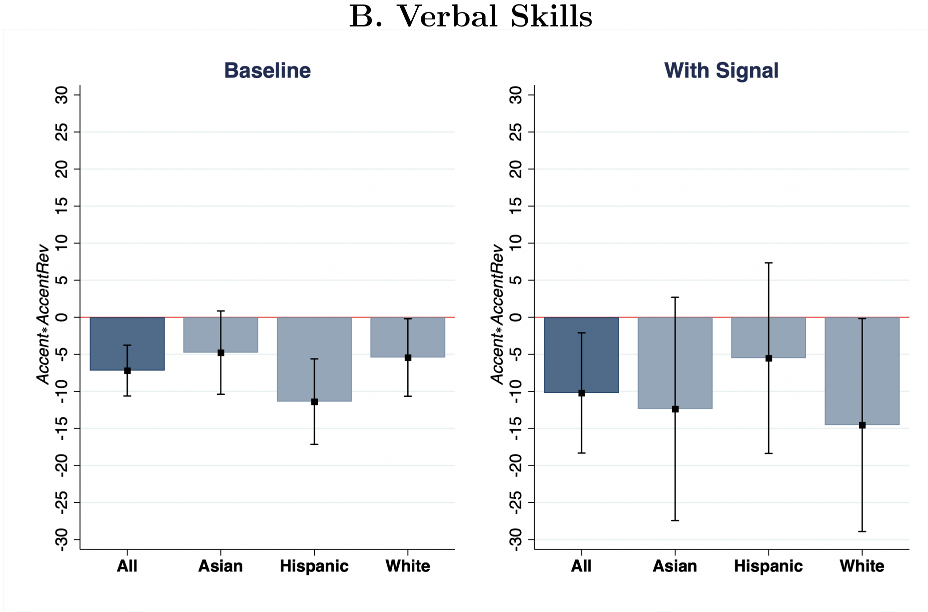
Baseline WTP. The statistics reported in the Figure 4.3 notes indicate that in the **No Signal** experiment, employers' WTP is consistent with the average test score being lower for Mathematics than for Verbal skills: Employers' average WTP for hiring workers for their performance in the Mathematics test is 85.15 cents, while for the Verbal test, this number is 90.89 cents. The average WTP is also consistent with racial/ethnic group stereotypes. That is, the average WTP is highest for Asians in both skills tests, while it is lowest for Hispanic workers. Moreover, when accents are revealed to employers, employers report a lower WTP for accented workers compared to non-accented counterparts across all subjects, relative to their assessments of workers when their accent is blind. Specifically, employers are willing to pay accented workers 3.0 cents less for their performance in the Mathematics test and 7.2 cents less in the Verbal test compared to workers who are otherwise observationally identical but who do not have an accent. Coefficient estimates are consistently negative across each different race/ethnicity category, with the exception of the Mathematics-White coefficient, which is zero and not significant. Comparing across skills, the differences between the WTP for Mathematics and Verbal skills are in line with beliefs in Figure 4.1. Precisely, the baseline

Figure 4.3: Willingness to Pay (in cents)

A. Mathematics Skills



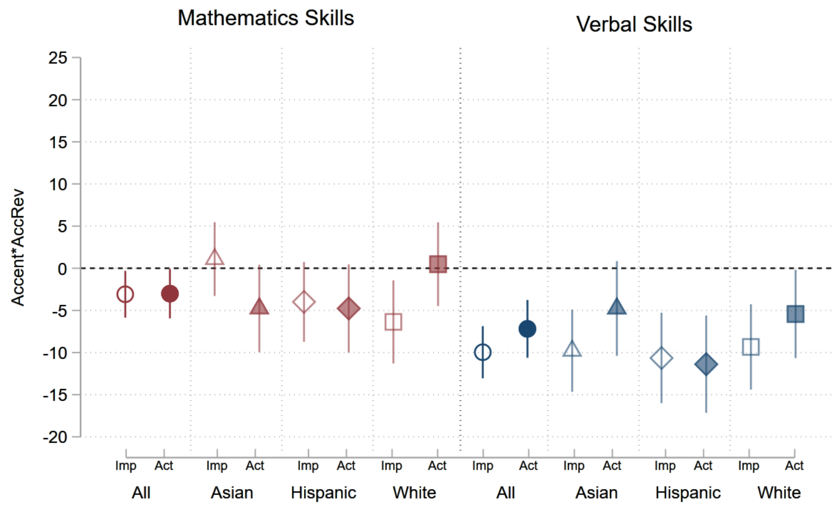
B. Verbal Skills



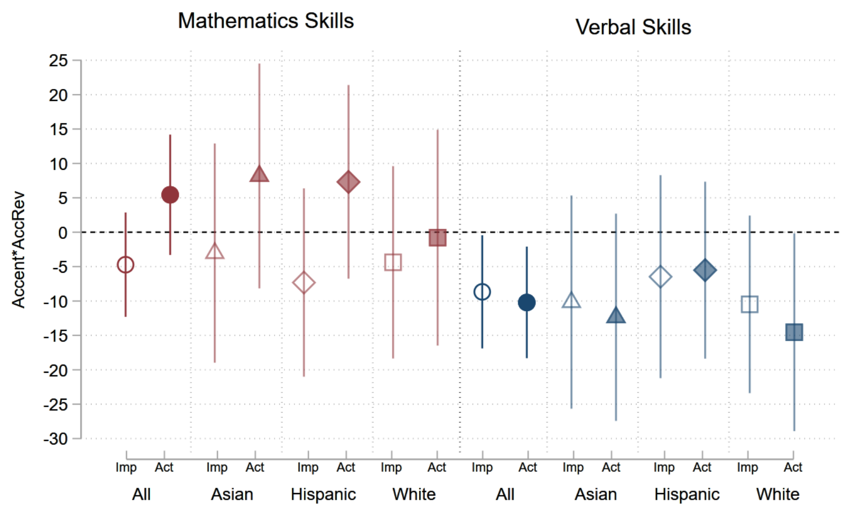
Notes: Estimated coefficients and 95% confidence intervals of *Accent* × *AccRev* in equation (4). For each panel, the first graph shows the results from the No Signal experiment (evaluations 13-24), while the second shows the results from the Signal Intervention experiment (evaluations 19-24). Covariates included in the estimations: *Accent*, *AccRev*, age, education, and race and ethnicity: Asian and Hispanic (the reference group is White). For the Signal Intervention experiment result, we also include a year dummy and the value of the signal (Random-5). For each skill, the information that follows is presented in the following order (All, Asian, Hispanic, White). Average WTP for Mathematics: $WTP_{NoSignal}$=(85.15; 93.82; 78.51; 83.13); WTP_{Signal}=(62.44; 78.98; 52.71; 55.18), Verbal: $WTP_{NoSignal}$=(90.89; 95.72; 85.86; 91.10); WTP_{Signal}=(63.62; 76.69; 56.25; 58.50). Observations Mathematics: $N_{NoSignal}$ =(5,796; 1,932; 1,932; 1,932), N_{Signal} =(1,770; 608; 657; 505); Verbal: $N_{NoSignal}$ =(4,800; 1,600; 1,600; 1,600), N_{Signal} =(1,764; 581; 682; 501).

Figure 4.4: Willingness to Pay (in cents), Implied by Beliefs vs. Actual

A. No Signal Experiment



B. Signal Intervention Experiment



Notes: Estimated coefficients and 95% confidence intervals of $Accent \times AccRev$ in equations (4) and (5). The top panel shows the estimation results from the No Signal experiment, where the hollow markers indicate the estimated WTP using evaluations 1-12 (Beliefs Implied WTP), while the filled markers indicate the estimated WTP using evaluations 13-24 (Actual WTP). The bottom panel shows the estimation results from the Signal Intervention experiment, where hollow markers indicate the estimated WTP using evaluations 13-18, where the Random-5 signal is present on the worker's profile (Beliefs Implied WTP), while the filled markers indicate the estimated WTP using evaluations 19-24, where the Random-5 signal is present on the worker's profile (Actual WTP). Observations Mathematics: $N_{NoSignal}$ =(5,796; 1,932; 1,932; 1,932), N_{Signal} =(1,770; 608; 657; 505); Verbal: $N_{NoSignal}$ =(4,800; 1,600; 1,600; 1,600), N_{Signal} =(1,764; 581; 682; 501).

WTP accent gap for Mathematics skills in Figure 4.3 is 2 times that of Verbal skills.[33]

The estimated accent gaps in WTP are commensurate with the implied WTP gaps based on the employers' beliefs, as demonstrated in Figure 4.4.[34] Specifically, the average accent gap in WTP implied by beliefs is approximately 3.1 cents for Mathematics skills and 10.0 cents for Verbal skills. These amounts are not statistically different than the average actual WTP gaps for each skill (3.0 cents for Mathematics and and 7.2 cents for Verbal skills). When analyzed separately across racial and ethnic groups, implied average average accent gaps in WTP for Mathematics skills are -1.1 cents for Asian workers, 4.0 cents for Hispanic workers and 6.3 cents for White workers. For Verbal skills, the implied gaps are 9.8 cents for Asian workers, 10.1 cents for Hispanic workers and 9.4 cents for White workers.

WTP in the presence of a performance signal. Figure 4.3 also shows the accent gap from the sample of employers who saw the signal of the workers' performance (**Signal Intervention** experiment). The results are in line with the beliefs, the performance signal affects the employers' WTP. This can be seen through the average WTP, which lowers when the signal is available in comparison to baseline across skills and race/ethnicity. Moreover, Figure 4.3 shows that the impact of a performance signal in the accent-based WTP gap depends on the skills test. For the Mathematics skills test, we do not find evidence that employers differ in their WTP between accented and non-accented workers. Notably, the sign for WTP in the presence of a signal in Figure 4.3 is positive for Mathematics, although the results are imprecise and cannot reject values such as -3.0, the accent-based WTP gap in the **No Signal** experiment.

In contrast, for the Verbal skills test we find that the accent-based WTP gap persists in the presence of the performance signal: Employers are willing to pay accented workers 10.2 cents less for their performance in the Verbal skills test. This WTP gap with the signal is qualitatively similar to that in the **No Signal** experiment, thus suggesting that the performance signal does not eliminate the accent-based WTP gap for Verbal skills. Moreover, the effect of the signal being different across skills tests is in line with our findings regarding employer beliefs on performance in Figure 4.2. While the performance signal has very little impact in the Verbal skills test on moderating employer beliefs about worker performance based on their accent, it is impactful in the Mathematics skills test.[35]

---

[33] We augment Equation 4 to include test score controls and find that a worker's test score is largely uncorrelated with the employer's reported baseline WTP. Results are displayed in Table A.6 in the appendix. Our findings are consistent with earlier findings in Tables A.2–A.3, which show test scores to be uncorrelated with employer beliefs about worker productivity.

[34] Estimation results are provided in Table A.8 and Table A.9.

[35] Accent×AccentRev×Signal coefficient is 40% and 4% of the Accent×AccentRev coefficient in Mathematics and Verbal skills, as demonstrated in Figure 4.2.

Broadly speaking, the results when the signal is available are estimated less precisely, as the standard errors more than double. While this difference could be due to differences in sample sizes, it could also be attributed to the signal creating more variation in WTP because more information is available to the employers: When the signal is unavailable, comparable workers are likely treated the same by employers. However, employers might treat otherwise comparable workers differently if their signal value is different (e.g., 0 vs 4), thus increasing the variation in their WTP.

We note that the effect of the signal on WTP is asymmetric across skills: it reduces the wage accent penalty for Mathematics but not for Verbal skills. This pattern is in line with our earlier result on the asymmetric effects of the signal on beliefs across the two skills and strengthens the case for the accent being perceived as a more informative signal than the one we provide (i.e. "Random-5") for Verbal skills.

High vs. Low Signal Values. If employers are utilizing the signals, a high signal value may lead the employer to infer the worker's test score to be high. The employer's WTP response should vary between skills. For Mathematics, the WTP with signal in Figure 4.3 suggests that the employers respond strongly to a high-type worker. Because the accent gap is in the opposite direction, these results may stem from the employer overcompensating by offering more to accented workers. In contrast, because we argue that the accent is a negative signal of English proficiency, we would expect that being a high-type worker would have no or even a negative effect on the employer WTP. To explore this possibility, we estimate:

$$\begin{aligned} WTP_{ie} = \alpha_0 + \alpha_1(Accent \times AccRev \times HighType)_{ie} + \alpha_2(Accent \times AccRev)_{ie} + \\ + \alpha_3(Accent \times HighType)_{ie} + \alpha_4(AccRev \times HighType)_{ie} + \alpha_5 Accent_i + \\ + \alpha_6 AccRev_e + \alpha_7 HighType_{ie} + \mathbf{X}'_{ie}\Gamma + \epsilon_{ie} \end{aligned} \tag{6}$$

where $HighType$ equals one if the worker's "Random-5" score is three or more, and zero otherwise. The coefficient of interest, α_1, displays the effect of being a $HighType$ on the pay accent gap. The results are in Table 4.1, showing that, first, employers are willing to pay more for workers who have a relatively high signal (i.e., a $HighType$). At the same time, employers have a significantly larger WTP for an accented $HighType$ worker in the Mathematics test when the accent is revealed. While not significant, the $Accent \times AccRev$ coefficient for Math in Table 4.1 is negative, suggesting that the positive effect in the presence of a signal, as seen in Figure 4.3, is due to the signal. Interestingly, the accented worker being a $HighType$ does not seem to have similar effects for Verbal skills. These results are in line with the interpretation that when the employer hears a worker's accent, the accent becomes the dominant signal of workers' English proficiency.

Table 4.1: Willingness to Pay (in cents) after Signal Information, Interaction with Skill Level

	(1) All	(2) Asian	(3) Hispanic	(4) White
Panel A. Mathematics Skills				
Accent × *AccRev* × *HighType*	20.338**	-5.347	28.849**	35.884*
	(8.323)	(17.223)	(13.866)	(18.661)
Accent × *AccRev*	-4.414	13.261	-3.931	-13.578
	(5.623)	(13.081)	(10.292)	(10.266)
HighType	27.739***	27.070***	24.505***	34.694***
	(4.799)	(10.310)	(8.834)	(9.148)
Average WTP	62.44	78.98	52.71	55.18
N	1,770	608	657	505
Panel B. Verbal Skills				
Accent × *AccRev* × *HighType*	-0.643	4.280	-8.820	6.693
	(7.127)	(13.961)	(14.032)	(14.045)
Accent × *AccRev*	-10.062*	-14.002	-1.551	-19.562**
	(5.248)	(10.614)	(8.228)	(8.432)
HighType	24.651***	29.690***	26.253***	16.780*
	(4.827)	(9.545)	(8.057)	(8.616)
Average WTP	63.62	76.69	56.25	58.50
N	1,764	581	682	501

Note. The table displays the estimated coefficients in equation (4), using the sample from the Signal Intervention experiment (evaluations 19-24). The following covariates are included in the estimations but are excluded from the table for brevity: *Accent*, *AccRev*, *AccRev* × *HighType*, *Accent* × *HighType*, age, education, race and ethnicity: Asian and Hispanic (the reference group is White), and a year dummy. Columns: All include the sample of the employer evaluations 19-24 for all workers; Asian, Hispanic and White include the subsample of the employer evaluations 19-24 that were Asian, Hispanic and White workers, respectively. Standard errors clustered at the employer level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Comparing the effect of the signal between skill tests in Table 4.1 also gives us some insight into how powerful accents might be in shaping beliefs about ability. If the accent penalty is driven by employers' lack of information about worker ability, such a penalty should decrease as the worker signal value increases. As shown in Table 4.1, the expected pattern is present for Mathematics but not for Verbal skills, suggesting that the existence of an accent might be more powerful than the signal for this specific skill.

Overall, the findings suggest that the accent gap is driven by inaccurate beliefs about worker performance, with employers relying on accent as a proxy for ability. In tasks requiring minimal English proficiency, such as the Mathematics test, a highly informative performance signal corrects the misperception and reduces the WTP accent gap. In contrast, in Verbal tasks, accent serves as a salient—though inaccurate—indicator of English proficiency. As a result, the WTP accent gap in Verbal skills persists even after employers receive the performance signal (Figure 4.3).

In the Online Appendix, we present a series of analyses that explore whether our results are heterogeneous across additional characteristics such as gender and education level for workers (Appendix Section A.4.1) and employers (Appendix Section A.4.2).

4.2 Robustness Checks

We further explore the robustness of our results through various analyses. First, we show our results when using an alternative accent measure in Section 4.2.1. Next, Section 4.2.2 displays how controlling for audio clip characteristics impacts the results. We follow by exploring the role of evaluation order and the sequence of workers who employers see in Section 4.2.3. We also control for employer characteristics in Section 4.2.4. Finally, Section 4.2.5 includes additional robustness checks regarding sample selection.

4.2.1 How does the specific Accent measure affect the results?

As described in Section 2.3, we measured a worker's *Accent* based on how their speech-to-speech recording's intonation and pronunciation scored in ELSA. A worker was said to have an accent if either of these scores fell below 90%. Otherwise they are non-accented. Alternatively, instead of using a dummy variable, we can use the minimum ELSA score to create a continuous variable to measure accent. To maintain the direction of the variable, we subtracted 100 from the minimum ELSA score to have a new continuous *Accent* measure that the higher it is, the further away the worker sounds from a native U.S. English speaker.

The results using this continuous accent measure are shown in Table A.12. We find qualitatively similar results. Employers perceive those with lower accent fluency in the English language to perform worse on both tasks. For the Verbal task, employers displayed a lower WTP to hire these workers.

4.2.2 Do worker preferences and audio clip characteristics affect the results?

In our experiment, the employers heard a processed version of an audio clip recorded by the worker. As described in Section 2, it was standardized to make word choice and grammar independent of the workers' English knowledge, and the content was kept constant across treatment arms. The content was also based on "filler" information that should not correlate with workers' performances.

One possible concern is that the content of the audio clip might actually affect the employers' results for two reasons. Employers may use the audio clip content to discriminate against workers based on their "filler" preferences. Although the preference for coffee over tea does not explain performance, employers who also prefer coffee might prefer this worker to a tea drinker. Our main results in Section 4.1 should ease our concerns because the content of the audio recording is kept constant between treatment arms (*Accent Revealed* and *Accent Blind*). Additionally, we replicate our main results controlling for a preference combination fixed effect and find qualitatively similar results (see Table A.13).

Secondly, while the audio clips were homogeneous, there could be small variations. For example, the length of the recording varied slightly due to potential scripts in the paragraph the worker read (i.e., due to the inclusion of "do not" instead of "do"). Also, variations in people's pacing in their speaking may make *Accent Revealed*'s audio clips longer for some and shorter for others. *Accent Blind*'s audio clips, created with text-to-speech, were more homogeneous in this respect.

Because there might be some concerns with these two variations across audio clips, we create two variables: (1) the length of the paragraph, which measures the number of letters in the audio clip paragraph; and (2) the duration of the audio clip. The latter depends on the treatment arm. We show that the results controlling for these two variables (see Table A.14) are qualitatively similar to our main results, thus alleviating concerns that small variations across audio clips would explain our main results.

4.2.3 Does the order of and choice of workers shown affect the results?

Order of Evaluation. Another possible concern is whether the order in which the employer evaluates the workers interferes with our estimation of the accent gap (e.g., due to factors such as learning or boredom). To explore this, the results in Table A.15 show our four main results when we include the order of evaluation within the experiment task as an additional variable.[36] The results in Table A.15 are qualitatively similar to our main results. If anything, the results are strengthened when adding the evaluation order as a control.

Sequence of Workers. Since the Worker Pool has a total of 72 workers, some employers were presented the same sequence of workers (i.e., the same workers in the same order) as other employers. We concentrate on the sequence of workers for which the employer evaluates

[36] The order of the evaluations goes from 1–12 in columns (1)–(2), 1–18 in column (3), and from 1–6 in column (4).

beliefs and reports their WTP (i.e., the first 24 workers for both experiments). To investigate the presence of sequence effects, we re-estimate the four main results while controlling for the 24-worker sequence as fixed effects. The results in Table A.16 are unchanged.

4.2.4 Do results change after controlling for Employer Characteristics

A remaining possible concern with our main results is that they could be driven by employer characteristics. To alleviate this concern, we explore the robustness of the results to the inclusion of the employer characteristics. To do so, we first control for the demographic characteristics of the employers. As we show in Table 2.3, the employers are similar across treatment arms. The fact that employers are being randomized should therefore mean that these controls do not affect our results. Nevertheless, we show that our results are robust to the inclusion of these characteristics (see Tables A.17–A.18). Alternatively, instead of controlling for employer characteristics, we control for employer fixed effects. As before, we show that these results are also qualitatively similar (see Table A.19).

4.2.5 Are results robust to different sample selection choices?

Finally, we provide three additional analyses with respect to the samples presented in our main results.

First, as discussed in Section 4.1, our baseline beliefs include the "prior-to-signal beliefs" collected in the Signal Intervention experiment and all of the beliefs in the **No Signal** experiment. The regression analysis includes a dummy variable that is equal to one if the observation is from the **No Signal** experiment. Although the information collected is the same, we still show the results for the two experiments, separately, in Table A.20 and Table A.21. These results are qualitatively similar for both experiments.

Second, for the **Signal Intervention** experiment, we collected the sample in two waves: late 2024 and early 2025.[37] For completeness, we show that the results are qualitatively similar when limiting the sample to the evaluations collected in 2024. See Table A.22 and Table A.24.

Lastly, we explore whether our results are due to outlier employers in terms of the time spent on the experiment. These employers may not have exerted enough effort and gone through the worker profiles too fast (i.e., spent too little time) or may not have paid enough attention or were distracted during the experiment (i.e., spent too much time). To explore this, we calculate the time the employer spent in the experiment. By experiment, then

[37] To control for the waves, our main analysis included a year fixed effect based on the year in which we collected the information.

exclude the observations from employers whose total time sent on the experiment is below or above the 5th and 95th percentiles for the total time distribution. This rule excludes around 8% of our sample. Our results with this subsample are in Table A.25.

In general, the results are qualitatively similar in all estimations for our main results, thus alleviating concerns regarding the results being driven by time-in-experiment outliers.

5 Conclusion

This paper studies the effect of having a foreign accent on employer beliefs about worker productivity and hiring decisions in an online lab experiment. Using AI software to disentangle the effect of accent on decisions from other related characteristics, such as English proficiency, we find that employers exhibit foreign accent bias in evaluating worker productivity. This bias is present across different skills and different race and ethnicity groups. Providing employers with a signal of worker performance does not mitigate the accent gap. That said, there is suggestive evidence that the signal partly reduces the gap when the accent may not be perceived as informative of the skill (e.g., Mathematics). These beliefs translate into decisions that affect workers, as we find that employers have lower willingness to pay for hiring accented workers relative to non-accented ones.

Our findings come with some caveats that speak to the generalizability of our findings and represent some of the possible directions for future research. In this study, the employer-employee voice communication is unidirectional, which abstracts away from the real-world interactions in workplaces and the labor market. The settings discussed in the meta-analysis of Spence *et al.* (2024) highlight that accents might be considered to hinder efficiency in communications or group dynamics in team-based work, and larger accent gaps might be expected. Moreover, real-world interactions tend to be longer and likely more dynamic, which could make taste-based considerations more salient and widen the accent gaps.

While simplified, our experiment design allows us to focus on the employer belief and preference biases in the absence of some channels, such as the need for communication as well as strategic concerns (i.e., WTP for hiring is not a wage offer). Notably, the audio clips in our experiment hold workers' English proficiency characteristics (e.g., grammar) constant, a feature that would not be possible within complex real-world interactions but it is key to identifying the effect of accent. Overall, this study presents novel evidence of discrimination due to a foreign accent in hiring decisions. In future research, we plan to build on these findings to further understand the accent penalty, paying close attention to the effect of interactions and its dynamic changes.

References

Acquisti, Alessandro, & Fong, Christina. 2020. An Experiment in Hiring Discrimination via Online Social Networks. *Management Science*, **66**(3), 1005–1024.

Adank, Patti, Evans, Bronwen G, Stuart-Smith, Jane, & Scott, Sophie K. 2009. Comprehension of Familiar and Unfamiliar Native Accents Under Adverse Listening Conditions. *Journal of Experimental Psychology: Human perception and performance*, **35**(2), 520.

Albrecht, Konstanze, Von Essen, Emma, Parys, Juliane, & Szech, Nora. 2013. Updating, Self-confidence, and Discrimination. *European Economic Review*, **60**, 144–169.

Baese-Berk, Melissa M, Bradlow, Ann R, & Wright, Beverly A. 2013. Accent-Independent Adaptation to Foreign Accented Speech. *The Journal of the Acoustical Society of America*, **133**(3), EL174–EL180.

Bartoš, Vojtěch, Bauer, Michal, Chytilová, Julie, & Matějka, Filip. 2016. Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition. *American Economic Review*, **106**(6), 1437–1475.

Becker, Gary S. 1957. *The Economics of Discrimination*. University of Chicago Press.

Becker, Gordon M., Degroot, Morris H., & Marschak, Jacob. 1964. Measuring Utility by a Single-Response Sequential Method. *Behavioral Science*, **9**(3), 226–232.

Bellemare, Charles, Goussé, Marion, Lacroix, Guy, & Marchand, Steeve. 2023. Physical Disability and Labor Market Discrimination: Evidence from a Video Résumé Field Experiment. *American Economic Journal: Applied Economics*, **15**(4), 452–476.

Bertrand, M., & Duflo, E. 2017. Chapter 8 - Field Experiments on Discrimination. *Pages 309–393 of:* Banerjee, Abhijit Vinayak, & Duflo, Esther (eds), *Handbook of Field Experiments*. Handbook of Economic Field Experiments, vol. 1. North-Holland.

Bertrand, Marianne, & Mullainathan, Sendhil. 2004. Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *The American Economic Review*, **94**(4).

Bohren, Aislinn, Haggag, Kareem, Imas, Alex, & Pope, Devin. 2023. Inaccurate Statistical Discrimination: An Identification Problem. *The Review of Economics and Statistics*, 1–45.

Bohren, J Aislinn, Imas, Alex, & Rosenberg, Michael. 2019. The Dynamics of Discrimination: Theory and Evidence. *American economic review*, **109**(10), 3395–3436.

Bradlow, Ann R, & Bent, Tessa. 2008. Perceptual Adaptation to Non-Native Speech. *Cognition*, **106**(2), 707–729.

Bureau of Labor Statistics, U.S. Department of Labor. 2023. *Foreign-Born Workers Were a Record High 18.1 Percent of the U.S. Civilian Labor Force in 2022*.

Burns, Justine, & Keswell, Malcolm. 2015. Diversity and the Provision of Public Goods: Experimental Evidence From South Africa. *Journal of Economic Behavior & Organization*, **118**, 110–122.

Caballero, Jonathan A., & Pell, Marc D. 2020. Implicit Effects of Speaker Accents and Vocally-Expressed Confidence on Decisions to Trust. *Decision*, **7**(4), 314–331.

Castillo, Marco, & Petrie, Ragan. 2010. Discrimination in the Lab: Does Information Trump Appearance? *Games and Economic Behavior*, **68**(1), 50–59.

Chan, Alex. 2022. Discrimination Against Doctors: A Field Experiment.

Chiswick, Barry R, & Miller, Paul W. 1995. The Endogeneity Between Language and Earnings: International analyses. *Journal of Labor Economics*, **13**(2), 246–288.

Coffman, Katherine B., Exley, Christine L., & Niederle, Muriel. 2021. The Role of Beliefs in Driving Gender Discrimination. *Management Science*, **67**(6), 3551–3569.

Coffman, Katherine B, Kostyshak, Scott, & Saygin, Perihan O. 2025. Choosing and Using Information in Evaluation Decisions. *arXiv preprint arXiv:2507.13798*.

Deprez-Sims, Anne-Sophie, & Morris, Scott B. 2010. Accents in the Workplace: Their Effects During a Job Interview. *International Journal of Psychology*, **45**(6), 417–426.

Derwing, Tracey M, & Munro, Murray J. 2009. Putting Accent in its Place: Rethinking Obstacles to Communication. *Language Teaching*, **42**(4), 476–490.

Eames, Taryn. 2024. Pronoun Disclosure and Hiring Discrimination: A Resume Audit Study. *Available at SSRN*.

Evsyukova, Yulia, Rusche, Felix, & Mill, Wladislaw. 2025. LinkedOut? A Field Experiment on Discrimination in Job Network Formation. *The Quarterly Journal of Economics*, **140**(1), 283–334.

Ewens, Michael, Tomlin, Bryan, & Wang, Liang Choon. 2014. Statistical Discrimination or Prejudice? A Large Sample Field Experiment. *Review of Economics and Statistics*, **96**(1), 119–134.

Floccia, Caroline, Butler, Joseph, Goslin, Jeremy, & Ellis, Lucy. 2009. Regional and Foreign Accent Processing in English: Can Listeners Adapt? *Journal of psycholinguistic research*, **38**, 379–412.

Fuertes, Jairo N, Gottdiener, William H, Martin, Helena, Gilbert, Tracey C, & Giles, Howard. 2012. A Meta-Analysis of the Effects of Speakers' Accents on Interpersonal Evaluations. *European Journal of Social Psychology*, **42**(1), 120–133.

Ge, Qi, & Wu, Stephen. 2024. How Do You Say Your Name? Difficult-To-Pronounce Names and Labor Market Outcomes. *American Economic Journal: Economic Policy*, **16**(4), 254–279.

Gluszek, Agata, & Dovidio, John F. 2010. Speaking with a Nonnative Accent: Perceptions of Bias, Communication Difficulties, and Belonging in the United States. *Journal of Language and Social Psychology*, **29**(2), 224–234.

Haaland, Ingar, Roth, Christopher, & Wohlfart, Johannes. 2023. Designing Information Provision Experiments. *Journal of Economic Literature*, **61**(1), 3–40.

Healy, Paul J. 2017. *Epistemic Game Theory Experiments: Utility Elicitation and Irrational Play*. Tech. rept. mimeo.

Hedegaard, Morten Størling, & Tyran, Jean-Robert. 2018. The Price of Prejudice. *American Economic Journal: Applied Economics*, **10**(1), 40–63.

Hossain, Tanjim, & Okui, Ryo. 2013. The Binarized Scoring Rule. *The Review of Economic Studies*, **80**(3), 984–1001.

Huang, Laura, Frideger, Marcia, & Pearce, Jone L. 2013. Political Skill: Explaining the Effects of Nonnative Accent on Managerial Hiring and Entrepreneurial Investment Decisions. *Journal of Applied Psychology*, **98**(6), 1005.

Isphording, Ingo E, & Sinning, Mathias. 2012. The Returns to Language Skills in the US Labor Market. *Ruhr Economic Paper*.

Laouénan, Morgane, & Rathelot, Roland. 2022. Can Information Reduce Ethnic Discrimination? Evidence from Airbnb. *American Economic Journal: Applied Economics*, **14**(1), 107–132.

Li, Mingsheng, & Campbell, Jacqui. 2009. Accessing Employment: Challenges Faced by Non-Native English-Speaking Professional Migrants. *Asian and Pacific Migration Journal*, **18**(3), 371–395.

Munro, Murray J, & Derwing, Tracey M. 1995. Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners. *Language Learning*, **45**(1), 73–97.

Munro, Murray J, Derwing, Tracey M, & Sato, Kyoko. 2006. Salient Accents, Covert Attitudes: Consciousness-Raising for Pre-Service Second Language Teachers. *Prospect: an Australian journal of TESOL*, **21**(1), 65–77.

Neumark, David. 2018. Experimental Research on Labor Market Discrimination. *Journal of Economic Literature*, **56**(3), 799–866.

Oreopoulos, Philip. 2011. Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Thirteen Thousand Resumes. *American Economic Journal: Economic Policy*, **3**(4), 148–171.

Park, Hyoeun. 2024. Stereotype-Confirming Race Discrimination Across Multiple Minorities.

Pew Research Center, Report. 2017. *The Partisan Divide on Political Values Grows Even Wider*. Tech. rept. Pew Research Center.

Rackstraw, Emma. 2022. Bias-Motivated Updating in the Labor Market.

Reuben, Ernesto, Sapienza, Paola, & Zingales, Luigi. 2014. How Stereotypes Impair Women's Careers in Science. *Proceedings of the National Academy of Sciences*, **111**(12), 4403–4408.

Rubin, Donald L. 1992. Nonlanguage Factors Affecting Undergraduates' Judgments of Nonnative English-Speaking Teaching Assistants. *Research in Higher education*, **33**(4), 511–531.

Ruggles, Steven, Flood, Sarah, Sobek, Matthew, Backman, Daniel, Chen, Annie, Cooper, Grace, Richards, Stephanie, Rodgers, Renae, & Schouweiler, Megan. 2024. *IPUMS USA: Version 15.0, American Community Survey 2022*.

Schmaus, Miriam, & Kristen, Cornelia. 2022. Foreign Accents in the Early Hiring Process: A Field Experiment on Accent-Related Ethnic Discrimination in Germany. *International Migration Review*, **56**(2), 562–593.

Schroeder, Juliana, & Epley, Nicholas. 2015. The Sound of Intellect: Speech Reveals a Thoughtful Mind, Increasing a Job Candidate's Appeal. *Psychological Science*, **26**(6), 877–891.

Shuck, Gail. 2004. Conversational Performance and the Poetic Construction of an Ideology. *Language in Society*, **33**(2), 195–222.

Shuck, Gail. 2006. Racializing the Nonnative English Speaker. *Journal of Language, Identity, and Education*, **5**(4), 259–276.

Spence, Jessica L, Hornsey, Matthew J, & Imuta, Kana. 2021. Something About the Way You Speak: A Meta-Analysis on Children's Linguistic-Based Social Preferences. *Child Development*, **92**(2), 517–535.

Spence, Jessica L., Hornsey, Matthew J., Stephenson, Eloise M., & Imuta, Kana. 2024. Is Your Accent Right for the Job? A Meta-Analysis on Accent Bias in Hiring Decisions. *Personality and Social Psychology Bulletin*, **50**(3), 371–386.

Tenzer, Helene, Pudelko, Markus, & Harzing, Anne-Wil. 2014. The Impact of Language Barriers on Trust Formation in Multinational Teams. *Journal of International Business Studies*, **45**, 508–535.

Tomlin, Bryan. 2024. Pronoun Usage and Gender Identity's Effects on Market Outcomes: Evidence From a Preregistered Field Experiment. *Economics Letters*, **236**.

van Wijngaarden, Sander J, Steeneken, Herman JM, & Houtgast, Tammo. 2001. The Effect of a Non-Native Accent in Dutch on Speech Intelligibility. *The Journal of the Acoustical Society of America*, **109**(5_Supplement), 2473–2473.

Wozniak, David, & MacNeill, Timothy. 2020. Racial Discrimination in the Lab: Evidence of Statistical and Taste-Based Discrimination. *Journal of Behavioral and Experimental Economics*, **85**, 101512.

Zuckerman, Miron, & Driver, Robert E. 1989. What Sounds Beautiful is Good: The Vocal Attractiveness Stereotype. *Journal of Nonverbal Behavior*, **13**(2), 67–82.

Online Appendix

A.1 Figures

Figure A.1: Example of Test Questions: Mathematics Skills

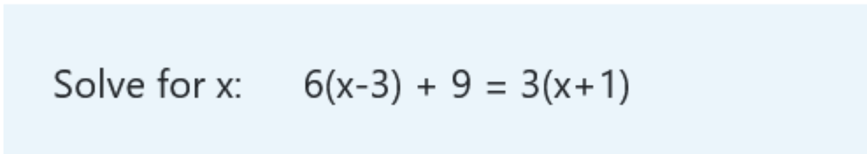


Figure A.2: Example of Test Questions: Verbal Skills

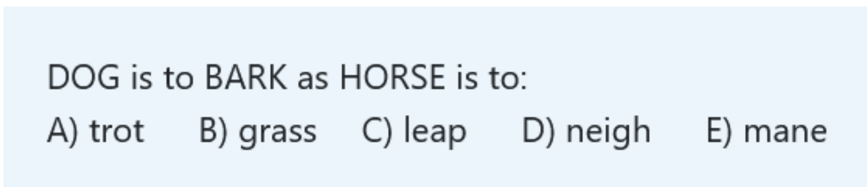
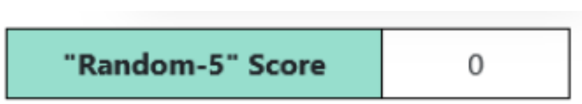


Figure A.3: Example of Signal



A.2 Example: Voice Recording and Accent

Measuring Accent

As described in subsection 2.3, we measure a worker's accent ELSA grade of a recording's pronunciation and intonation. Using the voice-transformed (**speech-to-speech**) recording as an example, **ELSA Pronunciation** score is 80%, and **ELSA Intonation** score is 77%.

The ELSA pronunciation score is based on how closely the speech in the recording matches the English sounds. ELSA highlights the errors that lead to the grade. At the same time, ELSA grades the audio clip intonation based on the number of words where intonation was not in line with what would be expected for the type of words. See Figure A.4 for the grade breakdown of the recording example. The letters (words) highlighted in red are the ones that ELSA highlighted as incorrectly pronounced (emphasized).

Figure A.4: ELSA Grade Detail

Note. This figure presents the ELSA pronunciation and intonation grade detail for the voice-transformed recording example.

A.3 Tables

Table A.1: Worker Sample Description by Accent Status

	(1) All	(2) Asian	(3) Hispanic	(4) White
		No Accent		
Age	33.19	28.83	31.33	39.42
	(9.21)	(6.78)	(7.19)	(10.26)
College or More	0.61	0.67	0.50	0.67
	(0.49)	(0.49)	(0.52)	(0.49)
Math Score	7.53	9.42	6.67	6.50
	(3.35)	(3.29)	(3.23)	(2.94)
Verbal Score	10.31	10.92	9.92	10.08
	(1.82)	(1.83)	(2.15)	(1.38)
Math Random-5 Signal	2.61	3.25	2.25	2.33
	(1.59)	(1.42)	(1.42)	(1.83)
Verbal Random-5 Signal	2.69	3.17	2.50	2.42
	(1.14)	(1.11)	(1.31)	(0.90)
N	36	12	12	12
		Accent		
Age	37.08	33.00	37.42	40.83
	(11.17)	(12.11)	(11.40)	(9.31)
College or More	0.58	0.83	0.50	0.42
	(0.50)	(0.39)	(0.52)	(0.51)
Math Score	6.61	9.25	4.83	5.75
	(3.69)	(3.39)	(3.30)	(3.02)
Verbal Score	9.50	10.08	9.67	8.75
	(1.80)	(2.15)	(1.44)	(1.60)
Math Random-5 Signal	2.33	3.50	1.67	1.83
	(1.76)	(1.68)	(1.61)	(1.47)
Verbal Random-5 Signal	2.39	2.67	2.42	2.08
	(1.08)	(1.30)	(0.90)	(1.00)
N	36	12	12	12

Note. This table presents the mean and standard deviation (in parenthesis) of the sample of workers by Accent Status. The columns refer to the worker's self-identified race/ethnicity. College or More is a variable equal to 1 if the worker reported educational attainment is a Bachelor's degree in college (4-year) or higher, and 0 otherwise; Math score and Verbal score represent the number of questions a worker answered correctly out of 15 in the Math and Verbal tests, respectively.

Table A.2: Baseline Beliefs over Worker Performance in Mathematics Test, Conditional on Worker Test Score

	(1) All	(2) Asian	(3) Hispanic	(4) White
Accent × *AccRev*	-0.327***	-0.173	-0.443***	-0.375***
	(0.079)	(0.131)	(0.133)	(0.136)
Accent	0.041	0.130	-0.031	-0.016
	(0.052)	(0.098)	(0.096)	(0.103)
AccRev	0.059	0.078	0.106	-0.005
	(0.167)	(0.203)	(0.177)	(0.183)
$TestScore$	0.024**	0.013	0.043***	-0.001
	(0.010)	(0.017)	(0.016)	(0.021)
Asian Worker	0.932***			
	(0.070)			
Hispanic Worker	-0.208***			
	(0.057)			
Worker: 35 or older	-0.124*	-0.234*	-0.018	-0.021
	(0.067)	(0.129)	(0.106)	(0.109)
Worker: Male	-0.007	0.022	-0.095	0.124
	(0.047)	(0.101)	(0.075)	(0.094)
Worker: Bachelor's Degree	1.870***	1.909***	1.793***	1.864***
	(0.082)	(0.152)	(0.116)	(0.126)
Year = 2025	0.306	0.335	0.329	0.255
	(0.261)	(0.306)	(0.277)	(0.283)
No Signal Experiment	-0.594**	-0.733***	-0.502**	-0.542**
	(0.238)	(0.277)	(0.255)	(0.256)
Constant	5.378***	6.414***	5.123***	5.250***
	(0.254)	(0.472)	(0.314)	(0.314)
Average Beliefs	8.40	9.45	7.69	8.07
N	9,336	3,112	3,112	3,112

Note. The table displays the estimated coefficients in equation (2). The estimations include evaluations 1-12 from both experiments. "Year = 2025" and "No Signal experiment" are dummy variables for the year of data collection and the first experiment. Columns: All includes the sample of the employer evaluations 1-12 for all workers; Asian, Hispanic and White include the subsample of the employer evaluations 1-12 that were Asian, Hispanic and White workers, respectively. Standard errors clustered at the employer level in parentheses. N represents the number of observations. * p<0.10, ** p<0.05, *** p<0.01.

Table A.3: Baseline Beliefs over Worker Performance in Verbal Test, Conditional on Worker Test Score

	(1) All	(2) Asian	(3) Hispanic	(4) White
Accent × *AccRev*	-0.662***	-0.703***	-0.692***	-0.587***
	(0.088)	(0.137)	(0.146)	(0.148)
Accent	-0.055	-0.037	-0.113	0.073
	(0.061)	(0.103)	(0.111)	(0.133)
AccRev	0.114	0.186	0.064	0.090
	(0.235)	(0.254)	(0.239)	(0.253)
TestScore	-0.004	-0.011	0.040	-0.002
	(0.021)	(0.044)	(0.035)	(0.046)
Asian Worker	0.232***			
	(0.071)			
Hispanic Worker	-0.359***			
	(0.063)			
Worker: 35 or older	0.030	0.061	-0.075	0.121
	(0.080)	(0.155)	(0.147)	(0.133)
Worker: Male	-0.145***	-0.091	-0.190***	-0.221**
	(0.041)	(0.096)	(0.067)	(0.095)
Worker: Bachelor's Degree	-0.145***	-0.091	-0.190***	-0.221**
	(0.041)	(0.096)	(0.067)	(0.095)
Year = 2025	-0.198	-0.460	-0.070	-0.063
	(0.340)	(0.373)	(0.346)	(0.353)
No Signal Experiment	-0.633**	-0.582*	-0.740**	-0.583*
	(0.314)	(0.350)	(0.318)	(0.323)
Constant	8.382***	8.881***	7.645***	7.952***
	(0.350)	(0.604)	(0.500)	(0.531)
Average Beliefs	9.43	9.87	8.95	9.45
N	8,328	2,776	2,776	2,776

Note. The table displays the estimated coefficients in equation (2). The estimations include evaluations 1-12 from both experiments. "Year = 2025" and "No Signal experiment" are dummy variables for the year of data collection and the first experiment. Columns: All includes the sample of the employer evaluations 1-12 for all workers; Asian, Hispanic and White include the subsample of the employer evaluations 1-12 that were Asian, Hispanic and White workers, respectively. Standard errors clustered at the employer level in parentheses. N represents the number of observations. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.4: Beliefs over Worker Performance in Mathematics Skills Test in the Presence of Informative Signal, Conditional on Worker Test Score

	(1) All	(2) Asian	(3) Hispanic	(4) White
Accent × *AccRev* × *Signal*	0.239	0.444	0.242	0.108
	(0.300)	(0.640)	(0.558)	(0.558)
Accent × *AccRev*	-0.523***	-0.577**	-0.774***	-0.249
	(0.130)	(0.231)	(0.210)	(0.207)
Accent × *Signal*	-0.086	-0.310	0.286	-0.446
	(0.186)	(0.449)	(0.371)	(0.368)
AccRev × *Signal*	-0.329	-0.903*	-0.083	-0.034
	(0.301)	(0.514)	(0.468)	(0.424)
Signal	-1.365***	-1.142***	-1.401***	-1.454***
	(0.212)	(0.394)	(0.322)	(0.271)
TestScore	0.153***	0.106***	0.201***	0.130***
	(0.016)	(0.025)	(0.024)	(0.031)
Accent	0.273***	0.373**	0.368**	0.109
	(0.083)	(0.172)	(0.147)	(0.148)
AccRev	0.565**	0.657**	0.600**	0.453
	(0.258)	(0.314)	(0.272)	(0.291)
Asian Worker	0.611***			
	(0.115)			
Hispanic Worker	-0.049			
	(0.087)			
Worker: 35 or older	-0.093	-0.288	0.198	-0.217
	(0.094)	(0.186)	(0.151)	(0.153)
Worker: Male	0.091	0.201	-0.112	0.172
	(0.064)	(0.144)	(0.114)	(0.136)
Worker: Bachelor's Degree	1.941***	1.810***	1.914***	1.943***
	(0.115)	(0.204)	(0.150)	(0.192)
Year = 2025	0.383	0.306	0.485*	0.347
	(0.256)	(0.317)	(0.262)	(0.267)
Constant	3.892***	5.219***	3.358***	4.255***
	(0.323)	(0.596)	(0.412)	(0.425)
Average Beliefs	8.15	9.27	7.49	7.72
N	5,310	1,733	1,778	1,799

Note. The table displays the estimated coefficients in equation (3). The estimations include evaluations 1-18 from the Signal Intervention experiment. A year dummy for the year of data collection is included. Columns: All include the sample of the employer evaluations 1-18 for all workers; Asian, Hispanic and White include the subsample of the employer evaluations 1-18 that were Asian, Hispanic and White workers, respectively. Standard errors clustered at the employer level in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A.5: Beliefs over Worker Performance in Verbal Skills Test in the Presence of Informative Signal, Conditional on Worker Test Score

	(1) All	(2) Asian	(3) Hispanic	(4) White
Accent × *AccRev* × *Signal*	0.029	0.044	0.228	-0.296
	(0.319)	(0.620)	(0.578)	(0.518)
Accent × *AccRev*	-0.601***	-0.711***	-0.604***	-0.474*
	(0.141)	(0.218)	(0.220)	(0.243)
Accent × *Signal*	0.031	-0.186	-0.310	0.651*
	(0.197)	(0.431)	(0.371)	(0.361)
AccRev × *AccRev*	0.003	-0.668	-0.072	0.771*
	(0.283)	(0.452)	(0.449)	(0.430)
Signal	-1.650***	-1.411***	-1.325***	-2.245***
	(0.198)	(0.340)	(0.317)	(0.296)
TestScore	0.139***	0.078	0.239***	0.100*
	(0.027)	(0.055)	(0.039)	(0.060)
Accent	-0.025	-0.059	-0.097	0.106
	(0.096)	(0.159)	(0.178)	(0.212)
AccRev	-0.303	-0.065	-0.347	-0.504
	(0.338)	(0.367)	(0.348)	(0.367)
Asian Worker	0.148			
	(0.099)			
Hispanic Worker	-0.292***			
	(0.086)			
Worker: 35 or older	-0.020	0.274	-0.216	0.121
	(0.097)	(0.195)	(0.186)	(0.161)
Worker: Male	-0.176***	-0.054	-0.301***	-0.152
	(0.051)	(0.142)	(0.109)	(0.127)
Worker: Bachelor's Degree	1.219***	1.298***	1.366***	1.405***
	(0.112)	(0.244)	(0.176)	(0.185)
Year = 2025	-0.184	-0.406	-0.095	-0.059
	(0.328)	(0.359)	(0.347)	(0.333)
Constant	7.277***	7.435***	6.224***	7.020***
	(0.433)	(0.720)	(0.585)	(0.735)
Average Beliefs	9.33	9.81	8.91	9.28
N	5,292	1,739	1,783	1,770

Note. The table displays the estimated coefficients in equation (3). The estimations include evaluations 1-18 from the Signal Intervention experiment. A year dummy for the year of data collection is included. Columns: All include the sample of the employer evaluations 1-18 for all workers; Asian, Hispanic and White include the subsample of the employer evaluations 1-18 that were Asian, Hispanic and White workers, respectively. Standard errors clustered at the employer level in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A.6: Baseline Willingness to Pay (in cents), Conditional on Test Score

	Mathematics Skills				Verbal Skills			
	(1) All	(2) Asian	(3) Hispanic	(4) White	(5) All	(6) Asian	(7) Hispanic	(8) White
Accent × *AccRev*	-3.014**	-4.765*	-4.810*	0.515	-7.196***	-4.770*	-11.520***	-5.401**
	(1.493)	(2.635)	(2.658)	(2.524)	(1.746)	(2.858)	(2.945)	(2.652)
Accent	-0.415	0.537	-0.392	-0.952	0.614	1.128	0.588	-0.541
	(0.989)	(1.869)	(1.982)	(1.698)	(1.153)	(2.266)	(2.135)	(2.100)
AccRev	0.592	2.869	0.600	-1.709	-0.143	-0.955	1.527	-0.794
	(3.854)	(4.238)	(4.123)	(4.218)	(4.405)	(4.754)	(4.622)	(4.751)
$TestScore$	0.084	-0.200	0.163	0.174	0.290	0.053	1.237*	0.288
	(0.211)	(0.378)	(0.352)	(0.484)	(0.442)	(0.899)	(0.724)	(0.893)
Asian Worker	3.837***				-0.405			
	(1.371)				(1.334)			
Hispanic Worker	-3.992***				-4.482***			
	(1.078)				(1.192)			
Worker: Age	-1.311	-0.812	-0.343	-2.503	-0.248	-0.865	0.760	-1.051
	(1.361)	(2.810)	(2.395)	(2.327)	(1.403)	(3.181)	(2.609)	(2.497)
Worker: Male	0.992	2.846	-1.359	0.820	-0.945	0.741	-1.719	-1.802
	(0.824)	(1.816)	(1.296)	(1.677)	(0.912)	(1.929)	(1.504)	(1.816)
Worker: Bachelor's Degree	28.029***	26.810***	28.291***	29.275***	22.784***	22.063***	26.250***	21.002***
	(1.894)	(3.323)	(2.613)	(2.680)	(1.900)	(4.150)	(2.929)	(2.969)
Constant	40.710***	44.817***	38.453***	41.153***	56.540***	57.707***	37.687***	62.482***
	(4.656)	(9.786)	(5.711)	(5.574)	(6.664)	(11.006)	(10.182)	(10.147)
N	5796	1932	1932	1932	4800	1600	1600	1600

Note. The table displays the estimated coefficients in equation (4), using the sample from the No Signal experiment (evaluations 13-24). Columns: All include the sample of the employer evaluations 13-24 for all workers; Asian, Hispanic and White include the subsample of the employer evaluations 13-24 that were Asian, Hispanic and White workers, respectively. Standard errors clustered at the employer level in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A.7: Willingness to Pay (in cents) in the Signal Intervention Experiment, Conditional on Test Score

	Mathematics Skills				Verbal Skills			
	(1) All	(2) Asian	(3) Hispanic	(4) White	(5) All	(6) Asian	(7) Hispanic	(8) White
Accent × *AccRev*	5.441	8.173	7.671	-0.598	-10.220**	-12.458	-5.481	-14.841**
	(4.446)	(8.318)	(7.149)	(7.997)	(4.115)	(7.638)	(6.557)	(7.294)
Accent	-3.997	-9.501	-6.596	4.842	3.537	7.424	2.206	11.366*
	(2.837)	(5.885)	(5.237)	(5.250)	(3.004)	(5.336)	(4.694)	(6.673)
AccRev	3.490	4.885	3.098	1.897	2.410	5.182	-2.429	6.547
	(5.383)	(7.881)	(6.808)	(7.244)	(4.727)	(7.418)	(6.139)	(5.887)
TestScore	-0.098	-1.005	2.463*	-3.025	1.074	4.052	-1.162	4.685*
	(0.882)	(1.537)	(1.407)	(1.954)	(1.176)	(2.732)	(1.814)	(2.507)
Asian Worker	11.532***				10.640***			
	(3.482)				(3.219)			
Hispanic Worker	-0.707				-2.355			
	(2.779)				(2.489)			
Worker: Age	2.022	9.659*	0.811	-1.650	4.560*	8.553*	1.184	6.106
	(2.605)	(5.493)	(4.169)	(5.080)	(2.535)	(4.680)	(3.909)	(4.890)
Worker: Male	2.289	1.348	6.206*	-5.410	-1.331	-1.795	3.746	-10.606***
	(1.683)	(5.968)	(3.513)	(4.109)	(1.525)	(4.218)	(3.322)	(3.956)
Worker: Bachelor's Degree	12.600***	17.258**	15.409***	18.121***	10.532***	2.945	10.246***	20.287***
	(2.552)	(6.684)	(4.339)	(4.962)	(2.300)	(6.386)	(3.685)	(5.073)
Worker: Signal Value	8.728***	13.863***	0.833	14.383***	9.595***	8.424*	11.861***	-0.394
	(2.003)	(3.917)	(3.419)	(3.503)	(2.269)	(4.931)	(2.721)	(5.509)
Year = 2025	9.711*	8.195	12.962**	6.363	12.464***	13.309**	12.899**	12.055**
	(5.279)	(7.147)	(6.017)	(6.421)	(4.708)	(6.200)	(5.220)	(5.687)
Constant	6.607	-5.051	-1.167	21.057	1.356	-9.827	14.411	-19.007
	(8.144)	(17.839)	(11.388)	(13.554)	(8.930)	(17.637)	(14.862)	(21.213)
N	1770	608	657	505	1764	581	682	501

Note. The table displays the estimated coefficients in equation (4), using the sample from the Signal Intervention experiment (evaluations 19-24). Columns: All include the sample of the employer evaluations 19-24 for all workers; Asian, Hispanic and White include the subsample of the employer evaluations 13-24 that were Asian, Hispanic and White workers, respectively. Standard errors clustered at the employer level in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.8: Implied WTP Gap vs. Actual WTP Gap for Verbal Skills: No Signal Experiment

	(1) All [Implied]	(2) All [Actual]	(3) Asian [Implied]	(4) Asian [Actual]	(5) Hispanic [Implied]	(6) Hispanic [Actual]	(7) White [Implied]	(8) White [Actual]
Accent × *AccRev*	-3.085**	-3.012**	1.089	-4.761*	-3.988*	-4.768*	-6.362**	0.484
	(1.414)	(1.493)	(2.228)	(2.636)	(2.410)	(2.660)	(2.509)	(2.527)
Accent	-0.356	-0.454	0.484	0.526	-1.488	-0.633	-0.617	-0.900
	(0.920)	(0.983)	(1.682)	(1.870)	(1.657)	(1.929)	(1.853)	(1.692)
AccRev	-3.406	0.594	-3.901	2.865	-2.492	0.593	-3.858	-1.689
	(3.026)	(3.854)	(3.698)	(4.238)	(3.229)	(4.122)	(3.298)	(4.216)
Asian Worker	13.287***	4.057***						
	(1.174)	(1.280)						
Hispanic Worker	-2.798***	-4.052***						
	(1.049)	(1.069)						
Worker: 35 or older	-2.260*	-1.417	-3.971*	-0.581	-1.285	-0.564	-0.570	-2.629
	(1.212)	(1.320)	(2.296)	(2.753)	(1.922)	(2.354)	(1.998)	(2.298)
Worker: Male	-0.104	1.135	-1.246	2.380	-0.675	-1.217	1.493	1.079
	(0.825)	(0.760)	(1.630)	(1.670)	(1.347)	(1.280)	(1.576)	(1.520)
Worker: Bachelor's Degree	24.438***	28.114***	24.330***	26.724***	22.915***	28.485***	23.933***	29.559***
	(1.471)	(1.854)	(2.806)	(3.325)	(2.137)	(2.556)	(2.151)	(2.504)
Constant	79.312***	41.074***	95.493***	43.485***	78.547***	39.301***	75.939***	41.562***
	(3.999)	(4.549)	(8.332)	(9.276)	(4.924)	(5.369)	(5.113)	(5.477)
Average WTP	115.50	85.15	129.56	93.82	106.22	78.51	110.71	83.13
N	5,796	5,796	1,932	1,932	1,932	1,932	1,932	1,932

Note. The estimations for Implied WTP (in cents) in equation (5) in the *odd numbered columns* include evaluations 1-12 (belief elicitation) from the No Signal experiment, Mathematics Skills treatment arm, multiplied by 14 to calculate the implied WTP based on elicited beliefs. *Even numbered columns* report estimations for Actual WTP (in cents) in equation (4) using evaluations 13-24 (hiring task) from the No Signal experiment. Columns: All includes the sample of the employer decisions for all workers; Asian, Hispanic and White include the subsample of the employer decisions for workers that were Asian, Hispanic and White workers, respectively. Standard errors clustered at the employer level in parentheses. N represents the number of observations. * p<0.10, ** p<0.05, *** p<0.01.

Table A.9: Implied WTP Gap vs. Actual WTP Gap for Verbal Skills: No Signal Experiment

	(1) All [Implied]	(2) All [Actual]	(3) Asian [Implied]	(4) Asian [Actual]	(5) Hispanic [Implied]	(6) Hispanic [Actual]	(7) White [Implied]	(8) White [Actual]
Accent × *AccRev*	-9.960***	-7.191***	-9.781***	-4.769*	-10.643***	-11.383***	-9.333***	-5.427**
	(1.572)	(1.744)	(2.479)	(2.858)	(2.728)	(2.935)	(2.577)	(2.660)
Accent	0.429	0.400	0.447	1.045	-0.107	0.393	2.534	-0.824
	(1.005)	(1.114)	(1.592)	(1.962)	(1.976)	(2.125)	(1.885)	(1.874)
AccRev	5.905	-0.149	5.271	-0.954	5.121	1.400	7.315	-0.802
	(4.511)	(4.405)	(4.893)	(4.752)	(4.567)	(4.638)	(4.828)	(4.749)
Asian Worker	1.344	-0.060						
	(1.233)	(1.283)						
Hispanic Worker	-6.601***	-4.375***						
	(1.125)	(1.202)						
Worker: 35 or older	-1.204	-0.160	-1.556	-0.826	-2.734	1.348	1.258	-0.955
	(1.433)	(1.409)	(2.804)	(3.066)	(2.750)	(2.620)	(2.530)	(2.549)
Worker: Male	-1.644**	-0.963	-0.760	0.782	-0.673	-1.689	-5.324***	-1.879
	(0.765)	(0.910)	(1.899)	(1.859)	(1.205)	(1.501)	(1.776)	(1.793)
Worker: Bachelor's Degree	15.164***	23.038***	14.552***	22.238***	14.554***	25.443***	19.177***	21.294***
	(1.524)	(1.902)	(3.533)	(3.438)	(2.434)	(2.837)	(2.692)	(2.715)
Constant	108.678***	58.870***	110.477***	57.894***	104.511***	50.115***	101.788***	64.852***
	(4.244)	(5.125)	(10.527)	(10.195)	(5.426)	(6.152)	(5.653)	(6.403)
Average WTP	127.50	90.90	133.20	95.72	120.97	85.86	128.33	91.10
N	4,800	4,800	1,600	1,600	1,600	1,600	1,600	1,600

Note. The estimations for Implied WTP (in cents) in equation (5) in the *odd numbered columns* include evaluations 1-12 (belief elicitation) from the No Signal experiment, Verbal Skills treatment arm, multiplied by 14 to calculate the implied WTP based on elicited beliefs. *Even numbered columns* report estimations for Actual WTP (in cents) in equation (4) using evaluations 13-24 (hiring task) from the No Signal experiment. Columns: All includes the sample of the employer decisions for all workers; Asian, Hispanic and White include the subsample of the employer decisions for workers that were Asian, Hispanic and White workers, respectively. Standard errors clustered at the employer level in parentheses. N represents the number of observations. * p<0.10, ** p<0.05, *** p<0.01.

Table A.10: Implied WTP Gap vs. Actual WTP Gap for Mathematics Skills: Signal Intervention Experiment

	(1) All [Implied]	(2) All [Actual]	(3) Asian [Implied]	(4) Asian [Actual]	(5) Hispanic [Implied]	(6) Hispanic [Actual]	(7) White [Implied]	(8) White [Actual]
Accent × *AccRev*	-4.716	5.446	-3.034	8.179	-7.315	7.316	-4.373	-0.786
	(3.854)	(4.446)	(8.094)	(8.301)	(6.958)	(7.151)	(7.103)	(7.966)
Accent	1.344	-3.981	2.550	-9.139	0.484	-6.386	0.533	5.068
	(2.599)	(2.828)	(5.740)	(5.925)	(5.101)	(5.228)	(5.116)	(5.280)
AccRev	3.793	3.487	-2.092	4.987	7.313	3.237	6.423	1.695
	(4.906)	(5.381)	(7.615)	(7.890)	(6.090)	(6.773)	(6.214)	(7.246)
Asian Worker	3.754	11.416***						
	(3.123)	(3.463)						
Hispanic Worker	3.472	-0.682						
	(2.492)	(2.707)						
Worker: 35 or older	-0.919	2.002	-5.399	9.276*	4.830	1.004	-4.159	-2.532
	(2.292)	(2.606)	(4.557)	(5.441)	(4.327)	(4.191)	(3.685)	(5.142)
Worker: Male	1.282	2.310	-4.077	2.226	-3.664	5.595	5.297	-6.017
	(1.736)	(1.638)	(4.666)	(5.657)	(3.635)	(3.484)	(4.079)	(4.162)
Worker: Bachelor's Degree	13.433***	12.620***	3.914	17.620***	9.480**	13.368***	20.607***	16.627***
	(2.929)	(2.518)	(5.851)	(6.624)	(4.557)	(4.032)	(5.145)	(4.863)
Worker Signal Value	13.197***	8.546***	15.000***	11.799***	14.833***	5.882***	10.151***	9.648***
	(1.106)	(1.035)	(2.056)	(2.161)	(1.448)	(1.386)	(1.483)	(1.623)
Year = 2025	8.726*	9.705*	5.330	8.233	12.431**	13.281**	8.174	6.184
	(4.714)	(5.275)	(6.903)	(7.143)	(5.144)	(6.040)	(5.070)	(6.436)
Constant	37.727***	6.360	69.159***	-9.335	41.010***	5.973	32.462***	17.334
	(6.560)	(7.420)	(16.094)	(16.104)	(8.531)	(10.188)	(10.208)	(13.401)
Average WTP	99.87	62.44	114.73	78.98	94.79	52.71	91.49	55.18
N	1,770	1,770	553	608	598	657	619	505

Note. The estimations for Implied WTP (in cents) in equation (5) in the *odd numbered columns* include evaluations 13-18 (belief elicitation, signal available) from the Signal Intervention experiment, Mathematics Skills treatment arm, multiplied by 14 to calculate the implied WTP based on elicited beliefs. *Even numbered columns* report estimations for Actual WTP (in cents) in equation (4) using evaluations 19-24 (hiring task) from the Signal Intervention experiment. Columns: All includes the sample of the employer decisions for all workers; Asian, Hispanic and White include the subsample of the employer decisions for workers that were Asian, Hispanic and White workers, respectively. Dummy variables for the year of data collection (Year = 2025) included. Standard errors clustered at the employer level in parentheses. N represents the number of observations. * p<0.10, ** p<0.05, *** p<0.01.

Table A.11: Implied WTP Gap vs. Actual WTP Gap for Verbal Skills: Signal Intervention Experiment

	(1) All [Implied]	(2) All [Actual]	(3) Asian [Implied]	(4) Asian [Actual]	(5) Hispanic [Implied]	(6) Hispanic [Actual]	(7) White [Implied]	(8) White [Actual]
Accent × *AccRev*	-8.670**	-10.211**	-10.157	-12.366	-6.462	-5.514	-10.484	-14.542**
	(4.181)	(4.121)	(7.875)	(7.651)	(7.497)	(6.531)	(6.558)	(7.297)
Accent	1.106	3.042	-0.382	5.594	-5.487	2.572	8.784*	6.074
	(2.750)	(3.014)	(5.388)	(5.395)	(5.581)	(4.666)	(4.789)	(5.818)
AccRev	-4.062	2.359	-10.140	4.883	-5.712	-2.343	3.878	6.049
	(5.364)	(4.719)	(7.399)	(7.369)	(6.844)	(6.092)	(6.692)	(5.866)
Asian Worker	-5.070*	10.818***						
	(2.932)	(3.215)						
Hispanic Worker	-3.428	-2.210						
	(2.806)	(2.469)						
Worker: 35 or older	-2.773	4.720*	0.660	9.326**	-0.833	0.581	-1.820	3.605
	(2.425)	(2.519)	(4.953)	(4.579)	(4.748)	(3.958)	(4.320)	(4.657)
Worker: Male	-0.810	-0.968	2.076	0.408	-0.127	3.423	-3.295	-8.190**
	(1.302)	(1.459)	(4.202)	(4.114)	(3.485)	(3.194)	(3.905)	(3.460)
Worker: Bachelor's Degree	11.641***	10.847***	18.563***	7.530	8.909**	10.330***	16.825***	18.387***
	(2.613)	(2.303)	(6.158)	(6.296)	(4.271)	(3.698)	(5.414)	(4.770)
Worker: Signal Value	12.261***	11.116***	11.420***	13.825***	12.983***	10.384***	9.929***	6.947**
	(1.307)	(1.379)	(2.191)	(2.713)	(1.763)	(1.779)	(2.849)	(2.736)
Year = 2025	-2.025	12.425***	-3.783	13.051**	-1.804	12.857**	-1.071	11.664**
	(5.141)	(4.711)	(6.146)	(6.195)	(6.384)	(5.219)	(5.910)	(5.670)
Constant	77.500***	6.984	58.743***	5.824	76.234***	7.790	69.411***	14.985
	(7.748)	(6.518)	(16.354)	(13.611)	(10.299)	(9.519)	(12.684)	(12.839)
Average WTP	115.68	63.62	121.23	76.69	111.88	56.25	114.31	58.50
N	1,764	1,764	563	581	607	682	594	501

Note. The estimations for Implied WTP (in cents) in equation (5) in the *odd numbered columns* include evaluations (belief elicitation, signal available) from the Signal Intervention experiment, Verbal Skills treatment arm, multiplied by 14 to calculate the implied WTP based on elicited beliefs. *Even numbered columns* report estimations for Actual WTP (in cents) in equation (4) using evaluations 19-24 (hiring task) from the Signal Intervention experiment. Columns: All includes the sample of the employer decisions for all workers; Asian, Hispanic and White include the subsample of the employer decisions for workers that were Asian, Hispanic and White workers, respectively. Dummy variables for the year of data collection (Year = 2025) included. Standard errors clustered at the employer level in parentheses. N represents the number of observations. * p<0.10, ** p<0.05, *** p<0.01.

A.3.1 Robustness Checks

Table A.12: Results using Continuous Accent Measure

	Beliefs		WTP	
	(1)	(2)	(3)	(4)
	No Signal	w/Signal	No Signal	w/Signal
Panel A. Mathematics Skills				
Accent × *AccRev*	-0.017**	-0.016	-0.273*	0.525*
	(0.007)	(0.011)	(0.141)	(0.285)
Accent	-0.004	-0.011	-0.042	-0.217
	(0.005)	(0.007)	(0.096)	(0.197)
AccRev	0.097	0.487*	2.274	0.006
	(0.183)	(0.279)	(4.174)	(5.881)
Average Outcome	8.40	8.15	85.15	62.44
N	9,336	5,310	5,796	1,770
Panel B. Verbal Skills				
Accent × *AccRev*	-0.039***	-0.028**	-0.561***	-0.393
	(0.009)	(0.013)	(0.157)	(0.296)
Accent	0.001	-0.009	0.027	0.153
	(0.006)	(0.010)	(0.113)	(0.242)
AccRev	0.228	-0.285	2.792	1.782
	(0.256)	(0.369)	(4.712)	(5.128)
Average Outcome	9.43	9.33	90.89	63.62
N	8,328	5,292	4,800	1,764

Note. The table displays the estimated coefficients of *Accent* × *AccRev*, *Accent* and *AccRev* from equation (1) in column (1), equation (3) in column (2), and equation (4) in columns (3)–(4). The following covariates are included in the estimations but are excluded from the table for brevity: Worker age, education, and race and ethnicity: Asian and Hispanic (the reference group is White). For the estimation of equation (1), we include a year and an experiment dummy. A year dummy is also included when estimating equations (3) and (4) for the Signal Intervention evaluations. The latter also includes the value of the signal (Random-5). Column (2) includes the following additional covariates: *Signal*, *AccRev* × *Signal*, *Accent* × *Signal*, and *Accent* × *AccRev* × *Signal*. Standard errors clustered at the employer level in parentheses. * p<0.10, ** p<0.05, ***p<0.01.

Table A.13: Results after Controlling for Worker Preferences Fixed Effects

	Beliefs		WTP	
	(1)	(2)	(3)	(4)
	No Signal	w/Signal	No Signal	w/Signal
<u>Panel A. Mathematics Skills</u>				
Accent × *AccRev*	-0.337***	-0.515***	-2.972**	5.634
	(0.079)	(0.129)	(1.502)	(4.493)
Accent	-0.027	0.056	-1.257	0.184
	(0.068)	(0.104)	(1.451)	(3.402)
AccRev	0.063	0.554**	0.607	3.463
Average Outcome	8.40	8.15	85.15	62.44
N	9,336	5,310	5,796	1,770
<u>Panel B. Verbal Skills</u>				
Accent × *AccRev*	-0.662***	-0.605***	-7.138***	-10.728***
	(0.088)	(0.142)	(1.780)	(4.117)
Accent	-0.032	-0.101	-0.288	5.212
	(0.079)	(0.111)	(1.493)	(3.469)
AccRev	0.115	-0.303	-0.047	2.707
Average Outcome	9.43	9.33	90.89	63.62
N	8,328	5,292	4,800	1,764

Note. The table displays the estimated coefficients of *Accent* × *AccRev*, *Accent* and *AccRev* from equation (1) in column (1), equation (3) in column (2), and equation (4) in columns (3)–(4). The following covariates are included in the estimations but are excluded from the table for brevity: Worker age, education, and race and ethnicity: Asian and Hispanic (the reference group is White), dummy variables for worker preferences. For the estimation of equation (1), we include a year and an experiment dummy. A year dummy is also included when estimating equations (3) and (4) for the Signal Intervention evaluations. The latter also includes the value of the signal (Random-5). Column (2) includes the following additional covariates: *Signal*, *AccRev* × *Signal*, *Accent* × *Signal*, and *Accent* × *AccRev* × *Signal*. Estimations also include workers' preferences fixed effects. Standard errors clustered at the employer level in parentheses. * p<0.10, ** p<0.05, ***p<0.01.

Table A.14: Results after Controlling for Audio Clip Characteristics

	Beliefs		WTP	
	(1)	(2)	(3)	(4)
	No Signal	w/Signal	No Signal	w/Signal
Panel A. Mathematics Skills				
Accent × *AccRev*	-0.226***	-0.305**	-1.405	5.097
	(0.079)	(0.127)	(1.601)	(4.572)
Accent	0.000	0.083	-0.806	-3.915
	(0.050)	(0.077)	(1.000)	(2.827)
AccRev	0.170	0.782***	2.359	2.903
	(0.174)	(0.271)	(3.884)	(5.636)
Clip duration	-0.068***	-0.140***	-1.101**	0.326
	(0.021)	(0.033)	(0.493)	(0.866)
Paragraph Length	-0.006	-0.009	0.022	0.173
	(0.007)	(0.009)	(0.164)	(0.278)
Average Outcome	8.40	8.15	85.15	62.44
N	9,336	5,310	5,796	1,770
Panel B. Verbal Skills				
Accent × *AccRev*	-0.529***	-0.441***	-3.911**	-10.416**
	(0.096)	(0.142)	(1.826)	(4.158)
Accent	-0.075	-0.162*	-0.297	3.222
	(0.058)	(0.095)	(1.142)	(2.991)
AccRev	0.267	-0.111	3.161	2.381
	(0.238)	(0.340)	(4.539)	(4.882)
Clip Duration	-0.093***	-0.121***	-2.123***	0.043
	(0.027)	(0.033)	(0.495)	(0.619)
Paragraph Length	0.013	0.004	0.084	0.605**
	(0.009)	(0.009)	(0.187)	(0.257)
Average Outcome	9.43	9.33	90.89	63.62
N	8,328	5,292	4,800	1,764

Note. The table displays the estimated coefficients of *Accent* × *AccRev*, *Accent*, *AccRev*, length of the preference paragraph and the duration of the audio clip from equation (1) in column (1), equation (3) in column (2), and equation (4) in columns (3)–(4). The following covariates are included in the estimations but are excluded from the table for brevity: Worker age, education, and race and ethnicity: Asian and Hispanic (the reference group is White). For the estimation of equation (1), we include a year and an experiment dummy. A year dummy is also included when estimating equations (3) and (4) for the Signal Intervention evaluations. The latter also includes the value of the signal (Random-5). Column (2) includes the following additional covariates: *Signal*, *AccRev* × *Signal*, *Accent* × *Signal*, and *Accent* × *AccRev* × *Signal*. Standard errors clustered at the employer level in parentheses. * p<0.10, ** p<0.05, ***p<0.01.

Table A.15: Results after Controlling for Evaluation Order

	Beliefs		WTP	
	(1)	(2)	(3)	(4)
	No Signal	w/Signal	No Signal	w/Signal
Panel A. Mathematics Skills				
Accent × *AccRev*	-0.324***	-0.498***	-3.119**	5.693
	(0.079)	(0.125)	(1.478)	(4.447)
Accent	0.012	0.101	-0.329	-4.313
	(0.050)	(0.077)	(0.977)	(2.823)
AccRev	0.058	0.555**	0.649	3.381
	(0.166)	(0.258)	(3.856)	(5.379)
Evaluation No.	0.017*	-0.003	0.864***	1.487***
	(0.009)	(0.013)	(0.157)	(0.508)
Average Outcome	8.40	8.15	85.15	62.44
N	9,336	5,310	5,796	1,770
Panel B. Verbal Skills				
Accent × *AccRev*	-0.662***	-0.590***	-7.178***	-10.166**
	(0.088)	(0.140)	(1.729)	(4.128)
Accent	-0.052	-0.137	0.391	2.948
	(0.058)	(0.094)	(1.103)	(3.028)
AccRev	0.114	-0.312	-0.155	2.332
	(0.235)	(0.339)	(4.409)	(4.725)
Evaluation No.	0.005	-0.000	0.694***	0.373
	(0.009)	(0.012)	(0.180)	(0.461)
Average Outcome	9.43	9.33	90.89	63.62
N	8,328	5,292	4,800	1,764

Note. The table displays the estimated coefficients of *Accent* × *AccRev*, *Accent*, *AccRev* and evaluation order from equation (1) in column (1), equation (3) in column (2), and equation (4) in columns (3)–(4). The following covariates are included in the estimations but are excluded from the table for brevity: Worker age, education, and race and ethnicity: Asian and Hispanic (the reference group is White). For the estimation of equation (1), we include a year and an experiment dummy. A year dummy is also included when estimating equations (3) and (4) for the Signal Intervention evaluations. The latter also includes the value of the signal (Random-5). Column (2) includes the following additional covariates: *Signal*, *AccRev* × *Signal*, *Accent* × *Signal*, and *Accent* × *AccRev* × *Signal*. Estimations also include the order of evaluation as control. Standard errors clustered at the employer level in parentheses. * p<0.10, ** p<0.05, ***p<0.01.

Table A.16: Results after Controlling for the Sequence of the First 24 Workers Fixed Effects

	Beliefs		WTP	
	(1)	(2)	(3)	(4)
	No Signal	w/Signal	No Signal	w/Signal
Panel A. Mathematics Skills				
Accent × *AccRev*	-0.325***	-0.497***	-2.995**	3.991
	(0.079)	(0.125)	(1.501)	(4.119)
Accent	0.020	0.093	-0.516	-2.129
	(0.050)	(0.077)	(0.989)	(2.314)
AccRev	0.044	0.542**	-0.113	4.064
	(0.164)	(0.254)	(3.740)	(5.152)
Average Outcome	8.40	8.15	85.15	62.44
N	9,336	5,310	5,796	1,770
Panel B. Verbal Skills				
Accent × *AccRev*	-0.663***	-0.592***	-7.147***	-12.009***
	(0.088)	(0.140)	(1.755)	(3.917)
Accent	-0.061	-0.145	0.272	4.227*
	(0.057)	(0.093)	(1.115)	(2.385)
AccRev	0.085	-0.418	1.392	2.290
	(0.225)	(0.309)	(4.234)	(4.562)
Average Outcome	9.43	9.33	90.89	63.62
N	8,328	5,292	4,800	1,764

Note. The table displays the estimated coefficients of *Accent* × *AccRev*, *Accent* and *AccRev* from equation (1) in column (1), equation (3) in column (2), and equation (4) in columns (3)–(4). The following covariates are included in the estimations but are excluded from the table for brevity: Worker age, education, and race and ethnicity: Asian and Hispanic (the reference group is White). For the estimation of equation (1), we include a year and an experiment dummy. A year dummy is also included when estimating equations (3) and (4) for the Signal Intervention evaluations. The latter also includes the value of the signal (Random-5). Column (2) includes the following additional covariates: *Signal*, *AccRev* × *Signal*, *Accent* × *Signal*, and *Accent* × *AccRev* × *Signal*. Estimations include the 24-workers sequence fixed effects. Standard errors clustered at the employer level in parentheses. * p<0.10, ** p<0.05, ***p<0.01.

Table A.17: Results after Controlling for Employer Characteristics: Mathematics Skills

	Beliefs		WTP	
	(1)	(2)	(3)	(4)
	No Signal	w/Signal	No Signal	w/Signal
Accent × *AccRev*	-0.325***	-0.497***	-3.011**	5.815
	(0.079)	(0.125)	(1.494)	(4.374)
Accent	0.016	0.098	-0.431	-4.132
	(0.050)	(0.077)	(0.985)	(2.813)
AccentRev	0.032	0.543**	1.520	3.352
	(0.164)	(0.248)	(3.807)	(5.209)
Employer: Age	0.015**	0.022**	0.198	0.553***
	(0.006)	(0.010)	(0.154)	(0.192)
Employer: Female	-0.202	-0.047	-3.439	-8.004*
	(0.160)	(0.235)	(3.712)	(4.615)
Employer: Asian	0.599	1.180	-11.027	12.971
	(0.390)	(0.750)	(8.734)	(14.015)
Employer: Hispanic	0.204	0.676*	6.829	13.419
	(0.332)	(0.386)	(10.333)	(9.774)
Employer: Black	-0.024	0.861	3.592	3.503
	(0.395)	(0.765)	(7.943)	(14.581)
Employer: White	0.057	1.274*	9.234	10.416
	(0.368)	(0.727)	(7.179)	(13.582)
Employer: Bachelor's Degree	-0.569***	-0.736***	10.500**	5.453
	(0.162)	(0.274)	(4.401)	(4.849)
Employer: Lives in City	-0.182	-0.229	2.869	2.110
	(0.178)	(0.258)	(4.027)	(5.003)
Emploer: Multilingual	0.056	1.034***	9.287*	11.769*
	(0.215)	(0.301)	(4.734)	(6.722)
Employer: Works Full Time	-0.294*	0.062	-1.439	0.279
	(0.162)	(0.248)	(3.954)	(4.651)
Average Outcome	8.40	8.15	85.15	62.44
N	9,336	5,310	5,796	1,770

Note. The table displays the estimated coefficients in equation (1) in column (1), equation (3) in column (2), and equation (4) in columns (3)–(4). The following covariates are included in the estimations but are excluded from the table for brevity: Worker age, education, and race and ethnicity: Asian and Hispanic (the reference group is White). For the estimation of equation (1), we include a year and an experiment dummy. A year dummy is also included when estimating equations (3) and (4) for the Signal Intervention evaluations. The latter also includes the value of the signal (Random-5). Column (2) includes the following additional covariates: *Signal*, *AccRev* × *Signal*, *Accent* × *Signal*, and *Accent* × *AccRev* × *Signal*. The estimations also control for employer characteristics in Table 2.3. Standard errors clustered at the employer level in parentheses. * p<0.10, ** p<0.05, ***p<0.01.

Table A.18: Results after Controlling for Employer Characteristics: Verbal Skills

	Beliefs		WTP	
	(1)	(2)	(3)	(4)
	No Signal	w/Signal	No Signal	w/Signal
Accent × *AccRev*	-0.662***	-0.590***	-7.187***	-9.362**
	(0.088)	(0.140)	(1.746)	(4.017)
Accent	-0.054	-0.142	0.382	2.533
	(0.058)	(0.094)	(1.114)	(2.875)
AccRev	0.090	-0.403	0.205	0.583
	(0.224)	(0.328)	(4.400)	(4.600)
Employer: Age	0.024**	0.022*	0.237	0.247
	(0.010)	(0.013)	(0.164)	(0.177)
Employer: Female	0.469**	0.322	-2.157	2.788
	(0.219)	(0.309)	(4.342)	(4.436)
Employer: Asian	1.546***	1.143*	-0.106	-2.234
	(0.505)	(0.586)	(15.380)	(7.713)
Employer: Hisp	0.270	-0.010	16.995	12.795
	(0.508)	(0.678)	(13.524)	(9.691)
Employer: Black	-0.366	-0.324	10.899	1.230
	(0.577)	(0.797)	(13.140)	(7.904)
Employer: White	0.159	0.549	5.860	11.852*
	(0.511)	(0.647)	(12.771)	(6.922)
Employer: Bachelor's Degree	-1.159***	-1.110***	8.727*	-4.296
	(0.234)	(0.335)	(4.913)	(4.714)
Employer: Lives in City	-0.422*	0.324	0.940	16.958***
	(0.249)	(0.347)	(4.537)	(5.252)
Employer: Multilingual	-0.049	-0.383	-5.600	3.803
	(0.310)	(0.440)	(5.907)	(6.270)
Employer: Works Full Time	0.161	0.357	2.057	2.432
	(0.228)	(0.329)	(4.532)	(4.385)
Average Outcome	9.43	9.33	90.89	63.62
N	8,328	5,292	4,800	1,764

Note. The table displays the estimated coefficients from equation (1) in column (1), equation (3) in column (2), and equation (4) in columns (3)–(4). The following covariates are included in the estimations but are excluded from the table for brevity: Worker age, education, and race and ethnicity: Asian and Hispanic (the reference group is White). For the estimation of equation (1), we include a year and an experiment dummy. A year dummy is also included when estimating equations (3) and (4) for the Signal Intervention evaluations. The latter also includes the value of the signal (Random-5). Column (2) includes the following additional covariates: *Signal*, *AccRev* × *Signal*, *Accent* × *Signal*, and *Accent* × *AccRev* × *Signal*. The estimations also control for employer characteristics in Table 2.3. Standard errors clustered at the employer level in parentheses. * p<0.10, ** p<0.05, ***p<0.01.

Table A.19: Results after Controlling for the Employer Fixed Effects

	Beliefs		WTP	
	(1)	(2)	(3)	(4)
	No Signal	w/Signal	No Signal	w/Signal
Panel A. Mathematics Skills				
Accent × *AccRev*	-0.325***	-0.497***	-2.995**	2.734
	(0.079)	(0.125)	(1.494)	(3.530)
Accent	0.020	0.093	-0.516	-1.499
	(0.050)	(0.077)	(0.984)	(2.050)
Average Outcome	8.40	8.15	85.15	62.44
N	9,336	5,310	5,796	1,770
Panel B. Verbal Skills				
Accent × *AccRev*	-0.663***	-0.592***	-7.147***	-5.966**
	(0.088)	(0.140)	(1.745)	(2.836)
Accent	-0.061	-0.144	0.272	1.157
	(0.057)	(0.092)	(1.109)	(1.868)
Average Outcome	9.43	9.33	90.89	63.62
N	8,328	5,292	4,800	1,764

Note. The table displays the estimated coefficient of *Accent* × *AccRev* and *Accent* from equation (1) in column (1), equation (3) in column (2), and equation (4) in columns (3)–(4). *AccRev* is not estimated due to the between-subjects assignment of this treatment variable. to The following covariates are included in the estimations but are excluded from the table for brevity: Worker age, education, and race and ethnicity: Asian and Hispanic (the reference group is White). For the estimation of equation (1), we include a year and an experiment dummy. A year dummy is also included when estimating equations (3) and (4) for the Signal Intervention evaluations. The latter also includes the value of the signal (Random-5). Column (2) includes the following additional covariates: *Signal*, *AccRev* × *Signal*, *Accent* × *Signal*, and *Accent* × *AccRev* × *Signal*. The estimations also control for employer fixed effects. Standard errors clustered at the employer level in parentheses. * p<0.10, ** p<0.05, ***p<0.01.

Table A.20: Baseline Beliefs over Workers' Performance Detail (No Signal Experiment)

	(1) All	(2) Asian	(3) Hispanic	(4) White
Panel A. Mathematics Skills				
Accent × *AccRev*	-0.220**	0.078	-0.285*	-0.454**
	(0.101)	(0.159)	(0.172)	(0.179)
Average Beliefs	8.25	9.25	7.59	7.91
N	5,796	1,932	1,932	1,932
Panel B. Verbal Skills				
Accent × *AccRev*	-0.711***	-0.699***	-0.760***	-0.667***
	(0.112)	(0.177)	(0.195)	(0.184)
Average Beliefs	9.11	9.51	8.64	9.17
N	4,800	1,600	1,600	1,600

Note. The table displays the estimated coefficient of *Accent* × *AccRev* in equation (1). The following covariates are included in the estimations but are excluded from the table for brevity: *Accent*,*AccRev*, age, education, and race and ethnicity: Asian and Hispanic (the reference group is White). Columns: All include the sample of the employers' evaluations 1-12 for all workers; Asian include the subsample of the employers' evaluations 1-12 that were Asian workers; Hispanic include the subsample of the employers' evaluations 1-12 that were Hispanic workers; White include the sample of the employers' evaluations 1-12 that were White workers. Standard errors clustered at the employer level in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.21: Baseline Beliefs over Worker Performance Detail (Signal Intervention Experiment)

	(1) All	(2) Asian	(3) Hispanic	(4) White
Panel A. Mathematics Skills				
Accent × *AccRev*	-0.501***	-0.612***	-0.656***	-0.249
	(0.125)	(0.225)	(0.207)	(0.202)
Average Beliefs	8.65	9.77	7.85	8.34
N	3,540	1,180	1,180	1,180
Panel B. Verbal Skills				
Accent × *AccRev*	-0.589***	-0.701***	-0.579***	-0.475*
	(0.140)	(0.217)	(0.220)	(0.244)
Average Beliefs	9.86	10.36	9.38	9.84
N	3,528	1,176	1,176	1,176

Note. The table displays the estimated coefficient of *Accent* × *AccRev* in equation (1). The following covariates are included in the estimations but are excluded from the table for brevity: *Accent*,*AccRev*, age, education, a year dummy, and race and ethnicity: Asian and Hispanic (the reference group is White). Columns: All include the sample of the employer evaluations 1-12 for all workers; Asian, Hispanic and White include the subsample of the employer evaluations 1-12 that were Asian, Hispanic and White workers, respectively. Standard errors clustered at the employer level in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.22: Baseline Beliefs over Worker Performance: Signal Intervention Experiment, 2024 Subsample

	(1) All	(2) Asian	(3) Hispanic	(4) White
Panel A. Mathematics Skills				
Accent × *AccRev*	-0.485***	-0.617**	-0.614**	-0.225
	(0.153)	(0.284)	(0.246)	(0.246)
Average Beliefs	8.55	9.65	7.75	8.24
N	2,304	768	768	768
Panel B. Verbal Skills				
Accent × *AccRev*	-0.736***	-0.500*	-0.982***	-0.708**
	(0.167)	(0.256)	(0.261)	(0.273)
Average Beliefs	9.93	10.52	9.39	9.87
N	2,292	764	764	764

Note. The table displays the estimated coefficient of *Accent* × *AccRev* in equation (1). The estimations include evaluations 1-12 from the Signal Intervention experiment sample collected in 2024. The following covariates are included in the estimations but are excluded from the table for brevity: *Accent*, *AccRev*, age, education, and race and ethnicity: Asian and Hispanic (the reference group is White). Columns: All include the sample of the employer evaluations 1-12 for all workers; Asian, Hispanic and White include the subsample of the employer evaluations 1-12 that were Asian, Hispanic and White workers, respectively. Standard errors clustered at the employer level in parentheses. N represents the number of observations. * p<0.10, ** p<0.05, *** p<0.01.

Table A.23: Response of Beliefs to Informative Signal: Signal Intervention Experiment, 2024 Subsample

	(1) All	(2) Asian	(3) Hispanic	(4) White
Panel A. Mathematics Skills				
Accent × *AccRev* × *Signal*	0.479	0.169	1.097	0.175
	(0.376)	(0.825)	(0.748)	(0.665)
Accent × *AccRev*	-0.486***	-0.608**	-0.615**	-0.227
	(0.153)	(0.284)	(0.247)	(0.246)
Signal	-1.581***	-1.370***	-1.708***	-1.650***
	(0.249)	(0.499)	(0.382)	(0.325)
Average Beliefs	8.01	9.16	7.32	7.58
N	3,456	1,124	1,151	1,181
Panel B. Verbal Skills				
Accent × *AccRev* × *Signal*	-0.074	-0.111	0.140	-0.280
	(0.410)	(0.820)	(0.744)	(0.621)
Accent × *AccRev*	-0.742***	-0.500*	-1.022***	-0.706**
	(0.167)	(0.258)	(0.260)	(0.273)
Signal	-1.645***	-1.463***	-1.217***	-2.286***
	(0.248)	(0.464)	(0.442)	(0.330)
Average Beliefs	9.39	9.94	8.93	9.31
N	3,438	1,130	1,161	1,147

Note. The table displays the estimated coefficients of *Accent* × *AccRev*, *Accent* and *AccRev* in equation (3). The estimations include evaluations 1-18 from the Signal Intervention experiment sample collected in 2024. The following covariates are included in the estimations but are excluded from the table for brevity: *Accent*, *AccRev*, *AccRev* × *Signal*, *Accent* × *Signal*, age, education, and race and ethnicity: Asian and Hispanic (the reference group is White). Columns: All include the sample of the employer evaluations 1-18 for all workers; Asian include the subsample of the employer evaluations 1-18 that were Asian, Hispanic and White workers, respectively. Standard errors clustered at the employer level in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A.24: Willingness to Pay (in cents) with Signal: Signal Intervention Experiment, 2024 Subsample

	(1) All	(2) Asian	(3) Hispanic	(4) White
Panel A. Mathematics Skills				
Accent × *AccRev*	5.813	5.676	6.036	3.872
	(5.194)	(9.405)	(8.119)	(9.235)
Average WTP	59.27	75.65	48.23	53.58
N	1,152	400	425	327
Panel B. Verbal Skills				
Accent × *AccRev*	-8.695*	-6.783	-6.899	-12.191
	(4.969)	(9.595)	(7.763)	(8.634)
Average WTP	59.48	72.79	51.51	54.85
N	1,146	377	435	334

Note. The table displays the estimated coefficient of *Accent* × *AccRev* in equation (4). The estimations include evaluations 19-24 from the Signal Intervention experiment sample collected in 2024. The following covariates are included in the estimations but are excluded from the table for brevity: *Accent*, *AccRev*, age, education, and race and ethnicity: Asian and Hispanic (the reference group is White). The controls also include the value of the signal (Random-5). Columns: All include the sample of the employer evaluations 19-24 for all workers; Asian, Hispanic and White include the subsample of the employer evaluations 19-24 that were Asian, Hispanic and White workers, respectively. Standard errors clustered at the employer level in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A.25: Results after Excluding Employers with Outlier Time in the Experiment

	Beliefs		WTP	
	(1)	(2)	(3)	(4)
	No Signal	w/Signal	No Signal	w/Signal
Panel A. Mathematics Skills				
Accent × *AccRev*	-0.320***	-0.511***	-3.205**	5.362
	(0.083)	(0.133)	(1.584)	(4.662)
Average Outcome	8.38	8.17	85.84	61.11
N	8,508	4,896	5,244	1,632
Panel B. Verbal Skills				
Accent × *AccRev*	-0.679***	-0.679***	-7.070***	-12.473***
	(0.092)	(0.146)	(1.777)	(4.210)
Average Outcome	9.49	9.52	90.83	63.87
N	7,644	4,770	4,464	1,590

Note. The table displays the estimated coefficient of *Accent* × *AccRev* from equation (1) in column (1), equation (3) in column (2), and equation (4) in columns (3)–(4). The sample excludes employers whose experiment time is below or above the 5th and the 95th percentiles. These percentiles are based on the total time within each experiment. The following covariates are included in the estimations but are excluded from the table for brevity: *Accent*, *AccRev*, age, education, and race and ethnicity: Asian and Hispanic (the reference group is White). For the estimation of equation (1), we include a year and an experiment dummy. A year dummy is also included when estimating equations (3) and (4) for the Signal Intervention evaluations. The latter also includes the value of the signal (Random-5). Column (2) includes the following additional covariates: *Signal*, *AccRev* × *Signal*, *Accent* × *Signal*, and *Accent* × *AccRev* × *Signal*. The estimations also control for employer fixed effects. Standard errors clustered at the employer level in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

A.4 Heterogeneity Analysis

In this section, we present a series of analyses that explore whether our results are heterogeneous across demographic characteristics. In Section A.4.1, we show the results by worker characteristics, and in Section A.4.2, we show the heterogeneity of the results by employer characteristics.

A.4.1 Worker characteristics

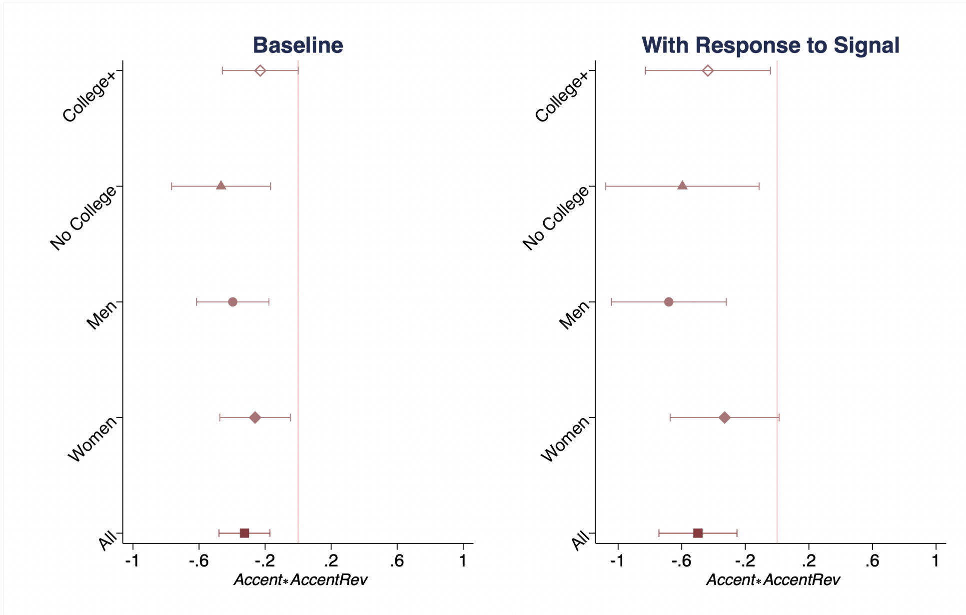
Figures A.5 and A.6 replicate our main results by the worker demographic characteristics. Figure A.5 presents the results for Math, and Figure A.6 presents the results for Verbal. Within each figure, Panel A presents the results for the beliefs over worker performance, with the baseline beliefs in the subfigure on the left, followed by the response to the information signal in the right subfigure. Panel B presents the results of baseline WTP (left subfigure) and with signals (right subfigure). For each figure, the first results present those for the complete sample for reference, followed by the results by gender, with results for women first and men second. The fourth and fifth points present the results by worker educational attainment, limited to workers with less than a college degree and those with a college degree or more, in that order.

Beliefs over worker productivity. As shown in A.5–Panel A, the accent gap in Mathematics skills treatment is similar across worker gender but is slightly larger for workers who have less than a college degree. While the results with the response to signal are similar, it qualitatively suggests that the information signal might reduce the accent gap for women and those with less than a college degree. Thus, the results suggest that the accent gap is driven in part by inaccurate beliefs about worker performance due to stereotypes based on their demographic characteristics.

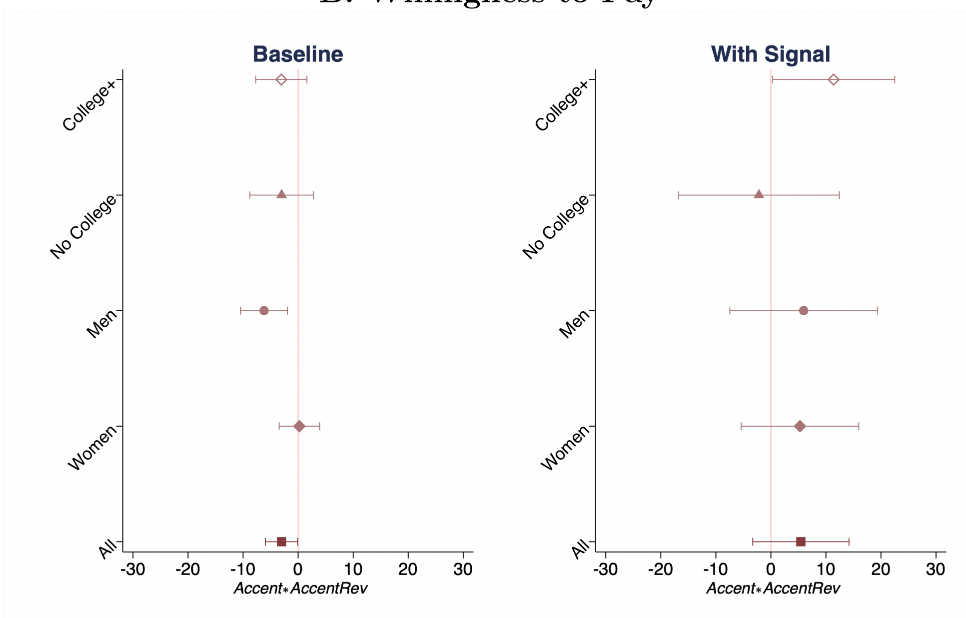
For Verbal skills, the baseline coefficients of interest in A.6–Panel A are very similar across worker characteristics. In line with our main results, the information signal has a muted effect on beliefs. This is largely true across the different worker groups.

Figure A.5: Heterogeneity by Worker Characteristics: Mathematics Skills

A. Beliefs over Worker Performance



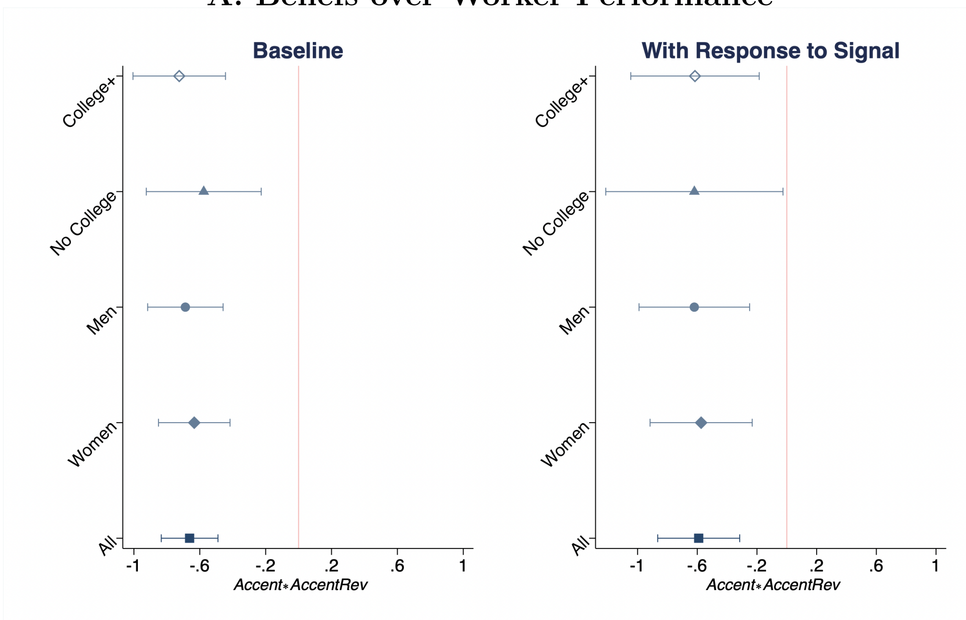
B. Willingness to Pay



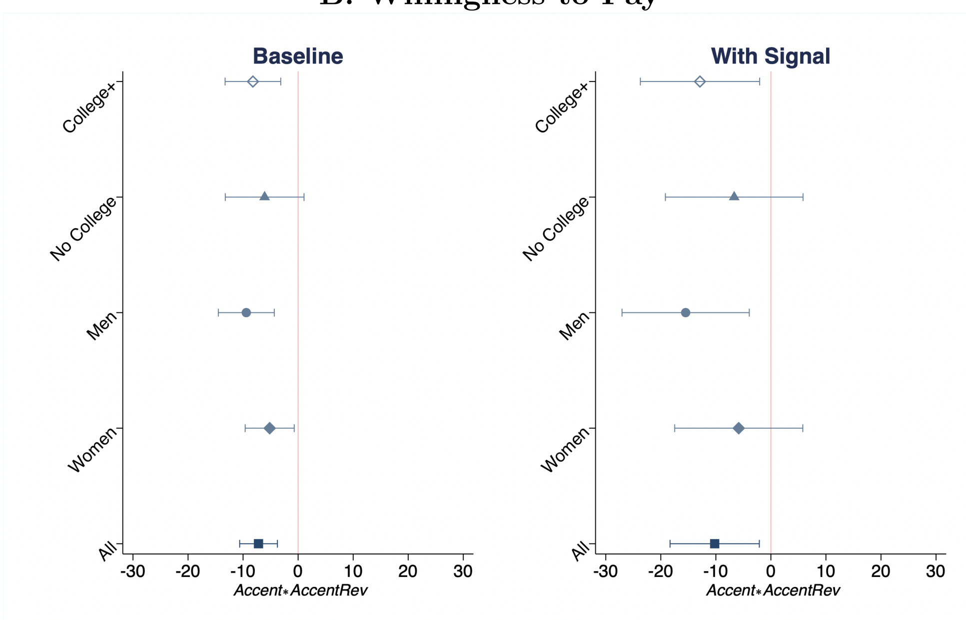
Notes: The figure displays the estimated coefficients and 95% confidence intervals of $Accent \times AccRev$ for the Mathematics skills by worker characteristics. Panel A shows the results of estimating equation (1) in the first graph (evaluations 1-12, both experiments), followed by the results of estimating equation (3) (evaluations 1-18, Signal Intervention experiment). Panel B presents the results from estimating equation (4), with the first and second graphs showing the results from the No Signal experiment (evaluations 13-24) and the Signal Intervention experiment (evaluations 19-24). The following covariates are included in the estimations: *Accent*, *AccRev*, age, education, and race and ethnicity: Asian and Hispanic (the reference group is White). For the estimation of equation (1), we include a year and an experiment dummy. A year dummy is also included when estimating equations (3) and (4) for the Signal Intervention evaluations. The latter also includes the value of the signal (Random-5).

Figure A.6: Heterogeneity by Worker Characteristics: Verbal Skills

A. Beliefs over Worker Performance



B. Willingness to Pay



Notes: The figure displays the estimated coefficients and 95% confidence intervals of *Accent* × *AccRev* for the Verbal skills by worker characteristics. Panel A shows the results of estimating equation (1) in the first graph (evaluations 1-12, both experiments), followed by the results of estimating equation (3) (evaluations 1-18, Signal Intervention experiment). Panel B presents the results from estimating equation (4), with the first and second graphs showing the results from the No Signal experiment (evaluations 13-24) and the Signal Intervention experiment (evaluations 19-24). The following covariates are included in the estimations: *Accent*, *AccRev*, age, education, and race and ethnicity: Asian and Hispanic (the reference group is White). For the estimation of equation (1), we include a year and an experiment dummy. A year dummy is also included when estimating equations (3) and (4) for the Signal Intervention evaluations. The latter also includes the value of the signal (Random-5).

Willingness to pay. The results in Figures A.5 and A.6 suggest that the accent gap in baseline WTP is more salient for men than women in both Mathematics and Verbal skills. We also see that there is a slightly larger accent gap in WTP for Verbal skills for individuals with a college degree or more. While the results are not significant for other groups in Mathematics, the results are in fact significant across the board for Verbal skills.

Moreover, once the signal is added, the results in Figure A.5 shows that the increase in WTP post-signal for accented individuals in Math is mainly driven by workers with a college or higher education. As discussed in Section 4.1.2, it appears likely that employers use the signal to overshoot their WTP for workers who they believe are "high skill." This interpretation is in line with the heterogeneous results, showing a positive and significant effect for higher educational attainment workers. Furthermore, the results for Verbal skills in Figure A.6 reveal that the persistent accent gap post-signal is more salient for men and college graduates. The fact that the signal has a more muted impact on WTP for male workers is in line with the gender stereotype that women are better at Verbal skills. Combining this with the result that men receive a positive WTP accent gap in Math, these overall results suggest that the worker's accent intensifies the gender stereotypes.

A.4.2 Employer characteristics

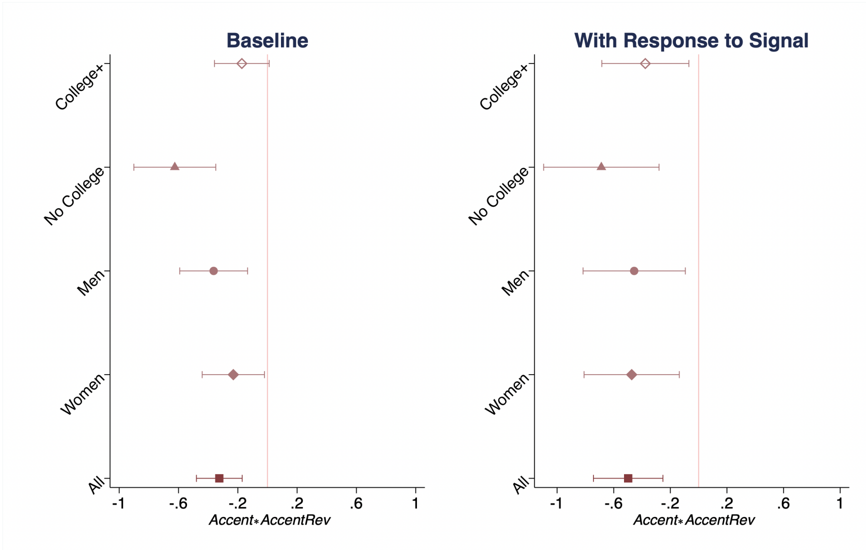
Figures A.7 and A.8 display the results when employer demographic characteristics are the focus. The figures format is the same as they were with worker characteristics, displaying the results for Math in Figure A.7 and for Verbal in Figure A.8. As with the workers, we are showing the results by the employer's gender and educational attainment.[38]

Beliefs over worker productivity. In Figures A.7 and A.8, we observe a more salient accent gap on the beliefs about worker performance when employers have an education of less than a college degree for Mathematics skills, and when they are women or have less than a college degree for Verbal skills. Therefore, women and less educated employers tend to believe that, when the worker's accent is revealed, accented workers perform worse than non-accented ones in the skills tests. Beliefs accent gaps are qualitatively similar when taking into account the observations when the signal is available.

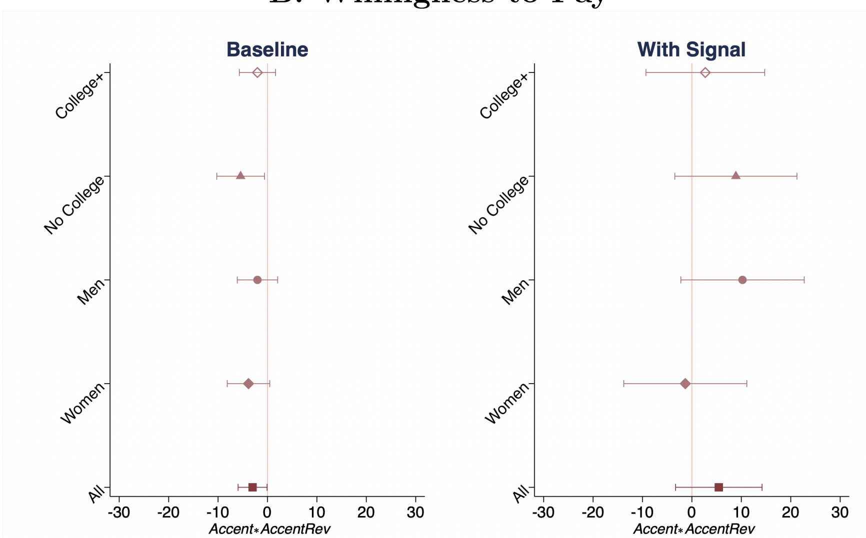
[38]While the workers are limited to individuals who reported their gender as being either male or female, the employers were not limited to it. The small number of employers who responded being Nonbinary/Nonconforming or Prefer not to say are then excluded from the analysis by gender.

Figure A.7: Heterogeneity by Employer Characteristics: Mathematics Skills

A. Beliefs over Worker Performance



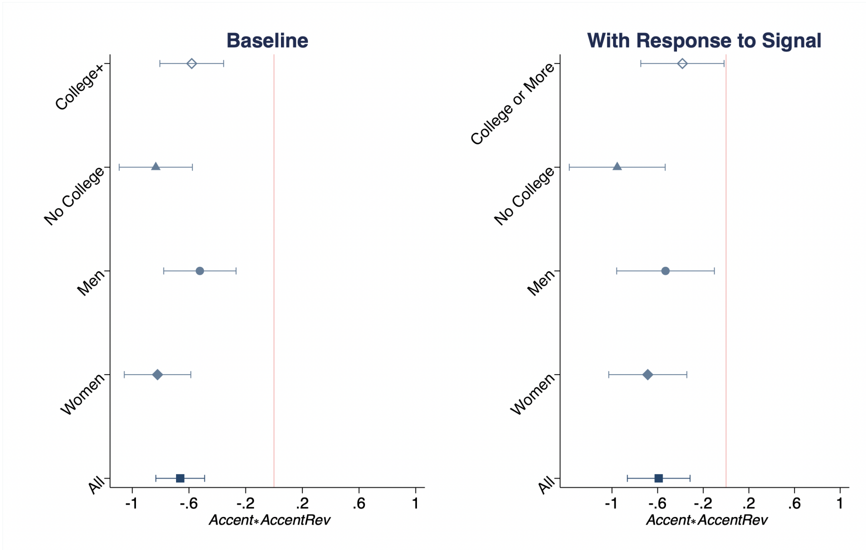
B. Willingness to Pay



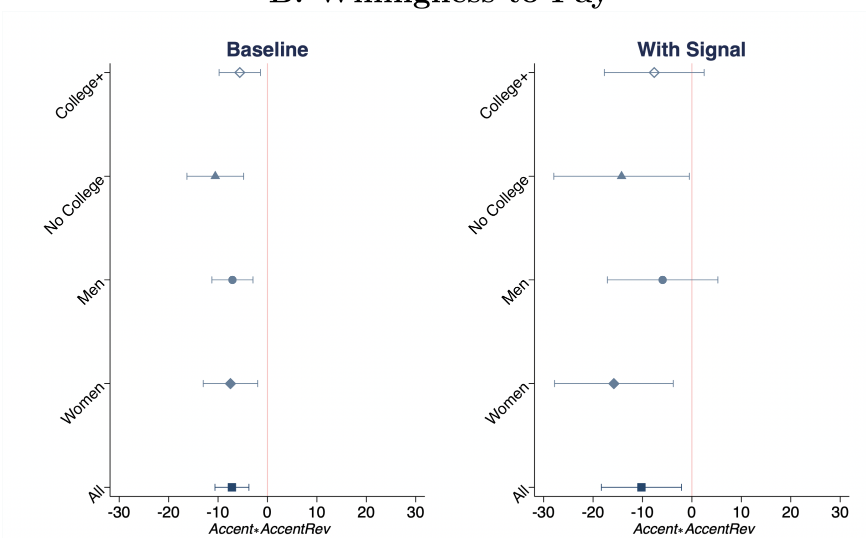
Notes: The figure displays the estimated coefficients and 95% confidence intervals of $Accent \times AccRev$ for the Mathematics skills by employer characteristics. Panel A shows the results of estimating equation (1) in the first graph (evaluations 1-12, both experiments), followed by the results of estimating equation (3) (evaluations 1-18, Signal Intervention experiment). Panel B presents the results from estimating equation (4), with the first and second graphs showing the results from the No Signal experiment (evaluations 13-24) and the Signal Intervention experiment (evaluations 19-24). The following covariates are included in the estimations: *Accent*, *AccRev*, worker age, education, and race and ethnicity: Asian and Hispanic (the reference group is White). For the estimation of equation (1), we include a year and an experiment dummy. A year dummy is also included when estimating equations (3) and (4) for the Signal Intervention evaluations. The latter also includes the value of the signal (Random-5).

Figure A.8: Heterogeneity by Employer Characteristics: Verbal Skills

A. Beliefs over Worker Performance



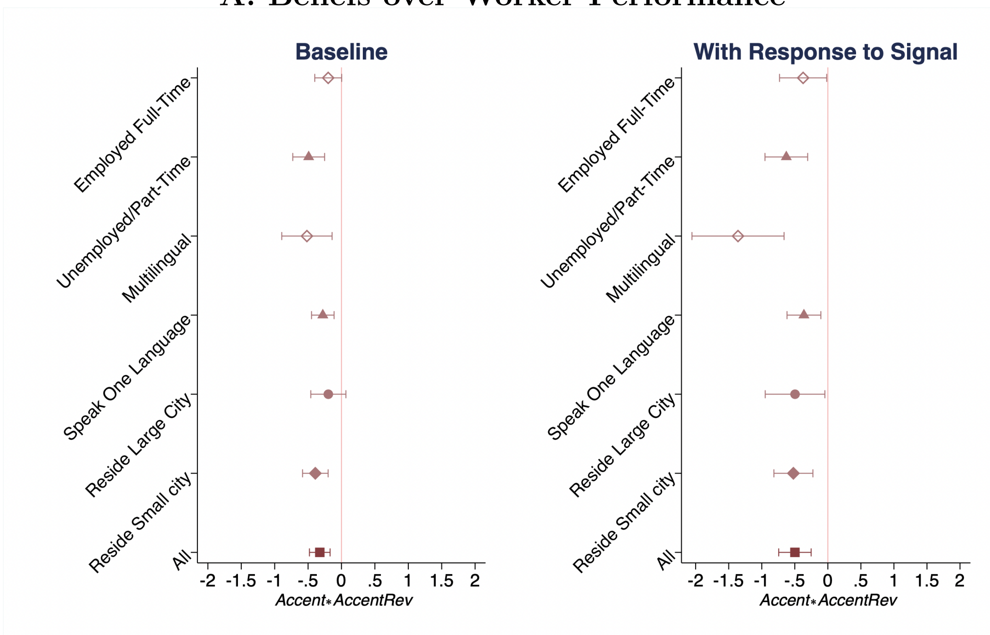
B. Willingness to Pay



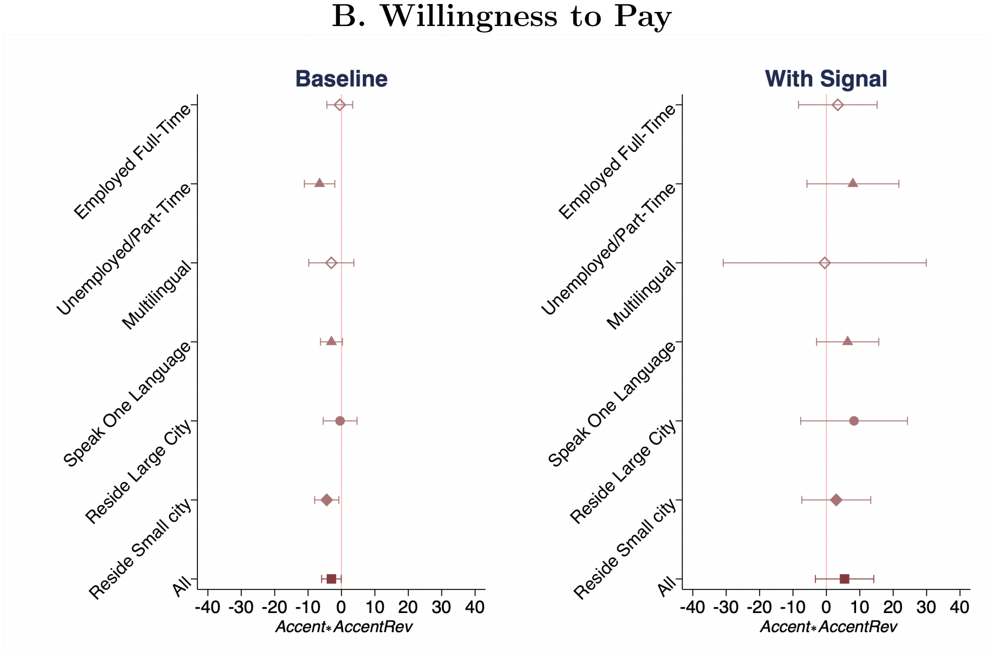
Notes: The figure displays the estimated coefficients and 95% confidence intervals of $Accent \times AccRev$ for the Verbal skills by employer characteristics. Panel A shows the results of estimating equation (1) in the first graph (evaluations 1-12, both experiments), followed by the results of estimating equation (3) (evaluations 1-18, Signal Intervention experiment). Panel B presents the results from estimating equation (4), with the first and second graphs showing the results from the No Signal experiment (evaluations 13-24) and the Signal Intervention experiment (evaluations 19-24). The following covariates are included in the estimations: *Accent*, *AccRev*, worker age, education, and race and ethnicity: Asian and Hispanic (the reference group is White). For the estimation of equation (1), we include a year and an experiment dummy. A year dummy is also included when estimating equations (3) and (4) for the Signal Intervention evaluations. The latter also includes the value of the signal (Random-5).

Figure A.9: Heterogeneity by Additional Employer Characteristics: Mathematics Skills

A. Beliefs over Worker Performance



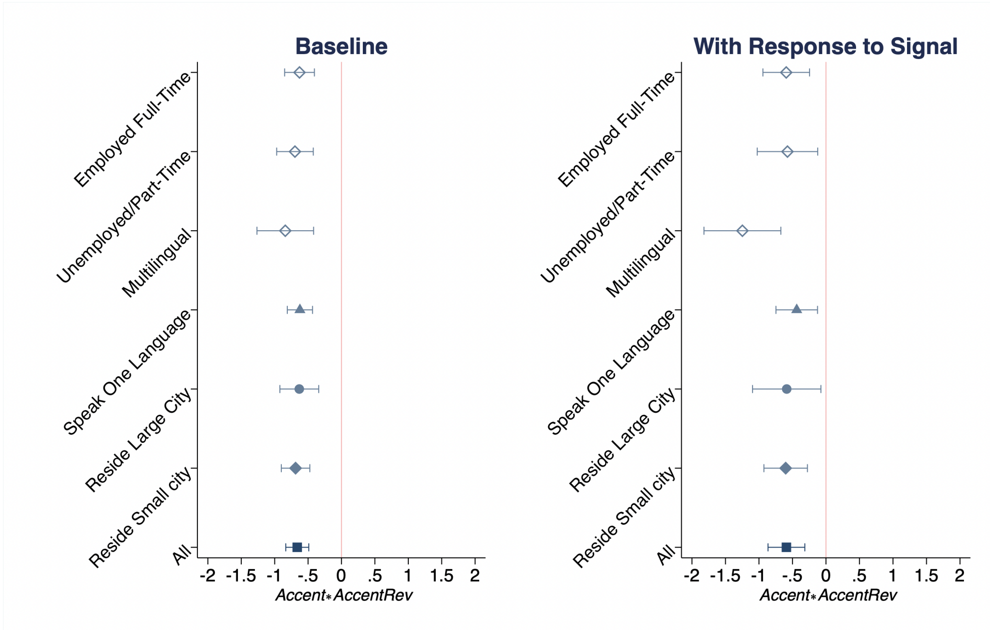
B. Willingness to Pay



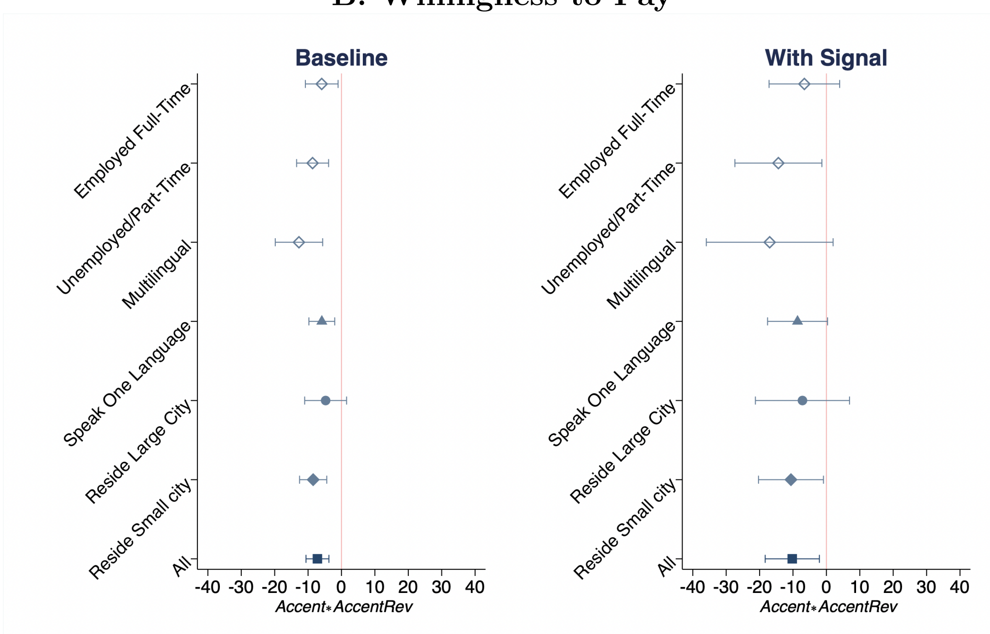
Notes: The figure displays the estimated coefficients and 95% confidence intervals of $Accent \times AccRev$ for the Mathematics skills by employer additional characteristics. Panel A shows the results of estimating equation (1) in the first graph (evaluations 1-12, both experiments), followed by the results of estimating equation (3) (evaluations 1-18, Signal Intervention experiment). Panel B presents the results from estimating equation (4), with the first and second graphs showing the results from the No Signal experiment (evaluations 13-24) and the Signal Intervention experiment (evaluations 19-24). All regressions include a full set of control variables. For the estimation of equation (1), we include a year and an experiment dummy. A year dummy is also included when estimating equations (3) and (4) for the Signal Intervention evaluations. The latter also includes the value of the signal (Random-5).

Figure A.10: Heterogeneity by Additional Employer Characteristics: Verbal Skills

A. Beliefs over Worker Performance



B. Willingness to Pay



Notes: The figure displays the estimated coefficients and 95% confidence intervals of $Accent \times AccRev$ for the Verbal skills by employer additional characteristics. Panel A shows the results of estimating equation (1) in the first graph (evaluations 1-12, both experiments), followed by the results of estimating equation (3) (evaluations 1-18, Signal Intervention experiment). Panel B presents the results from estimating equation (4), with the first and second graphs showing the results from the No Signal experiment (evaluations 13-24) and the Signal Intervention experiment (evaluations 19-24). All regressions include a full set of control variables. For the estimation of equation (1), we include a year and an experiment dummy. A year dummy is also included when estimating equations (3) and (4) for the Signal Intervention evaluations. The latter also includes the value of the signal (Random-5).

Willingness to pay. In terms of the employer WTP, the results in Figures A.7 and A.8 show that the employers who believe accented workers perform worse than non-accented workers (employers who are women and those less than college educated) are also driving the baseline WTP accent gaps in both Mathematics and Verbal skills tests. These same employer characteristics, after observing the signal, are responsible for the WTP accent gaps for performance in Verbal skills test.

Here, we concentrate on looking at the heterogeneity by employer gender and educational attainment. Results exploring additional characteristics are available in Figures A.9 andA.10. Overall, results are qualitatively similar across different dimensions. Nevertheless, one could highlight that employers who might be more exposed to accents (i.e., those who reside in large cities) tend to penalize accented workers less relative to non-accented workers.

A.5 Experimental Instructions

No Signal Experiment, Mathematics Skills Accent Revealed treatment. ***[Modifications for No Signal Experiment, Mathematics Skills Accent Blind treatment]***

General Instructions

Thank you for participating in this study, which will take approximately 30 minutes.

All information collected will be anonymous, ensuring that nobody can link your identity to your decisions. To maintain this privacy, do not share your decisions with Prolific users.

Studies run by [...] do NOT use deception.This means that everything in the instructions is true,including the rules and the payoffs, and you will evaluate real people.

You will need to answer understanding questions correctly to proceed. Additionally, attention check questions will be asked after listening to audio recordings. **Please ensure you have access to a well functioning speaker and/or headphones in a quiet environment.** If you cannot meet these requirements, please return the study. **Failing more than 5 attention checks** will terminate your participation.

Background

In a prior study, participants from Prolific ("workers") completed a survey and a 15-question test measuring their mathematics skills, focusing on arithmetic and algebra. The average score was 7.1 points out of 15. Here is an example of a question:

Solve for x: $6(x-3) + 9 = 3(x+1)$

The workers **all currently live in the U.S.** A worker's score on the test equals the number of questions they answered correctly and determined their earnings ($0.10 times their score).

Your decisions and payoffs

In this study, you will make predictions after reviewing the profiles of **24 different workers** and answer survey questions about your demographics and decisions. The workers will be randomly assigned and presented to you in **two different parts.**

After finishing all parts, you will receive $4.50 as completion payment. Additionally, your prediction regarding one worker will be randomly selected as the **decision-that-counts.** You will receive any amount you earn (up to $4.20) for that decision as **bonus.**

Q. Which of the following statements is true?

- ○ Your bonus payment will be determined based on your decisions regarding multiple workers.
- ○ There is only one part in this study.
- ○ The test workers took had 10 questions.
- ○ You will review the profile of 10 workers in total.
- ○ If you fail more than 5 attention checks, your participation from the study will be terminated.

Part 1

In Part 1, you will be randomly assigned **12 workers.** Your task is to predict how many questions each worker correctly answered out of **15 in the test.** As a reminder,

- The workers currently live in the U.S.
- The workers took a **math skills** test. An example for the questions in the test:

Solve for x: 6(x-3) + 9 = 3(x+1)

- The average score on the test was **7.1 points** out of 15.

Before you make your decision for each worker:

- You will see the Curriculum Vitae (CV), which may help evaluate their skills.
- The CV information is self-reported by the worker.
- You will listen to an audio recording of the worker stating his/her preferences for beverages, desserts, and colors.

Please note that **for privacy reasons,** the voices in the audio files you will listen to were slightly altered using a software, but the content, **pronunciation and intonation of the original recordings have been preserved.**

[*Please note that **for privacy reasons,** the audio clips you will listen to were generated using a software, so, the voices in the audio clips do not belong to the workers. However, the clips reflect the workers' preferences.*]

To enter your predictions, you will first respond to attention check questions based on the worker's profile and the audio recording.

Your Payoff

If **your prediction** about a worker's score is:

- **Correct:** Your payoff will be $4.20.
- **Within 3 points:** You will have a chance to win $4.20 with a higher chance the closer your prediction is to the worker's actual score.

You will submit your response on a slider that ranges from 0 to 15. **We are using a payment rule that gives you the best chance of winning $4.20 when you report your most accurate guess.** If you would like to view the details of this payment rule, click the bar below and scroll down.

Q. Which of the following statements is true?

- ○ Multiple workers in this part will be chosen for your bonus payment.
- ○ Some workers live outside the U.S.
- ○ For workers' privacy, the voices in the audio recordings are digitally altered versions of the original voices of the workers. The altered recordings still contain the workers' other speech characteristics such as pronunciation and intonation. [*For workers' privacy, the audio clips for workers were created using a software but they reflect the workers' preferences.*]
- ○ Your chance of earning $4.20 is the same regardless of what you submit as your prediction.

Click here and scroll down to review the rules that determine your payoff

We will use your response in the slider to calculate how much your prediction differs from the actual score of the worker. If your prediction is within 3 points of the worker's correct score, then you will have a chance of winning $4.20. Specifically, your response will determine which of the following $4.20-prize bets you will enter:

Bet #	Difference (Your Prediction vs. Worker's Actual Score)		Your Chance of Winning $4.20
0	0		100% chance
1	1		89% chance
2	2		56% chance
3	3		1% chance
4	4		0% chance
5	5		0% chance
.	.	.	.
.	.	.	.
14	14		0% chance
15	15		0% chance

Note that after we have your answer, we will calculate your prediction's deviation from the worker's actual score to determine which bet you will enter. We will also draw a random whole number (integer) between 1 and 100 for you, where each number (including 1 and 100) is equally likely to be drawn.

- If the bet you enter gives you a larger than zero chance of winning \$4.20 and if your random number is:
  - Smaller than or equal to the % chance of winning in the bet you enter, you will win \$4.20.
  - Larger than the % chance of winning in the bet you enter, your payoff will be \$0.
- If the bet you enter gives you a zero chance of winning \$4.20, your payoff will be \$0.

Example. Suppose that you predict the score to be 12 and the worker's actual score is 13. Then, we would calculate your prediction's deviation as 1 and enter you in Bet 1:

Bet #	Deviation (Your Prediction, Worker's Actual Score)		Your Chance of Winning \$4.20
0	0		100% chance
1	1		89% chance
2	2		56% chance
3	3		1% chance
4	4		0% chance
5	5		0% chance
.	.	.	.
.	.	.	.
14	14		0% chance
15	15		0% chance

In Bet 1 since your chance of winning \$4.20 is 89%, your payoff depends on how your random integer compares to 89. Specifically, if your random integer is smaller than or equal to 89, then your payoff is \$4.20 and \$0 otherwise.

Part 2

You are randomly **assigned 12 new workers, all residing in the U.S.**

Your task is to **submit the highest wage you are willing to pay to "hire" each worker.** If you hire a worker and that worker is chosen as your worker-that-counts, they will be given a bonus (the wage that is chosen based on your decision). In return, you will receive money based on their test score. Specifically:

- **If you hire a worker:**
  - You will earn **14 cents** for each question the worker answered correctly in the test and pay the worker's wage:

    → Your Profit = (\$0.14 * Worker's Test Score) - Wage

→ Worker's Bonus = Wage

- **If you do NOT hire a worker:**
  - You will NOT earn any additional revenue and also NOT pay anything to the worker

    → Your Profit = $0
    → Worker's Bonus = $0

How will we determine the wage?

You will submit your response on a slider that ranges from $0 to $2.10 in 7-cent increments.

Note that the **wage** is **not** necessarily going to be the amount you submitted: we will set the worker's wage as a random multiple of 7-cents between 0*and*$2.10(*witheachvaluebeingequallylikelytobedrawn*)*andcompareittowhatyousubmitted* :

- If the wage is the **higher** amount, because you indicated to us that you are not willing to pay it, you will NOT hire the worker and will earn zero profit.
- If the wage is the **lower** amount, you will hire the worker **at the wage** and collect the profit you have earned.

With the above rule, submitting a high amount as the highest wage you are willing to pay to the worker increases your chances of hiring the worker, but at the same time, it can lead you to hire at a wage you are not willing to pay to the worker. So, to avoid undesirable outcomes – such as not hiring the worker at a wage you are willing to pay or hiring the worker at a wage higher than what you are willing to pay – you should submit exactly what the worker is worth to you, not lower or higher.

Please also note that if this part is chosen for your bonus payment, we are going to automatically give you $2.10 in addition to the profit you make with your hiring decision, so that there is no way you end up owing us any money at the end of the study.

As a summary:

- You will report us the highest wage you are willing to pay to each worker.
- **The actual wage for each worker will be determined randomly** and you will hire the worker **only if** this random wage is smaller than the highest wage you are willing to pay to that worker.

- If you hire the worker:

  → Your Profit = ($0.14 * Worker's Test Score) - Wage)
  → Worker's Bonus = Wage

- If you do not hire the worker:

  → Your Profit = $0
  → Worker's Bonus = $0

Just to make sure you understand, imagine that you submitted $0.49 as the highest wage you are willing to pay for a worker whose test score is 10.

Q1. If the randomly generated wage is $0.07, how much would you have to pay to the worker? ____ Hint: Think about whether the wage you entered as your highest willing to pay allows you to hire this worker, and if so, how much you would need to pay them as wage.

Q2. If the randomly generated wage is $0.07,how much profit would you make? ____ Hint: Think about whether the wage you entered as your highest willing to pay allows you to hire this worker. If not, your profit will be $0. It it allows you to hire the worker, calculate how much money you would earn given their test score (14 cents*score) and subtract their random wage from that.

Q3. If the randomly generated wage is $1.05, how much would you have to pay to the worker? ____ Hint: If the random wage is larger than the amount you submit, you do not pay it and do not hire the worker.

Click here and scroll down if you would like to examine the details and explanation of the hiring and wage determination:

We will use your willingness to pay response in the slider to fill in your responses to each of the following 31 questions about hiring:

Q#	**If the wage is:**	**Option A: Hire**		**Option B: Don't Hire**
0	**$0.00**, would you rather:	Pay **$0.00** and earn revenue = $0.14*(test score)	or	Don't pay anything and not hire this worker
1	**$0.07**, would you rather:	Pay **$0.07** and earn revenue = $0.14*(test score)	or	Don't pay anything and not hire this worker
2	**$0.14**, would you rather:	Pay **$0.14** and earn revenue = $0.14*(test score)	or	Don't pay anything and not hire this worker
3	**$0.21**, would you rather:	Pay **$0.21** and earn revenue = $0.14*(test score)	or	Don't pay anything and not hire this worker
4	**$0.28**, would you rather:	Pay **$0.28** and earn revenue = $0.14*(test score)	or	Don't pay anything and not hire this worker
5	**$0.35**, would you rather:	Pay **$0.35** and earn revenue = $0.14*(test score)	or	Don't pay anything and not hire this worker
.	.	.	.	.
.	.	.	.	.
29	**$2.03**, would you rather:	Pay **$2.03** and earn revenue = $0.14*(test score)	or	Don't pay anything and not hire this worker
30	**$2.10**, would you rather:	Pay **$2.10** and earn revenue = $0.14*(test score)	or	Don't pay anything and not hire this worker

Note that you are actually asked to choose either Option A (Pay $x wage and hire this worker) or Option B (Don't pay anything and not hire this worker) for each question. After we have your answers to all 31 questions, we will randomly draw one question("question-that-counts") and your payoff will be determined based on the option you chose on that one question. Each wage (question) is equally likely to be chosen for determining your payoff. Note that you have no incentive to lie on any question, because if that question gets picked for your payoff, then you would end up with the option you want less.

We assume that you might choose the Option A in at least the first few questions but at some point switch to choosing Option B. So, to save time, we ask you to just tell us at which wage (question) you would switch. We then fill out your answers to all 31 questions based on your switch point: choosing Option A for all questions before and at your switch point, and Option B for all questions after your switch point.We will still draw one question randomly for determining your payoff. Again, if you lie about your switch point you might end up with an option that you want less.

Example. Suppose that you chose $0.21. Then we would fill in your answers as:

Q#	If the wage is:	Option A		Option B
0	$0.00, would you rather:	Pay $0.00 and earn revenue = $0.14*(test score)	or	Don't pay anything and not hire this worker
1	$0.07, would you rather:	Pay $0.07 and earn revenue = $0.14*(test score)	or	Don't pay anything and not hire this worker
2	$0.14, would you rather:	Pay $0.14 and earn revenue = $0.14*(test score)	or	Don't pay anything and not hire this worker
3	$0.21, would you rather:	Pay $0.21 and earn revenue = $0.14*(test score)	or	Don't pay anything and not hire this worker
4	$0.28, would you rather:	Pay $0.28 and earn revenue = $0.14*(test score)	or	Don't pay anything and not hire this worker
5	$0.35, would you rather:	Pay $0.35 and earn revenue = $0.14*(test score)	or	Don't pay anything and not hire this worker
.	.	.	.	.
.	.	.	.	.
29	$2.03, would you rather:	Pay $2.03 and earn revenue = $0.14*(test score)	or	Don't pay anything and not hire this worker
30	$2.10, would you rather:	Pay $2.10 and earn revenue = $0.14*(test score)	or	Don't pay anything and not hire this worker

A.6 Instructions about signal in the Signal Intervention Experiment

(After participants submit their beliefs about 12 workers in Part 1, they receive the following message about their future decisions.)

A subset of 5 questions from the 15 on the test has been chosen.

For the remaining 6 workers that you will evaluate in this part, you will see each worker's score out of these 5 questions.

These selected questions, called the “Random-5”, **are the same for each worker.** Click below to see Worker 13’s profile.

A.7 Experiment Interface Screenshots

Figure A.11: Baseline Beliefs Elicitation

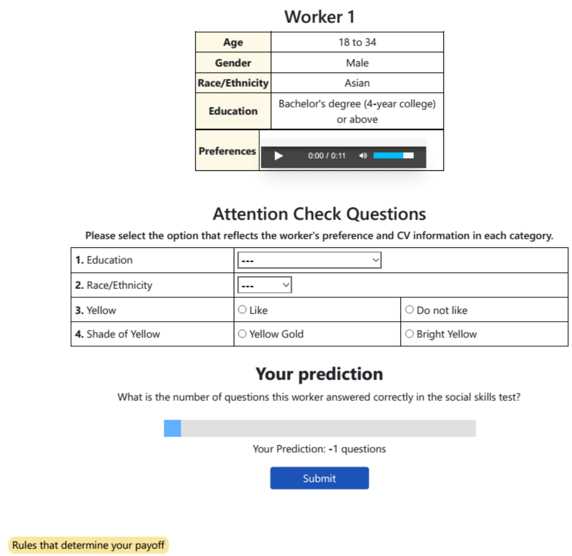


Figure A.12: Belief Elicitation with Random-5 Signal: Signal Intervention Experiment

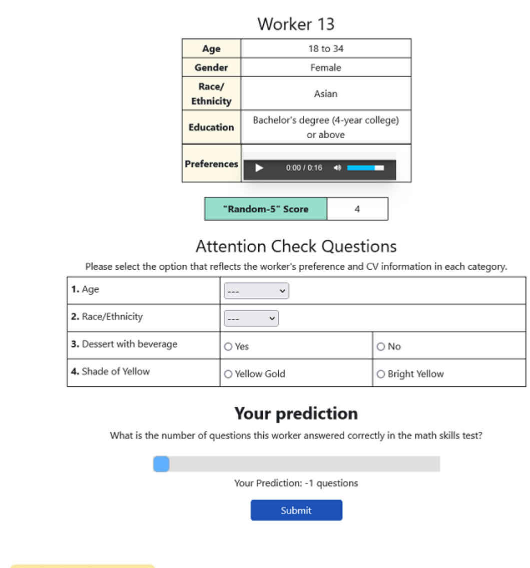


Figure A.13: WTP Elicitation: No Signal Experiment

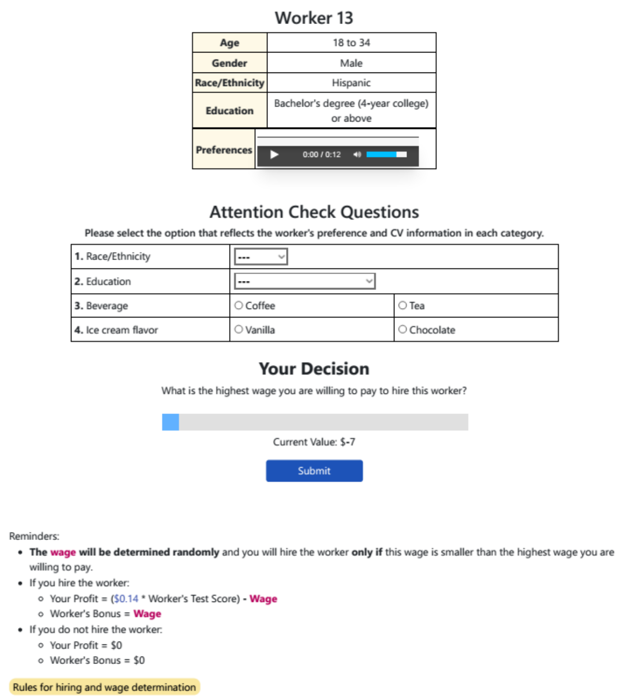


Figure A.14: WTP Elicitation: Signal Intervention Experiment