

7528787-discriminated-but-not-discouraged-investment-responses-to-algorithmic-discrimination

SUMMARY: This study examines how individuals adapt their investment behavior when an algorithm evaluates groups differently. Participants were randomly assigned to an advantaged or disadvantaged identity group and were asked to invest to improve their chances of being hired by an algorithm with different prior beliefs across groups. Disadvantaged participants invested 17.7% more than advantaged participants, indicating no discouragement in investment despite facing less favorable evaluation. The study then tested two interventions: "Role Model" treatments providing information about the behavior and outcomes of same-group predecessors, and a "No Risk" treatment that conditioned payment of the investment cost on being hired. The "Role Model" treatments had limited effects, while the "No Risk" treatment increased investment among disadvantaged participants by 24%. This effect was significantly larger than the corresponding effect among advantaged participants. The findings suggest that individuals can respond to algorithmic disadvantage by increasing rather than reducing investment, and that changing the economic consequences of investment generates a stronger behavioral response than same-group social information.

METHOD: The study used a pre-registered online experiment with 553 participants. Participants were randomly assigned to an advantaged or disadvantaged group and were asked to invest to improve their chances of being hired by an algorithm with different prior beliefs across groups. The "Role Model" and "No Risk" interventions were tested to see if they could increase investment among disadvantaged participants.

KEY FINDINGS:

- Disadvantaged participants invested 17.7% more than advantaged participants.
- "Role Model" treatments had limited effects.
- "No Risk" treatment increased investment among disadvantaged participants by 24%.
- The effect was significantly larger among disadvantaged participants than among advantaged participants.

IMPLICATION FOR TRADING: The findings suggest that individuals can respond to algorithmic disadvantage by increasing rather than reducing investment. Practitioners can use these insights to design interventions that change the economic consequences of investment, such as offering guarantees or incentives, to generate stronger behavioral responses. However, same-group social information may not be as effective in encouraging investment in the face of algorithmic disadvantage.

Discriminated but not Discouraged: Investment Responses to Algorithmic Discrimination*

Yung-Shiang Jasmine Yang[†]

Abstract

We study how individuals adapt their investment when an algorithm evaluates groups differently. In a pre-registered online experiment ($N = 553$), participants are randomly assigned to an advantaged or disadvantaged identity group and choose how much to invest to improve their probability of being hired by an algorithm that holds different prior beliefs across groups. Disadvantaged participants invest 17.7% more than advantaged participants, showing no evidence of discouragement in investment despite facing less favorable evaluation. We then test two interventions: “Role Model” treatments providing information about the behavior and outcomes of same-group predecessors, and a “No Risk” treatment that conditions payment of the investment cost on being hired. The “Role Model” treatments have limited effects, whereas “No Risk” increases investment among disadvantaged participants by 24%. This effect is significantly larger than the corresponding effect among advantaged participants. Our findings show that individuals can respond to algorithmic disadvantage by increasing rather than reducing investment, and that changing the economic consequences of investment generates a stronger behavioral response than same-group social information.

JEL classification: C91, D91, J71

Keywords: algorithmic bias, statistical discrimination, unequal opportunity, investment behavior, algorithmic hiring, online experiment

*I thank Emmanuelle Auriol, Thomas Buser, Haochi Chen, Vessela Daskalova, Paolo Falco, Marco Fongoni, Diogo Geraldes, Emma N. Guirado Forment, Oliver Hauser, Aditi Kharb, Boon Han Koh, Kevin Kloiber, Manuel E. Lago, Lucie Martin, Ivan Pastine, Deirdre Robertson, Mikhail Safronov, Margaret Samahita, Nikita Sangwan, Nora Strecker, Vasilisa Werner, Kun Wu, Bin Zhang, Yang Zhong, participants at the 2025 European Doctoral Group in Economics (EDGE) Conference, the University College Dublin School of Economics PhD and Research & Reading Group Seminars, the 2026 Irish Postgraduate and Early Career Economics (IPECE) Workshop, the 2026 Irish Economic Association (IEA) Annual Conference, the 50th International Association for Research in Economic Psychology (IAREP) Conference, the 3rd Annual Conference of the Network of Behavioural and Experimental Economists based in the UK (BEE UK), the 1st Minorities in Economics Mentoring Conference, and the EEA ESEM Congress 2026 for detailed feedback, helpful suggestions and stimulating conversations. All study procedures of the experiment described were approved by the ethics committee at University College Dublin (081-HS-LR-25-Yang-Samahita). The experiment and hypotheses were pre-registered on AsPredicted under ID #243245 [<https://aspredicted.org/f8jk-5xns.pdf>]. Funding from University College Dublin School of Economics and the Irish Economic Association is gratefully acknowledged.

[†]School of Economics, University College Dublin, Belfield, Dublin 4, Ireland; E-mail: yung-shiang.yang@ucdconnect.ie

1 Introduction

Algorithms increasingly shape access to economic opportunities, particularly in hiring.¹ Employers use data-driven tools to screen applicants, predict job performance, and support hiring decisions (Ajunwa & Greene 2019, Bogen & Rieke 2018). Such systems offer the prospect of more consistent decision-making and may reduce some forms of human bias (Houser 2019). At the same time, they can generate unequal opportunities when predictions depend on group membership or on data that reflect existing inequalities (Barocas & Selbst 2016, Binns 2018). This concern is closely related to statistical discrimination, in which group-level beliefs affect the evaluation of otherwise similar individuals (Phelps 1972, Arrow et al. 1973). Importantly, these beliefs need not be accurate: inaccurate group-level beliefs can themselves generate discriminatory treatment (Bohren et al. 2025), while non-Bayesian updating can amplify disadvantage even when informative individual signals are available (Campos-Mercade & Mengel 2024). Algorithmic systems may embed such disparities in decision rules that can be applied repeatedly and at scale. A prominent example is a recruiting algorithm developed and subsequently abandoned by Amazon, which was trained on historical applications and learned to penalize indicators associated with women (Hamilton 2018, Vincent 2018).

Unequal evaluation matters not only because it affects who is ultimately selected, but also because individuals can change their behavior in anticipation of how they will be evaluated. A worker who expects to face a less favorable hiring rule must decide whether improving her qualifications remains worthwhile. Existing theories and evidence point to competing behavioral responses. Statistical discrimination can become self-reinforcing when disadvantaged workers respond to lower expected returns by investing less in education or skills (Lundberg & Startz 1983, Coate & Loury 1993); experimentally induced non-Bayesian discrimination can similarly deter education investment among high-ability disadvantaged workers (Campos-Mercade & Mengel 2024). Discriminatory treatment can also reduce subsequent effort and performance (Glover et al. 2017, Heiserman & Simpson 2023). Yet disadvantage need not lead to disengagement. Individuals may instead invest more in an attempt to compensate for unfavorable evaluation. Perceived discrimination has been associated with greater subsequent investment in education and training (Caputo 2002, 2007), while qualitative evidence documents pressure among disadvantaged individuals to work “twice as hard” or to “over-effort” in response to stereotypes and unequal treatment (Evans et al. 2019, Isik et al. 2021, Gutman & Younas 2025). Whether unequal evaluation discourages investment or instead induces greater investment is therefore an empirical question with direct implications for how discrimination shapes the accumulation of skills and the persistence of inequality.

How individuals respond to unequal evaluation may, however, depend on whether the evaluator is human or algorithmic. A growing literature shows that individuals perceive algorithmic and human evaluators differently: workers form different beliefs about how the two use performance and personal characteristics (Fumagalli

¹We use hiring algorithms as the running example throughout the paper and experiment, although algorithmic decision-making extends to other high-stakes domains. For example, see Lee & Floridi (2021) on machine-learning algorithms in credit risk evaluation and mortgage lending, Kleinberg et al. (2018) on algorithmic predictions in bail decisions, and Cai et al. (2019) on machine learning in medical diagnosis.

et al. 2022), applicants often perceive fully algorithmic recruitment as less fair than human or human-assisted procedures (Lavanchy et al. 2023), and otherwise comparable discrimination can generate different moral reactions when attributed to an algorithm rather than a person (Bigman et al. 2023). Recent work further shows that these perceptions can translate into labor-market behavior: awareness of bias in a hiring algorithm can discourage applications from the disadvantaged group (Ip 2025), while disclosure of AI-based job matching can reduce jobseekers' willingness to participate (Bana & Boudreau 2025). Yet these studies primarily examine whether individuals choose an algorithmic evaluator or whether they enter an algorithmically mediated opportunity. Much less is known about what happens once individuals are already subject to unequal algorithmic evaluation. We ask whether individuals adapt their costly investment in response to an algorithm that evaluates groups differently: does algorithmic disadvantage discourage investment, thereby potentially reinforcing initial inequality, or does it induce individuals to invest more in an attempt to overcome the disadvantage?

To answer this question, we conduct a pre-registered incentivized online experiment ($N = 553$) that simulates a stylized hiring process with algorithmic screening. Participants are randomly assigned to either a disadvantaged group or an advantaged group and learn that the hiring algorithm evaluates the two groups differently. Before investment, participants in both groups have an equal 50% probability of becoming a top performer. Participants then decide how much to invest in improving their hiring prospects. Higher investment increases the probability of becoming a top performer, which in turn increases the probability of receiving a more favorable pre-employment test score.

The hiring algorithm does not observe participants' true productivity directly. Instead, it observes the pre-employment test score, which provides a noisy signal of productivity, and evaluates this signal differently depending on group membership. Consequently, participants assigned to the disadvantaged group face lower hiring probabilities than otherwise equivalent participants in the advantaged group. Participants are informed that the algorithm favors the advantaged group, as well as how investment affects productivity and how test scores translate into group-specific hiring decisions. Hiring outcomes nevertheless remain probabilistic because both productivity and test scores are stochastic.

We next examine whether individuals' responses to unequal algorithmic evaluation can be influenced by two distinct interventions. The first provides information about the behavior and outcomes of previous participants from the same group. In the "Successful Role Model" and "Unsuccessful Role Model" treatments, participants observe the investment decisions and hiring outcomes of five predecessors from their own identity group.² The second intervention changes the economic consequences of investment. In the "No Risk" treatment, participants pay their investment cost only if they are hired, reducing the downside consequences of investing while also increasing its expected payoff.³

²The treatment labels "Successful Role Model" and "Unsuccessful Role Model" are retained for consistency with our pre-registration. Because participants observe relatively sparse information about anonymous same-group predecessors rather than richer biographical narratives, we interpret these interventions as providing same-group social information with potential role-model and social-comparison channels.

³The treatment is motivated by the logic of income-contingent financing, under which repayment depends on subsequent economic outcomes (Friedman 1955, Chapman 2006).

We find that participants assigned to the disadvantaged group invest 17.7% more than participants assigned to the advantaged group. The “Role Model” treatments have little effect on investment in our main specification, whereas the “No Risk” treatment increases investment among disadvantaged participants by approximately 24%. Moreover, the effect of “No Risk” is significantly stronger among disadvantaged than advantaged participants. These findings indicate that changing the economic consequences of investment generates a substantially stronger behavioral response than the same-group social-information interventions in our setting. To provide additional descriptive evidence on the factors participants considered when making their decisions, we elicit self-reported motivations in a post-experiment questionnaire. Group identity is the most frequently reported motivation, followed by the potential bonus. Among participants exposed to predecessor information, a nontrivial share also report considering predecessors’ investment decisions or hiring outcomes. Exploratory analyses further show that the association between reporting group identity as a motivation and investment differs significantly between disadvantaged and advantaged participants.

Our study contributes to several strands of literature. First, we contribute to the literature on statistical discrimination and behavioral responses to unequal evaluation. Classical models of statistical discrimination show how group membership can affect decisions when evaluators use group-level information to infer otherwise unobserved individual productivity (Arrow et al. 1973, Fang & Moro 2011). More recent work emphasizes that discriminatory outcomes may also arise when group-level beliefs are themselves inaccurate or distorted. Bordalo et al. (2016), for example, show how representativeness can generate systematically distorted stereotypes, while Bohren et al. (2025) formalize inaccurate statistical discrimination and show that it can be difficult to distinguish from taste-based or accurate statistical discrimination. Relatedly, Campos-Mercade & Mengel (2024) show that non-Bayesian updating can amplify discrimination against disadvantaged groups, and Patty & Penn (2023) highlight the importance of behavioral feedback between algorithmic decision rules and the characteristics individuals choose to develop.

Our design captures a stylized form of inaccurate statistical discrimination: participants in the two identity groups have identical underlying productivity distributions, but the hiring algorithm applies less favorable prior beliefs to the Blue group. Rather than studying how such beliefs arise, how they are learned, or how evaluators use them, we exogenously impose the group-specific priors and study costly investment decisions by those subject to the resulting unequal evaluation. In this respect, our paper is related to models in which discrimination feeds back into workers’ incentives to acquire human capital (Lundberg & Startz 1983, Coate & Loury 1993) and to experimental work studying investment under statistical discrimination (Dianat et al. 2022), but differs in focusing on behavioral responses to an explicitly inaccurate and transparent algorithmic evaluation rule.

The resulting investment response runs counter to the discouragement mechanism emphasized in self-fulfilling-prophecy models, in which lower expected returns to investment can reduce skill acquisition among disadvantaged groups (Lundberg & Startz 1983, Coate & Loury 1993), and to evidence that discriminatory treatment can reduce worker effort or performance (Glover et al. 2017, Heiserman & Simpson 2023). Instead, disadvantaged participants invest more than advantaged participants. This pattern is consistent with evidence that perceived discrimination can be

associated with greater subsequent investment in education and training (Caputo 2002, 2007) and with qualitative evidence of “over-efforting” or “twice-as-hard” responses among disadvantaged groups (Evans et al. 2019, Isik et al. 2021, Gutman & Younas 2025). Because our design does not directly identify the psychological mechanism underlying the higher investment of Blue participants, we interpret these findings as consistent with, rather than direct evidence of, a compensatory or over-efforting response to disadvantage.

Second, our study contributes to the growing literature on how individuals perceive and respond to algorithmic decision-makers, particularly in hiring and personnel selection. A substantial body of work documents negative applicant reactions to algorithmic evaluation. AI-based selection procedures are often perceived as less procedurally or interactionally fair than human evaluation (Acikgoz et al. 2020), and algorithmic personnel decisions may be viewed as overly reductionistic because they appear to omit qualitative or contextual information (Newman et al. 2020). Similarly, applicants tend to perceive fully algorithm-driven recruitment as less fair than either human-only or human-assisted procedures (Lavanchy et al. 2023). These perceptions also shape preferences over evaluators. Workers distinguish between human and algorithmic recruiters according to how they believe each uses performance and personal information (Fumagalli et al. 2022), while aversion to algorithmic hiring depends on features such as gender profiling and workers’ self-confidence (Dargnies et al. 2024). Algorithmic discrimination can also be judged differently from otherwise comparable discrimination by human decision-makers (Bigman et al. 2023).

More recent work examines whether algorithmic evaluation changes labor-market behavior rather than only perceptions or evaluator preferences. Informing prospective applicants that interviews will be evaluated by AI can reduce application and job-pursuit intentions (Mirowska 2020), while awareness of potential gender bias in hiring algorithms can deter women from applying for competitive positions (Ip 2025). In field settings, disclosure that AI is used to match workers to jobs can likewise reduce market participation (Bana & Boudreau 2025). Our study examines a different behavioral margin. Individuals in our experiment do not choose whether to be evaluated by an algorithm or whether to enter the market; algorithmic evaluation is exogenously imposed. We instead study how much costly investment individuals undertake after entering a known, unequal algorithmic evaluation environment. This allows us to examine behavioral adaptation to algorithmic discrimination before the hiring decision is made, complementing work on evaluator preferences, application decisions, and market participation.

Third, we contribute to the literature on interventions designed to alter behavior under unequal opportunity. One class of interventions changes the economic incentives surrounding investment or selection. Income-contingent financing can reduce the downside risk associated with human-capital investment by tying repayment to future income (Chapman 2006), while affirmative-action policies can alter the returns to investment for disadvantaged individuals and, in experimental settings, increase study effort and performance among those targeted by the policy (Cotton et al. 2022). Related laboratory evidence shows that hiring subsidies can substantially reduce statistical discrimination by changing firms’ incentives to hire disadvantaged workers (Dianat et al. 2022). Our “No Risk” treatment operates on a different margin: rather than subsidizing evaluators or changing hiring thresholds, it directly changes the consequences of investment for the individual by waiving the

investment cost when the participant is not hired.

A second class of interventions provides information about the behavior or outcomes of relevant others. Peer information can affect decisions through social learning and imitation: investors respond to information about peers' financial choices (Bursztyn et al. 2014), imitate observed investment decisions (Delfino et al. 2016), and alter behavior when predecessor choices provide information about what others have done (Cavatorta et al. 2024). These effects need not translate into consequential action, however; recent field evidence shows that peer investment information can increase engagement without increasing actual investment take-up (Becker et al. 2026).

Closely related role-model and social-comparison interventions emphasize the motivational consequences of observing successful or unsuccessful comparison targets. Exposure to a fictitious successful labor-market entrant can increase inspiration, identification, and proactive career behavior relative to exposure to an unsuccessful target (Buunk et al. 2007), while positive and negative role models can motivate behavior through distinct success-seeking and failure-avoidance channels (Lockwood et al. 2005). In field settings, exposure to successful same-gender role models can alter consequential education and labor-market choices among underrepresented groups: successful female economics graduates increase women's subsequent enrollment in economics and their likelihood of majoring in the subject (Porter & Serra 2020), while female role-model interventions have also been shown to increase job applications and task performance (Martini & Urueña 2025). At the same time, role-model interventions are not uniformly effective. Among marginalized ethnic minorities, Narciso et al. (2026) find only limited and transient behavioral effects despite some evidence that social relatability matters.

The remainder of the paper proceeds as follows. Section 2 describes the experimental design. Section 3 presents the main results. Section 4 concludes.

2 Experimental Design

The experiment has two objectives. First, we test whether participants randomly assigned to a disadvantaged group invest differently from those assigned to an advantaged group under unequal algorithmic evaluation. Second, we test the effect of two classes of interventions: information about previous in-group participants ("Successful Role Model" and "Unsuccessful Role Model") and a financial intervention that reduces downside risk exposure associated with investment ("No Risk"). We begin by describing the Control treatment, which establishes the benchmark environment. We then introduce the design of the treatments before describing the experimental implementation.

Control treatment: Does random group assignment affect investment?

Figure 1 summarizes the sequence of the Control treatment. In the experiment, participants act as job seekers while the hiring algorithm acts as the potential employer. Below, we describe in detail the information available to participants and the hiring algorithm, the decisions made by each party, and the resulting payoffs.

The Participant's Information. Participants are randomly assigned to either the Blue or Green identity group, denoted by $d_i \in \{b, g\}$. Before investment, all participants have the same baseline probability of becoming a top performer, regardless of group identity:

$$P(y_i = 1 \mid d_i = b) = P(y_i = 1 \mid d_i = g) = 0.5,$$

where $y_i \in \{0, 1\}$ denotes the participant's subsequently realized productivity type, with $y_i = 1$ denoting a top performer and $y_i = 0$ an average performer.

Participants are informed that, despite the identical baseline productivity probabilities across groups, the hiring algorithm believes that Green participants are more likely to be top performers than Blue participants.⁴ Participants are also informed that the algorithm's hiring decisions are based on a pre-employment test score generated probabilistically from their subsequently realized productivity, as summarized in Figure 1.

The Participant's Choice. Participants receive an initial endowment I and choose how much of it to invest, I_i , before taking the pre-employment test. Each additional unit of investment increases the probability of becoming a top performer by 1.5 percentage points and therefore increases the probability of receiving a more favorable test score according to the group-blind testing technology shown in Figure 1. Investment is costly and deducted from the participant's endowment. If hired by the algorithm, participants receive wage ω ; otherwise, they receive zero.

The Hiring Algorithm's Information and Choice. The hiring algorithm observes each participant's identity group and pre-employment test score, but not their productivity type. The algorithm receives payoff $\pi > 0$ from hiring a top performer, payoff $-\omega < 0$ from hiring an average performer, and zero from not hiring. Throughout the experiment, we set $\pi = 1.5\omega$, implying that the algorithm maximizes expected payoffs by hiring participants whose posterior probability of being a top performer is no less than an exogenously imposed threshold of 0.4. The derivation of this threshold is provided in Appendix A.1.

The resulting hiring rule is summarized in Figure 1. Participants receiving score A are always hired, whereas those receiving score C are never hired. For the intermediate score B, hiring depends on group identity: Green participants are hired whereas Blue participants are not, reflecting the algorithm's less favorable prior beliefs about the Blue group.

After the hiring decision is made, participants learn whether they are hired and receive their corresponding payoff.

⁴The algorithm's group-based beliefs are intended to capture a stylized form of inaccurate statistical discrimination rather than taste-based discrimination. In our setting, the algorithm has no intrinsic preference over groups and seeks only to maximize expected hiring payoffs; group identity enters solely through beliefs about productivity. These beliefs are exogenously imposed to isolate behavioral responses to algorithmic statistical discrimination, abstracting from how such beliefs arise in practice (e.g., through biased training data or historical disparities).

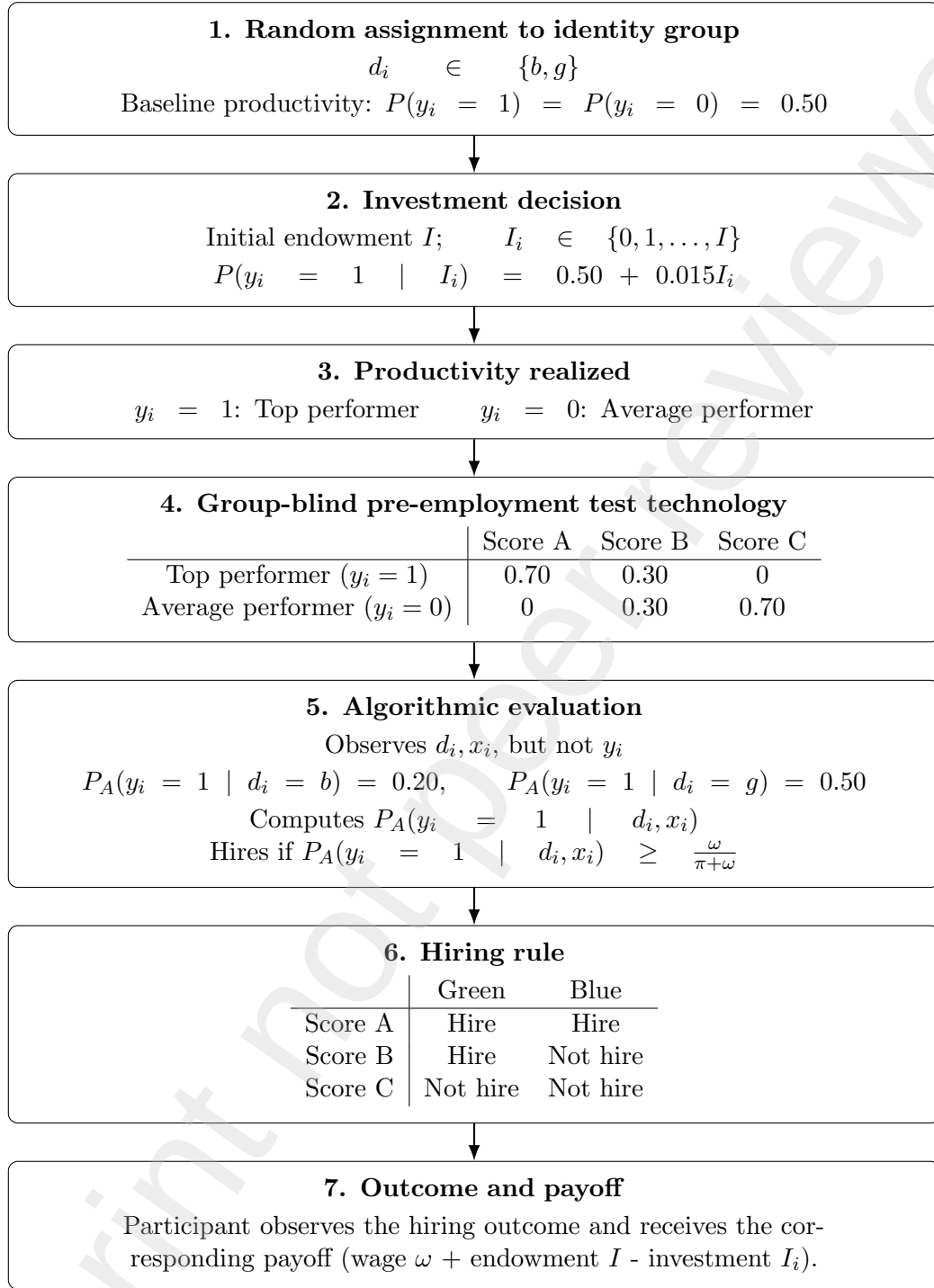


Figure 1: Sequence of the Control treatment

Note: $P_A(\cdot)$ denotes the hiring algorithm's beliefs. Participants in both identity groups have the same baseline probability of becoming a top performer. Investment increases this probability before productivity is realized and the group-blind testing technology generates a noisy test score. The hiring algorithm then combines the test score with its group-specific prior beliefs to form a posterior belief about productivity. Participants are informed that the algorithm evaluates Green participants more favorably than Blue participants and are shown the resulting hiring rule, but are not shown the numerical prior beliefs. The derivation of the algorithm's hiring rule is provided in Appendix A.1.

Treatment arms: How can participants be encouraged to invest?

The experimental procedure is identical to that in the Control treatment except for the following variations.

“Role Model” treatments: Prior to making their investment decisions, participants are shown the investment decisions and outcomes of five pilot-study participants from their own identity group. Participants are randomly assigned to one of the following two information conditions: the “Successful Role Model” treatment or the “Unsuccessful Role Model” treatment. In the “Successful Role Model” treatment, participants observe previous in-group participants who invested positive amounts and were hired. In the “Unsuccessful Role Model” treatment, participants observe previous in-group participants who invested positive amounts but were not hired. Figure A1 presents screenshots of all four possible information screens shown to participants. The decisions and outcomes displayed in these screens are based on actual pilot-study participants.⁵

“No Risk” treatment: Prior to making investment decisions, participants are informed that their chosen investment cost is deducted from their endowment only if they are hired, in contrast to the Control treatment where investment costs are deducted regardless of the hiring outcome. Importantly, this treatment both reduces the downside risk associated with investment and increases its expected payoff.⁶ Because Blue participants face a lower probability of being hired, conditioning the investment cost on hiring also generates a larger increase in expected payoff for Blue than for Green participants at any given positive investment level. Figure A2 presents screenshots of the participants’ decision screens with the “No Risk” reminder.

Implementation

The study was conducted in August 2025 on Prolific. Participants were recruited to approximate national representativeness of the U.S. population across age, gender, and ethnicity.⁷ Figure 2 illustrates the treatment assignment. The experiment combines between- and within-subject components to improve statistical efficiency while allowing comparisons across treatment conditions. Based on an assumed effect size of 0.2,⁸ a significance level of $\alpha = 0.05$, and power of 0.80, the pre-registered

⁵The “Role Model” examples were selected from pilot-study participants within the same identity group who invested positive amounts. To improve comparability across the “Successful Role Model” and “Unsuccessful Role Model” treatments, the selected examples were chosen to display broadly similar investment levels across the two conditions within each group. Participants were informed that the examples came from previous participants in their own group, but were not informed about the selection procedure.

⁶See Appendix A.2 for the derivation of these effects.

⁷Prolific constructs representative samples by matching participant subgroups to U.S. Census Bureau population shares. See <https://researcher-help.prolific.com/en/article/95c345>, accessed 02-17-2026.

⁸The assumed effect size reflects a conservative benchmark, chosen in the absence of closely comparable experimental estimates in algorithmic hiring settings, and is intended to ensure sensitivity to modest treatment effects.

power analysis indicated that approximately 200 observations per treatment would provide sufficient statistical power.

To achieve the target sample size, participants were randomly assigned to one of two identity groups (Blue or Green) and one of four treatment sequences: (1) first round: Control; second round: random assignment to “No Risk,” “Successful Role Model,” or “Unsuccessful Role Model”; (2) first round: “No Risk”; second round: Control; (3) first round: “Successful Role Model”; second round: “Unsuccessful Role Model”; and (4) first round: “Unsuccessful Role Model”; second round: “Successful Role Model.”⁹ Across the two decision rounds, this yields 276 observations in the Control treatment, 182 in the “No Risk” treatment, 325 in the “Successful Role Model” treatment, and 323 in the “Unsuccessful Role Model” treatment. Before making any investment decisions, participants receive information about the equal underlying productivity distribution, the algorithm’s more favorable evaluation of Green participants, the testing technology, and the resulting group-specific hiring rule. They then complete a series of comprehension checks,¹⁰ followed by a practice round and two incentivized decision rounds.

The experiment concludes with a short demographic questionnaire. We collect information on participants’ age, gender, education, employment status, income, risk attitude, trust in major technology companies, and prior experience interacting with AI or algorithmic decision-makers. The full survey is provided in Appendix B.

Payoffs are denominated in Experimental Currency Units (ECUs), where 100 ECUs equal USD 1. At the beginning of each decision round, participants receive an endowment of 20 ECUs. Participants hired by the algorithm receive a wage of 100 ECUs. In the Control and information treatments, investment costs are deducted regardless of the hiring outcome, whereas in the “No Risk” treatment investment costs are deducted only if the participant is hired. One of the two formal rounds is randomly selected to determine the bonus payment. Participants receive a participation fee in addition to their performance-based bonus, yielding an average total payment of USD 2.37. The median completion time is 9 minutes and 21 seconds.

Hypotheses

The purpose of our hypotheses is twofold: (1) to examine whether participants facing disadvantaged algorithmic evaluation invest differently from advantaged participants, and (2) to examine the effects of each treatment on investment decisions across groups. We pre-registered all hypotheses in the AsPredicted registry under ID #243245 [<https://aspredicted.org/f8jk-5xns.pdf>].

With Hypothesis 1, we examine whether participants assigned to the disadvantaged (Blue) group invest differently from those assigned to the advantaged (Green)

⁹Participants never experience both the “No Risk” treatment and either of the two information treatments. This separation avoids potential contamination between an intervention that changes the payoff structure (“No Risk”) and interventions that provide additional information (the two “Role Model” treatments). Because the resulting design combines between- and within-subject treatment variation, we additionally examine the robustness of the results to potential round and carryover effects in Section 3.

¹⁰In the “Role Model” treatments, participants complete additional comprehension checks on the investment decisions and hiring outcomes of the previous participants before making their own investment decision.

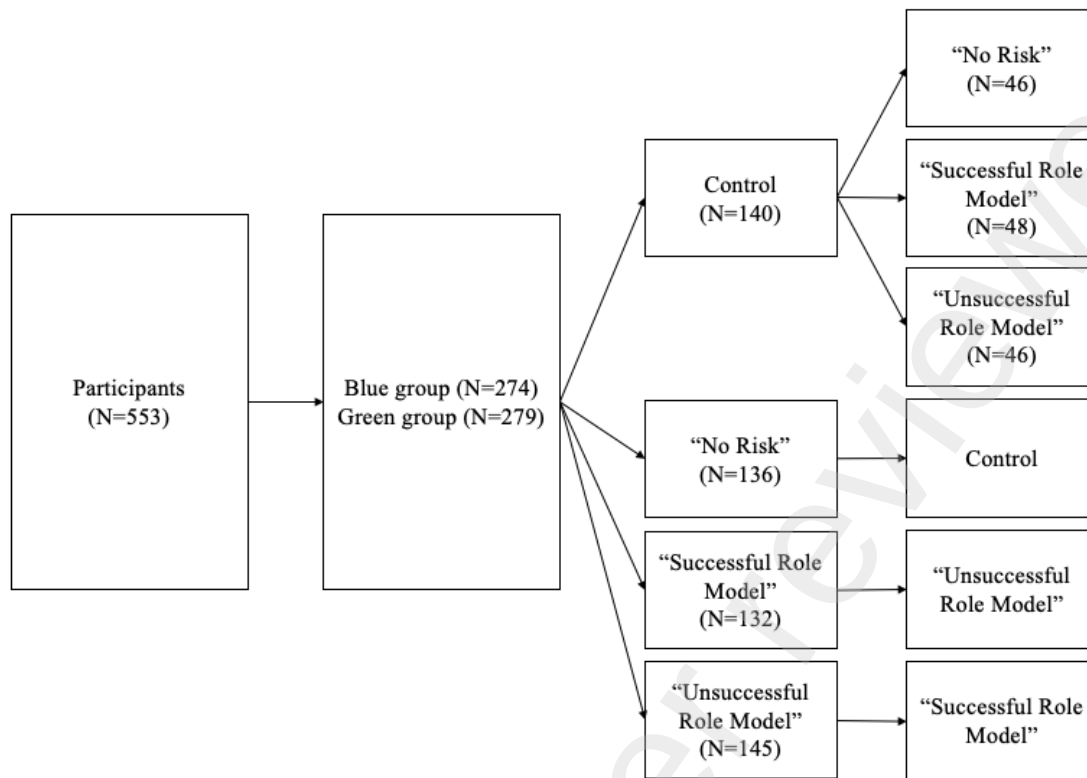


Figure 2: Treatment assignment

group under unequal algorithmic evaluation. Under standard expected-payoff maximization, participants in both groups maximize expected earnings by investing the full amount regardless of group assignment (see Appendix A.2 for a derivation of expected payoffs). Expected-payoff maximization therefore predicts no difference in investment across groups. Existing behavioral accounts, however, suggest competing directional responses. On the one hand, anticipated discrimination may discourage investment and contribute to self-fulfilling effects (see, for example, Glover et al. 2017). On the other hand, disadvantaged individuals may respond by increasing investment or effort in an attempt to compensate for their disadvantage (Caputo 2002). Because these mechanisms imply effects in opposite directions, we do not impose a directional alternative. We therefore evaluate the hypothesis below using a two-sided test of the difference in investment between Blue and Green participants.

Hypothesis 1 (Group Effect in Control): *In the Control treatment, participants in the disadvantaged Blue group invest the same as participants in the advantaged Green group.*

With Hypothesis 2, we examine the effects of the information treatments. Information about others' choices and outcomes can influence subsequent decisions through social learning, imitation, and social comparison. In investment settings, individuals have been shown to respond to information about peers' choices (Bursztyn et al. 2014, Delfino et al. 2016), while research on positive and negative comparison targets shows that observing successful or unsuccessful others can affect motivation and subsequent behavior (Lockwood et al. 2004, Buunk et al. 2007). In our experiment, participants assigned to the "Successful Role Model" treatment

observe five previous in-group participants who invested positive amounts and were subsequently hired, whereas participants assigned to the “Unsuccessful Role Model” treatment observe five previous in-group participants who invested positive amounts but were not hired. Because participants are informed about the hiring rule and the payoff consequences of their investment before making their decisions, these treatments test whether same-group social information about predecessors’ behavior and realized outcomes affects investment beyond the incentives implied by the current decision environment. The design does not separately identify whether any such response operates through social learning, social comparison, or a role-model mechanism.

Hypothesis 2a (“Successful Role Model” Treatment Effect): *Within each group (Blue or Green), participants in the “Successful Role Model” treatment will invest more than those in the Control treatment.*

Hypothesis 2b (“Unsuccessful Role Model” Treatment Effect): *Within each group (Blue or Green), participants in the “Unsuccessful Role Model” treatment will invest less than those in the Control treatment.*

With Hypotheses 3 and 4, we examine the effect of the “No Risk” treatment and compare it with both the Control and the information treatments. Unlike the information treatments, which provide information about previous participants while leaving the incentive structure unchanged, the “No Risk” treatment directly modifies the economic consequences of investment. Specifically, it increases the expected payoff from investing and reduces payoff variance relative to the Control treatment (see Appendix A.2). Because Blue participants face a lower probability of being hired, the increase in expected payoff generated by the treatment is larger for Blue than for Green participants at any given positive investment level. Nevertheless, the treatment improves the economic consequences of investment for both groups. Although expected-payoff maximization predicts full investment already in Control, observed investment may fall below this benchmark because participants respond to downside exposure and other behavioral considerations. By reducing downside exposure and improving expected payoffs for positive investment, the “No Risk” treatment is therefore expected to increase investment among participants whose choices fall below the expected-payoff benchmark. Furthermore, because the “No Risk” treatment directly changes participants’ economic incentives whereas the information treatments operate through social information without altering the payoff structure, we expect investment to be higher under the “No Risk” treatment than under either information treatment.

Hypothesis 3 (“No Risk” Treatment Effect): *Within each group (Blue or Green), participants in the “No Risk” treatment will invest more than those in the Control treatment.*

Hypothesis 4 (“No Risk” vs. “Role Model”): *Within each group (Blue or Green), participants in the “No Risk” treatment will invest more than those in either “Role Model” treatment.*

Finally, with Hypothesis 5, we examine whether the effectiveness of the interventions differs across the disadvantaged and advantaged groups. For the information treatments, all observed predecessors belong to the participant's own experimentally induced identity group. Evidence from role-model research suggests that similar or identity-relevant comparison targets can be particularly consequential when individuals face disadvantage or stigma (Morgenroth et al. 2015), while experimental evidence on social learning shows that individuals place greater weight on information from predecessors who share an experimentally induced group identity (Zou & Xu 2023). We therefore expect information about the behavior and outcomes of same-group predecessors to be especially salient for participants facing less favorable evaluation.

For the “No Risk” treatment, there is an additional economic rationale for heterogeneous effects. Blue participants face lower hiring probabilities and, at positive investment levels, greater payoff variance than Green participants. Moreover, because the investment cost is waived when a participant is not hired, the “No Risk” treatment produces a larger increase in expected payoff for Blue participants at any given positive investment level. Thus, a stronger behavioral response among Blue participants may reflect both their disadvantaged position and the objectively larger change in the investment environment induced by the treatment. We therefore hypothesize that the effects of both the “Role Model” and “No Risk” treatments will be stronger among Blue participants.

Hypothesis 5 (Heterogeneity Analysis): *The effects of the “Role Model” and “No Risk” treatments differ across the two groups, with stronger treatment effects among the disadvantaged Blue group.*

3 Results

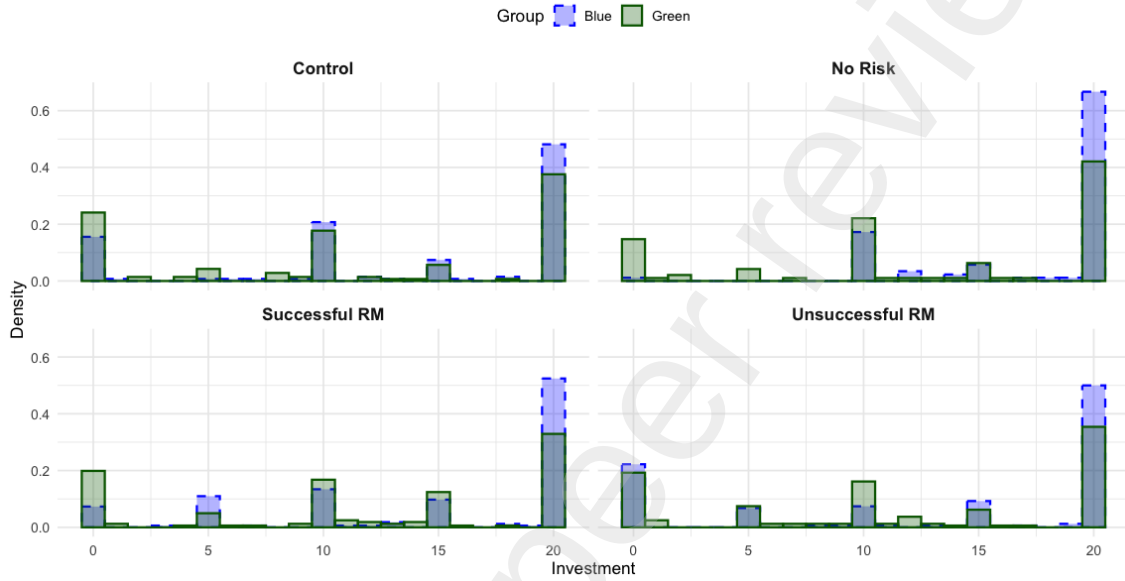
Summary Statistics

Table A1 presents summary statistics. Participants are, on average, 45.8 years old and roughly split between male and female. Approximately 54.5% hold at least a Bachelor's degree, and 69.3% are employed, with the median income falling within the bracket “Greater than or equal to 50,000 USD and less than 75,000 USD.” Participants scored an average of 6.5 on the 0–10 risk attitude scale, and 5.7 on the 0–10 trust scale in response to the question: “In general, how much do you trust major technology companies to make decisions that affect you in a fair and unbiased way?” When asked, “Have you ever applied for a job or other opportunity where you believe your application was screened or reviewed by an AI system or algorithm?” around 35.8% of participants selected the response: “Yes, I have submitted an application where I believe my application was screened by an AI system or algorithm.” Out of the participants who selected yes, replying to the follow-up question: “If so, in which contexts have you interacted with the AI system or algorithm? Select all that apply.” 96% reported that they had interacted with an AI system or algorithm during job applications or through hiring platforms; 35.4% reported having had such interaction in the context of credit scoring or loan applications. We report the summary statistics by identity group (Table A2) and by first-round treatment assignment (Table A3). Across both tables, we observe no systematic differences in

demographic characteristics, prior experience with AI or algorithm, trust in major tech companies, or risk attitudes.

Figure 3 presents the distribution of investment decisions by group across treatments. Investments cluster around 0, 10, and 20 ECUs for both groups. Visual inspection suggests that investment distributions differ across treatments, particularly for the Blue group. The following subsections formally examine these treatment effects.

Figure 3: Investment distribution by group



Note: The figure shows the distribution of investment choices for the Blue and Green groups separately across the Control, No Risk, Successful Role Model (RM), and Unsuccessful Role Model (RM) treatments. Investment ranges from 0 to 20 ECU. The vertical axis reports density.

Effect of Group Advantage on Investment

To test Hypothesis 1, we estimate the following OLS regression using data from the Control treatment:

$$Invest_i = \beta_0 + \beta_1 Blue_i + X_i' \gamma + \varepsilon_i$$

where $Invest_i$ is the outcome variable that denotes the investment amount of participant i ; β_0 is the investment level of the Green group; $Blue_i$ is a dummy variable that equals 1 if the participant is in the Blue group, and 0 if the participant is in the Green group. X_i' is a vector of pre-registered controls including round fixed effects, demographics, risk preferences, trust in major tech companies and prior experience with AI or algorithm. Standard errors are clustered at the individual level. Table 1 shows the regression results. In Column 1, we report our main specification controlling only for round fixed effects, in Column 2, we include demographic controls, and in Column 3, risk attitude, prior experience with AI or algorithm and trust in major technology companies are also controlled for.

Across all three specifications, we observe a statistically significant positive effect of being assigned to the disadvantaged (Blue) group on investment. In our

main specification (Column 1), Green participants invest 12.4 ECUs out of 20 on average, while Blue participants invest 2.2 ECUs more. This corresponds to an increase of approximately 17.7% relative to the Green group. The estimated group effect remains stable after controlling for demographic characteristics, risk attitudes, prior experience with AI or algorithmic decision-makers, and trust in major technology companies, suggesting that the result is not driven by observable participant characteristics.

The positive group effect is robust to the pre-registered specification excluding participants whose completion time falls outside the 10th and 90th percentiles (Table A5, Appendix). We additionally estimate the model using only first-round observations, thereby isolating behavior before participants receive feedback from a previous decision round. The estimated coefficient remains positive and similar in magnitude to the main specification, although it is no longer statistically significant. Because this specification relies on only 140 observations from the Control treatment, it has substantially lower statistical power than the pooled analysis. Table A7 reports the corresponding regression results. One additional pre-registered exclusion criterion involved removing participants who failed the attention checks. However, due to programming constraints on the survey platform, failed attempts were not recorded once participants eventually answered correctly. Consequently, this exclusion criterion could not be implemented ex post, and all participants who successfully completed the attention checks are retained in the analysis.

Overall, we find no evidence of investment discouragement among disadvantaged participants; instead, their higher investment is consistent with a compensatory response to unequal evaluation, although our design does not identify the underlying mechanism.

Result 1 (H1 rejected): *In the Control treatment, participants in the disadvantaged (Blue) group invest significantly differently from participants in the advantaged (Green) group.*

“Role Model” & “No Risk” Treatment Effects

Before examining treatment effects separately for the disadvantaged and advantaged groups, we first estimate the average treatment effects across the full sample. Table 2 shows that, relative to the Control treatment, the “No Risk” treatment increases investment by approximately 2.2 ECUs (around 11% of the initial endowment), and this effect remains highly statistically significant across all specifications. In contrast, neither the “Successful Role Model” nor the “Unsuccessful Role Model” treatment has a statistically significant average effect on investment. The estimated treatment effects remain largely unchanged after the inclusion of demographic controls, suggesting that the results are not driven by observable participant characteristics.

We next examine whether these average treatment effects differ across disadvantaged and advantaged participants by testing the remaining hypotheses. To test Hypotheses 2, 3, 4 and 5, we estimate the following OLS regression:

Table 1: Effect of group (Blue vs. Green) on investment in control treatment

	<i>Dependent variable:</i>		
	Investment		
	(1)	(2)	(3)
Blue Group (vs Green)	2.199** (0.928)	2.409** (0.984)	2.184** (1.015)
Constant	12.415*** (0.820)	7.961*** (2.668)	8.458** (3.788)
Round Control	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes
Risk Attitude	No	No	Yes
Prior Experience	No	No	Yes
Trust in Tech	No	No	Yes
Observations	276	276	276
R ²	0.041	0.136	0.235

Note: The dependent variable is the number of investment units allocated (ranging from 0 to 20 inclusive). All models include round fixed effects. Column (2) adds demographic controls: age (continuous), gender (binary), education level (ordered factor), employment status (binary), and income bracket (ordered factor). Column (3) further includes individual-level covariates: risk attitude (self-reported on a 0–10 scale), prior experience with algorithmic screening (binary), and trust in major technology companies (0–10 scale). Robust standard errors, clustered at the participant level, are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table 2: Overall treatment effects on investment

	<i>Dependent variable:</i>		
	Investment		
	(1)	(2)	(3)
Successful Role Model	0.815 (0.590)	0.888 (0.600)	0.713 (0.596)
Unsuccessful Role Model	-0.157 (0.625)	-0.084 (0.633)	-0.277 (0.621)
No Risk	2.279*** (0.528)	2.249*** (0.534)	2.236*** (0.536)
Constant	12.865*** (0.462)	12.818*** (1.522)	14.451*** (2.248)
Round Fixed Effects	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes
Risk Attitude	No	No	Yes
Prior Experience	No	No	Yes
Trust in Tech	No	No	Yes
Observations	1,106	1,106	1,106
R ²	0.018	0.042	0.083

Note: The dependent variable is investment (ECUs). Control is the omitted treatment category. All specifications include round fixed effects. Columns (2) and (3) additionally control for demographic characteristics, while Column (3) further controls for risk attitude, prior experience with AI or algorithmic decision-makers, and trust in major technology companies. Standard errors clustered at the participant level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

$$\begin{aligned} Invest_i = & \beta_0 + \beta_1 SRM_i + \beta_2 URM_i + \beta_3 NR_i + \beta_4 Blue_i + \beta_5 (SRM_i \times Blue_i) \\ & + \beta_6 (URM_i \times Blue_i) + \beta_7 (NR_i \times Blue_i) + X_i' \gamma + \varepsilon_i. \end{aligned}$$

where $SRM_i = 1$ if participant i is assigned to the “Successful Role Model” treatment, $URM_i = 1$ if participant i is assigned to the “Unsuccessful Role Model” treatment, $NR_i = 1$ if participant i is assigned to the “No Risk” treatment, and $Blue_i = 1$ if participant i belongs to the disadvantaged Blue group (0 otherwise). X_i' denotes the vector of pre-registered controls. Standard errors are clustered at the participant level. The coefficients of interest are β_1 and $\beta_1 + \beta_5$ (H2a), β_2 and $\beta_2 + \beta_6$ (H2b), and β_3 and $\beta_3 + \beta_7$ (H3), which identify the treatment effects within the Green and Blue groups, respectively. The interaction terms β_5 , β_6 , and β_7 test whether treatment effects differ across groups (H5). Table 3 reports the results from our main specification together with the linear combination tests.¹¹

We first consider the two information treatments. Table 3 shows that neither the “Successful Role Model” nor the “Unsuccessful Role Model” treatment has a statistically significant effect on investment for either group in our main specification. These findings are unchanged under the pre-registered specification excluding participants whose completion time falls outside the 10th and 90th percentiles (Table A6, Appendix).

We additionally estimate the model using only first-round observations, thereby isolating behavior before participants receive feedback from a previous decision round. In the first-round specification (Table A8, Appendix), the “Unsuccessful Role Model” treatment reduces investment by approximately 2 ECUs among Green participants. The interaction between the treatment and the Blue group is small and statistically insignificant, indicating that the treatment effect is not significantly different for disadvantaged participants. By contrast, the “Successful Role Model” treatment remains statistically insignificant throughout.

Comparing the pooled and first-round specifications suggests that information about unsuccessful predecessors may initially discourage investment, but that this effect attenuates when participants have gained experience with the evaluation environment. This pattern is however not robust across specifications. Overall, we do not find robust evidence that providing information about previous participants’ behavior and outcomes changes investment decisions in our experimental environment. More broadly, these findings suggest that when participants already have experience with the algorithmic evaluation rule, social information has limited effects on investment decisions.

Result 2 (no support for H2a and H2b): *Within each group (Blue or Green), participants in the “Role Model” treatments do not invest significantly differently than those in the Control treatment in our main specification.*

We next examine the effect of the “No Risk” treatment. As discussed above, this intervention changes both the downside exposure associated with investing and the expected payoffs. In our main specification, the “No Risk” treatment substantially increases investment among participants in the Blue group. Relative to the Control

¹¹Table A4 (Appendix) reports treatment effects under all three pre-registered specifications.

treatment, Blue participants invest approximately 3.5 additional ECUs, corresponding to an increase of around 24%. By contrast, we do not observe a statistically significant effect among Green participants.

The positive treatment effect among Blue participants is robust across alternative specifications. When observations with completion times outside the 10th and 90th percentiles are excluded (Table A6, Appendix), the estimated effect remains positive and statistically significant. The same pattern emerges when the analysis is restricted to first-round observations only (Table A8, Appendix), indicating that the result is not driven by participants updating their behavior after observing the outcome of an earlier decision round. Although the completion-time specification also yields a positive treatment effect among Green participants, this result is not replicated in either the main specification or the first-round analysis. We therefore interpret the evidence as indicating that the “No Risk” treatment is robust among disadvantaged participants, while evidence of a positive response among advantaged participants is less conclusive.

Result 3 (partial support for H3): *The “No Risk” treatment increases investment robustly among participants in the disadvantaged (Blue) group. Evidence of a positive treatment effect among the advantaged (Green) participants is weaker and not robust across specifications.*

The results above suggest that the “No Risk” treatment has a different effect size from the two information treatments. Comparing the effectiveness of these treatments directly, among Green participants, investment does not differ significantly across the three treatments; among Blue participants, however, investment under the “No Risk” treatment is significantly higher than under both the “Successful Role Model” and “Unsuccessful Role Model” treatments. These comparisons suggest that changing the economic consequences of investment is more effective than providing information about previous participants’ behavior and outcomes in encouraging investment under unequal algorithmic evaluation.

As shown in Appendix A.2, for any positive investment level, the “No Risk” treatment both increases expected payoff and reduces payoff variance relative to the Control treatment. Despite these improved incentives, Green participants do not systematically increase investment. By contrast, Blue participants respond strongly by investing significantly more.

Result 4 (partial support for H4): *Among the disadvantaged (Blue) participants, investment under the “No Risk” treatment is significantly higher than under either information treatment. No corresponding differences are detected among the advantaged (Green) participants.*

Result 5 (partial support for H5): *The treatment effect of the “No Risk” intervention is significantly stronger for the disadvantaged (Blue) participants than for the advantaged (Green) participants, whereas the effects of the information treatments do not differ across groups.*

Since the experiment combines between- and within-subject treatment variation, we additionally examine whether the treatment heterogeneity results are sensitive to

Table 3: Treatment effects by group and linear combination tests

	Regression coefficients		Linear combination tests	
	Estimate	Std. Error	Estimate	Std. Error
Main effects				
Blue Group	2.317**	(0.929)		
Successful RM (Green)	0.503	(0.847)	0.503	(0.847)
Unsuccessful RM (Green)	0.152	(0.889)	0.152	(0.889)
No Risk (Green)	1.216	(0.750)	1.216	(0.750)
Interactions				
Blue $\times$ Successful RM	0.551	(1.145)	1.055	(0.785)
Blue $\times$ Unsuccessful RM	-0.672	(1.232)	-0.520	(0.856)
Blue $\times$ No Risk	2.264**	(1.046)	3.480***	(0.717)
Constant	11.744***	(0.678)		
Observations			1,106	
R ²			0.052	

Note: The dependent variable is the number of investment units allocated (ranging from 0 to 20 inclusive). All models include round fixed effects. The first two columns report OLS regression coefficients from Column (1) of Table A4. The final two columns show linear combination tests comparing each treatment condition to the Control group within Blue and Green participants. Standard errors are clustered at the participant level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

the repeated-decision component of the design. Comparing estimates based only on first-round decisions with those from the pooled specification, we find no significant differences in the Blue–Green treatment-effect heterogeneity for any of the three treatments (Table A9). We also allow treatment effects to vary across rounds. The treatment-by-round interactions are jointly significant, reflecting stronger second-round responses to the “No Risk” and “Unsuccessful Role Model” treatments. However, none of the corresponding Blue $\times$ treatment $\times$ Round 2 interactions is individually significant, and they are also jointly insignificant (Table A10).

As additional checks on potential carryover, we examine second-round behavior conditional on participants’ first-round decisions and outcomes. Among participants assigned to the Control treatment in the first round, assignment to “No Risk” in the second round is associated with higher investment even after conditioning on first-round investment, aligning with the result we find in the main specification reported in Table 2. We find no evidence that being hired in the first round systematically predicts subsequent investment adjustments or differentially affects responses to the second-round treatment; the interactions between first-round hiring and second-round treatment are jointly insignificant. These results suggest that, although participants exhibit some adjustment across rounds, the treatment responses do not appear to be explained by feedback from previous hiring outcomes (Table A11).

Taken together, we find that the magnitude of some treatment responses varies across rounds, but find no evidence that the differential treatment responses between Blue and Green participants vary systematically across rounds or are driven by pooling repeated decisions. In particular, the stronger response of the Blue participants to the “No Risk” treatment is similar when estimated using only first-round decisions and when pooling both rounds.

Exploratory Analyses

As an exploratory exercise, we examine whether treatment responses vary with participants' demographic characteristics and psychological traits. We consider heterogeneity by age, gender, education, employment status, income, risk attitudes, trust in major technology companies, and prior experience with algorithmic screening. For each characteristic, we interact the corresponding measure with the treatment indicators and progressively add demographic and additional individual-level controls. The full regression results are reported in Tables A12–A19. Across these specifications, the positive effect of assignment to the disadvantaged (Blue) group and the positive effect of the “No Risk” treatment remain stable, indicating that the main findings are robust to accounting for these observable participant characteristics.

Several patterns emerge from these analyses. First, older participants tend to invest more: a one-year increase in age is associated with approximately 0.05–0.06 additional investment units, although age does not significantly moderate any of the treatment effects (Table A12). By contrast, we find no systematic evidence that the effect of the “No Risk” treatment varies with income or self-reported risk attitudes (Tables A15 and A17). The latter is noteworthy given that the treatment changes both expected payoffs and downside exposure: its behavioral effect does not appear to vary systematically with participants' reported willingness to take risks.

We find some evidence that responses to the information treatments vary with participants' income and trust in major technology companies. Higher income is associated with a smaller response to the “Successful Role Model” treatment in the fully controlled specification (Table A15). Participants with greater trust in major technology companies also respond more positively to the “Unsuccessful Role Model” treatment, with the interaction statistically significant after demographic and additional individual-level controls are included (Table A18). We find no systematic treatment heterogeneity by age, gender, education, employment status, prior experience with algorithmic screening, or risk attitudes (Tables A12–A19).

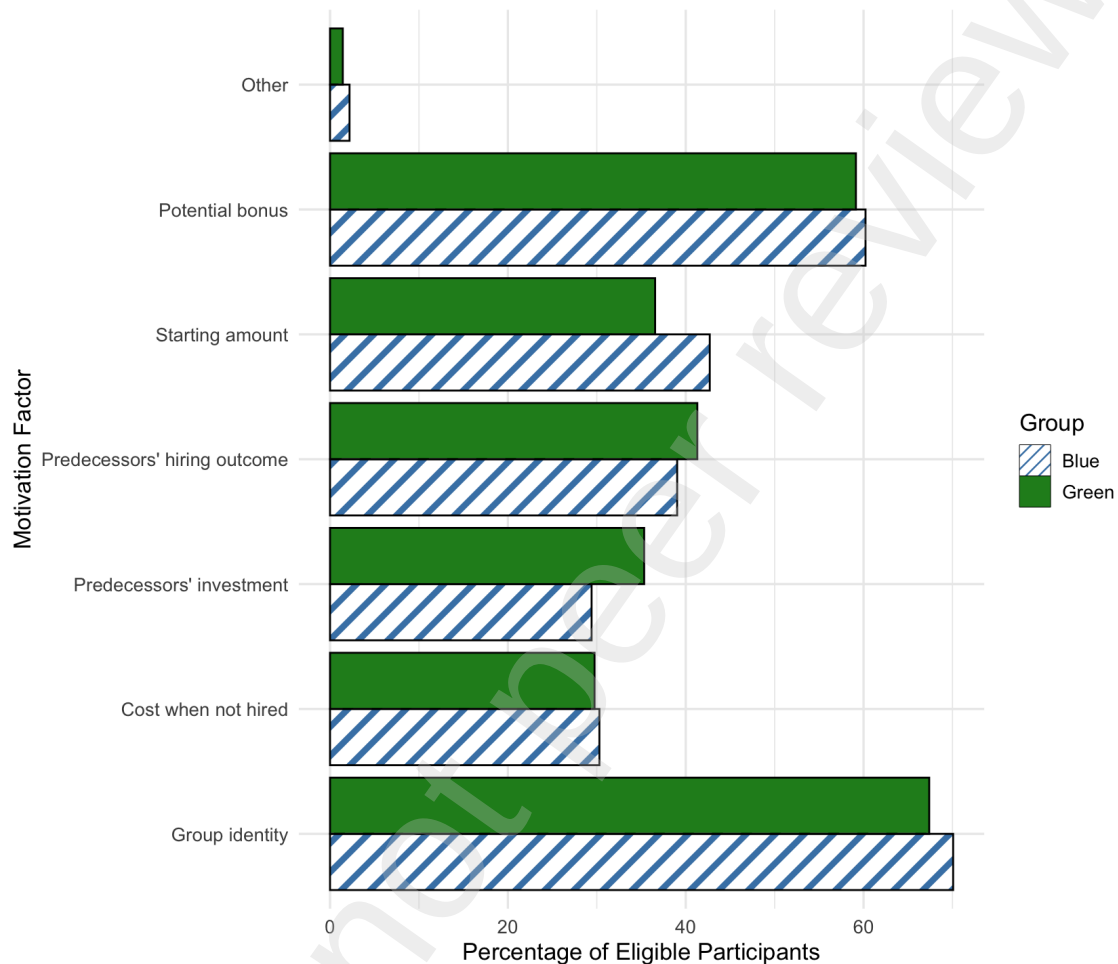
To provide additional descriptive evidence on the motivation behind participants' investment decisions, we ask participants after the experiment which factors influenced their choices. Participants can select all that apply from a set of options presented in random order, including their group identity, the investment cost when not hired, previous participants' investment decisions and hiring outcomes, the initial endowment, the potential hiring bonus, and an open-ended “Other” option. The options concerning previous participants' investment decisions and hiring outcomes are displayed only to participants who experienced at least one of the “Role Model” treatments. Figure 4 reports the percentage of eligible participants selecting each factor, separately by group.

Group identity is the most frequently reported factor for both Blue and Green participants, followed by the potential hiring bonus. This suggests that more than half of the participants across groups are attentive both to their experimentally assigned group and to the financial incentives embedded in the decision environment. Information about predecessors is also reported by a nontrivial share of participants who are eligible to observe it, particularly information about whether predecessors are hired.

Overall, the self-reported motivation indicates that the group-based evaluation environment is salient to participants' decisions, alongside financial considerations. Exploratory analyses further show that the association between reporting group

identity as a motivation and investment differs significantly across groups: relative to Green participants, the association is approximately 3.1–3.4 investment units more positive among Blue participants, depending on the specification (Table A20). This pattern is consistent with participants responding differently to the salience of group identity depending on whether their assigned group is disadvantaged or advantaged, and is compatible with a compensatory interpretation of the higher investment observed among Blue participants.

Figure 4: Self-reported motivation for investment



Note: The figure reports the percentage of eligible participants in each group who select each factor in response to the question: “In the experiment, what factor(s) below motivated your investment choice? Select all that apply.” Participants could select multiple responses, and response options are presented in random order. The available options are: “My group identity (whether I was assigned to the Green or Blue group),” “Whether I have to pay for the investment cost when I’m not hired,” “How much previous participants invested,” “Whether previous participants were hired,” “The starting amount I had available to invest (20 ECUs),” “The potential bonus I could get if hired (100 ECUs),” and “Other (please specify in the textbox below).” The options concerning previous participants’ investment and hiring outcomes are displayed only to participants who experience at least one of the “Role Model” treatments; percentages for these two categories are therefore calculated relative to the number of participants in each group who are eligible to select them. Percentages for all other categories are calculated relative to the number of participants in each group who answer the motivation question.

4 Conclusion

This paper examines how individuals respond to unequal algorithmic evaluation and whether different interventions can encourage self-investment under algorithmic discrimination. We find that participants assigned to the disadvantaged group invest more than participants assigned to the advantaged group, despite knowing before their investment decisions that their group faces less favorable evaluation. Thus, we find no evidence of investment discouragement; if anything, the higher investment of disadvantaged participants is consistent with a compensatory response to unequal evaluation. Among the interventions we study, modifying the economic consequences of investment substantially increases investment among disadvantaged participants, with a significantly larger effect than among advantaged participants, whereas providing information about previous participants' behavior and outcomes does not generate robust changes.

The limited effects of the information treatments should be interpreted in the context of this experimental environment rather than as evidence that social-comparison interventions are generally ineffective. One possible explanation is that, because participants are informed of the group-specific hiring rule and how investment affects their hiring prospects, observing the behavior and outcomes of previous participants may have provided relatively little additional information about their own prospects. At the same time, the limited treatment effects do not appear to simply reflect participants disregarding the information: a nontrivial share of participants exposed to the information treatments report predecessors' behavior or outcomes as relevant to their investment decisions. Thus, while participants appear to attend to this information, it does not translate into robust changes in investment in our setting.

By contrast, changing the economic consequences of investment generates a substantially stronger behavioral response in our setting. The "No Risk" treatment reduces downside exposure while increasing expected payoffs, leading disadvantaged participants to invest substantially more. Because disadvantaged participants are less likely to be hired, they benefit more from conditioning the investment cost on hiring: the protection applies more often and generates a larger increase in expected payoff. This may help explain why their response is stronger, while advantaged participants show no robust increase in investment. Our design, however, does not identify whether the response is driven primarily by reduced downside exposure or higher expected payoffs.

Our experiment necessarily abstracts from several features of real-world investment and hiring decisions. Investment is represented by a costly allocation decision rather than a real-effort task, and participants face an exogenously specified and known relationship between investment and performance. This structure allows us to isolate investment responses to unequal evaluation, but abstracts from channels such as confidence, self-efficacy, identity threat, endogenous beliefs about ability, and uncertainty about the productivity of effort, all of which may shape responses to discrimination outside the experimental environment. Future work could examine whether the behavioral responses documented here persist when investment requires real effort and when these psychological and informational channels can operate.

A further limitation is that the hiring algorithm applies a fixed decision rule and is evaluated in isolation, without a human-recruiter benchmark. This static structure was deliberately chosen to maintain experimental control and isolate behav-

ioral responses to unequal algorithmic evaluation. In practice, algorithmic systems may update based on incoming data or operate alongside human decision makers (Tambe et al. 2019, Marabelli et al. 2021). The experiment also abstracts from competition among applicants: all participants who meet the qualification threshold may be hired, eliminating strategic responses to others' investment decisions. Future research could therefore compare responses to equivalent disadvantage originating from human and algorithmic evaluators, examine settings in which algorithmic learning and individual behavior co-evolve, and introduce relative-performance or rank-based selection.

Finally, our design studies a relatively transparent decision environment. Participants are informed that the algorithm evaluates Green participants more favorably than Blue participants, observe the resulting group-specific hiring rule, and know how investment affects performance, although they do not observe the numerical prior beliefs underlying the algorithm. In many real-world settings, individuals may be uncertain about whether unequal evaluation exists, its magnitude, or whether additional investment can overcome it. Our results therefore provide a benchmark for behavior when the consequences of unequal algorithmic evaluation are transparent. This setting is also relevant as emerging AI regulation places greater emphasis on transparency and information provision for high-risk AI systems, including those used in employment and recruitment.¹² Future research could examine how investment changes when individuals must instead form beliefs about the existence and severity of algorithmic discrimination.

¹²For example, the EU Artificial Intelligence Act imposes transparency and information requirements on high-risk AI systems, including requirements that such systems be sufficiently transparent to enable deployers to interpret their outputs and use them appropriately, and that providers supply information concerning system capabilities, limitations, and human oversight measures (Regulation (EU) 2024/1689, Article 13).

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the author used ChatGPT (OpenAI) to assist with HTML coding for the Qualtrics survey, to identify grammatical errors in the manuscript, and to group open-text survey responses into broad thematic categories for descriptive analysis. The author reviewed and edited all outputs as needed and takes full responsibility for the content of the published article.

References

- Acikgoz, Y., Davison, K. H., Compagnone, M. & Laske, M. (2020), 'Justice perceptions of artificial intelligence in selection', *International Journal of Selection and Assessment* **28**(4), 399–416.
- Ajunwa, I. & Greene, D. (2019), Platforms at work: Automated hiring platforms and other new intermediaries in the organization of work, *in* 'Work and labor in the digital age', Vol. 33, Emerald Publishing Limited, pp. 61–91.
- Arrow, K., Ashenfelter, O. & Rees, A. (1973), 'Discrimination in labor markets', *The Theory of Discrimination* pp. 3–33.
- Bana, S. & Boudreau, K. (2025), 'Ai-enabled job markets & market participation: A field experiment on how ai shapes jobseekers' expectations of competition', *Available at SSRN 5437196*.
- Barocas, S. & Selbst, A. D. (2016), 'Big data's disparate impact', *Calif. L. Rev.* **104**, 671.
- Becker, M. G., Stolper, O. & Walter, A. (2026), 'Investing by example: Leveraging peer information in digital banking', *Journal of Behavioral and Experimental Economics* p. 102646.
- Bigman, Y. E., Wilson, D., Arnestad, M. N., Waytz, A. & Gray, K. (2023), 'Algorithmic discrimination causes less moral outrage than human discrimination.', *Journal of Experimental Psychology: General* **152**(1), 4.
- Binns, R. (2018), Fairness in machine learning: Lessons from political philosophy, *in* 'Conference on fairness, accountability and transparency', PMLR, pp. 149–159.
- Bogen, M. & Rieke, A. (2018), 'Help wanted: An examination of hiring algorithms, equity, and bias', *Upturn*, December **7**.
- Bohren, J. A., Haggag, K., Imas, A. & Pope, D. G. (2025), 'Inaccurate statistical discrimination: An identification problem', *Review of Economics and Statistics* **107**(3), 605–620.
- Bordalo, P., Coffman, K., Gennaioli, N. & Shleifer, A. (2016), 'Stereotypes', *The Quarterly Journal of Economics* **131**(4), 1753–1794.
- Bursztyn, L., Ederer, F., Ferman, B. & Yuchtman, N. (2014), 'Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions', *Econometrica* **82**(4), 1273–1301.
- Buunk, A. P., Peiró, J. M. & Griffioen, C. (2007), 'A positive role model may stimulate career-oriented behavior 1', *Journal of Applied Social Psychology* **37**(7), 1489–1500.
- Cai, C. J., Reif, E., Hegde, N., Hipp, J., Kim, B., Smilkov, D., Wattenberg, M., Viegas, F., Corrado, G. S., Stumpe, M. C. et al. (2019), Human-centered tools for coping with imperfect algorithms during medical decision-making, *in* 'Proceedings of the 2019 chi conference on human factors in computing systems', pp. 1–14.

- Campos-Mercade, P. & Mengel, F. (2024), 'Non-bayesian statistical discrimination', *Management Science* **70**(4), 2549–2567.
- Caputo, R. K. (2002), 'Discrimination and human capital: A challenge to economic theory and social justice', *J. Soc. & Soc. Welfare* **29**, 105.
- Caputo, R. K. (2007), 'Perceived work-related discrimination by women: Implications for social justice and affirmative action', *Journal of Policy Practice* **6**(2), 5–22.
- Cavatorta, E., Guarino, A. & Huck, S. (2024), 'Social learning with partial and aggregate information: Experimental evidence', *Games and Economic Behavior* **146**, 292–307.
- Chapman, B. (2006), 'Income contingent loans for higher education: International reforms', *Handbook of the Economics of Education* **2**, 1435–1503.
- Coate, S. & Loury, G. C. (1993), 'Will affirmative-action policies eliminate negative stereotypes?', *The American Economic Review* pp. 1220–1240.
- Cotton, C. S., Hickman, B. R. & Price, J. P. (2022), 'Affirmative action and human capital investment: Evidence from a randomized field experiment', *Journal of Labor Economics* **40**(1), 157–185.
- Dargnies, M.-P., Hakimov, R. & Kübler, D. (2024), 'Aversion to hiring algorithms: Transparency, gender profiling, and self-confidence', *Management Science* .
- Delfino, A., Marengo, L. & Ploner, M. (2016), 'I did it your way. an experimental investigation of peer effects in investment choices', *Journal of Economic Psychology* **54**, 113–123.
- Dianat, A., Echenique, F. & Yariv, L. (2022), 'Statistical discrimination and affirmative action in the lab', *Games and Economic Behavior* **132**, 41–58.
- Evans, D. P., Sales, J. M., Krause, K. H. & Del Rio, C. (2019), 'You have to be twice as good and work twice as hard: a mixed-methods study of perceptions of sexual harassment, assault and women's leadership among female faculty at a research university in the usa', *Global health, epidemiology and genomics* **4**, e6.
- Fang, H. & Moro, A. (2011), 'Theories of statistical discrimination and affirmative action: A survey', *Handbook of social economics* **1**, 133–200.
- Friedman, M. (1955), 'The role of government in education'.
- Fumagalli, E., Rezaei, S. & Salomons, A. (2022), 'Ok computer: Worker perceptions of algorithmic recruitment', *Research Policy* **51**(2), 104420.
- Glover, D., Pallais, A. & Pariente, W. (2017), 'Discrimination as a self-fulfilling prophecy: Evidence from french grocery stores', *The Quarterly Journal of Economics* **132**(3), 1219–1260.
- Gutman, L. M. & Younas, F. (2025), 'Understanding the awarding gap through the lived experiences of minority ethnic students: An intersectional approach', *British Educational Research Journal* **51**(2), 990–1008.

- Hamilton, I. A. (2018), 'Why it's totally unsurprising that amazon's recruitment ai was biased against women', *Business Insider*. Retrieved November 11.
- Heiserman, N. & Simpson, B. (2023), 'Discrimination reduces work effort of those who are disadvantaged and those who are advantaged by it', *Nature Human Behaviour* **7**(11), 1890–1898.
- Houser, K. A. (2019), 'Can ai solve the diversity problem in the tech industry: Mitigating noise and bias in employment decision-making', *Stan. Tech. L. Rev.* **22**, 290.
- Ip, E. (2025), 'Fair ai in hiring: Experimental evidence on how biased hiring algorithms and different debiasing methods affect the quality and diversity of applicants', *Behavioral Science & Policy* **11**(1), 44–54.
- Isik, U., Wouters, A., Verdonk, P., Croiset, G. & Kusurkar, R. A. (2021), "“as an ethnic minority, you just have to work twice as hard.” experiences and motivation of ethnic minority students in medical education', *Perspectives on medical education* **10**(5), 272–278.
- Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J. & Mullainathan, S. (2018), 'Human decisions and machine predictions', *The quarterly journal of economics* **133**(1), 237–293.
- Lavanchy, M., Reichert, P., Narayanan, J. & Savani, K. (2023), 'Applicants' fairness perceptions of algorithm-driven hiring procedures', *Journal of Business Ethics* **188**(1), 125–150.
- Lee, M. S. A. & Floridi, L. (2021), 'Algorithmic fairness in mortgage lending: from absolute conditions to relational trade-offs', *Minds and Machines* **31**(1), 165–191.
- Lockwood, P., Marshall, T. C. & Sadler, P. (2005), 'Promoting success or preventing failure: Cultural differences in motivation by positive and negative role models', *Personality and Social Psychology Bulletin* **31**(3), 379–392.
- Lockwood, P., Sadler, P., Fyman, K. & Tuck, S. (2004), 'To do or not to do: Using positive and negative role models to harness motivation', *Social Cognition* **22**(4), 422–450.
- Lundberg, S. J. & Startz, R. (1983), 'Private discrimination and social intervention in competitive labor market', *The American economic review* **73**(3), 340–347.
- Marabelli, M., Newell, S. & Handunge, V. (2021), 'The lifecycle of algorithmic decision-making systems: Organizational choices and ethical challenges', *The Journal of Strategic Information Systems* **30**(3), 101683.
- Martini, C. A. & Urueña, V. (2025), 'Can role models influence female students' decision to participate in the labor market? evidence from a field experiment', *Economic Development and Cultural Change* **73**(2), 875–907.
- Mirowska, A. (2020), 'Ai evaluation in selection: Effects on application and pursuit intentions.', *Journal of Personnel Psychology* **19**(3), 142.

- Morgenroth, T., Ryan, M. K. & Peters, K. (2015), 'The motivational theory of role modeling: How role models influence role aspirants' goals', *Review of general psychology* **19**(4), 465–483.
- Narciso, G., Newman, C. & Tarp, F. (2026), 'Using role models to inspire marginalized groups: A cautionary tale', *Journal of Economic Behavior & Organization* **248**, 107665.
- Newman, D. T., Fast, N. J. & Harmon, D. J. (2020), 'When eliminating bias isn't fair: Algorithmic reductionism and procedural justice in human resource decisions', *Organizational Behavior and Human Decision Processes* **160**, 149–167.
- Patty, J. W. & Penn, E. M. (2023), 'Algorithmic fairness and statistical discrimination', *Philosophy Compass* **18**(1), e12891.
- Phelps, E. S. (1972), 'The statistical theory of racism and sexism', *The american economic review* **62**(4), 659–661.
- Porter, C. & Serra, D. (2020), 'Gender differences in the choice of major: The importance of female role models', *American Economic Journal: Applied Economics* **12**(3), 226–254.
- Tambe, P., Cappelli, P. & Yakubovich, V. (2019), 'Artificial intelligence in human resources management: Challenges and a path forward', *California Management Review* **61**(4), 15–42.
- Vincent, J. (2018), 'Amazon reportedly scraps internal ai recruiting tool that was biased against women', *The Verge* **10**.
- Zou, W. & Xu, X. (2023), 'Ingroup bias in a social learning experiment', *Experimental economics* **26**(1), 27–54.

A Appendix

A.1 Algorithmic Hiring Rule Derivation

Let $\pi > 0$ denote the hiring algorithm's net payoff from hiring a top performer, $-\omega < 0$ the payoff from hiring an average performer, and 0 the payoff from not hiring. Let $P_A(\cdot)$ denote the hiring algorithm's beliefs. After observing the participant's group (d_i) and test score (x_i), the hiring algorithm hires the participant if and only if the expected payoff from hiring is weakly positive:

$$EU(\text{hire}) = P_A(y_i = 1 \mid d_i, x_i)\pi + [1 - P_A(y_i = 1 \mid d_i, x_i)](-\omega) \geq 0.$$

Thus, a participant is hired if:

$$P_A(y_i = 1 \mid d_i, x_i) \geq \frac{\omega}{\pi + \omega}.$$

In the experiment, we exogenously impose that the net benefit of hiring a top performer (π) is 1.5 times the wage (ω) paid to hire a participant. Substituting $\pi = 1.5\omega$ into the expression above yields a hiring threshold of 40%. The algorithm therefore hires a participant whenever its posterior probability that the participant is a top performer is at least 40%.

Under the group-blind testing technology described in Figure 1, a test score of A can only be generated by a top performer. Hence, $P_A(y_i = 1 \mid d_i, x_i = A) = 1$, and the participant is hired regardless of group identity. Similarly, a score of C can only be generated by an average performer, so $P_A(y_i = 1 \mid d_i, x_i = C) = 0$, and the participant is not hired regardless of group identity.

A score of B, by contrast, is equally likely for top and average performers: $P(x_i = B \mid y_i = 1) = P(x_i = B \mid y_i = 0) = 0.30$. The score therefore provides no additional information about productivity, and the algorithm's posterior remains equal to its group-specific prior. For a Green participant,

$$P_A(y_i = 1 \mid d_i = g, x_i = B) = 0.50,$$

which exceeds the 40% hiring threshold, so the participant is hired. For a Blue participant,

$$P_A(y_i = 1 \mid d_i = b, x_i = B) = 0.20,$$

which falls below the threshold, so the participant is not hired.

A.2 Marginal Payoff and Variance Derivation

Given the testing technology and hiring threshold described in Section 2 and Appendix A.1, let $p_0 \in \{p_g, p_b\}$ denote a participant's baseline probability of being hired without investment. A Green participant's baseline probability of being hired is $p_g = 0.65$: with probability 0.5 the participant becomes a top performer and is hired with probability 1, while with probability 0.5 the participant becomes an average performer and is hired only upon receiving a score of B, which occurs with probability 0.3. A Blue participant is hired only upon receiving a score of A, giving a baseline hiring probability of $p_b = 0.35$.

Recall that investment I_i increases a participant's probability of becoming a top performer by 1.5 percentage points per unit invested. Incorporating the endogenous investment decision, a Green participant's probability of being hired can therefore be expressed as

$$(0.5 + 0.015I_i) \times 1 + (1 - 0.5 - 0.015I_i) \times 0.3 = 0.65 + 0.0105I_i.$$

For a Blue participant, who is hired only upon receiving a score of A, the probability of being hired is

$$(0.5 + 0.015I_i) \times 0.7 = 0.35 + 0.0105I_i.$$

Thus, participant i 's probability of being hired can be written as a function of their baseline hiring probability p_0 and investment:

$$p = P(\text{hired}) = p_0 + 0.0105I_i.$$

The expected payoff from investing I_i in the Control treatment, denoted by μ_C , is

$$\begin{aligned}\mu_C &= p(\omega + I - I_i) + (1 - p)(I - I_i) \\ &= I - I_i + p\omega.\end{aligned}$$

Because investment increases the hiring probability by 0.0105 per unit, the marginal expected payoff from investment is

$$\frac{\partial \mu_C}{\partial I_i} = -1 + 0.0105\omega.$$

The marginal expected payoff from investment is therefore identical across groups, since investment increases the hiring probability by the same amount for Blue and Green participants.

Although the marginal expected payoff from investment is identical across groups, participants face different payoff risk because their hiring probabilities differ. In the Control treatment, participants receive $\omega + I - I_i$ if hired and $I - I_i$ otherwise. Therefore,

$$\begin{aligned}VAR_C &= p(\omega + I - I_i - \mu_C)^2 + (1 - p)(I - I_i - \mu_C)^2 \\ &= \omega^2 p(1 - p).\end{aligned}$$

At zero investment, payoff variance is identical across groups because $p_g = 0.65$ and $p_b = 0.35$ imply the same value of $p(1 - p)$. However, investment increases both groups' hiring probabilities by the same amount while Blue participants begin from a

lower baseline hiring probability. Consequently, for every positive feasible investment level, Blue participants face higher payoff variance than Green participants.

This difference motivates the “No Risk” treatment, in which participants incur their investment cost only if they are hired. Their expected payoff is therefore

$$\begin{aligned}\mu_{NR} &= p(\omega + I - I_i) + (1 - p)I \\ &= I + p(\omega - I_i),\end{aligned}$$

and payoff variance is

$$\begin{aligned}VAR_{NR} &= p(\omega + I - I_i - \mu_{NR})^2 + (1 - p)(I - \mu_{NR})^2 \\ &= (\omega - I_i)^2 p(1 - p).\end{aligned}$$

Comparing payoff variances across treatments gives

$$\begin{aligned}VAR_C - VAR_{NR} &= [\omega^2 - (\omega - I_i)^2] p(1 - p) \\ &= I_i(2\omega - I_i)p(1 - p).\end{aligned}$$

For any positive investment $I_i \leq \omega$ and any non-degenerate hiring probability $0 < p < 1$, this expression is strictly positive. Thus, whenever the investment cost does not exceed the hiring wage, the “No Risk” treatment reduces payoff variance relative to the Control treatment.

The “No Risk” treatment also increases expected payoff relative to the Control treatment:

$$\mu_{NR} - \mu_C = I_i(1 - p) > 0$$

for $I_i > 0$ and $p < 1$. Importantly, the magnitude of this expected-payoff increase depends on the participant’s probability of being hired. Since Blue participants have a lower hiring probability than Green participants at every investment level,

$$p_b(I_i) < p_g(I_i),$$

the expected-payoff gain from the “No Risk” treatment is larger for Blue participants:

$$\begin{aligned}(\mu_{NR} - \mu_C)_b - (\mu_{NR} - \mu_C)_g &= I_i [(1 - p_b) - (1 - p_g)] \\ &= I_i(p_g - p_b) \\ &= 0.30I_i > 0\end{aligned}$$

for any positive investment. Thus, the “No Risk” treatment simultaneously reduces payoff variance and increases expected payoff for both groups, while providing a larger expected-payoff improvement to participants in the disadvantaged Blue group because they face a lower probability of being hired.

A.3 Appendix Figures

Figure A1: “Successful Role Models” and “Unsuccessful Role Models” information screens

For your reference, before you make your investment decision, here is some information about previous **Blue participants who have completed this task:**

Participant	Group	Investment (ECU)	Test Score	Hired?
Participant A	Blue	20	A	Yes
Participant B	Blue	20	A	Yes
Participant C	Blue	15	A	Yes
Participant D	Blue	10	A	Yes
Participant E	Blue	5	A	Yes

For your reference, before you make your investment decision, here is some information about previous **Blue participants who have completed this task:**

Participant	Group	Investment (ECU)	Test Score	Hired?
Participant A	Blue	20	B	No
Participant B	Blue	20	C	No
Participant C	Blue	15	B	No
Participant D	Blue	13	C	No
Participant E	Blue	10	C	No

For your reference, before you make your investment decision, here is some information about previous **Green participants who have completed this task:**

Participant	Group	Investment (ECU)	Test Score	Hired?
Participant A	Green	20	A	Yes
Participant B	Green	20	B	Yes
Participant C	Green	15	A	Yes
Participant D	Green	12	A	Yes
Participant E	Green	11	A	Yes

For your reference, before you make your investment decision, here is some information about previous **Green participants who have completed this task:**

Participant	Group	Investment (ECU)	Test Score	Hired?
Participant A	Green	20	C	No
Participant B	Green	11	C	No
Participant C	Green	10	C	No
Participant D	Green	10	C	No
Participant E	Green	10	C	No

Note: Top left: “Successful Role Models” screen for the Blue group; top right: “Successful Role Models” screen for the Green group; bottom left: “Unsuccessful Role Models” screen for the Blue group; bottom right: “Unsuccessful Role Models” screen for the Green group.

Figure A2: “No Risk” treatment reminders on decision page

How much ECU would you like to invest to improve your probability of becoming a top performer? Please input below any whole number from 0 to 20.

Reminder: You are in the **Blue Group**. The hiring algorithm only hires a **Blue** applicant with a score of A.

Your investment cost is only deducted from your earnings if you are hired. If not, you keep the full base bonus of 20 ECU.

How much ECU would you like to invest to improve your probability of becoming a top performer? Please input below any whole number from 0 to 20.

Reminder: You are in the **Green Group**. The hiring algorithm hires a **Green** applicant with either a score of A or B.

Your investment cost is only deducted from your earnings if you are hired. If not, you keep the full base bonus of 20 ECU.

Note: left: “No Risk” reminder for the Blue group; right: “No Risk” reminder for the Green group.

A.4 Appendix Tables

Table A1: Summary statistics

Variable	Summary (Mean, SD)
Age	45.79 (15.62)
Risk attitude	6.50 (2.55)
Trust in major tech companies	5.72 (2.40)
Male	49.4%
Employed	69.3%
Prior experience with AI or algorithm	35.8%
Education: Some high school or less	1.1%
Education: High school diploma or GED	13.4%
Education: Some college, but no degree	16.6%
Education: Associates or technical degree	14.3%
Education: Bachelor's degree	35.7%
Education: Graduate or professional degree (MA, MS, MBA, PhD, JD, MD, etc.)	18.8%
Income: Less than 25k	24.2%
Income: 25-50k	25.5%
Income: 50-75k	19.5%
Income: 75-100k	14.6%
Income: 100-125k	4.9%
Income: Greater than or equal to 125k	11.4%
Observations	553

Note: Age is measured in years as a continuous variable. Risk attitude is recorded on a 0–10 scale, with higher scores indicating greater willingness to take risks. Trust in major tech companies is based on responses to the question: “In general, how much do you trust major technology companies to make decisions that affect you in a fair and unbiased way?”, also on a 0–10 scale. Male, Employed, and Prior experience with AI or algorithm are binary variables coded as 1 if the participant identifies as male, reports being employed (full-time or part-time), or has previously submitted an application believed to be screened by an AI system or algorithm.

Table A2: Summary statistics by group

Variable	Blue	Green	Test differences
Age	46.24	45.34	(0.502)
Risk attitude	6.51	6.48	(0.888)
Trust in major tech companies	5.61	5.83	(0.276)
Male	48.5%	50.2%	(0.701)
Employed	69.3%	69.5%	(0.961)
Prior experience with AI or algorithm	36.1%	35.5%	(0.874)
Education: Some high school or less	1.5%	0.7%	(0.447)
Education: High school diploma or GED	14.6%	12.2%	(0.479)
Education: Some college, but no degree	16.4%	16.8%	(0.985)
Education: Associates or technical degree	12.8%	15.8%	(0.376)
Education: Bachelor's degree	35.4%	36.2%	(0.915)
Education: Graduate or professional degree (MA, MS, MBA, PhD, JD, MD, etc.)	19.3%	18.3%	(0.833)
Income: Less than 25k	25.5%	22.9%	(0.538)
Income: 25-50k	25.5%	25.4%	(1.000)
Income: 50-75k	16.1%	22.9%	(0.053)
Income: 75-100k	13.5%	15.8%	(0.526)
Income: 100-125k	5.5%	4.3%	(0.658)
Income: Greater than or equal to 125k	13.9%	8.6%	(0.068)
Observations	274	279	

Note: Age is measured in years as a continuous variable. Risk attitude is recorded on a 0–10 scale, with higher scores indicating greater willingness to take risks. Trust in major tech companies is based on responses to the question: “In general, how much do you trust major technology companies to make decisions that affect you in a fair and unbiased way?”, also on a 0–10 scale. Male, Employed, and Prior experience with AI or algorithm are binary variables coded as 1 if participants identify as male, report being employed (full-time or part-time), or have previously submitted an application believed to be screened by an AI system or algorithm. The column labeled Test differences reports p-values from two-sample t-tests for continuous variables and difference-in-means tests for binary variables, and from chi-squared or Fisher’s exact tests for categorical variables (Education and Income).

Table A3: Summary statistics by first-round treatment assignment

Variable	Control	Successful RM	Unsuccessful RM	No Risk	Test differences
Age	46.90	47.14	45.01	44.15	(0.317)
Risk attitude	6.40	6.32	6.58	6.68	(0.644)
Trust in major tech companies	5.74	5.49	5.81	5.83	(0.641)
Male	50.7%	50.8%	53.1%	42.6%	(0.326)
Employed	63.6%	65.9%	75.2%	72.8%	(0.111)
Prior experience with AI or algorithm	38.6%	36.4%	33.1%	35.3%	(0.812)
Education: Associates or technical degree	17.9%	18.2%	10.3%	11%	(0.106)
Education: Bachelor's degree	32.1%	34.8%	37.2%	39%	(0.661)
Education: Graduate or professional degree (MA, MS, MBA, PhD, JD, MD, etc.)	15.7%	18.9%	19.3%	21.3%	(0.690)
Education: High school diploma or GED	12.9%	11.4%	14.5%	14.7%	(0.836)
Education: Some college, but no degree	20%	15.9%	16.6%	14%	(0.596)
Education: Some high school or less	1.4%	0.8%	2.1%	0%	(0.494)
Income: 100-125k	5.7%	6.8%	5.5%	1.5%	(0.186)
Income: 25-50k	26.4%	31.1%	24.1%	20.6%	(0.253)
Income: 50-75k	22.9%	10.6%	20.7%	23.5%	(0.027)
Income: 75-100k	12.1%	15.9%	13.8%	16.9%	(0.679)
Income: Greater than or equal to 125k	8.6%	7.6%	14.5%	14%	(0.149)
Income: Less than 25k	24.3%	28%	21.4%	23.5%	(0.633)
Observations	140	132	145	136	

Note: Summary statistics are reported at the participant level by randomized first-round treatment assignment. Age is measured in years. Risk attitude is measured on a 0–10 scale, with higher values indicating greater willingness to take risks. Trust in major technology companies is measured on a 0–10 scale, with higher values indicating greater trust in their ability to make fair and unbiased decisions. Male, Employed, and Prior experience with AI or algorithm are binary indicators equal to one for participants who identify as male, are employed full-time or part-time, or have previously submitted an application they believed was screened by an AI system or algorithm, respectively. The column “Test differences” reports p -values from one-way ANOVA F-tests of equality across the four randomized first-round treatment groups for continuous and binary variables, and from chi-squared tests (or Fisher’s exact tests when expected cell counts are small) for categorical variables.

Table A4: Treatment effects by group

	<i>Dependent variable:</i>		
	Investment		
	(1)	(2)	(3)
Blue Group	2.317** (0.929)	2.172** (0.944)	2.014** (0.943)
Successful RM (Green)	0.503 (0.847)	0.481 (0.854)	0.286 (0.861)
Unsuccessful RM (Green)	0.152 (0.889)	0.087 (0.895)	−0.130 (0.909)
No Risk (Green)	1.216 (0.750)	1.212 (0.760)	1.207 (0.763)
Blue x Successful RM	0.551 (1.145)	0.735 (1.175)	0.834 (1.176)
Blue x Unsuccessful RM	−0.672 (1.232)	−0.405 (1.257)	−0.303 (1.261)
Blue x No Risk	2.264** (1.046)	2.206** (1.054)	2.187** (1.047)
Constant	11.744*** (0.678)	11.445*** (1.591)	12.897*** (2.284)
Round Control	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes
Risk Attitude	No	No	Yes
Prior Experience	No	No	Yes
Trust in Tech	No	No	Yes
Observations	1,106	1,106	1,106
R ²	0.052	0.074	0.112

Note: The dependent variable is the number of investment units allocated (ranging from 0 to 20 inclusive). All models include round fixed effects. Column (2) adds demographic controls: age (continuous), gender (binary), education level (ordered factor), employment status (binary), and income bracket (ordered factor). Column (3) further includes individual-level covariates: risk attitude (self-reported on a 0–10 scale), prior experience with algorithmic screening (binary), and trust in major technology companies (0–10 scale). Robust standard errors, clustered at the participant level, are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A5: Effect of group (Blue vs. Green) on investment in control treatment (excluding participants with duration time outside the 10th to 90th percentile)

	<i>Dependent variable:</i>		
	Investment		
	(1)	(2)	(3)
Blue Group (vs Green)	2.311** (1.038)	2.666** (1.116)	2.640** (1.145)
Constant	12.409*** (0.891)	7.327** (2.864)	7.831* (4.264)
Round Control	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes
Risk Attitude	No	No	Yes
Prior Experience	No	No	Yes
Trust in Tech	No	No	Yes
Observations	219	219	219
R ²	0.051	0.197	0.307

Note: This table reports OLS regressions estimating the effect of being assigned to the Blue group (relative to the Green group) on investment decisions in the Control treatment only. The sample excludes participants with completion times below the 10th percentile or above the 90th percentile, following the pre-registered robustness specification. The dependent variable is the number of investment units allocated (ranging from 0 to 20 inclusive). All models include round fixed effects. Column (2) adds demographic controls: age (continuous), gender (binary), education level (ordered factor), employment status (binary), and income bracket (ordered factor). Column (3) further includes individual-level covariates: risk attitude (self-reported on a 0–10 scale), prior experience with algorithmic screening (binary), and trust in major technology companies (0–10 scale). Robust standard errors, clustered at the participant level, are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A6: Treatment effects by group (excluding participants with duration time outside the 10th to 90th percentile)

	<i>Dependent variable:</i>		
	Investment		
	(1)	(2)	(3)
Blue Group	2.435** (1.039)	2.251** (1.053)	2.121** (1.037)
Successful RM (Green)	0.971 (0.948)	0.914 (0.950)	0.799 (0.954)
Unsuccessful RM (Green)	0.617 (0.970)	0.493 (0.973)	0.388 (0.983)
No Risk (Green)	1.727** (0.816)	1.756** (0.821)	1.643** (0.824)
Blue x Successful RM	0.336 (1.286)	0.374 (1.307)	0.560 (1.305)
Blue x Unsuccessful RM	-1.449 (1.381)	-1.336 (1.399)	-1.149 (1.387)
Blue x No Risk	2.137* (1.169)	2.025* (1.180)	2.192* (1.173)
Constant	11.681*** (0.742)	10.172*** (1.790)	11.454*** (2.440)
Round Control	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes
Risk Attitude	No	No	Yes
Prior Experience	No	No	Yes
Trust in Tech	No	No	Yes
Observations	886	886	886
R ²	0.057	0.093	0.142

Note: This table reports OLS regressions estimating treatment effects by group, excluding participants with completion times below the 10th percentile or above the 90th percentile, following the pre-registered robustness specification. The dependent variable is the number of investment units allocated (ranging from 0 to 20 inclusive). All models include round fixed effects. Column (2) adds demographic controls: age (continuous), gender (binary), education level (ordered factor), employment status (binary), and income bracket (ordered factor). Column (3) further includes individual-level covariates: risk attitude (self-reported on a 0–10 scale), prior experience with algorithmic screening (binary), and trust in major technology companies (0–10 scale). Robust standard errors, clustered at the participant level, are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A7: Effect of group (Blue vs. Green) on investment in control treatment using only first-round data

	<i>Dependent variable:</i>		
	Investment		
	(1)	(2)	(3)
Blue Group (vs Green)	1.939 (1.259)	1.967 (1.384)	2.167 (1.413)
Constant	12.554*** (0.921)	8.558*** (3.063)	5.265 (4.805)
Demographic Controls	No	Yes	Yes
Risk Attitude	No	No	Yes
Prior Experience	No	No	Yes
Trust in Tech	No	No	Yes
Observations	140	140	140
R ²	0.017	0.154	0.358

Note: This table reports OLS regressions estimating the effect of group assignment (Blue vs. Green) on investment in the Control treatment, using only first-round observations as specified in the pre-registration. The dependent variable is the number of investment units allocated (ranging from 0 to 20 inclusive). Column (2) adds demographic controls: age (continuous), gender (binary), education level (ordered factor), employment status (binary), and income bracket (ordered factor). Column (3) further includes individual-level covariates: risk attitude (self-reported on a 0–10 scale), prior experience with algorithmic screening (binary), and trust in major technology companies (0–10 scale). Robust standard errors, clustered at the participant level, are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A8: Treatment effects using only first-round data

	<i>Dependent variable:</i>		
	Investment		
	(1)	(2)	(3)
Blue Group	1.939 (1.209)	1.458 (1.245)	1.422 (1.243)
Successful RM (Green)	1.087 (1.256)	0.739 (1.289)	0.568 (1.297)
Unsuccessful RM (Green)	-1.878 (1.213)	-2.079* (1.248)	-2.439* (1.264)
No Risk (Green)	-0.475 (1.205)	-0.653 (1.243)	-0.895 (1.271)
Blue x Successful RM	0.302 (1.733)	1.208 (1.800)	0.987 (1.806)
Blue x Unsuccessful RM	0.357 (1.693)	0.825 (1.757)	1.213 (1.765)
Blue x No Risk	3.048* (1.726)	3.305* (1.789)	3.261* (1.810)
Constant	12.554*** (0.885)	12.491*** (1.895)	12.562*** (2.591)
Demographic Controls	No	Yes	Yes
Risk Attitude	No	No	Yes
Prior Experience	No	No	Yes
Trust in Tech	No	No	Yes
Observations	553	553	553
R ²	0.069	0.099	0.154

Note: This table reports OLS regressions estimating treatment effects by group using only first-round observations, following the pre-registered robustness specification. The dependent variable is the number of investment units allocated (ranging from 0 to 20 inclusive). Column (2) adds demographic controls: age (continuous), gender (binary), education level (ordered factor), employment status (binary), and income bracket (ordered factor). Column (3) further includes individual-level covariates: risk attitude (self-reported on a 0–10 scale), prior experience with algorithmic screening (binary), and trust in major technology companies (0–10 scale). Robust standard errors, clustered at the participant level, are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A9: First-round and pooled treatment effect heterogeneity

Treatment	First round	Pooled	Pooled - First round	Bootstrap SE	p-value
Successful Role Model	0.302	0.551	0.249	1.478	0.866
Unsuccessful Role Model	0.357	-0.672	-1.028	1.428	0.472
No Risk	3.048	2.264	-0.785	1.248	0.530

Note: The first-round and pooled estimates are based on the specifications reported in Tables A8 and A4, respectively. The first two numeric columns report the difference between the treatment effect for Blue and Green participants for each treatment. The third numeric column reports the pooled estimate minus the corresponding first-round estimate. Bootstrap standard errors and p -values are obtained from participant-level resampling and test whether the pooled and first-round Blue-Green treatment-effect differences are equal.

Table A10: Treatment effects across decision rounds

	<i>Dependent variable:</i>	
	Investment	
	Treatment $\times$ Round	Group $\times$ Treatment $\times$ Round
	(1)	(2)
Round 2	-2.098** (0.931)	-2.356* (1.348)
Blue $\times$ Round 2		0.530 (1.863)
Successful RM $\times$ Round 2	-0.559 (1.262)	-0.810 (1.857)
Unsuccessful RM $\times$ Round 2	2.881** (1.348)	3.761** (1.866)
No Risk $\times$ Round 2	4.066*** (1.541)	5.909** (2.545)
Blue $\times$ Successful RM $\times$ Round 2		0.473 (2.526)
Blue $\times$ Unsuccessful RM $\times$ Round 2		-1.775 (2.699)
Blue $\times$ No Risk $\times$ Round 2		-3.408 (3.127)
Group $\times$ Treatment Effects	Yes	Yes
Treatment $\times$ Round Effects	Yes	Yes
Group $\times$ Treatment $\times$ Round Effects	No	Yes
Joint test: Treatment $\times$ Round	$p = 0.002$	
Joint test: Blue $\times$ Treatment $\times$ Round		$p = 0.488$
Observations	1,106	1,106
R ²	0.066	0.068

Note: The dependent variable is the number of investment units allocated (ranging from 0 to 20 inclusive). Green, Control, and Round 1 are the reference categories. Column (1) allows the effects of the Successful Role Model, Unsuccessful Role Model, and No Risk treatments to differ between Round 1 and Round 2. Column (2) additionally allows these round-specific treatment effects to differ between Blue and Green participants. The regressions also include the main effects of group and treatment and the group-by-treatment interactions; these coefficients are omitted from the table for brevity. Robust standard errors clustered at the participant level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Round 2 investment and potential carryover check

	<i>Dependent variable:</i>	
	Round 2 Investment	
	Control in Round 1	All Participants
	(1)	(2)
Round 1 Investment	0.614*** (0.073)	0.551*** (0.041)
Blue	-0.199 (0.967)	0.512 (0.610)
Unsuccessful RM in Round 2	1.210 (1.319)	
No Risk in Round 2	3.407*** (1.157)	
Hired in Round 1		-0.039 (1.450)
Hired in Round 1 $\times$ Successful RM in Round 2		-1.195 (1.724)
Hired in Round 1 $\times$ Unsuccessful RM in Round 2		-0.374 (1.867)
Hired in Round 1 $\times$ No Risk in Round 2		-1.665 (1.944)
Round 1 Treatment	Control	All Treatments
Treatment-Sequence Controls	No	Yes
Joint test: Hiring R1 $\times$ Treatment R2		$p = 0.793$
Observations	140	552
R ²	0.448	0.304

Note: The dependent variable is investment in Round 2. Column (1) restricts the sample to participants assigned to the Control treatment in Round 1 and estimates Round 2 investment as a function of Round 1 investment, group assignment, and Round 2 treatment. Successful Role Model is the omitted Round 2 treatment category. Column (2) uses all participants for whom the Round 1 hiring outcome was recorded and examines whether the Round 1 hiring outcome predicts Round 2 investment or moderates the response to the Round 2 treatment, conditional on Round 1 investment, group assignment, and treatment sequence. One participant is excluded from Column (2) because the hiring outcome was not recorded. Control is the omitted Round 2 treatment category for the hiring-by-treatment interactions. Treatment-sequence coefficients in Column (2) are included in the regression but omitted from the table for brevity. The reported joint test evaluates whether the three interactions between Round 1 hiring and Round 2 treatment are jointly equal to zero. Robust standard errors are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A12: Treatment effects by age

	<i>Dependent variable:</i>		
	Investment		
	(1)	(2)	(3)
Blue	2.654*** (0.547)	2.598*** (0.558)	2.489*** (0.547)
Successful RM	0.757 (0.581)	0.846 (0.591)	0.693 (0.587)
Unsuccessful RM	−0.211 (0.618)	−0.129 (0.627)	−0.292 (0.616)
No Risk	2.323*** (0.518)	2.295*** (0.522)	2.283*** (0.523)
Age	0.061** (0.031)	0.053* (0.032)	0.063** (0.032)
Age × Successful RM	−0.061 (0.037)	−0.058 (0.038)	−0.058 (0.038)
Age × Unsuccessful RM	−0.045 (0.038)	−0.045 (0.038)	−0.044 (0.038)
Age × No Risk	−0.020 (0.034)	−0.021 (0.035)	−0.016 (0.035)
Round Fixed Effects	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes
Additional Controls	No	No	Yes
Observations	1,106	1,106	1,106
R ²	0.054	0.068	0.106

Note: The dependent variable is the number of investment units allocated (ranging from 0 to 20 inclusive). Age is measured in years and mean-centered. The omitted categories are the Green group and the Control treatment. All models include round fixed effects. Column (2) adds demographic controls for gender, education level, employment status, and income bracket. Column (3) further includes controls for risk attitude, prior experience with algorithmic screening, and trust in major technology companies. Control-variable coefficients are omitted from the table for brevity. Robust standard errors clustered at the participant level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A13: Treatment effects by gender

	<i>Dependent variable:</i>		
	Investment		
	(1)	(2)	(3)
Blue	2.657*** (0.548)	2.568*** (0.558)	2.470*** (0.548)
Successful RM	−0.117 (0.777)	−0.043 (0.778)	−0.078 (0.781)
Unsuccessful RM	−1.233 (0.865)	−1.143 (0.873)	−1.178 (0.870)
No Risk	1.551** (0.682)	1.579** (0.683)	1.557** (0.681)
Gender	−1.349 (0.934)	−1.397 (0.949)	−1.537 (0.945)
Gender × Successful RM	1.859 (1.148)	1.907* (1.150)	1.662 (1.158)
Gender × Unsuccessful RM	2.133* (1.235)	2.139* (1.238)	1.876 (1.237)
Gender × No Risk	1.626 (1.058)	1.514 (1.071)	1.531 (1.082)
Round Fixed Effects	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes
Additional Controls	No	No	Yes
Observations	1,106	1,106	1,106
R ²	0.052	0.068	0.107

Note: The dependent variable is the number of investment units allocated (ranging from 0 to 20 inclusive). Gender is coded as 1 if the participant identifies as male and 0 otherwise. The omitted categories are the Green group and the Control treatment. All models include round fixed effects. Column (2) adds demographic controls for age, education level, employment status, and income bracket. Column (3) further includes controls for risk attitude, prior experience with algorithmic screening, and trust in major technology companies. Control-variable coefficients are omitted from the table for brevity. Robust standard errors clustered at the participant level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A14: Treatment effects by education

	<i>Dependent variable:</i>		
	Investment		
	(1)	(2)	(3)
Blue	2.691*** (0.548)	2.584*** (0.557)	2.484*** (0.547)
Successful RM	0.775 (0.582)	0.853 (0.593)	0.713 (0.588)
Unsuccessful RM	-0.184 (0.619)	-0.123 (0.628)	-0.275 (0.618)
No Risk	2.321*** (0.518)	2.310*** (0.520)	2.296*** (0.520)
Education	0.189 (0.355)	0.080 (0.379)	0.036 (0.378)
Education × Successful RM	-0.241 (0.443)	-0.180 (0.450)	-0.158 (0.445)
Education × Unsuccessful RM	0.112 (0.453)	0.168 (0.460)	0.234 (0.452)
Education × No Risk	-0.789* (0.409)	-0.772* (0.412)	-0.761* (0.407)
Round Fixed Effects	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes
Additional Controls	No	No	Yes
Observations	1,106	1,106	1,106
R ²	0.051	0.062	0.100

Note: The dependent variable is the number of investment units allocated (ranging from 0 to 20 inclusive). Education is coded as an ordinal variable ranging from 1 (some high school or less) to 6 (graduate or professional degree) and is mean-centered. The omitted categories are the Green group and the Control treatment. All models include round fixed effects. Column (2) adds demographic controls for age, gender, employment status, and income bracket. Column (3) further includes controls for risk attitude, prior experience with algorithmic screening, and trust in major technology companies. Control-variable coefficients are omitted from the table for brevity. Robust standard errors clustered at the participant level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A15: Treatment effects by income

	<i>Dependent variable:</i>		
	Investment		
	(1)	(2)	(3)
Blue	2.651*** (0.549)	2.664*** (0.549)	2.544*** (0.542)
Successful RM	0.766 (0.581)	0.755 (0.581)	0.578 (0.579)
Unsuccessful RM	−0.197 (0.618)	−0.214 (0.619)	−0.395 (0.610)
No Risk	2.288*** (0.522)	2.297*** (0.521)	2.323*** (0.522)
Income	0.231 (0.212)	0.260 (0.213)	0.337* (0.197)
Income × Successful RM	−0.465* (0.254)	−0.443* (0.252)	−0.525** (0.238)
Income × Unsuccessful RM	−0.328 (0.274)	−0.322 (0.271)	−0.403 (0.263)
Income × No Risk	0.028 (0.237)	0.029 (0.238)	0.075 (0.253)
Round Fixed Effects	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes
Additional Controls	No	No	Yes
Observations	1,106	1,106	1,106
R ²	0.052	0.062	0.102

Note: The dependent variable is the number of investment units allocated (ranging from 0 to 20 inclusive). Income is measured using the participant's gross annual income bracket, coded as an ordinal variable ranging from 1 (less than 25,000 USD) to 13 (300,000 USD or more), and is mean-centered. The omitted categories are the Green group and the Control treatment. All models include round fixed effects. Column (2) adds demographic controls for age, gender, education, and employment status. Column (3) further includes controls for risk attitude, prior experience with algorithmic screening, and trust in major technology companies. Control-variable coefficients are omitted from the table for brevity. Robust standard errors clustered at the participant level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A16: Treatment effects by employment

	<i>Dependent variable:</i>		
	Investment		
	(1)	(2)	(3)
Blue	2.648*** (0.549)	2.564*** (0.558)	2.453*** (0.547)
Successful RM	0.745 (1.107)	0.597 (1.110)	0.438 (1.109)
Unsuccessful RM	−0.029 (1.144)	−0.161 (1.147)	−0.283 (1.147)
No Risk	3.721*** (0.933)	3.650*** (0.925)	3.740*** (0.921)
Employment	−0.169 (1.041)	−0.469 (1.092)	−0.429 (1.080)
Employment × Successful RM	0.054 (1.284)	0.368 (1.295)	0.388 (1.300)
Employment × Unsuccessful RM	−0.224 (1.356)	0.067 (1.365)	0.013 (1.366)
Employment × No Risk	−2.070* (1.137)	−1.997* (1.133)	−2.163* (1.127)
Round Fixed Effects	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes
Additional Controls	No	No	Yes
Observations	1,106	1,106	1,106
R ²	0.052	0.068	0.107

Note: The dependent variable is the number of investment units allocated (ranging from 0 to 20 inclusive). Employment is coded as 1 if the participant reports working full-time or part-time and 0 if the participant reports being unemployed and looking for work, a homemaker or stay-at-home parent, a student, or retired. The omitted categories are the Green group and the Control treatment. All models include round fixed effects. Column (2) adds demographic controls for age, gender, education, and income. Column (3) further includes controls for risk attitude, prior experience with algorithmic screening, and trust in major technology companies. Control-variable coefficients are omitted from the table for brevity. Robust standard errors clustered at the participant level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A17: Treatment effects by risk attitude

	<i>Dependent variable:</i>		
	Investment		
	(1)	(2)	(3)
Blue	2.654*** (0.549)	2.570*** (0.558)	2.471*** (0.553)
Successful RM	0.789 (0.583)	0.872 (0.592)	0.878 (0.589)
Unsuccessful RM	-0.185 (0.619)	-0.113 (0.627)	-0.107 (0.621)
No Risk	2.311*** (0.522)	2.296*** (0.524)	2.314*** (0.524)
Risk attitude	0.116 (0.186)	0.109 (0.189)	0.158 (0.195)
Risk attitude × Successful RM	0.020 (0.221)	0.047 (0.223)	0.087 (0.221)
Risk attitude × Unsuccessful RM	0.020 (0.248)	0.061 (0.249)	0.083 (0.247)
Risk attitude × No Risk	-0.088 (0.200)	-0.095 (0.201)	-0.092 (0.200)
Round Fixed Effects	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes
Additional Controls	No	No	Yes
Observations	1,106	1,106	1,106
R ²	0.050	0.067	0.089

Note: The dependent variable is the number of investment units allocated (ranging from 0 to 20 inclusive). Risk attitude is measured using the participant's response on a 0–10 scale to the question: "In general, how willing are you to take risks?" (0 = completely unwilling to take risks; 10 = very willing to take risks) and is mean-centered. The omitted categories are the Green group and the Control treatment. All models include round fixed effects. Column (2) adds demographic controls for age, gender, education, employment status, and income. Column (3) further includes controls for prior experience with algorithmic screening and trust in major technology companies. Control-variable coefficients are omitted from the table for brevity. Robust standard errors clustered at the participant level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A18: Treatment effects by trust in major tech companies

	<i>Dependent variable:</i>		
	Investment		
	(1)	(2)	(3)
Blue	2.686*** (0.547)	2.596*** (0.557)	2.538*** (0.545)
Successful RM	0.757 (0.581)	0.837 (0.589)	0.627 (0.588)
Unsuccessful RM	-0.192 (0.617)	-0.125 (0.623)	-0.349 (0.618)
No Risk	2.346*** (0.524)	2.342*** (0.528)	2.293*** (0.529)
Trust	-0.267 (0.191)	-0.339* (0.196)	-0.417** (0.195)
Trust × Successful RM	0.349 (0.234)	0.437* (0.241)	0.461* (0.238)
Trust × Unsuccessful RM	0.480* (0.249)	0.557** (0.252)	0.594** (0.247)
Trust × No Risk	-0.001 (0.213)	-0.027 (0.215)	-0.033 (0.215)
Round Fixed Effects	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes
Additional Controls	No	No	Yes
Observations	1,106	1,106	1,106
R ²	0.053	0.072	0.098

Note: The dependent variable is the number of investment units allocated (ranging from 0 to 20 inclusive). Trust is measured using the participant's response on a 0–10 scale to the question: "In general, how much do you trust major technology companies to make decisions that affect you in a fair and unbiased way?" (0 = no trust at all; 10 = complete trust) and is mean-centered. The omitted categories are the Green group and the Control treatment. All models include round fixed effects. Column (2) adds demographic controls for age, gender, education, employment status, and income. Column (3) further includes controls for risk attitude and prior experience with algorithmic screening. Control-variable coefficients are omitted from the table for brevity. Robust standard errors clustered at the participant level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A19: Treatment effects by prior experience with AI/algorithms

	<i>Dependent variable:</i>		
	Investment		
	(1)	(2)	(3)
Blue	2.660*** (0.549)	2.574*** (0.557)	2.522*** (0.549)
Successful RM	0.624 (0.720)	0.743 (0.730)	0.582 (0.717)
Unsuccessful RM	−0.102 (0.788)	−0.008 (0.799)	−0.176 (0.780)
No Risk	2.255*** (0.643)	2.232*** (0.649)	2.217*** (0.653)
Prior experience	−0.866 (0.959)	−0.834 (0.982)	−1.154 (0.996)
Prior experience × Successful RM	0.374 (1.195)	0.298 (1.203)	0.554 (1.207)
Prior experience × Unsuccessful RM	−0.273 (1.260)	−0.298 (1.262)	−0.078 (1.255)
Prior experience × No Risk	0.118 (1.108)	0.127 (1.111)	0.109 (1.113)
Round Fixed Effects	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes
Additional Controls	No	No	Yes
Observations	1,106	1,106	1,106
R ²	0.051	0.068	0.101

Note: The dependent variable is the number of investment units allocated (ranging from 0 to 20 inclusive). Prior experience is coded as 1 if the participant reports having previously submitted an application that they believe was screened or reviewed by an AI system or algorithm, and 0 otherwise. The omitted categories are the Green group and the Control treatment. All models include round fixed effects. Column (2) adds demographic controls for age, gender, education, employment status, and income. Column (3) further includes controls for risk attitude and trust in major technology companies. Control-variable coefficients are omitted from the table for brevity. Robust standard errors clustered at the participant level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A20: Self-Reported Motivations and Investment

	<i>Dependent variable:</i>	
	Investment	
	Group Identity	All Motivations
	(1)	(2)
Group identity	-1.894** (0.853)	-2.001*** (0.766)
Blue × Group identity	3.072*** (1.189)	3.409*** (1.092)
Cost when not hired		-4.289*** (0.827)
Blue × Cost when not hired		1.864 (1.160)
Starting amount		1.777** (0.808)
Blue × Starting amount		-0.729 (1.066)
Potential bonus		2.306*** (0.783)
Blue × Potential bonus		1.420 (1.075)
Other		-6.413** (2.976)
Blue × Other		8.980** (4.010)
Group × Treatment Controls	Yes	Yes
Round Fixed Effects	Yes	Yes
Observations	1,106	1,106
R ²	0.062	0.166

Note: The dependent variable is the number of investment units allocated (ranging from 0 to 20 inclusive). Motivation indicators equal 1 if the participant selected the corresponding factor in the post-experiment multiple-response motivation question and 0 otherwise. Column (1) examines the association between reporting group identity as a motivation and investment differs between Blue and Green participants. Column (2) additionally includes the other motivation factors that were available to all participants and their interactions with group assignment. Green is the reference group. Both specifications include group-by-treatment controls and round fixed effects. Robust standard errors clustered at the participant level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

B Full Survey



Welcome & Consent

Welcome and Study Information

You are invited to take part in a short research study on decision-making, conducted by a researcher from the School of Economics at University College Dublin.

What you'll do:

- You will be given some information about a scenario and then you will be asked to make a number of decisions
- You have the chance to earn an additional bonus depending on your decisions
- Answer a brief demographic questionnaire

Participation details:

- You must be at least **18 years old**
- The study takes about **10 minutes**
- On top of the participation fee, you will earn a bonus between **\$0 USD** to **\$1.2 USD** based on your decisions

Data & Privacy:

- Your responses are **anonymous** and **confidential**
- No personal identifiers (e.g. name, email, IP) will be collected
- Data will be stored securely and handled according to GDPR and university policies

Voluntary Participation:

- Participation is voluntary — you can withdraw at any time before submitting
- After submission, data is anonymized and cannot be withdrawn

Questions?

You can contact the researcher at yung-shiang.yang@ucdconnect.ie.

By clicking the button below, you confirm that:

- You are at least 18 years old
- You understand your participation is voluntary
- You may withdraw at any time before submission

I consent, begin the study

I do not consent, I do not wish to participate

Prolific ID

What is your Prolific ID? Please note that this response should auto-fill with the correct ID

`${e://Field/PROLIFIC_PID}`

Payment Information

Payment Information

About ECU:

You'll earn and spend points called **ECU (Experimental Currency Units)**, **100 ECU = \$1.0 USD**, this is used to calculate your bonus.

You will start with:

20 ECU + a chance to earn an additional **100 ECU** based on your decision and outcome in a randomly selected round in the experiment.

Make your decisions carefully.

Scenario Instructions - Control

Scenario Overview

- You are applying for a job where a **hiring algorithm** decides who gets hired.

- There are two groups of applicants: **Green** and **Blue**. You've been assigned to the **#{e://Field/GroupColor}**.
-

Goal of the Algorithm

- Each applicant will either become a **top** or an **average** performer in the job.
 - The algorithm only wants to hire applicants who will become **top performers** in the job, but it can't tell who they are with certainty.
 - Therefore, it asks that every applicant take a **pre-employment test** — a signal, though not a perfect one, of future performance.
-

Algorithm's Belief

- The algorithm uses both your **test score** and your **group identity** to decide whether to hire you.
 - It believes **Green** applicants are more likely to become top performers than **Blue** applicants — even though in reality, **everyone starts with a 50%** chance of becoming a top performer.
-

Investment Option

- You can invest any amount from your base bonus of **20 ECU** to boost your chance of becoming a **top performer**.
 - **1 ECU = +1.5%** boost. So, **10 ECU = +15%**; **20 ECU = +30%**.
 - Your total chance of becoming a top performer = **50% + investment boost**.
-

The Pre-Employment Test

- If you're **a top performer**, you have 70% chance of scoring **A** (the highest score), 30% chance of scoring **B**, and 0% chance of scoring **C** (the lowest score).
 - If you're **an average performer**, you have 0% chance of scoring **A**, 30% chance of scoring **B**, and 70% chance of scoring **C**.
-

The Algorithm's Hiring Rule

- **Green:** Hired if your score is **A or B**
 - **Blue:** Hired only if your score is **A**
-

Bonus

- **Base bonus:** 20 ECU
 - **If hired:** +100 ECU
 - **Important: Your investment cost is deducted from the base bonus whether you're hired or not.**
-

Examples

- Invest 10 ECU & hired: Bonus = $20 - 10 + 100 = \mathbf{110 \text{ ECU}}$
 - Invest 10 ECU & not hired: Bonus = $20 - 10 + 0 = \mathbf{10 \text{ ECU}}$
 - Invest 0 ECU & hired: Bonus = $20 - 0 + 100 = \mathbf{120 \text{ ECU}}$
-

 **You'll complete 1 practice and 2 formal rounds. Your bonus depends on 1 randomly chosen formal round — so decide carefully!**

Scenario Instructions - No Risk

Scenario Overview

- You are applying for a job where a **hiring algorithm** decides who gets hired.
- There are two groups of applicants: **Green** and **Blue**. You've been assigned to the **#{e://Field/GroupColor}**.

Goal of the Algorithm

- Each applicant will either become a **top** or an **average performer** in the job.
- The algorithm only wants to hire applicants who will become **top performers** in the job, but it can't tell who they are with certainty.
- Therefore, it asks that every applicant take a **pre-employment test** — a signal, though not a perfect one, of future performance.

Algorithm's Belief

- The algorithm uses both your **test score** and your **group identity** to decide whether to hire you.
- It believes **Green** applicants are more likely to become top performers than **Blue** applicants — even though in reality, **everyone starts with a 50%** chance of becoming a top performer.

Investment Option

- You can invest any amount from your base bonus of **20 ECU** to improve your chance of becoming a top performer.
- **1 ECU = +1.5%** chance. So, **10 ECU = +15%**; **20 ECU = +30%**.

- Your total chance of becoming a top performer = **50% + investment boost.**
-

The Pre-Employment Test

- If you're a top performer, you have 70% chance of scoring **A** (the highest score), 30% chance of scoring **B**, and 0% chance of scoring **C** (the lowest score).
 - If you're an average performer, you have 0% chance of scoring **A**, 30% chance of scoring **B**, and 70% chance of scoring **C**.
-

Hiring Rule

- **Green:** Hired if your score is **A or B**
 - **Blue:** Hired only if your score is **A**
-

Bonus

- **Base bonus:** 20 ECU
 - **If hired:** +100 ECU
 - **Important: Your investment cost is only deducted from the base bonus if you're hired.**
-

Examples

- Invest 10 ECU & hired: Bonus = $20 - 10 + 100 = \mathbf{110 \text{ ECU}}$
 - Invest 10 ECU & not hired: Bonus = $20 - 0 + 0 = \mathbf{20 \text{ ECU}}$
 - Invest 0 ECU & hired: Bonus = $20 - 0 + 100 = \mathbf{120 \text{ ECU}}$
-

 **You'll complete 1 practice and 2 formal rounds. Your bonus depends on 1 randomly chosen formal round — so decide carefully!**

Attention Check

Attention Check

To ensure you've understood the instructions, please answer the following questions. (You will not be able to proceed unless you answer correctly.)

Please drag and drop the following events into the correct order based on how the study works:

Test score is simulated

You are randomly assigned to a group

You decide how much to invest

The hiring algorithm makes a hiring decision

What is your chance of becoming a top performer in the job if you invest the maximum amount of 20 ECU?

(Hint: Everyone starts with a 50% chance of becoming a top performer. Each 1 ECU you invest increases your chance by 1.5 percentage points.)

65%

70%

75%

80%

Suppose, following the investment, you end up becoming a top performer. What are your chances of scoring A, B, C, respectively?

70%, 30%, 0%

0%, 30%, 70%

100%, 0%, 0%

Suppose you're a **Green applicant with a score of B. Will you be hired?**

*(Hint: The algorithm hires a **Green** applicant with either an **A** or **B**, only hires a **Blue** applicant with an **A**.)*

Yes

No

Suppose you're hired following your investment of 20 ECU. What will be the total amount of your bonus?

(Hint: You start off with a base bonus of 20 ECU, which you can invest at your discretion. You also get an additional bonus of 100 ECU if you're hired.)

120 ECU

100 ECU

0 ECU

20 ECU

Blue SRM

For your reference, before you make your investment decision, here is some information about previous **Blue** participants who have completed this task:

Participant	Group	Investment (ECU)	Test Score	Hired?
Participant A	Blue	20	A	Yes
Participant B	Blue	20	A	Yes
Participant C	Blue	15	A	Yes
Participant D	Blue	10	A	Yes
Participant E	Blue	5	A	Yes



Attention Check

To ensure you've read the information, please answer the following questions. (You will not be able to proceed unless you answer correctly.)

The previous participants shown above invested ____ ECU to improve their chances of becoming top performers.

some

zero

The previous participants shown above were ____ by the algorithm.

hired

not hired

Green SRM

For your reference, before you make your investment decision, here is some information about previous **Green participants who have completed this task:**

Participant	Group	Investment (ECU)	Test Score	Hired?
Participant A	Green	20	A	Yes
Participant B	Green	20	B	Yes
Participant C	Green	15	A	Yes

Participant	Group	Investment (ECU)	Test Score	Hired?
Participant D	Green	12	A	Yes
Participant E	Green	11	A	Yes

Attention Check

To ensure you've read the information, please answer the following questions. (You will not be able to proceed unless you answer correctly.)

The previous participants shown above invested ____ ECU to improve their chances of becoming top performers.

some

zero

The previous participants shown above were ____ by the algorithm.

hired

not hired

Blue URM

For your reference, before you make your investment decision, here is some information about previous **Blue participants who have completed this task:**

Participant	Group	Investment (ECU)	Test Score	Hired?
Participant A	Blue	20	B	No
Participant B	Blue	20	C	No
Participant C	Blue	15	B	No
Participant D	Blue	13	C	No
Participant E	Blue	10	C	No

Attention Check

To ensure you've read the information, please answer the following questions. (You will not be able to proceed unless you answer correctly.)

The previous participants shown above invested ____ ECU to improve their chances of becoming top performers.

some

zero

The previous participants shown above were ____ by the

algorithm.

hired

not hired

Green URM

For your reference, before you make your investment decision, here is some information about previous **Green** participants who have completed this task:

Participant	Group	Investment (ECU)	Test Score	Hired?
Participant A	Green	20	C	No
Participant B	Green	11	C	No
Participant C	Green	10	C	No
Participant D	Green	10	C	No
Participant E	Green	10	C	No

Attention Check

To ensure you've read the information, please answer the following questions. (You will not be able to proceed unless you answer correctly.)

The previous participants shown above invested ____ ECU to improve their chances of becoming top performers.

some

zero

The previous participants shown above were ____ by the algorithm.

hired

not hired

Practice Round

Practice Round

This round is just for practice and will not affect your final bonus. It is designed to help you understand how the investment and test process work.

How much ECU would you like to invest to improve your probability of becoming a top performer? Please input below any whole number from 0 to 20.

Reminder: You are in the \${e://Field/GroupColor}.

\${e://Field/GroupReminder}

Your investment cost is deducted from your earnings whether you're hired or not.

Practice Simulation

Thanks for making your decision!

Your test result is now simulating...



Click  to reveal your results.

Practice Results

Practice Round Result

Here's how your decision played out in the practice round (this will not affect your final bonus):

 **You invested:** $\${e://Field/PracticeInvestment}$

 **Your probability of becoming a top performer after investment is:** $\${e://Field/PracticeProb}\%$

 **Your test score:** $\${e://Field/PracticeTestScoreLabel}$

 **Hiring decision:** You are $\${e://Field/PracticeHired}$ by the hiring algorithm.

 **Your (hypothetical) bonus in the practice round:** $20 - \${e://Field/PracticeInvestmentDisplay} + \${e://Field/PracticeWage} = \${e://Field/PracticeEarnings}$ ECU

Note: This was just a practice round. You'll now complete **2 formal rounds**, and your bonus will be based on *one randomly selected formal round*.

Formal Round

✓ Formal Round 1 of 2

This is the first of two formal rounds. If this is the randomly chosen round, your decision here WILL affect your final bonus.

How much ECU would you like to invest to improve your probability of becoming a top performer? Please input below any whole number from 0 to 20.

Reminder: You are in the \${e://Field/GroupColor}.

\${e://Field/GroupReminder}

Simulation

Thanks for making your decision!

Your test result is now simulating...



Click  to reveal your results.

Results

Result of Round 1 of 2

This was the first of two formal rounds. If this is the randomly chosen round, the outcome below WILL determine your final bonus.

 **You invested:** $\${e://Field/Investment}$

 **Your probability of becoming a top performer after investment is:** $\${e://Field/Prob}\%$

 **Your test score:** $\${e://Field/TestScoreLabel}$

 **Hiring decision:** You are $\${e://Field/Hired}$ by the hiring algorithm.

 **Your bonus:** $20 - \${e://Field/InvestmentDisplay} + \${e://Field/Wage} = \${e://Field/Earnings}$ ECU

You've completed the first formal round. You'll now proceed to the second and final formal round.

Green LRM 2

For your reference, before you make your investment decision in the final round, here is some information about previous **Green participants who have completed this task:**

Participant	Group	Investment (ECU)	Test Score	Hired?
Participant A	Green	20	A	Yes
Participant B	Green	20	B	Yes
Participant C	Green	15	A	Yes
Participant D	Green	12	A	Yes
Participant E	Green	11	A	Yes

Attention Check

To ensure you've read the information, please answer the following questions. (You will not be able to proceed unless you answer correctly.)

The previous participants shown above invested ____ ECU to improve their chances of becoming top performers.

some

zero

The previous participants shown above were ____ by the algorithm.

hired

not hired

Blue LRM 2

For your reference, before you make your investment decision in the final round, here is some information about previous **Blue** participants who have completed this task:

Participant	Group	Investment (ECU)	Test Score	Hired?
Participant A	Blue	20	A	Yes
Participant B	Blue	20	A	Yes
Participant C	Blue	15	A	Yes
Participant D	Blue	10	A	Yes
Participant E	Blue	5	A	Yes

Attention Check

To ensure you've read the information, please answer the following questions. (You will not be able to proceed unless you answer correctly.)

The previous participants shown above invested ____ ECU to improve their chances of becoming top performers.

some

zero

The previous participants shown above were ____ by the algorithm.

hired

not hired

Green URM 2

For your reference, before you make your investment decision in the final round, here is some information about previous **Green** participants who have completed this task:

Participant	Group	Investment (ECU)	Test Score	Hired?
Participant A	Green	20	C	No
Participant B	Green	11	C	No
Participant C	Green	10	C	No
Participant D	Green	10	C	No
Participant E	Green	10	C	No



Attention Check

To ensure you've read the information, please answer the following questions. (You will not be able to proceed unless you answer correctly.)

The previous participants shown above invested ____ ECU to improve their chances of becoming top performers.

some

zero

The previous participants shown above were ____ by the algorithm.

hired

not hired

Blue URM 2

For your reference, before you make your investment decision in the final round, here is some information about previous **Blue participants who have completed this task:**

Participant	Group	Investment (ECU)	Test Score	Hired?
Participant A	Blue	20	B	No
Participant B	Blue	20	C	No
Participant C	Blue	15	B	No

Participant	Group	Investment (ECU)	Test Score	Hired?
Participant D	Blue	13	C	No
Participant E	Blue	10	C	No

Attention Check

To ensure you've read the information, please answer the following questions. (You will not be able to proceed unless you answer correctly.)

The previous participants shown above invested ____ ECU to improve their chances of becoming top performers.

some

zero

The previous participants shown above were ____ by the algorithm.

hired

not hired

Second Round No Risk Instructions

Attention: Change in Rules

In the following round, your investment cost is **only deducted if you're hired**. Your investment cost is **not deducted if you're not hired**. The hiring rule from the hiring algorithm remains the same.

Second Round Control Instructions

Attention: Change in Rules

In the following round, your investment cost **WILL be deducted whether you're hired or not**. The hiring rule from the hiring algorithm remains the same.

Second Formal Round

Formal Round 2 of 2

This is the second and final formal round. If this is the randomly chosen round, your decision here WILL affect your final bonus.

How much ECU would you like to invest to improve your probability of becoming a top performer? Please input below any whole number from 0 to 20.

Reminder: You are in the $\${e://Field/GroupColor}$.
 $\${e://Field/GroupReminder}$

Second Simulation

Thanks for making your decision!

Your test result is now simulating...



Click  to reveal your results.

Second Results

Result of Round 2 of 2

This was the second and final formal round. If this is the randomly chosen round, the outcome below WILL determine your final bonus.

 **You invested:** $\${e://Field/Investment2}$

 **Your probability of becoming a top performer after investment is:** $\${e://Field/Prob2}\%$

 **Your test score:** $\${e://Field/TestScoreLabel2}$

 **Hiring decision:** You are $\${e://Field/Hired2}$ by the hiring algorithm.

 **Your bonus:** $20 - \{e://Field/InvestmentDisplay2\} + \{e://Field/Wage2\} = \{e://Field/Earnings2\}$ ECU

This concludes the decision-making part of the study. You'll now answer a short demographic questionnaire.

Demographics

Please answer the following short questions.

Your responses are anonymous and will be used for research purposes only. You will not be personally identifiable from your answers.

What is your age?

(Please enter in years)

What is your gender?

Male

Female

Non-binary / third gender

Prefer not to say

What is the highest level of education you have completed?

Some high school or less

High school diploma or GED

Some college, but no degree

Associates or technical degree

Bachelor's degree

Graduate or professional degree (MA, MS, MBA, PhD, JD, MD, etc.)

Prefer not to say

What best describes your employment status?

Working full-time

Working part-time

Unemployed and looking for work

A homemaker or stay-at-home parent

Student

Retired

Other (please specify in the textbox below)

Please estimate your gross annual salary.

Less than 25,000 USD

Greater than/equal to 25,000 USD and less than 50,000 USD

Greater than/equal to 50,000 USD and less than 75,000 USD

Greater than/equal to 75,000 USD and less than 100,000 USD

Greater than/equal to 100,000 USD and less than 125,000 USD

Greater than/equal to 125,000 USD and less than 150,000 USD

Greater than/equal to 150,000 USD and less than 175,000 USD

Greater than/equal to 175,000 USD and less than 200,000 USD

Greater than/equal to 200,000 USD and less than 225,000 USD

Greater than/equal to 225,000 USD and less than 250,000 USD

Greater than/equal to 250,000 USD and less than 275,000 USD

Greater than/equal to 275,000 USD and less than 300,000 USD

Greater than/equal to 300,000 USD

Prefer not to say

In general, how willing are you to take risks?

0 = Completely unwilling to take risks 10 = Very willing to take risks

0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐

Have you ever applied for a job or other opportunity where you believe your application was screened or reviewed by an AI system or algorithm?

(For example: resume filtering by applicant tracking systems used by companies and recruitment agencies; or credit checks by systems like FICO Score.)

Yes, I have submitted an application where I believe my application was screened by an AI system or algorithm

No, I have submitted applications but do not believe an AI system or an algorithm was ever used to screen them

I have never applied for a job or an opportunity that would involve screening

Not sure

Prefer not to say

If so, in which contexts have you interacted with the AI system or algorithm?

Select all that apply:

Job applications or hiring platforms

Credit scoring or loan applications

Other (please specify in the textbox below)

In general, how much do you trust major technology companies to make decisions that affect you in a fair and unbiased way?

0 = No trust at all 10 = Complete trust

0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐

In the experiment, what factor(s) below motivated your investment choice? Select all that apply.

My group identity (whether I was assigned to the **Green** or **Blue** group)

Whether I have to pay for the investment cost when I'm not hired

How much previous participants invested

Whether previous participants were hired

The starting amount I had available to invest (20 ECU)

The potential bonus I could get if hired (100 ECU)

Other (please specify in the textbox below)

Please let us know if you experienced any issue in the study, for example if anything was unclear or didn't work (optional).

End of Survey Message

Thank You for Participating!

This study investigates how people make decisions about investing in their skills when evaluated by a hiring algorithm. The hiring process in this experiment was simulated using fictional groups (**Green** and **Blue**) to study how perceived bias might affect decision-making. These group assignments were entirely hypothetical and not tied to any real-world identities.

Your bonus payment was calculated using Experimental Currency Units (ECU). The conversion rate is: **100 ECU = \$1 USD**. The final bonus amount (if any) will be transferred to your Prolific account in the next few weeks.

If you have any questions about the study or would like to learn more about the research, feel free to contact the researcher at: yung-shiang.yang@ucdconnect.ie

Please click the button below to be redirected back to Prolific and register your submission.

Alternatively, you can enter the following completion code to register your submission: **CS8K8PPI**