Identity as Information:

Perceived Risk-Preference Similarity and Informational Influence in a Laboratory Experiment on Decisions under Risk

Soong Kit Wong

University of Wisconsin–Madison

August 2026

Abstract

Many consequential choices must be made before people can fully evaluate how each option will perform for them. In these settings, people may use other people's choices in place of direct evidence, but only when those people are seen as relevant counterfactuals. This paper tests whether shared identity affects whose choices people use when direct payoff information is unavailable. In an incentivized experiment, 226 participants first choose between lotteries on their own and later make choices after seeing decisions recorded earlier by other participants, whom I call referents. Across repeated blocks, I vary whether the referent shares the participant's gender and whether participants can see the lottery probabilities and possible payoffs. When lottery details are hidden, participants are more likely to copy the referent. This increase is larger for same-gender referents. The gender pattern is consistent with a perceived-similarity mechanism: random assignment to a same-gender referent increases perceived similarity in risk preference before any referent choice is revealed, and perceived similarity predicts exact copying when lottery details are hidden. The results show that identity shapes whose choices are treated as informative.

1 Introduction

Many consequential choices involve uncertainty about how the alternatives would work for the decision-maker. A person choosing a school, neighborhood, migration destination, or occupation may need information from others to help evaluate the options. Choices involving experience goods provide one example: important aspects of quality are learned through use (Nelson, 1970). For these goods, learning which alternative fits best through one's own experience may require trying numerous, if not all, alternatives. Others' experiences can help with evaluation before committing to a choice. The broader problem is not limited to experience goods. Trying alternatives oneself can be costly, slow, or difficult to reverse when it involves relocation, training, or substantial time. In other cases, the alternatives are mutually exclusive: a person who has a child, for example, can never observe how their life would have unfolded had they not done so.

Social learning can help people evaluate alternatives when information about their likely outcomes is incomplete. A decision-maker can observe what someone else chose and use that action as evidence about the suitability of available options (Foster and Rosenzweig, 1995; Manski, 2000). Yet the same action need not be equally informative to every observer. If the best option depends on risk preference, constraints, or other individual characteristics, the observer wants a *personal counterfactual*: a person whose informed choice is relevant to the observer's own decision. The person being observed serves as the *referent*. The informational problem is not only whether to learn from others, but whose choices to treat as useful evidence.

Consider women deciding whether to study economics. Exposure to women who have chosen the field can change students' decisions: Porter and Serra (2020) find that visits from successful female economics alumni increased women's probability of majoring in economics by 8 percentage points, against a 9 percent baseline. Male students attended the same visits, but the authors find no statistically significant effects on their economics enrollment or major choice. This contrast suggests that the identity of the person sharing an experience may

matter for how it is received.[1]

Such an intervention can operate through several channels. A student might learn about opportunities, revise her beliefs about how well economics fits her preferences, feel a stronger sense of belonging, or be inspired by the speaker. An increase in enrollment alone cannot distinguish between these explanations. Separating informational influence is therefore central to understanding whether exposure changes students' evaluation of economics, their desire to identify with the field, or both. The distinction guides policy design: if the effect operates through learning, both the information conveyed and the perceived relevance of its source may matter; if it operates through belonging or inspiration, supplying the same facts need not reproduce it. Higher observed demand for economics among women may reflect latent demand: students learn that the major fits preferences and needs they already had. Alternatively, role models may create new demand by changing what students value.

Throughout the paper, I use *referent influence* as the umbrella term for behavioral responses to the observed referent. Referent influence can arise through two channels. *Normative influence* includes affinity, trust, conformity, attention, and utility from matching another person. An *informational channel* arises when the referent's informed choice helps the observer infer which option fits their own preferences. Identity can affect both. The mechanism studied here is that identity acts as a noisy proxy for preference similarity: an identity cue changes the observer's belief about the referent's risk preference, which changes the weight placed on the referent's choice. Identity need not reveal preferences accurately to affect choice; it need only change whether the referent appears to be a useful counterfactual.

This paper asks: when direct payoff information is missing, does social identity affect whose observed choices influence decisions under risk? The experiment has two parts. In Part 1, subjects make lottery choices alone, providing an untreated behavioral baseline. In Part 2, each subject completes eight blocks with prerecorded referents. The subject sees the referent's identity, reports perceived similarity in risk preference, observes the referent's stored choice,

[1]See their Section III.D and Appendix Tables A4–A5. Without randomized speaker gender, the contrast cannot separate gender matching from different responses to the same information.

and then chooses. Payoff-visible blocks preserve the subject's ability to evaluate the lotteries directly. Payoff-blind blocks hide the lottery probabilities and outcomes, increasing the value of information in the referent's informed choice. Referent gender and nationality vary across blocks. Nationality refers to the country where a participant reported growing up.

The primary outcome is *exact copying*: whether the subject and referent choose the same option. Hiding lottery information increases exact copying. Same-gender matching matters primarily when payoffs are hidden (a double difference): its effect on exact copying is 7.0 percentage points larger than when payoffs are visible. As robustness checks, I use two measures of changes from subjects' initial choices: *switching toward the referent* among trials with initial disagreement, and *net shift*, which subtracts initial agreement from final agreement and accounts for switches both toward and away from the referent. Both reproduce the qualitative pattern, showing that it is not limited to final agreement that was already present before exposure.

The belief evidence provides a mechanism-consistent chain. Randomized same-gender assignment raises pre-choice perceived similarity between the subject's and the referent's risk preferences. Perceived similarity, in turn, predicts exact copying much more strongly when payoff information is hidden. The pattern varies with the subject's own revealed risk preference in a way consistent with beliefs about gender and risk preferences: the same-gender minus different-gender difference in perceived risk-preference similarity grows with risk aversion among women and falls with risk aversion among men. Nationality provides a contrast: shared nationality does not increase referent influence, but the referent's country shapes beliefs about whether the referent is more or less risk-taking than the subject (Section 6.5).

The paper contributes to two literatures. First, it adds to work separating social learning from other sources of correlated behavior (Manski, 2000; Bursztyn et al., 2014; Bayer et al., 2022). Prerecorded referent choices and randomized information conditions sidestep simultaneity and common-shock problems that arise when peers choose together. Second, it adds an informational mechanism to identity economics. Existing work shows that identity

can enter preferences and valuations (Akerlof and Kranton, 2000; Hackel et al., 2018); this paper shows that identity can also change whose actions are treated as evidence.

The paper proceeds as follows. Section 2 positions the experiment in the social learning and identity literatures. Section 3 presents the conceptual framework. Section 4 explains the design and measurement. Section 5 describes the data and estimands. Section 6 reports results, and Sections 7 and 8 discuss interpretation and policy implications.

2 Related Literature

2.1 Separating the Channels of Social Influence

Similar behavior can reflect common circumstances, sorting, shared preferences, or responses to other people (Manski, 2000). Even when exposure to a peer is randomized, identifying an effect does not by itself reveal why it occurs. A decision can respond to information conveyed by someone else's behavior or to normative influence: the value of matching that behavior.

Bursztyn et al. (2014) address this distinction by separating a peer's desire to purchase a financial asset from possession of the asset. Their randomization identifies learning from the purchase decision separately from possession-related social utility. The two experiments use different treatments but share a common approach: creating an explicit counterfactual for each component of social influence, rather than inferring the channel from correlated choices or a respondent's description of their motivation.

This paper applies the same logic to identify how shared identity changes informational influence; Section 4.5 explains the experimental comparison.

2.2 Identity and Information Use

Identity can enter utility through group prescriptions and identity-congruent action (Akerlof and Kranton, 2000). It can also affect responses to information. Berger et al. (2018) show that shared identity promotes herding in an information-cascade game. In a computer-mediated

trivia task, Lee (2005) varies substantive versus placebic information and the represented partner's sex; responses to informational content depend on source sex. Wabitsch (2025) provides experimental evidence that messenger nationality affects the use of information in inflation forecasts. These studies establish that source identity can matter for learning and information processing; the connection itself is not new.

The object of learning here differs from an objectively correct trivia answer or a common inflation realization. A lottery choice reflects both information about the alternatives and the chooser's preferences. The same informed choice may therefore convey different implications to observers with different preferences. Taylor-Davies et al. (2025) directly study this problem: in virtual decision environments, group cues guide selective learning when they reveal preference alignment. Their artificial group structure, simulated agents, and assigned utility correspondence make this a close conceptual predecessor. The present design instead uses natural identity match, participants' own preferences, and actual earlier choices, and varies whether direct lottery information is available to the observer.

The contribution is to separate informational influence from normative influence and identify how the isolated informational component varies with shared identity. This distinguishes identity's role in learning about personal fit from normative influence.

2.3 Preference Similarity and the Relevance of Others' Experience

The relevance of another person's experience depends on how its determinants transfer to the observer. Foster and Rosenzweig (1995) provide foundational evidence of learning from neighbors in agriculture. Munshi (2004) shows why heterogeneous conditions can limit that learning: outcomes on another farm are less informative when important determinants of performance differ and cannot be adequately accounted for. This distinction between access to an experience and its applicability has a counterpart in preference-dependent decisions.

Yaniv et al. (2011) study advice in matters of taste, where similarity between the adviser and recipient can determine how useful the advice is. Lenoir (2015) examines heritage

congruence in recommendations across product settings that differ in preference sensitivity. This work motivates the idea that identity may provide a clue about personal fit. This paper examines whether shared gender makes a referent's risk preferences seem more similar to the subject's own. It then asks whether perceived similarity predicts copying more strongly when direct information about the lotteries is missing.

3 Conceptual Framework

3.1 Choice with Direct and Social Information

In the framework, *direct alignment value* denotes normative influence: utility from matching a referent's choice apart from the information it conveys.

In block b and lottery pair t, subject i chooses option $a \in \{A, B\}$ after seeing that referent j chose k_{jt}. Let $B_{ib} = 1$ when the lottery details are hidden, $G_{ij} = 1$ when the subject and referent share gender, and S_{ij} denote the subject's perceived similarity to the referent in risk preference. A compact representation of the subject's utility from option a is

$$U_{ijbt}(a) = \underbrace{(1 - B_{ib})Q_{it}(a; \theta_i) + \mathbf{1}\{a = k_{jt}\}B_{ib}\omega S_{ij}}_{\substack{\text{choosing the option that fits} \\ \text{subject } i\text{'s risk preference}}} + \underbrace{\mathbf{1}\{a = k_{jt}\}\nu G_{ij}}_{\text{direct alignment value}} + \varepsilon_{ijbta}. \tag{1}$$

Here θ_i denotes the subject's risk preference, and $Q_{it}(a; \theta_i)$ is the subject's direct evaluation of how well option a fits that preference from the displayed probabilities and outcomes. It is available when payoffs are visible and set to zero when those details are hidden. The indicator $\mathbf{1}\{a = k_{jt}\}$ equals one for the option chosen by the referent. The parameter ω weights perceived similarity in the informational term, while ν measures direct alignment value when the referent shares the subject's gender. The first underbraced component captures utility from choosing the option that fits the subject's risk preference. With payoffs visible, that fit is

assessed directly through $Q_{it}(a;\theta_i)$. With hidden payoffs, $B_{ib}=1$ activates the informational term $B_{ib}\omega S_{ij}$: the referent's informed choice points to an option that appears to fit the subject better when the referent seems similar in the preference dimension relevant to the choice. This is information used to satisfy the subject's risk preference, not a separate taste for copying. The second underbraced component, νG_{ij}, represents direct alignment value and operates under both payoff-visibility conditions. The final term, ε_{ijbta}, collects unobserved option-specific influences. Equation 1 organizes the mechanisms; it is not estimated as a structural utility model.

3.2 Identity as a Noisy Proxy

Let θ_i and θ_j denote the subject's and referent's risk preferences. The subject observes the referent's identity cues X_j, but not θ_j, and forms the expectation

$$\widehat{\theta}_j = \mathbb{E}[\theta_j \mid X_j], \tag{2}$$

where $\widehat{\theta}_j$ is the perceived risk preference associated with the referent's identity cues. Perceived similarity rises as the distance between θ_i and $\widehat{\theta}_j$ shrinks. One continuous representation is

$$S_{ij} = \exp\{-\kappa|\theta_i - \widehat{\theta}_j|\}, \qquad \kappa > 0. \tag{3}$$

The parameter κ controls how quickly similarity falls with perceived distance. The experiment elicits a coarse 0–2 version of S_{ij}. Identity need not predict the referent's true preference well; it can affect choice by changing the subject's expectation of preference fit.

3.3 Comparative Statics and Predictions

Figure 1 traces the two paths from gender match to exact copying. Let ΔQ_{it} be the directly evaluated value of the referent's option minus that of the alternative. This difference can be

negative. Let $Copy_{ibt}$ indicate choosing the referent's option. Equation 1 gives its systematic utility advantage:

$$V_{ijbt} = (1 - B_{ib})\Delta Q_{it} + B_{ib}\omega S_{ij} + \nu G_{ij}. \tag{4}$$

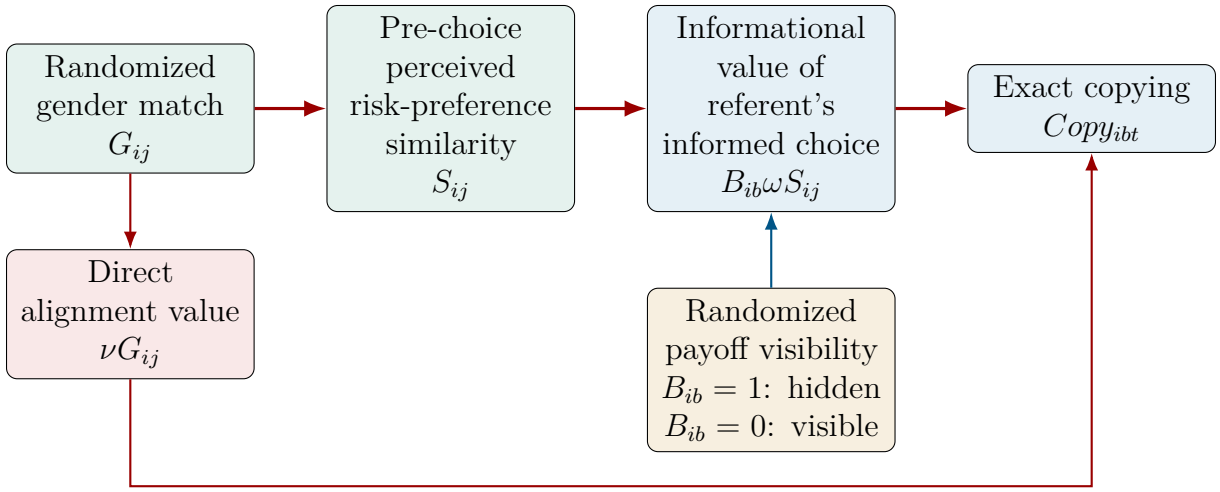


Figure 1: Identity-conditioned social learning. The upper path operates through perceived risk-preference similarity, S_{ij}, with informational value $B_{ib}\omega S_{ij}$. The lower path contributes direct alignment value, νG_{ij}, under either visibility condition. Exact copying is the primary outcome; switching toward the referent and net shift are robustness checks.

3.4 Predictions

Gender match. Let $\Delta_G S$ be the change in perceived similarity from same-gender assignment. Holding direct evaluation fixed, gender match changes the utility advantage by ν with visible details and by $\nu + \omega\Delta_G S$ with hidden details. The upper path therefore adds $\omega\Delta_G S$ to the gender effect when information is hidden. This increment is positive if shared gender increases perceived similarity; the lower path, νG_{ij}, operates in both conditions. This motivates testing whether the same-gender copying gap is larger when details are hidden.

Payoff visibility. Moving from $B_{ib} = 0$ to $B_{ib} = 1$ changes the utility advantage by $\omega S_{ij} - \Delta Q_{it}$. Hiding details makes copying more attractive when the informational value of the referent's informed choice exceeds the advantage from direct evaluation, but not for every subject and lottery. Appendix A relates these utility comparisons to copying probabilities.

Perceived similarity. Along the upper path, $\partial V_{ijbt}/\partial S_{ij} = B_{ib}\omega$. For $\omega > 0$, greater similarity makes copying more attractive when details are hidden; its effect is zero in the model's full-information benchmark. Under a monotone choice rule, this predicts a more positive similarity-copying relationship with hidden than visible details.

Identity and perceived similarity. Equation 3 implies that an identity cue raises S_{ij} when it brings perceived referent risk preference, $\widehat{\theta}_j$, closer to the subject's own, θ_i. If subjects associate women with greater risk aversion and men with greater risk tolerance, the same-gender similarity gap should be larger among more risk-averse women and more risk-tolerant men. A nationality cue likewise raises similarity when it brings the referent's perceived preference closer to the subject's own, not simply because nationality is shared.

4 Experimental Design

4.1 Participant Timeline

The experiment has two choice phases. Part 1 contains no referent and establishes each subject's baseline lottery choices and Part 1 risk-taking index (see Appendix L.4 for construction). Those choices also supply prerecorded decisions to the referent pool. Part 2 contains eight referent blocks. Figure 2 gives the complete participant sequence.

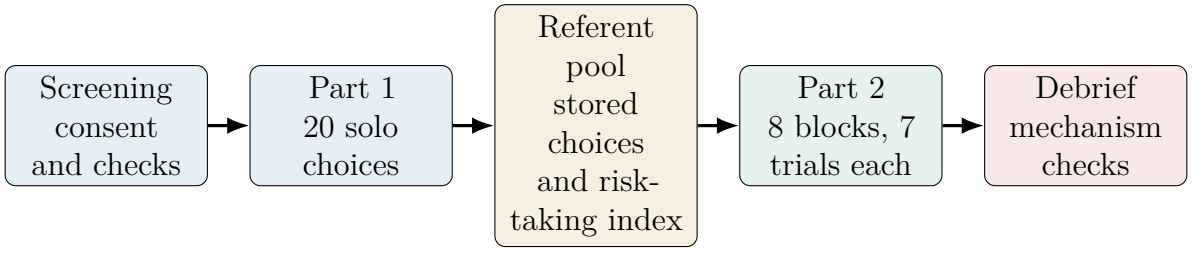


Figure 2: Participant timeline. Part 1 is completed without social information. Part 2 randomizes identity and information conditions across repeated blocks.

4.2 Part 1: Baseline Choice

Subjects first complete two practice choices and 20 incentivized pairwise lottery choices. Each lottery has three monetary outcomes and associated probabilities. The main pairs present risk-return tradeoffs without a dominated option. Figure 3 shows the Part 1 decision screen.

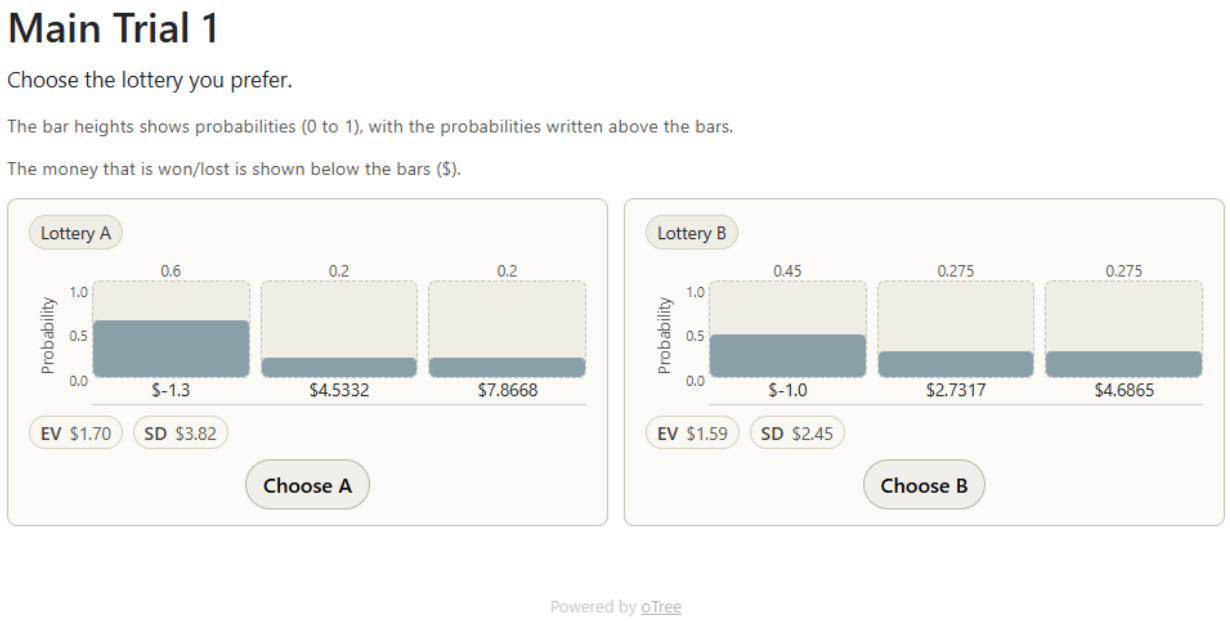


Figure 3: Part 1 baseline-choice screen. Subjects choose between two fully described lotteries before observing any referent.

Part 1 serves two purposes. First, it measures the subject's choice before any social exposure. Second, it produces the Part 1 risk-taking index and authentic stored choices from which later referents can be selected. Because the same lottery pairs reappear in Part 2, the baseline distinguishes pre-existing agreement from movement after observing a referent.

4.3 Part 2: Referents, Beliefs, and Choice

At the start of each Part 2 block, the subject sees one prerecorded referent's profile. The displayed profile includes gender, nationality (country of upbringing), class year, and a neutral shape label. Gender and nationality are the identity dimensions used in the analysis; class year and shape are not treatments. The referent is another participant who previously faced

the same lotteries alone and with full information. The interaction is asynchronous: subjects observe choices recorded in Part 1. No realized referent payoff is displayed.

Before any referent choice is shown, the subject answers: "How does this referent compare to you in their riskiness?" Responses range from "much more risky than me" to "much less risky than me." The pre-choice perceived risk-preference similarity score is

$$Similarity_{ib} = 2 - |Belief_{ib} - 3|, \tag{5}$$

so exact perceived alignment receives 2, adjacent responses receive 1, and distant responses receive 0. The measure concerns similarity in risk preference, not general social closeness. The block's identity condition is already fixed, but the referent's choices have not yet been shown and the subject has not made any block choice. Figure 4 shows the elicitation and the referent profile displayed alongside it.

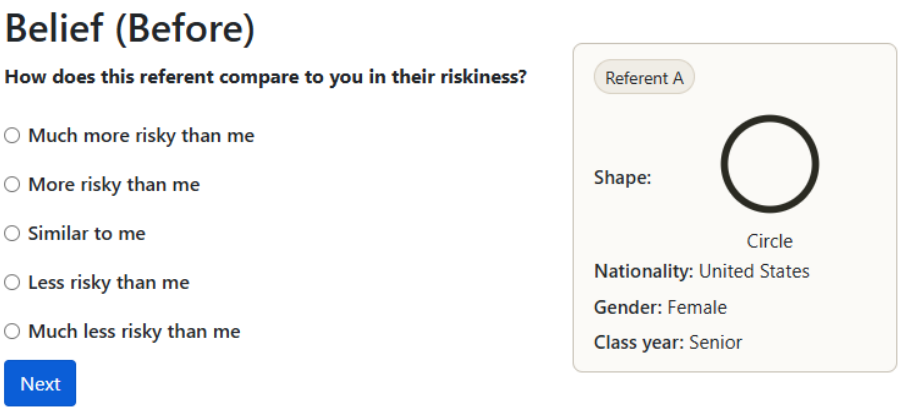


Figure 4: Pre-choice similarity elicitation. The response scale appears at left and the displayed referent profile at right.

On each of seven trials, the subject observes the referent's stored choice and then chooses between the same two lottery labels. In *payoff-visible* blocks, the probabilities and monetary outcomes remain available. Within visible blocks, a separate assignment determines whether expected value and standard deviation are displayed in addition to probabilities and outcomes. This overlay is controlled in robustness analyses but is not the main information treatment.

Figure 5 shows a payoff-visible trial with the overlay displayed.

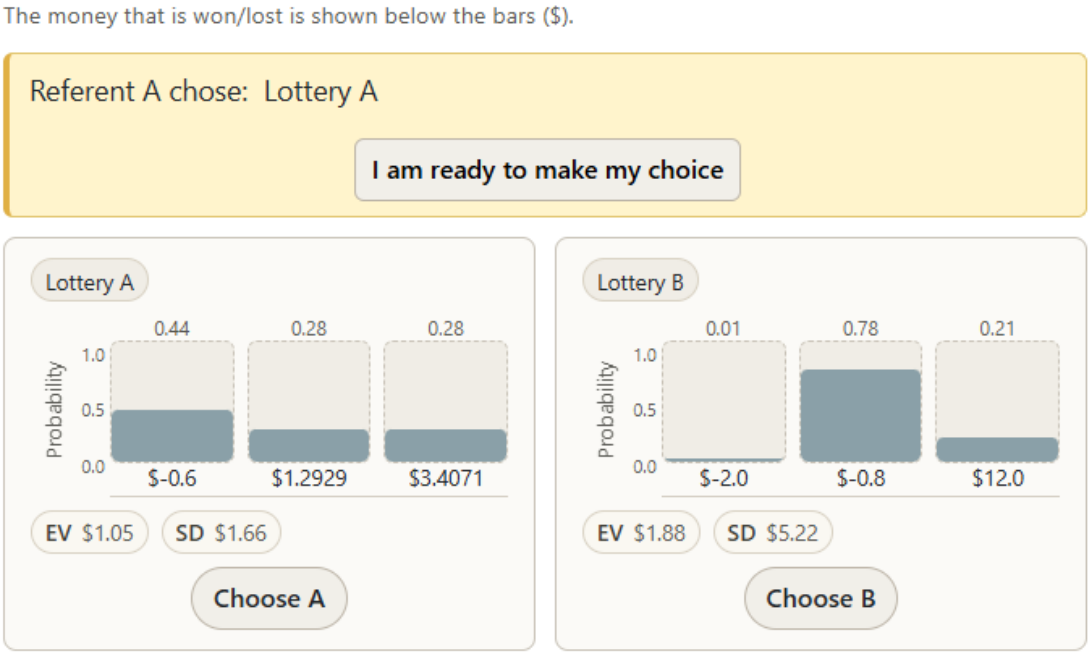


Figure 5: Payoff-visible Part 2 choice area. The referent's stored choice appears above the lotteries, whose probabilities and monetary outcomes remain visible. This example also displays the expected-value and standard-deviation overlay.

In *payoff-blind* blocks, the lottery details are removed; subjects still see the referent's identity and informed choice. Figure 6 shows the corresponding payoff-blind screen.

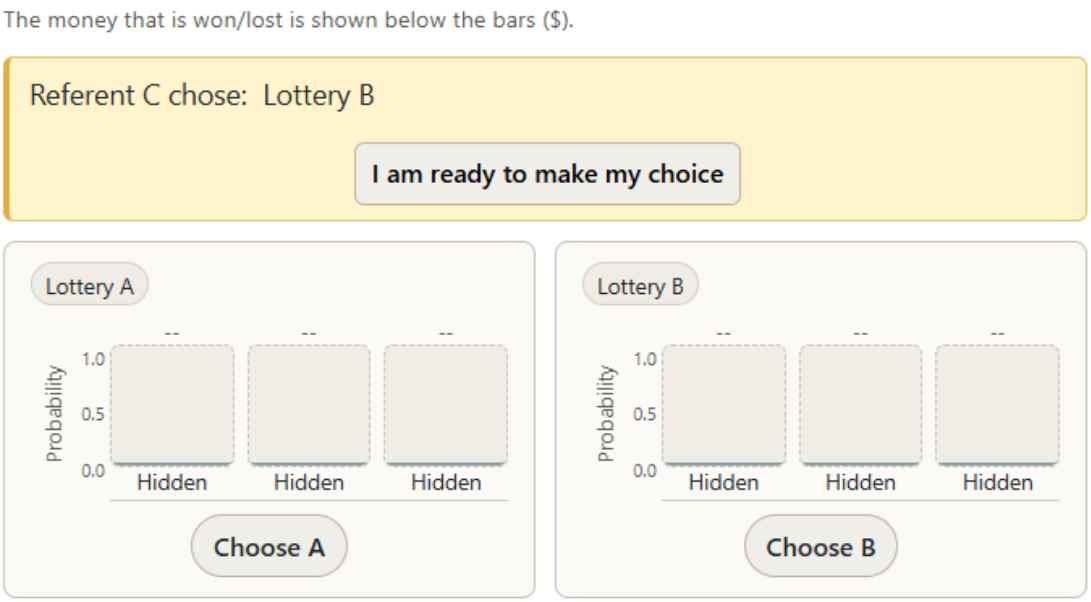


Figure 6: Payoff-blind Part 2 choice area. The referent's informed choice remains visible while the lottery probabilities and monetary outcomes are hidden.

Part 2 also contains two dominance checks, one in a randomly selected payoff-visible block in each half of the task. The referent's choice is hidden on these trials. The checks assess comprehension and attention and are excluded from all social-learning outcomes.

Figure 7 separates events at the block boundary from events repeated on each trial. At the block boundary, the experiment first assigns the nationality-match $\times$ payoff-visibility condition and then selects a referent, maintaining gender balance within the condition and randomizing the targeted risk tail. The subject sees the selected referent's profile and reports similarity once before completing seven substantive trials.

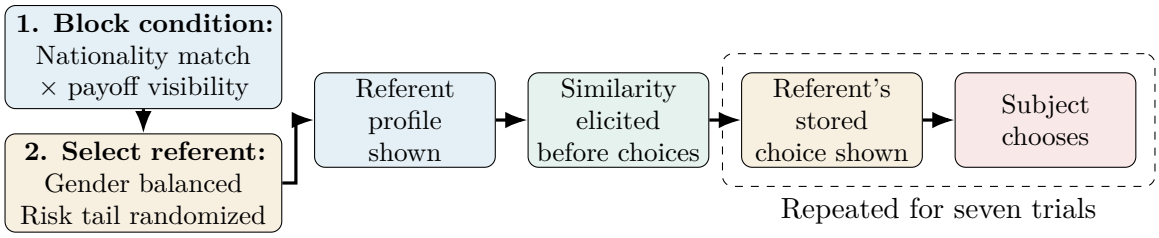


Figure 7: Timing within a referent block. Similarity is elicited before referent-choice exposure and before the subject's block choices.

4.4 Identity Assignment and Referent Selection

The nationality treatment varies whether the referent shares the subject's nationality. Crossing nationality match with payoff visibility produces four within-subject conditions. A Williams sequence counterbalances the four conditions across eight blocks, so each condition appears twice. Referent gender is randomized and balanced within each nationality-by-visibility condition: the first gender is randomized, and the opposite gender is assigned on that condition's second occurrence. Thus, each subject experiences one same-gender and one different-gender referent within every nationality-by-visibility cell.

Referents are selected after the block conditions are assigned. The algorithm restricts the pool to the assigned nationality and excludes the subject's own record. It then randomly targets the bottom 30 percent or top 30 percent of the distribution of the Part 1 risk-taking index among eligible referents. Selecting from the tails maximizes behavioral contrast and statistical power: with only seven trials per block, mean-risk or unrestricted referents would leave less scope to detect movement toward or away from a referent. Within the targeted tail, the algorithm applies the gender-balancing rule, avoids reusing a referent when possible, and samples among the least-used eligible records.

Contingency rules covered unavailable nationality, risk-tail, or support pools, but none was needed in the completed analysis sample. All 1,808 blocks used an eligible top- or bottom-tail referent, and all 904 participant-by-condition cells achieved one male and one female referent. Appendix F records the selection steps in implementation order.

4.5 Identification

Randomized identity assignment identifies whether identity changes pre-choice perceived similarity and supplies exogenous variation in those beliefs. Randomized payoff visibility identifies whether referent influence changes when direct lottery information is removed.

To identify how shared gender changes informational influence, I compare the same-gender effect on exact copying when lottery information is hidden with the corresponding effect when it is visible. This double difference uses the visible-information gender effect as the benchmark for the copying response due to direct alignment value. Its informational interpretation requires that hiding lottery information leave that response unchanged.

The similarity mechanism is supported when these patterns line up: identity changes similarity before choice, and similarity predicts later referent influence most strongly when direct information is unavailable.

Subjects with the same gender- and nationality-match assignments can still report different levels of risk-preference similarity to their referents. These differences in reported similarity are measured rather than randomized, so their association with copying may partly reflect trust, attention, affinity, or other unobserved factors. The estimated similarity slope is therefore consistent with a similarity mechanism, but does not by itself identify the causal effect of perceived similarity. A reduction in the estimated same-gender effect after controlling for perceived similarity therefore cannot be read as the amount of that effect caused by similarity. Likewise, the remaining estimate does not isolate the influence of other channels.

5 Data and Empirical Strategy

5.1 Recruitment and Analysis Sample

The recruitment target was 300 subjects completing both parts, ideally split evenly between participants who grew up in the United States and China. Budget and semester timing

constrained recruitment. Subjects were recruited from the UW–Madison BRITE Lab pool and the broader undergraduate population through a university research-email service. The analysis is restricted to participants who reported growing up in the United States or China.

Of 267 subjects who completed Part 1, 232 started Part 2 and 226 completed all eight treatment blocks. The completed sample contains 207 United States participants and 19 China participants. Primary analyses retain completed Part 2 participants, valid referent assignments, and non-dominance trials. This yields 12,656 trial observations for the primary exact-copying outcome and the net-shift robustness check. The conditional switch-toward-referent robustness sample contains 5,758 trials on which the subject's Part 1 choice differed from the referent's stored choice. Because each completed participant experiences all four nationality-match × payoff-visibility conditions, the same participants appear in every treatment cell. Participant characteristics are therefore mechanically balanced across conditions, making conventional between-group covariate balance tests uninformative. The relevant checks are differential completion, treatment-sequence assignment, and the characteristics of the referents realized in each condition. Table 2 reports completion balance, and Appendix Table 16 reports assignment and referent checks.

Table 1: Recruitment, Completion, and Analysis Samples

Sample or stage	Count
Target sample: completed both parts	300
Completed Part 1	267
Started Part 2	232
Completed Part 2	226
United States	207
China	19
Exact-copying and net-shift observations	12,656
Conditional switch-toward robustness observations	5,758

Notes: The trial samples exclude dominance checks and require completion of all eight Part 2 blocks and a valid referent condition. Nationality is based on the country in which the participant reported growing up.

Table 2: Part 1 Characteristics by Part 2 Completion

	Did not complete Part 2	Completed Part 2	Standardized difference	p-value
Grew up in China	0.171	0.084	−0.293	0.162
Female	0.732	0.708	−0.052	0.754
Numeracy concepts correct, first attempt	0.549	0.558	0.031	0.854
General recruitment pool	0.927	0.912	−0.054	0.733
Part 1 comprehension mistakes	7.244	5.186	−0.309	0.136
Part 1 trials completed within 2 seconds	0.027	0.028	0.015	0.914
Part 1 median response time, seconds	10.441	10.261	−0.032	0.875
Part 1 risk-taking index	0.511	0.456	−0.252	0.130
Participants	41	226		

Notes: The first column combines 35 Part 1 completers who never started Part 2 and six who started but did not complete all eight blocks. The difference is completed minus not completed. The final column reports two-sided tests of equal means; every Holm-adjusted p-value across the eight rows equals 1.000. A joint linear-probability model of Part 2 completion using the core covariates gives $F = 1.22$ and $p = 0.296$. Failure to reject does not establish equality on unobserved characteristics.

5.2 Outcomes

For subject i, block position b, lottery pair t, and referent j, let C^0_{it} be the subject's Part 1 baseline choice, R_{jt} the referent's stored choice, and C_{ibt} the subject's Part 2 choice. The primary outcome is exact copying:

$$Copy_{ibt} = \mathbf{1}\{C_{ibt} = R_{jt}\}. \tag{6}$$

It records final agreement with the referent on every primary-sample trial, regardless of whether the subject's baseline choice already matched.

I use two baseline-referenced outcomes as robustness checks:

$$ShiftToward_{ibt} = \mathbf{1}\{C_{ibt} = R_{jt} \mid C^0_{it} \neq R_{jt}\}, \tag{7}$$

$$NetShift_{ibt} = \mathbf{1}\{C_{ibt} = R_{jt}\} - \mathbf{1}\{C^0_{it} = R_{jt}\}. \tag{8}$$

Switching toward the referent measures exact copying conditional on initial disagreement, so it records an active change in choice but excludes trials with baseline agreement. Net shift uses the full sample: $+1$ denotes movement toward the referent, -1 movement away, and 0 no change in alignment. It is a change in agreement, not a copying probability. These checks assess whether the primary results also appear in changes from baseline choices; Appendix B provides further details.

5.3 Treatment Effects and Mechanism Tests

For the primary outcome, $Y_{ibt} = Copy_{ibt}$, I estimate both identity dimensions jointly:

$$\begin{aligned} Y_{ibt} = {} & \beta_G \text{SameGender}_{ib} + \beta_{GB}(\text{SameGender}_{ib} \times \text{Blind}_{ib}) \\ & + \beta_B \text{Blind}_{ib} + \beta_N \text{SameNat}_{ib} + \beta_{NB}(\text{SameNat}_{ib} \times \text{Blind}_{ib}) \\ & + \alpha_i + \delta_b + \lambda_t + \varepsilon_{ibt}. \end{aligned} \tag{9}$$

where α_i are participant fixed effects, δ_b block-position fixed effects, and λ_t lottery-pair fixed effects. Standard errors are clustered by participant. The coefficient β_{GB} measures whether the effect of same-gender assignment changes when payoffs are hidden, while β_{NB} measures the corresponding change in the same-nationality effect. This joint specification reflects the crossed design and estimates both comparisons using the same observations. I apply the same specification to switching toward the referent and net shift as robustness checks, using the initial-disagreement sample for the former and the full sample for the latter.

The block-level perceived-similarity model replaces Y_{ibt} with the perceived similarity score and omits lottery-pair fixed effects. The mechanism model adds $Similarity_{ib}$ and $Similarity_{ib} \times Blind_{ib}$ to Equation 9. The latter interaction tests whether the pre-choice similarity measure is more predictive of exact copying when direct information is restricted. I repeat the choice model with the two robustness outcomes.

The gender-risk model interacts same-gender assignment, participant gender, and standardized Part 1 risk aversion, including all lower-order terms. Risk aversion is the negative standardized Part 1 risk-taking index, so higher values indicate greater risk aversion. This model tests whether the effect of gender match on perceived similarity depends on the participant's own revealed risk preference.

6 Results

6.1 Payoff Blindness and Same-Gender Referent Influence

The first result establishes that subjects use the referent's choice when direct option information is restricted. In the different-gender comparison cell, payoff blindness increases exact copying, the primary outcome, by 7.8 percentage points.

Same-gender assignment matters primarily under payoff blindness. The same-gender-by-hidden interaction is 7.0 percentage points for exact copying (Table 3). Same-gender level effects under visible payoffs are small and statistically insignificant. This pattern is consistent with an informational role for identity, although direct alignment value may also increase when payoffs are hidden.

The two robustness outcomes reproduce this pattern (Table 3). In the different-gender comparison cell, payoff blindness increases switching toward the referent by 26.6 percentage points and net shift by 5.3 points. Their same-gender-by-hidden interactions are 7.6 and 9.7 percentage points, respectively. These changes from baseline show that the copying results are not limited to pre-existing agreement with the referent.

Table 3: Payoff Visibility, Gender Match, and Referent Influence

	Primary outcome	Robustness checks	
	Exact copy	Switch toward referent	Net shift
Same gender × payoff blind	0.070***	0.076***	0.097***
	(0.021)	(0.027)	(0.028)
Same gender	0.024	0.016	0.003
	(0.015)	(0.019)	(0.016)
Payoff blind	0.078***	0.266***	0.053**
	(0.018)	(0.025)	(0.021)
Participant, block-position, pair FE	Yes	Yes	Yes
Observations	12,656	5,758	12,656
Participants	226	226	226
Outcome mean	0.667	0.571	0.122

Notes: Linear models with participant-clustered standard errors. The payoff-blind estimates in the final row are hidden-minus-visible differences for different-gender referents, averaged over nationality assignment. Models include the nationality terms described in Equation 9. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

6.2 Same-Gender Assignment Changes Perceived Similarity

Randomized same-gender assignment increases the pre-choice perceived risk-preference similarity score by 0.195 points on the 0–2 scale with visible payoffs (SE = 0.047, $p < 0.001$). With hidden payoffs, the effect is 0.215 points (SE = 0.040, $p < 0.001$). The same-gender-by-blind interaction is only 0.019 (SE = 0.050). The perception difference is therefore present before payoff visibility becomes relevant to the subject's trial-level choice. Same nationality, by contrast, does not increase perceived alignment.

6.3 Similarity and Referent Influence Under Payoff Blindness

Perceived risk-preference similarity has little relationship with exact copying when lottery details are visible: the coefficient is 0.011 (SE = 0.016). Under payoff blindness, the slope is approximately 0.156, the sum of the visible coefficient and a 0.145 interaction. The raw pattern is visible in Figure 8: copying rises from 48 percent to 84 percent across the 0–2 similarity scale when payoffs are hidden, but remains nearly flat under visible payoffs.

The same pattern appears in the two baseline-referenced robustness checks. The similarity-by-blind interaction equals 0.096 for switching toward the referent and 0.096 for net shift (Table 4). The net-shift result is useful because it counts switches away from the referent as decreases in alignment rather than grouping them with all other non-copying choices.

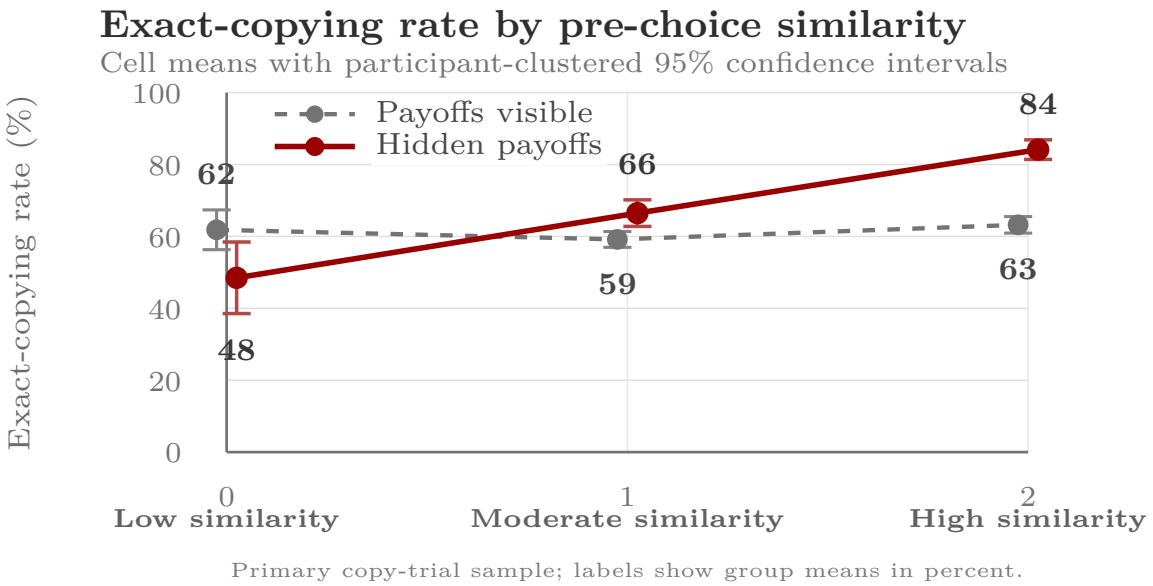


Figure 8: Exact copying by pre-choice perceived risk-preference similarity and payoff visibility. Points show group means with 95 percent confidence intervals clustered by participant and are offset horizontally for readability.

Table 4: Gender Match, Perceived Similarity, and Referent Influence

	Beliefs	Primary	Robustness checks	
	Similarity score	Exact copy	Switch toward referent	Net shift
Panel A: Identity changes perceived similarity				
Same gender	0.195***			
	(0.047)			
Same gender × payoff blind	0.019			
	(0.050)			
Panel B: Similarity and residual identity terms in choice models				
Similarity score, visible payoffs		0.011	0.036*	0.042***
		(0.016)	(0.020)	(0.016)
Additional similarity slope, hidden payoffs		0.145***	0.096***	0.096***
		(0.023)	(0.032)	(0.026)
Same-gender effect, visible payoffs		0.021	0.007	−0.005
		(0.016)	(0.019)	(0.015)
Additional same-gender effect, hidden payoffs		0.040**	0.058**	0.076***
		(0.020)	(0.027)	(0.027)
Participant and block-position FE	Yes	Yes	Yes	Yes
Lottery-pair FE	No	Yes	Yes	Yes
Observations	1,808	12,656	5,758	12,656
Participants	226	226	226	226

Notes: Linear models with participant-clustered standard errors. Panel A is estimated at the participant-block level. Panel B is estimated at the trial level and includes the listed main effects and interactions, as well as nationality-by-information terms. Similarity is elicited before referent choices are shown. "Additional" denotes the hidden-minus-visible difference. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Including measured similarity does not eliminate the same-gender-by-hidden interaction. The remaining estimate is 4.0 percentage points for exact copying. The corresponding estimates in the robustness checks are 5.8 points for switching toward the referent and 7.6 points for net shift. These are residual same-gender associations: the portions not statistically accounted for by the measured similarity terms.

6.4 Gender, Risk Preference, and Perceived Similarity

The estimates are consistent with gendered beliefs about risk preference. Among men, a one-standard-deviation increase in risk aversion reduces the same-gender minus different-gender similarity difference by 0.126 points (SE = 0.043, $p = 0.004$). Equivalently, the same-gender differential is larger among more risk-tolerant men. Among women, a one-standard-deviation increase in risk aversion increases the same-gender differential by 0.166 points (SE = 0.041, $p < 0.001$). The female-minus-male difference in these slopes is 0.292 points ($p < 0.001$).

These estimates show why "same gender" is not a universal similarity rule. Gender is used as a clue about risk preference, and the relevance of that clue depends on the subject's own revealed risk preference. The estimates concern perceived similarity, where the pattern is precise. Risk aversion does not significantly moderate the remaining same-gender exact-copying coefficient after the similarity terms are included.

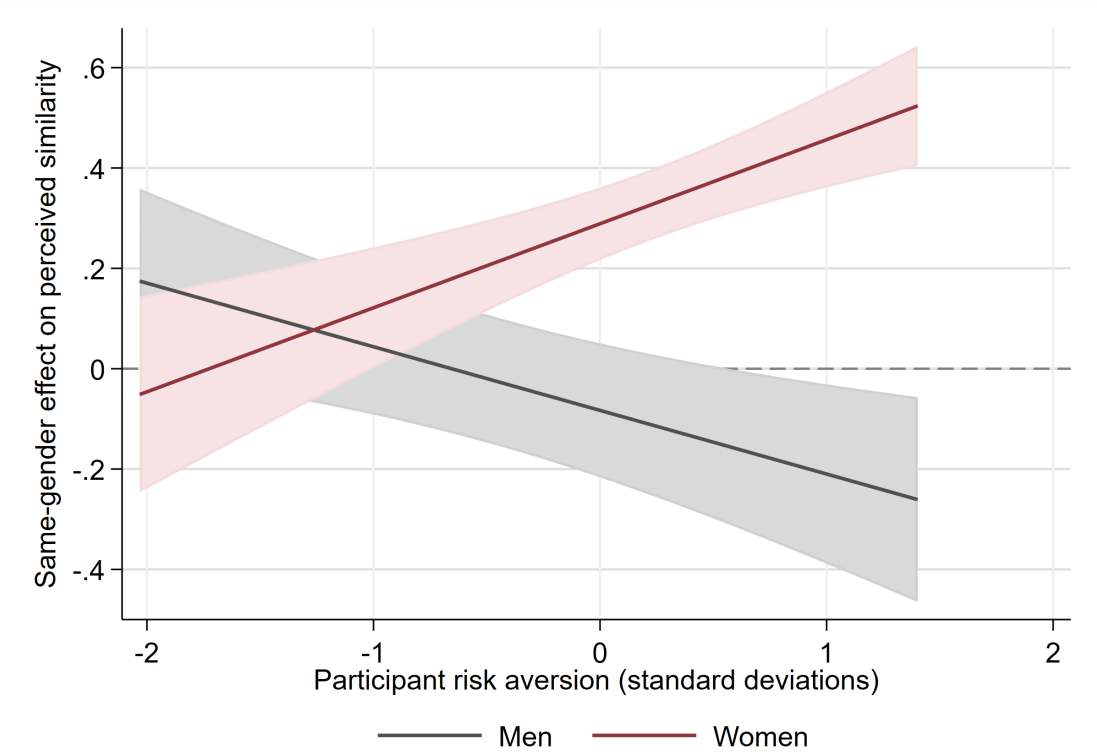


Figure 9: Estimated effect of same-gender assignment on perceived risk-preference similarity across participant risk aversion. Lines average over payoff-visibility conditions; shaded bands are 95 percent confidence intervals. The horizontal range covers the 5th through 95th percentiles of the participant risk-aversion distribution.

6.5 Nationality and Referent Country

In contrast to gender, nationality does not operate through same-nationality matching. The clearer pattern is attached to the referent's country: under hidden payoffs, exact copying is 75.3 percent for China referents and 69.5 percent for U.S. referents. The net-shift robustness check gives corresponding changes of 21.3 and 13.3 percentage points. These gaps nearly disappear when payoffs are visible. Both U.S. and China participants copy China referents more under hidden payoffs, so the result is not driven by one participant group copying its own nationality.

The referent pools do not differ detectably in actual risk-taking, but subjects perceive China referents as less risk-taking than U.S. referents. After controlling for the other displayed

profile traits, the China-minus-U.S. difference in this pre-choice belief is 0.278 points on the 1-to-5 scale, where higher values indicate that the referent is perceived as less risk-taking relative to the subject. Among more risk-averse subjects, a larger China-minus-U.S. "less risky" belief gap predicts greater China-minus-U.S. exact-copying gap under hidden payoffs: the risk-aversion-by-belief-gap interaction is 6.17 percentage points. For example, holding other covariates fixed, a one-point rather than zero-point belief gap predicts a 1.92-percentage-point larger China-minus-U.S. exact-copying gap for a participant at mean risk aversion and an 8.09-percentage-point larger gap for a participant one standard deviation above the mean. The 6.17-percentage-point difference between these predicted changes is the interaction. The corresponding interaction is small and statistically insignificant when payoffs are visible. This pattern is consistent with risk-preference matching, in which a referent perceived as less risk-taking is more informative to a more risk-averse subject. Appendix Tables 11 and 14 report the supporting country-specific and risk-belief estimates.

The China and U.S. referent pools differ in class-year composition and reuse frequency because the China pool is much smaller. Consequently, the estimated country difference is a referent-country and pool effect, not a pure nationality label effect. Appendix Table 16 reports these checks. Only 19 completed participants grew up in China, so comparisons between participant-nationality subgroups remain imprecise.

7 Discussion

The experiment shows that identity can change how people use a fixed social signal. When subjects can evaluate the lotteries directly, perceived similarity has little relationship with referent influence. When the probabilities and outcomes are hidden, subjects rely heavily on the referent, same-gender matching increases referent influence, and the pre-choice similarity score becomes strongly predictive. The baseline-referenced robustness checks show that these results include actual movement after exposure rather than only pre-existing agreement.

Taken together, these results support perceived preference similarity as one mechanism, but do not establish how much of the same-gender effect operates through it.

Several alternatives remain. Hidden payoffs may increase direct alignment value because subjects feel less able to justify an independent choice. Shared gender may increase attention, trust, or affinity independently of risk-preference beliefs. The similarity score is coarse and symmetric: believing that the referent is more risky and less risky are treated as equally dissimilar. It may also capture perceived averageness, stereotype use, midpoint responding, or differences in how subjects interpret the response categories. The nationality imbalance limits subgroup precision and makes the China-referent pattern difficult to generalize beyond this two-country student sample.

The design also has external-validity limits. Lottery choices isolate a preference-dependent information problem, but schooling, migration, housing, and occupational choices involve richer constraints, networks, and endogenous exposure. The experiment does not show that identity changes awareness, feasibility, or the considered set. Its external implication is more modest: when direct evaluation is difficult, identity can change which observed choices are treated as relevant evidence. In field settings, this mechanism can make information use, rather than tastes, look group-specific.

8 Policy Implications

The experiment identifies how shared identity changes informational influence, with implications for both interpreting observed demand and designing information interventions. A choice reflects the decision-maker's beliefs and constraints as well as preferences. Group differences in choices may reflect whose experience people find informative, not only differences in preferences. An option with low take-up may therefore appeal to more people when they have better and more personally relevant information. When people lack evidence about personal fit, the referent's informed choice may help them evaluate it. If information from

certain sources seems more relevant, providing the same facts to everyone may not produce the same response. This is a reason to test information barriers, not to presume that all low take-up reflects unmet demand.

Neighborhood choice illustrates the distinction. Aliprantis et al. (2024) document racial differences in neighborhood socioeconomic status that income and wealth do not explain, and interpret their evidence in terms of racial homophily and neighborhood supply. The present experiment motivates an additional question: does apparent preference for familiar neighborhoods also reflect uncertainty about life elsewhere and whose experience households consider informative? A household might value better schools but remain unsure whether an unfamiliar neighborhood would meet its needs. Hearing from a family with similar priorities could change that assessment without changing the household's underlying preferences. The experiment thus points to informational influence as a potential explanation for part of the observed racial homophily, motivating a direct test of this mechanism in neighborhood choice.

For a housing-mobility program, the implication is to consider information and source relevance alongside access. Detailed neighborhood information and opportunities to hear from residents with varied backgrounds and priorities could complement rental assistance. Darrah and DeLuca (2014) describe how participants in the Baltimore Mobility Program revised their evaluations of neighborhoods after counseling and residential experience. Their qualitative evidence does not isolate learning from peers or identity matching; it shows why initial choices need not describe how households would evaluate options after further exposure. Limited experience with a type of neighborhood can make it difficult to assess whether living there would meet a household's needs. An informational-influence perspective suggests tailoring the content to these uncertainties and choosing sources whose experiences households find personally relevant. Information provision is not a substitute for addressing affordability, discrimination, housing supply, or the loss of existing support networks. Appendix M retains the detailed neighborhood evidence and other field comparisons.

The women-in-economics example raises the same distinction. The role-model intervention

in Porter and Serra (2020) increased women's major choice, but does not establish that preference-relevant learning caused the response; the authors emphasize evidence consistent with inspiration. Outreach could reveal latent demand by helping students recognize that economics fits goals they already have. It could instead create demand by changing what students value, including their desire to identify with the field. These possibilities imply different interventions: clearer information about courses and careers, tailored to students' differing needs and highlighting relevant experiences from people with whom they share an identity, may address the first. The second may depend on belonging or inspiration. Enrollment alone cannot distinguish them.

A concrete policy test would vary factual information and its source separately. Within a common assistance or outreach program, participants could receive standard or more detailed information, crossed with comparable accounts from different messengers. Holding factual content fixed across messengers would test whether the source changes the response to that content; varying direct information would test whether the messenger effect depends on what participants can evaluate themselves. Measures of knowledge, expected personal fit, perceived similarity, and belonging before and after exposure would help interpret subsequent choices. Applying the laboratory interpretation also requires that additional information leave the choice response due to direct alignment value unchanged.

The policy objective is better-informed choice, not copying or take-up for its own sake. Shared identity need not indicate actual preference similarity, and greater influence need not improve outcomes. Evaluations should therefore consider satisfaction, persistence, costs, and alternatives forgone alongside participation. The laboratory estimate neither predicts the size of a housing or education intervention nor establishes that demographic matching is always the best way to select messengers. It instead motivates testing both what people can learn and whose experience helps them evaluate their options.

9 Conclusion

When key information is missing, people may use others' choices and experiences in place of direct evidence when evaluating their options. This paper shows that shared identity can shape whose choice they treat as informative. In a controlled experiment with decisions under risk, concealing lottery probabilities and possible payoffs increases reliance on referents, especially when the referent shares the participant's gender. Shared gender also raises perceived similarity in risk preferences before the referent's choice is revealed. When payoffs are hidden, perceived similarity predicts greater exact copying; this relationship largely disappears with visible payoffs. The switching-toward and net-shift robustness checks also show stronger similarity associations under hidden payoffs.

Together, these patterns suggest that identity can act as a noisy guide to whose experience is relevant when objective information is scarce. They do not show that perceived similarity is the only channel, or that every shared identity is informative. Identity can therefore shape decisions not only through tastes and norms, but also by changing whose behavior people treat as evidence. For policy, this motivates testing information content and source relevance together before interpreting low take-up as weak demand. Whether that improves decisions in housing or education remains a question for field evaluation.

References

Akerlof, G. A., and Kranton, R. E. (2000). Economics and identity. *Quarterly Journal of Economics*, 115(3), 715–753.

Aliprantis, D., Carroll, D. R., and Young, E. R. (2024). What explains neighborhood sorting by income and race? *Journal of Urban Economics*, 141, 103508. doi:10.1016/j.jue.2022.103508.

Bayer, P., Casey, M. D., McCartney, W. B., Orellana-Li, J., and Zhang, C. S. (2022).

Distinguishing causes of neighborhood racial change: A nearest neighbor design. NBER Working Paper 30487.

Bergman, P., Chan, E., and Kapor, A. (2020). Housing search frictions: Evidence from detailed search data and a field experiment. NBER Working Paper 27209.

Berger, S., Feldhaus, C., and Ockenfels, A. (2018). A shared identity promotes herding in an information cascade game. *Journal of the Economic Science Association*, 4(1), 63–72.

Braaksma, M. A. H., Rijlaarsdam, G., and van den Bergh, H. (2002). Observational learning and the effects of model-observer similarity. *Journal of Educational Psychology*, 94(2), 405–415.

Bruch, E. E., and Feinberg, F. (2019). Decision-making processes in social contexts. *Annual Review of Sociology*, 45, 207–227.

Bryan, G., Chowdhury, S., and Mobarak, A. M. (2014). Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh. *Econometrica*, 82(5), 1671–1748.

Bursztyn, L., Ederer, F., Ferman, B., and Yuchtman, N. (2014). Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. *Econometrica*, 82(4), 1273–1301.

Coffman, L. C., Featherstone, C. R., and Kessler, J. B. (2017). Can social information affect what job you choose and keep? *American Economic Journal: Applied Economics*, 9(1), 96–117.

Collette, S., Pauli, W. M., Bossaerts, P., and O'Doherty, J. P. (2017). Neural computations underlying inverse reinforcement learning in the human brain. *eLife*, 6, e29718.

Corradini, V., and Idoux, C. (2024). Overcoming racial gaps in school preferences: The role of peer diversity on school choice. Working paper.

Darrah, J., and DeLuca, S. (2014). "Living here has changed my whole perspective": How escaping inner-city poverty shapes neighborhood and housing choice. *Journal of Policy Analysis and Management*, 33(2), 350–384. doi:10.1002/pam.21758.

Eckel, C. C., and Grossman, P. J. (2008). Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *Journal of Economic Behavior & Organization*, 68(1), 1–17.

Ferraro, P. J., and Price, M. K. (2013). Using nonpecuniary strategies to influence behavior: Evidence from a large-scale field experiment. *Review of Economics and Statistics*, 95(1), 64–73. doi:10.1162/REST_a_00344.

Foster, A. D., and Rosenzweig, M. R. (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of Political Economy*, 103(6), 1176–1209.

Hackel, L. M., Coppin, G., Wohl, M. J. A., and Van Bavel, J. J. (2018). From groups to grits: Social identity shapes evaluations of food pleasantness. *Journal of Experimental Social Psychology*, 74, 270–280.

Lee, E.-J. (2005). When placebic information differs from real information: Cognitive and motivational bases of mindful reactions to informational social influence. *Communication Research*, 32(5), 615–645. doi:10.1177/0093650205279351.

Lenoir, A. S. I. (2015). *Are You Talking to Me? Addressing Consumers in a Globalised World*. Doctoral dissertation, Erasmus University Rotterdam, Chapter 4. Erasmus University repository.

Manski, C. F. (2000). Economic analysis of social interactions. *Journal of Economic Perspectives*, 14(3), 115–136.

Munshi, K. (2004). Social learning in a heterogeneous population: Technology diffusion in the Indian Green Revolution. *Journal of Development Economics*, 73(1), 185–213. doi:10.1016/j.jdeveco.2003.03.003.

Munshi, K., and Myaux, J. (2006). Social norms and the fertility transition. *Journal of Development Economics*, 80(1), 1–38.

Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy*, 78(2), 311–329.

Porter, C., and 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.

Tajfel, H. (1979). Individuals and groups in social psychology. *British Journal of Social and Clinical Psychology*, 18(2), 183–190.

Taylor-Davies, M., Bramley, N. R., and Lucas, C. G. (2025). A rational framework for group-based selective social learning. *Open Mind*, 9, 677–708. doi:10.1162/opmi_a_00205.

Wabitsch, A. (2025). The messenger matters. Working paper. doi:10.2139/ssrn.5238987.

Yaniv, I., Choshen-Hillel, S., and Milyavsky, M. (2011). Receiving advice on matters of taste: Similarity, majority influence, and taste discrimination. *Organizational Behavior and Human Decision Processes*, 115(1), 111–120.

A Framework and Interpretation

Let ΔQ_{it} be the directly evaluated value of the referent's option minus that of the alternative. This difference can be negative. Equation 1 implies the systematic utility difference

$$V_{ijbt} = (1 - B_{ib})\Delta Q_{it} + B_{ib}\omega S_{ij} + \nu G_{ij}. \tag{10}$$

Let η be the other option's utility shock minus the referent option's shock, with a differentiable cumulative distribution function F. Conditional on the systematic advantage in Equation 10, the probability of copying is $F(V_{ijbt})$. For an interior choice probability,

$$\frac{\partial P(Copy = 1)}{\partial S} = B_{ib}\omega F'(V_{ijbt}). \tag{11}$$

Perceived similarity therefore raises copying under hidden information when $\omega > 0$ and the shock density is positive. In the full-information benchmark, its slope is zero. This is a benchmark prediction, not a restriction that the estimated visible-information similarity slope must equal zero.

For fixed ΔQ, changing visibility from $B = 0$ to $B = 1$ changes the utility advantage by $\omega S - \Delta Q$. Copying rises when the preference-fit evidence supplied by the referent is stronger than the advantage obtained from direct evaluation. The model does not predict a positive visibility effect for every subject or lottery.

Let $\Delta_G S$ be the change in perceived similarity associated with same-gender assignment. Holding direct evaluation fixed, the same-gender utility difference is ν with full information and $\nu + \omega\Delta_G S$ with hidden information. Their difference, $\omega\Delta_G S$, is the additional informational utility associated with gender match.

This index comparison does not automatically become the same additive decomposition in observed probabilities. With a nonlinear F, the probability effect of a fixed increment in direct alignment value, ν, depends on the starting value of V. Consequently, direct alignment

value that is constant across visibility conditions is not alone sufficient to make its probability effect cancel. The interpretation of the experimental double difference instead uses the restriction stated in Section 4.5 on the mean choice response attributable to direct alignment value. The linear probability specification estimates that response contrast directly; it does not estimate F, ω, or ν as structural parameters.

B Alternative Outcomes and Specifications

For the gender-focused estimates reported in this appendix, the specification is

$$\begin{aligned} Y_{ibt} = \beta_G SameGender_{ib} + \beta_{GB}(SameGender_{ib} \times Blind_{ib}) \\ + \beta_B Blind_{ib} + \alpha_i + \delta_b + \lambda_t + \varepsilon_{ibt}. \end{aligned} \tag{12}$$

Here α_i, δ_b, and λ_t denote participant, block-position, and lottery-pair fixed effects. The nationality-adjusted specification below corresponds to the main-text model.

B.1 Baseline-Referenced Outcomes

Let C_{it}^0 denote subject i's Part 1 choice for lottery pair t, C_{ibt} the Part 2 choice, and R_{ibt} the displayed referent choice. The conditional switch-toward measure is

$$Shift_{ibt} = \mathbf{1}\{C_{ibt} = R_{ibt}\}, \qquad \text{for trials with } C_{it}^0 \neq R_{ibt}. \tag{13}$$

It asks whether a subject switches to the referent when they initially disagreed. Restricting the sample in this way is useful descriptively, but the retained trials depend on the referent's choice.

The net-shift measure retains every primary trial:

$$NetShift_{ibt} = \mathbf{1}\{C_{ibt} = R_{ibt}\} - \mathbf{1}\{C_{it}^0 = R_{ibt}\}. \tag{14}$$

It equals +1 for a switch toward the referent, −1 for a switch away, and zero when the choice is unchanged. Its mean is exact copying minus baseline agreement. A negative observation can reflect deliberate use of the referent's choice rather than absence of referent influence. Neither measure treats the baseline choice as a welfare optimum.

Table 5: Baseline-Referenced Gender Effects

Contrast	Switch toward	Net shift
Same-gender effect, visible	1.56	0.34
	(1.87)	(1.56)
Same-gender effect, hidden	9.13	10.07
	(2.16)	(2.36)
Hidden minus visible	7.56	9.73
	(2.65)	(2.79)
p-value, hidden minus visible	0.0047	0.0006
Observations	5,758	12,656
Participants	226	226

Notes: Estimates and standard errors are in percentage points. Both columns use the gender-focused specification in Equation 12, with participant, block-position, and lottery-pair fixed effects and participant-clustered standard errors. The switch-toward sample contains only trials with baseline disagreement.

B.2 Nationality-Adjusted Specification

For comparison with the nationality analyses, I add nationality match and its interaction with payoff visibility:

$$Y_{ibt} = \beta_G SameGender_{ib} + \beta_{GB}(SameGender_{ib} \times Blind_{ib}) + \beta_B Blind_{ib} + \beta_N SameNat_{ib} + \beta_{NB}(SameNat_{ib} \times Blind_{ib}) + \alpha_i + \delta_b + \lambda_t + \varepsilon_{ibt}. \quad (15)$$

The added terms leave the exact-copying same-gender estimates unchanged to the reported precision (Table 6). The visibility coefficient alone in this expanded model refers to a different-gender, different-nationality referent. To compare the same quantity across models, the final row averages the visibility effect over nationality matching while holding gender match at zero.

Table 6: Exact Copying: Gender-Focused and Nationality-Adjusted Models

Contrast	Gender-focused	Nationality-adjusted
Same-gender effect, visible	2.41	2.41
	(1.52)	(1.52)
Same-gender effect, hidden	9.45	9.45
	(1.68)	(1.68)
Hidden minus visible	7.05	7.05
	(2.07)	(2.07)
Visibility effect, different gender	7.83	7.83
	(1.82)	(1.82)

Notes: Estimates and participant-clustered standard errors are in percentage points. Both models use 12,656 trials from 226 participants and include participant, block-position, and lottery-pair fixed effects. The nationality-adjusted final row is $\beta_B + \frac{1}{2}\beta_{NB}$, reflecting the balanced nationality-match design; it is not the coefficient on *Blind* alone.

B.3 Ordinal and Directional Belief Controls

Treating the 0–2 similarity score as cardinal imposes a common incremental slope between adjacent categories. I therefore replace the linear score with indicators for all three similarity categories, each interacted with visibility. A second specification uses all five original directional responses, again allowing their relationship with copying to differ by visibility. Table 7 shows that the same-gender interaction remaining after these controls is similar across codings.

Table 7: Same-Gender Interaction under Alternative Belief Controls

Belief controls	Estimate	SE	p-value
Linear 0–2 score	3.99	2.03	0.0503
Three unrestricted similarity categories	4.04	2.02	0.0461
Five original directional categories	3.80	1.99	0.0568

Notes: Each row reports the same-gender-by-hidden-payoff interaction in exact copying, in percentage points. All models include same-gender and visibility terms, participant, block-position, and lottery-pair fixed effects, and participant-clustered standard errors; they use 12,656 trials from 226 participants. Every belief term is interacted with visibility. These are conditional associations, not estimates of a direct causal effect.

The stability of the coefficient indicates that the remaining same-gender association is not simply removed by allowing nonlinearity in the recorded belief response. It does not identify a separate identity mechanism: measurement error and channels not captured by the single

item can also leave a conditional same-gender association.

B.4 Similarity Controls for Copying and Net Shift

The nationality-adjusted specification provides the same comparison for net shift as for exact copying. Table 8 reports the same-gender terms after controlling for either the linear similarity score or its unrestricted categories. Allowing unequal spacing between categories leaves both results similar. As in Table 7, the remaining coefficients are conditional associations, not a decomposition of the randomized identity effect into direct and indirect effects.

Table 8: Belief Coding and Remaining Same-Gender Associations

Outcome and contrast	Linear 0–2 score	Categories
Hidden-minus-visible same-gender interaction (pp)		
Exact copying	4.00 (2.03)	4.06 (2.02)
Net shift	7.65 (2.72)	7.77 (2.70)
Same-gender association with hidden payoffs (pp)		
Exact copying	6.14 (1.51)	6.16 (1.51)
Net shift	7.13 (2.22)	7.17 (2.22)

Notes: Participant-clustered standard errors in parentheses. All models include participant, block-position, and lottery-pair fixed effects, same-nationality and same-nationality-by-visibility terms, and visibility interactions with every belief term. These added nationality terms distinguish this specification from Table 7. The categorical interaction has $p = 0.046$ for copying and $p = 0.004$ for net shift. All models use 12,656 trials and 226 participants.

B.5 Independent Random-Choice Benchmark

Independent random choices provide a benchmark for the magnitude of the headline interaction. The simulation holds the menus, referent choices, assignments, and regression design fixed and replaces all 12,656 subject choices with independent 50–50 draws between the two options. Across 100,000 replications, the same-gender-by-hidden-payoff copying interaction has mean approximately zero, standard deviation 1.79 percentage points, and a central 95 percent range of $[-3.51, 3.51]$ points. Five replications produce a contrast at least as extreme as the observed 7.05-point estimate. The observed regression's participant-clustered

standard error is 2.07 points and its p-value is 0.00077; this is separate from the simulation tail frequency.

Table 9: Identity Interactions Relative to Independent Random Choice

Hidden-minus-visible same-gender interaction	Observed (pp)	Null mean (pp)	Null SD (pp)
Exact copying	7.05	0.00	1.79
Switch toward referent	7.58	0.00	2.68
Net shift	9.73	2.68	1.79

Notes: The regressions include participant, block-position, and lottery-pair fixed effects and nationality-match terms. Switching toward the referent uses the 5,758 trials with initial disagreement. Its simulation tail probability is 0.00468. Net shift equals exact copying minus initial agreement on every trial. Initial agreement is held fixed in the simulation, so copying and net shift have identical simulation dispersion; their observed-minus-null-mean contrasts are both 7.05 points, up to rounding.

This benchmark rejects independent coin-flipping as an account of the observed interaction. It is not randomization inference over treatment assignments and does not test every model of noisy, inattentive, or noninformational choice. The pooled raw rates also describe different objects: exact copying is 66.72 percent, repetition of the subject's own Part 1 choice is 60.30 percent, and switching toward the referent conditional on initial disagreement is 57.05 percent. None is itself a treatment effect.

B.6 Participant-Level Changes in Choice Profiles

For each participant, I count disagreements with referents across the 56 primary Part 2 trials and subtract this from the count implied by repeating Part 1 choices. This reduction is the sum of trial-level net shift. It is positive for 176 of 226 subjects (77.9 percent; Figure 10). This is distributional evidence on the existing outcome, not a separate measure of latent preference distance.

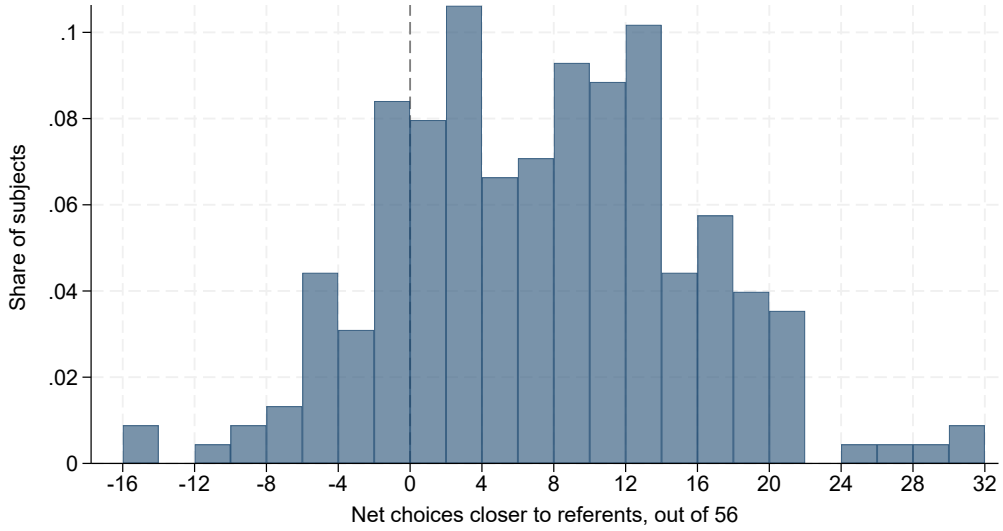


Figure 10: Participant-level reductions in disagreements with referents across 56 choices. Positive values indicate closer alignment than repeating Part 1 choices. Each observation is one participant.

C Nationality and Referent-Country Results

C.1 Original Nationality-Match Specification

The nationality-only specification was preregistered before data collection:

$$Y_{ibt} = \beta_1 SameNat_{ib} + \beta_2 Blind_{ib} + \beta_3 (SameNat_{ib} \times Blind_{ib}) + \alpha_i + \delta_b + \lambda_t + \varepsilon_{ibt}. \quad (16)$$

The estimated interaction is -4.6 percentage points (standard error 2.1 percentage points, $p = 0.027$), opposite to the predicted positive interaction. Adding the same-gender terms in Equation 15 leaves this estimate and its statistical inference unchanged.

Table 10: Same-Nationality Effects by Payoff Visibility

Exact-copying contrast	Estimate (pp)	SE	p-value
Same nationality, visible payoffs	−0.08	1.52	0.960
Same nationality, hidden payoffs	−4.66	1.45	0.001
Difference: hidden minus visible	−4.58	2.06	0.027

Notes: The joint identity specification includes same-gender and same-gender-by-visibility terms, participant, block-position, and lottery-pair fixed effects, and participant-clustered standard errors. There are 12,656 trials from 226 participants. These are nationality-match contrasts, not China-minus-U.S. referent contrasts.

Table 10 separates the interaction into its two components. Same-nationality assignment has essentially no estimated effect with visible payoffs and reduces copying by 4.66 percentage points with hidden payoffs.

C.2 Referent Country and Directional Risk Beliefs

The results do not show a positive same-nationality informational effect. The clearer pattern is attached to the referent's country: under hidden payoffs, exact copying is 75.3 percent for China referents and 69.5 percent for U.S. referents. The corresponding net shifts are 21.3 and 13.3 percentage points. These gaps nearly disappear when payoffs are visible. Both U.S.

and China participants copy China referents more under hidden payoffs, so the result is not driven by one participant group copying its own nationality.

The referent pools do not differ detectably in actual risk-taking, but subjects perceive China referents as less risk-taking than U.S. referents. After controlling for the other displayed profile traits, the China-minus-U.S. difference in this pre-choice belief is 0.278 points on the 1-to-5 scale, where higher values indicate that the referent is perceived as less risk-taking relative to the subject. Among more risk-averse subjects, a larger China-minus-U.S. "less risky" belief gap predicts greater China-minus-U.S. exact-copying gap under hidden payoffs: the risk-aversion-by-belief-gap interaction is 6.17 percentage points. For example, holding other covariates fixed, a one-point rather than zero-point belief gap predicts a 1.92-percentage-point larger China-minus-U.S. exact-copying gap for a participant at mean risk aversion and an 8.09-percentage-point larger gap for a participant one standard deviation above the mean. The 6.17-percentage-point difference between these predicted changes is the interaction. The corresponding interaction is small and statistically insignificant when payoffs are visible. This pattern is consistent with risk-preference matching, in which a referent perceived as less risk-taking is more informative to a more risk-averse subject. Tables 11 and 14 report the supporting country-specific and risk-belief estimates.

The China and U.S. referent pools differ in class-year composition and reuse frequency because the China pool is much smaller. Consequently, the estimated country difference is a referent-country and pool effect, not a pure nationality label effect. Table 16 reports these checks. Only 19 completed participants grew up in China, so comparisons between participant-nationality subgroups remain imprecise.

C.3 Referent-Pool Comparability and Cross-Condition Precision

Table 12 reports the actual-choice comparisons behind the distinction between perceived and observed risk-taking. Neither comparison detects a difference between the two country pools. These estimates do not establish equivalence, particularly given the wide interval implied by

Table 11: Nationality Cues, Beliefs, and Referent Influence

	U.S. referent	China referent	Difference
Panel A: Hidden-payoff behavior			
Exact copying	0.695	0.753	0.058***
			(0.020)
Switch toward referent	0.702	0.746	0.042*
			(0.023)
Net shift	0.133	0.213	0.080***
			(0.026)
Panel B: Visible-payoff behavior			
Exact copying	0.609	0.611	0.002
			(0.024)
Switch toward referent	0.423	0.418	−0.003
			(0.025)
Net shift	0.064	0.079	0.014
Panel C: Pre-choice beliefs, China minus U.S.			
Rated "less risky than me," hidden			0.278***
			(0.071)
Perceived similarity, hidden			0.050
			(0.053)
Rated "less risky than me," visible			0.231***
			(0.075)
Perceived similarity, visible			0.066
			(0.054)

Notes: The first two columns in Panels A and B are raw rates. Difference estimates come from trial-level linear models with participant, trial-position, and lottery-pair fixed effects. Exact-copy and switch-toward standard errors are two-way clustered by participant and referent; the hidden net-shift standard error comes from the corresponding net-shift model. Panel C uses block-level regressions with gender, class year, and shape controls and two-way clustered standard errors. The analysis contains 226 participants, 1,808 blocks, and 12,656 primary trials. $^{*}p < 0.10$, $^{***}p < 0.01$.

the high-risk-option standard error.

Table 12: Actual Risk Choices of China and U.S. Referents

China-minus-U.S. contrast	Estimate	SE	p-value
Part 1 risk-taking index	0.031	0.067	0.639
High-risk option selection rate (pp)	−0.3	7.4	0.973

Notes: The index comparison uses 201 unique referents. The option-selection comparison uses their displayed choices in 12,656 primary trials. The index is behavioral and is not a structural risk-aversion parameter.

For exact copying, the adjusted China-minus-U.S. referent contrast is 5.8 percentage points with hidden payoffs (SE 2.0) and 0.2 points with visible payoffs (SE 2.4). Their difference is 5.6 points (SE 3.0, $p < 0.10$), so the cross-condition contrast is less precise than the hidden-payoff estimate alone. The model includes participant, trial-position, and lottery-pair fixed effects and two-way clustering by participant and referent. These estimates describe the country-pool comparison, not the same-nationality coefficient in Table 10.

Likewise, a precisely estimated risk-aversion-by-belief-gap interaction with hidden payoffs and an imprecise one with visible payoffs do not by themselves establish that the two interactions differ. Table 14 reports both, along with the implied belief-gap slopes at low, mean, and high risk aversion.

C.4 Gender–Nationality Interactions

The nationality-adjusted specification in Equation 15 includes both identity dimensions and their interactions with payoff blindness but imposes additivity between gender and nationality. I relax that restriction by fully interacting payoff blindness, same-gender assignment, and same-nationality assignment. Table 13 reports how the hidden-minus-visible change in the same-gender effect differs by nationality matching.

For exact copying, the hidden-minus-visible same-gender interaction is 3.1 percentage points with a different-nationality referent and 11.0 points with a same-nationality referent. The 7.8-point difference has $p = 0.054$, while the two added higher-order terms are jointly

Table 13: Nationality Match and the Payoff-Blind Same-Gender Interaction

	Exact copying	Switch toward referent	Net shift
Hidden-minus-visible change in the same-gender effect			
Different-nationality referent	0.031	0.021	0.077
	(0.028)	(0.035)	(0.035)
Same-nationality referent	0.110	0.130	0.117
	(0.030)	(0.039)	(0.039)
Difference: same minus different nationality	0.078	0.109	0.040
	(0.040)	(0.052)	(0.048)
p-value for difference	0.054	0.038	0.410
Joint p-value for added higher-order terms	0.153	0.098	0.676
Observations	12,656	5,758	12,656
Participants	226	226	226

Notes: Linear models fully interact payoff blindness, same-gender assignment, and same-nationality assignment and include participant, block-position, and lottery-pair fixed effects. Standard errors in parentheses are clustered by participant. The joint test covers the same-gender-by-same- nationality term and the three-way interaction. The switch-toward outcome is defined only when the participant's Part 1 choice differs from the referent's choice.

insignificant ($p = 0.153$). The conditional switch-toward contrast is larger, but net shift shows a smaller and imprecise difference. The evidence therefore suggests possible complementarity between gender and nationality matching, but it does not displace the pooled 7.0-point exact-copying interaction as the main estimate.

C.5 Two Codings of the Pre-Choice Belief

Table 14 reports the risk-belief estimates. Because the directional belief gap and symmetric similarity gap are two codings of the same pre-choice response, the final specification asks whether the direction of the belief contains predictive information beyond its distance from perceived similarity. When both codings enter together, the directional interaction is 3.39 percentage points (standard error 2.21 percentage points, $p = 0.126$). The estimate is imprecise; the single belief item does not allow the directional and symmetric components to be cleanly distinguished.

Table 14: Risk Perception and China-Referent Exact Copying

Estimand	Estimate	SE	p-value
China minus U.S. effect on "less risky than me," hidden	0.278	0.071	< 0.001
Risk aversion $\times$ less-risky belief gap, hidden exact copying	0.0617	0.0205	0.0029
Belief-gap slope at +1 SD risk aversion, hidden exact copying	0.0809	0.0278	0.0040
Belief-gap slope at mean risk aversion, hidden exact copying	0.0192	0.0206	0.352
Belief-gap slope at -1 SD risk aversion, hidden exact copying	-0.0425	0.0302	0.160
Risk aversion $\times$ belief gap, visible exact-copying placebo	0.0147	0.0237	0.534
Risk aversion $\times$ directional belief gap, hidden; both codings included	0.0339	0.0221	0.126

Notes: The directional belief uses the original 1–5 scale, not the symmetric 0–2 similarity score. Higher values indicate that the referent is perceived as less risk-taking relative to the subject. The belief gap and exact-copying gap are each the participant's China-minus-U.S. difference. The first row uses 904 hidden blocks and controls for displayed gender, class year, and shape, with two-way clustering by participant and referent. The remaining rows use 226 participant-level gaps, HC2 standard errors, and participant risk aversion standardized from Part 1. The focused-family Holm-adjusted p-value for the hidden interaction is 0.0058. The final row adds the symmetric similarity gap and its interaction with risk aversion. Both gaps are derived from the same pre-choice response.

D Assignment and Referent Checks

Table 15 compares Part 1 characteristics of the 226 participants who completed Part 2 and the 41 who did not. Table 16 documents implementation in the completed sample.

Table 15: Part 1 Characteristics by Part 2 Completion

	Did not complete Part 2	Completed Part 2	Standardized difference	p-value
Grew up in China	0.171	0.084	−0.293	0.162
Female	0.732	0.708	−0.052	0.754
Numeracy concepts correct, first attempt	0.549	0.558	0.031	0.854
General recruitment pool	0.927	0.912	−0.054	0.733
Part 1 comprehension mistakes	7.244	5.186	−0.309	0.136
Part 1 trials completed within 2 seconds	0.027	0.028	0.015	0.914
Part 1 median response time, seconds	10.441	10.261	−0.032	0.875
Part 1 risk-taking index	0.511	0.456	−0.252	0.130
Participants	41	226		

Notes: The first column combines 35 Part 1 completers who never started Part 2 and six who started but did not complete all eight blocks. The difference is completed minus not completed. The final column reports two-sided tests of equal means; every Holm-adjusted p-value across the eight rows equals 1.000. A joint linear-probability model of Part 2 completion using the core covariates gives $F = 1.22$ and $p = 0.296$. Failure to reject does not establish equality on unobserved characteristics.

Table 16: Within-Subject Assignment and Referent-Selection Checks

Check	Count	Joint p-value
Panel A: Assignment and implementation		
Williams sequence 0	58	
Williams sequence 1	61	
Williams sequence 2	57	
Williams sequence 3	50	
Participants receiving two blocks in every nationality-by-visibility cell	226 of 226	
Participant-condition cells with one female and one male referent	904 of 904	
Blocks using an eligible assigned risk tail	1,808 of 1,808	
Blocks requiring a contingency fallback	0 of 1,808	
Panel B: Referent characteristics across assigned conditions		
Same gender		1.000
Female referent		1.000
Part 1 risk-taking index		0.502
High-risk option selection rate		0.986
Mean chosen expected value		0.969
Mean chosen standard deviation		0.975
Freshman indicator		0.012
Sophomore indicator		0.002
Junior indicator		< 0.001
Senior indicator		< 0.001
Referent reuse count		< 0.001

Notes: Panel B reports the participant-clustered joint test of payoff blindness, same-nationality assignment, and their interaction in 1,808 block-level observations. Referent country is omitted because it is mechanically determined by the assigned nationality-match condition and the participant's country. Class-year composition and reuse differ across conditions because the China referent pool is smaller and more senior-heavy than the U.S. pool. These differences make referent country a country-and-pool treatment rather than a pure identity-label treatment. Main country comparisons use participant/referent clustering, and the belief estimates control for displayed profile characteristics.

E Gender-Risk Contrasts

The baseline choices differ on average but overlap substantially across gender. Women select the higher-standard-deviation option on an average of 8.61 of the 20 Part 1 choices; men do so on 10.38. The standardized mean difference is 0.41 (Hedges' g). Across all male–female participant pairs, the probability that the man has the higher index, counting ties as one-half, is 0.60. Gender is therefore an imperfect clue about an individual's behavior, not a division into uniformly risk-averse women and risk-tolerant men. The index also reflects risk-return tradeoffs, as described in Section 4.

Figure 11 displays this overlap. The higher-SD option also has the higher expected value in 18 of the 20 comparisons, so the distribution describes behavior over risk-return tradeoffs rather than a pure measure of risk aversion.

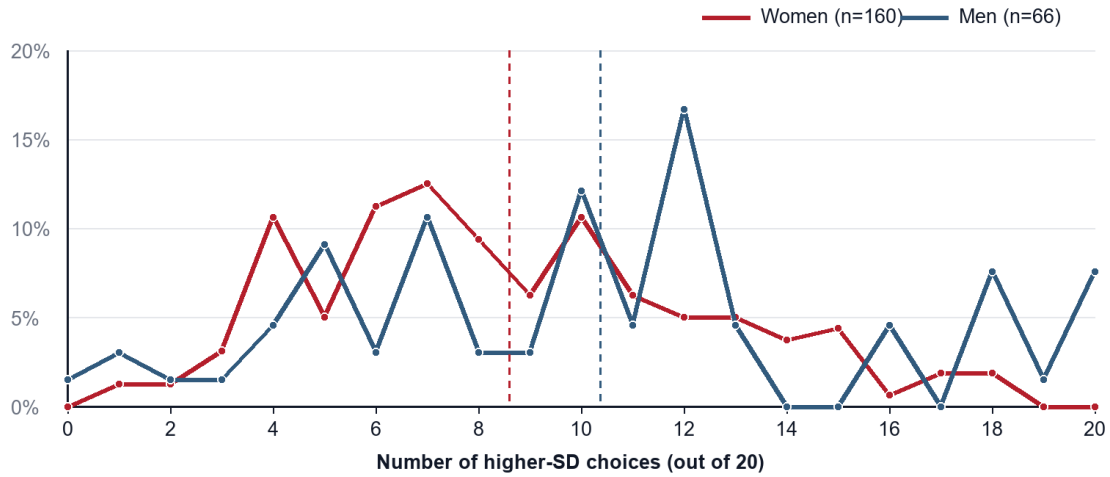


Figure 11: Distribution of the number of higher-standard-deviation options selected in Part 1, by gender. The sample contains 160 women and 66 men; each subject makes 20 choices.

The estimates are consistent with gendered beliefs about risk preference. Among men, a one-standard-deviation increase in risk aversion reduces the same-gender minus different-gender similarity difference by 0.126 points (SE = 0.043, $p = 0.004$). Equivalently, the same-gender differential is larger among more risk-tolerant men. Among women, a one-standard-deviation increase in risk aversion increases the same-gender differential by 0.166 points (SE = 0.041, $p < 0.001$). The female-minus-male difference in these slopes is 0.292

points ($p < 0.001$).

These estimates show why "same gender" is not a universal similarity rule. Gender is used as a clue about risk preference, and the relevance of that clue depends on the subject's own revealed risk preference. The estimates concern perceived similarity, where the pattern is precise. Risk aversion does not significantly moderate the remaining same-gender exact-copying coefficient after the similarity terms are included.

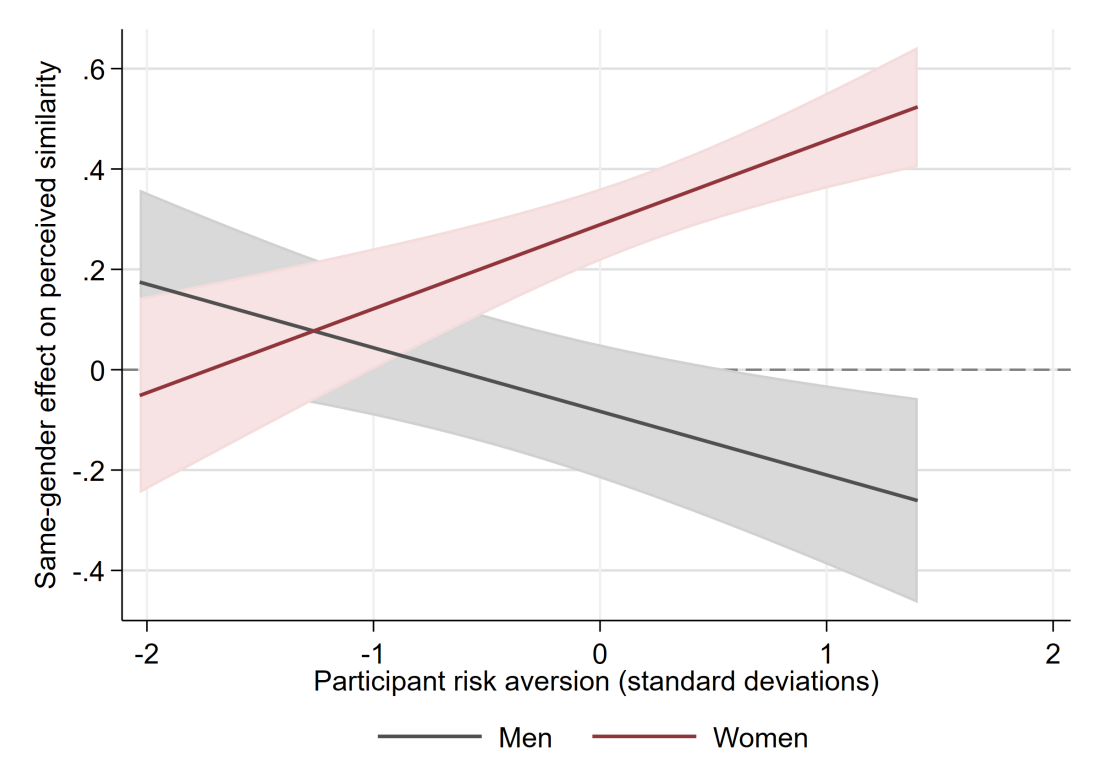


Figure 12: Estimated effect of same-gender assignment on perceived risk-preference similarity across participant risk aversion. Lines average over payoff-visibility conditions; shaded bands are 95 percent confidence intervals. The horizontal range covers the 5th through 95th percentiles of the participant risk-aversion distribution.

Table 17: Gender-Risk Moderation

Contrast	Estimate	SE	p-value
Risk-aversion slope of the same-gender minus different-gender similarity gap			
Men	−0.126	0.043	0.004
Women	0.166	0.041	< 0.001
Women minus men	0.292	0.059	< 0.001
Same-gender exact-copying slope per +1 SD risk aversion, conditional on similarity			
Men	0.019	0.016	0.251
Women	−0.004	0.016	0.774
Women minus men	−0.023	0.023	0.305

Notes: Risk aversion is the negative standardized Part 1 risk-taking index, a behavioral proxy rather than an estimated structural parameter. Similarity estimates use 1,808 participant-block observations; exact-copying estimates use 12,656 trials. Models include all lower-order terms, applicable fixed effects, and participant-clustered standard errors. The conditional exact-copying panel does not identify a causal direct effect.

F Referent Selection and Implementation Details

The selection algorithm operated within an already assigned block condition:

1. Restrict the pool to the assigned nationality and exclude the subject's own Part 1 record.
2. Restrict to the randomly assigned top or bottom 30 percent of the distribution of the Part 1 risk-taking index among eligible referents.
3. Apply the referent-gender requirement generated by the within-cell balancing rule.
4. Avoid a previously used referent for the same subject when possible.
5. Randomly select among the least-used eligible referents and record the realized assignment.

The top/bottom targeting was chosen before implementation to increase behavioral contrast and power. No nationality, risk-tail, or support fallback was required in the completed analysis sample. All 1,808 blocks used an eligible target-tail referent, and all 904 participant-by-condition cells achieved the intended one-male, one-female balance.

G Measurement Details

The pre-choice question is: "How does this referent compare to you in their riskiness?" Table 18 gives the coding.

Table 18: Perceived Risk-Preference Similarity Coding

Response value	1	2	3	4	5
Label	Much more	More	Similar	Less	Much less
Similarity score	0	1	2	1	0

Notes: "More" and "less" refer to the referent's risk taking relative to the subject. Symmetric coding treats equal distances above and below the subject as equally dissimilar.

H Order, Carryover, and Belief Updating

Because subjects complete eight referent blocks, treatment effects could change as subjects gain experience with the task or could depend on the immediately preceding treatment. Table 19 examines both possibilities. The same-gender-by-hidden interaction in exact copying is 6.9 percentage points in Blocks 1–4 and 7.2 points in Blocks 5–8. The estimated early-to-late change is 0.2 points ($p = 0.966$), and a joint test does not reject equality across all eight block positions ($p = 0.974$). Dropping the first block produces a 7.6-point estimate. The corresponding early-to-late change in the nationality interaction is also small and imprecise ($p = 0.554$).

Restricting the sample to Blocks 2–8 permits direct controls for the preceding block. Adding the prior block's payoff-visibility, gender-match, and nationality-match assignments leaves the current same-gender-by-hidden interaction essentially unchanged for both exact copying and net shift. Allowing this interaction to vary with all three preceding assignments yields joint p-values of 0.603 and 0.913, respectively. These tests detect neither systematic order change nor one-block treatment carryover, although they cannot rule out more diffuse or longer-lived changes.

Table 19: Order and One-Block Carryover Checks

	Estimate	SE	p-value
Panel A: Same-gender-by-hidden interaction in exact copying			
Blocks 1–4	6.9 pp	3.2 pp	0.031
Blocks 5–8	7.2 pp	3.6 pp	0.051
Change from Blocks 1–4 to Blocks 5–8	0.2 pp	5.5 pp	0.966
Drop Block 1	7.6 pp	2.4 pp	0.002
Joint test of equality across eight positions			0.974
Panel B: Prior-treatment controls, Blocks 2–8			
	Exact copying	Net shift	
Baseline interaction	7.6 pp	11.2 pp	
	(2.4 pp)	(3.0 pp)	
Interaction with prior-treatment controls	7.6 pp	11.0 pp	
	(2.4 pp)	(3.0 pp)	
Joint test: interaction varies with prior treatments	0.603	0.913	

Notes: Entries in Panel A are same-gender-by-hidden interactions from exact-copying linear probability models. Panel B reports this interaction before and after adding indicators for the preceding block's randomized payoff, gender-match, and nationality-match treatments; the final row jointly tests interactions with those three prior treatments. All models include gender and nationality match, their visibility interactions, participant, block-position, and lottery-pair fixed effects, and cluster standard errors by participant. Panel A uses 12,656 trials from 226 participants, except the drop-Block-1 estimate (11,074 trials). Panel B uses 11,074 trials from 226 participants.

H.1 Belief Updating and Generalization to Other Referents

Visible-payoff blocks reveal seven choices made by the current referent. Comparing beliefs before and after those choices tests whether subjects update their view of that person. Comparing the next referent's initial rating tests whether that experience generalizes to a person whose choices have not yet been shown. The exposure contrast is the randomly assigned top versus bottom Part 1 risk pool.

For this check alone, the directional item is reversed to $6 - B$, where B is the original 1–5 response in Table 18. Higher values therefore indicate more perceived risk-taking, and current-referent updating is the post-choice minus pre-choice rating. This is not the symmetric 0–2 similarity score. Table 20 reports a 0.64-point larger update after observing a high-risk rather than low-risk referent, but small and imprecise differences in ratings of subsequent referents.

Table 20: Risk-Pool Exposure, Belief Updating, and Subsequent Ratings

High-minus-low risk-pool contrast	Unadjusted	Adjusted	Observations
Current referent: post-minus-pre rating	0.637 (0.074)	0.773 (0.064)	904
Immediately next referent: pre-choice rating	0.053 (0.067)	0.013 (0.072)	797
Next referent of the same gender: pre-choice rating	0.073 (0.079)	0.139 (0.155)	452

Notes: Estimates are points on the reversed 1–5 scale; participant-clustered standard errors are in parentheses. All checks include 226 participants. Current-referent updating uses visible-payoff blocks, where post-choice beliefs are elicited. The immediately-next row requires a subsequent block after that exposure. The same-gender row starts with each gender's first visible referent and measures the next referent of that gender. Adjusted models include participant fixed effects, the current pre-choice belief, and relevant current and subsequent block/profile controls. Unadjusted 95 percent intervals are $[0.491, 0.783]$, $[-0.078, 0.184]$, and $[-0.084, 0.229]$; the latter two have $p = 0.426$ and $p = 0.360$.

Updating is clear for the observed person. Neither subsequent-rating comparison detects generalization, including when the next referent shares that person's gender. The adjusted same-gender estimate is particularly imprecise (95 percent CI $[-0.167, 0.445]$), so the results do not establish that spillovers are absent. They complement the order and carryover checks without requiring the assumption that beliefs remain fixed throughout the experiment.

I Baseline Reliability and Experimental Payments

Part 1 provides a useful but imperfect measure of each subject's decisions under risk. The Part 1 risk-taking index has standardized Cronbach's $\alpha = 0.803$, and its split-half reliability is 0.726 before and 0.841 after the Spearman–Brown correction. A Part 1 choice agrees with the same subject's later payoff-visible choice on 68.4 percent of trials. Among participant–lottery-pair cells with a unique modal payoff-visible choice, the Part 1 choice agrees with that mode 72.5 percent of the time. Table 21 reports the full diagnostics.

Table 21: Reliability of Part 1 Choices

Measure	Estimate	95% CI	Units
Standardized Cronbach's α	0.803		226 participants
Split-half correlation	0.726		226 participants
Spearman–Brown reliability	0.841		226 participants
Part 1 vs. payoff-visible trial agreement	68.4%	[66.1, 70.6]%	6,328 trials
Part 1 vs. payoff-visible modal agreement	72.5%	[69.8, 75.2]%	1,395 cells
Cohen's κ, Part 1 vs. payoff-visible trial	0.365	[0.321, 0.409]	6,328 trials

Notes: The split-half statistic compares two ten-item halves of the Part 1 choice set. Modal agreement excludes participant–pair cells tied across repeated payoff-visible choices. Because those choices occur after exposure to referents, the modal benchmark may itself incorporate prior social information.

Choices carried real stakes. One Part 1 decision and one decision from each of the eight Part 2 blocks were randomly selected for payment. Participants also received completion payments of \$2 and \$6, could earn up to \$1 for belief accuracy, and 185 early participants received a \$2 supplement. Referent choices did not affect the payoffs attached to the participant's chosen lottery.

For the belief bonus, one of the eight blocks was randomly selected. In a payoff-blind block, the pre-choice report was scored; in a payoff-visible block, the pre- or post-choice report was selected with equal probability. A correct report earned \$1 and an incorrect report earned \$0. For scoring, responses 1–2 denoted a more risk-taking referent, response 3 similarity, and responses 4–5 a less risk-taking referent. The correct category was "more" or "less" when the referent's Part 1 risk-taking index was at least 0.5 points above or below the subject's index, respectively, and "similar" otherwise.

Across the 226-person analysis sample, total payments averaged $21.98 (median $20.83), with a range from $2.81 to $64.64; the instructions described expected total payment as approximately $16. All payments were delivered through Tremendous.

Figure 13 shows the payment distribution. The positive observed minimum is not a guaranteed payment floor: a selected lottery could produce a loss of up to $2.

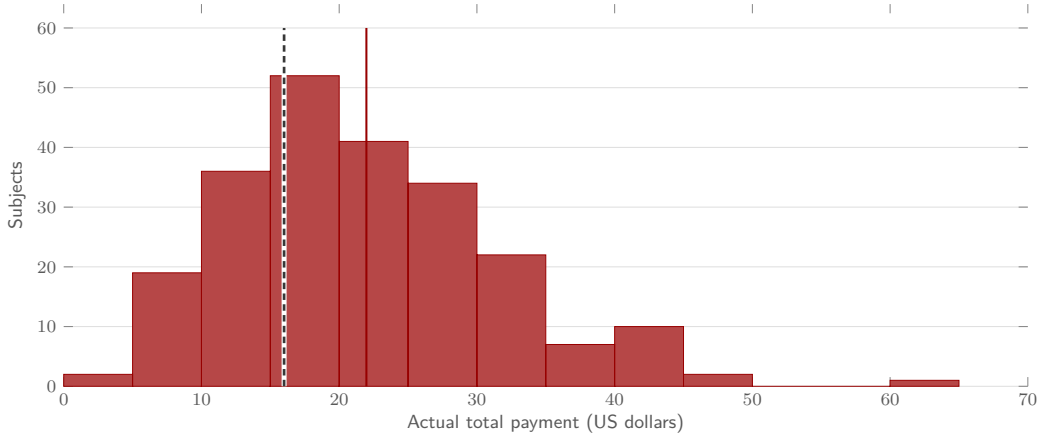


Figure 13: Actual total payments for the 226 completed participants. The gray dashed line marks the approximately $16 described in the instructions; the red line marks the actual mean of $21.98.

I.1 Choice Consistency and Sensitivity to Referents

Part 1 contains 20 distinct unordered lottery comparisons rather than repeated trials of the same comparison. Within-Part 1 test–retest reliability is therefore unavailable. Instead, I measure the minimum number of binary choices that must be reversed to make the observed comparisons consistent with a transitive strict ranking of all ten lotteries. Unobserved comparisons may be completed in whichever way requires the fewest reversals. This is a revealed-choice consistency measure, not a diagnosis of inattention or a structural error rate.

The mean is 1.08 reversals (SD 0.98). Of the 226 subjects, 31.9 percent require none and 29.2 percent require at least two. The best-fitting transitive ranking agrees with 94.6 percent of choices on average. Greater inconsistency predicts more changes both toward and away from referents, but not a detectable increase in exact copying or net shift (Table 22).

Table 22: Part 1 Inconsistency and Part 2 Choice Changes

Association per +1 SD inconsistency	Estimate (pp)	SE	p-value
Exact copying	0.41	0.85	0.631
Initial disagreement with referent	0.35	0.59	0.552
Switch toward referent, unconditional	1.98	0.58	< 0.001
Switch away from referent, unconditional	1.23	0.54	0.023
Switch toward referent, initial disagreement	3.97	1.08	< 0.001
Net shift	0.76	0.93	0.415

Notes: Models adjust for the treatment conditions, block position, and lottery pair, with participant-clustered standard errors. They use between-participant variation in the standardized inconsistency measure; participant fixed effects would absorb its level. All outcomes use 12,656 trials except switching toward the referent conditional on initial disagreement, which uses 5,758 trials. The measure is observed, not randomly assigned.

A separate participant-fixed-effects model allows treatment effects to vary with this measure. The same-gender-by-hidden-payoff copying interaction changes by -0.14 percentage points per additional SD of inconsistency (SE 1.81, $p = 0.937$; 95 percent CI $[-3.70, 3.42]$). Thus, the headline interaction is not detectably concentrated among less consistent subjects. The hidden-payoff same-gender effect itself is 7.81 points at one SD below mean inconsistency (SE 2.50) and 11.11 points at one SD above it (SE 2.00); this level comparison is different from moderation of the hidden-minus-visible interaction.

For net shift, the analogous interaction with inconsistency is -5.21 points (SE 2.69, $p = 0.054$), with Holm-adjusted $p = 0.272$ across six outcomes. Counting cycles among the 15 fully observed three-lottery comparisons gives a similar qualitative pattern of changes in both directions and no precise net-shift moderation ($p = 0.131$). The evidence is consistent with noisier baseline choices predicting more subsequent switching, without showing that they explain the primary copying result.

J Signal Usefulness and Payout Benchmarks

The experiment does not contain a payoff-blind condition without a referent signal. I therefore evaluate signal usefulness against explicit policy benchmarks rather than interpret the comparisons below as treatment effects of signal access. In payoff-blind trials, observed choices match the subject's own Part 1 choice 2.24 percentage points more often than a 50–50 random-choice policy. An always-copy policy yields alignment 5.12 percentage points above the same random-choice benchmark, or 2.88 points above observed choices. Neither policy produces a precisely estimated expected-dollar gain in the payoff-blind trials. Preference alignment and expected earnings need not move together because higher-risk lotteries in this choice set also tend to have higher expected values.

The realized-payout benchmark redraws the paid option within each of the eight Part 2 blocks under a 50–50 random policy. Observed Part 2 payouts were \$0.21 lower per participant than the simulated random-policy total, with a simulation interval from −\$1.43 to +\$0.99 and a two-sided Monte Carlo p-value of 0.735. This is a policy comparison, not a literal ex post payout loss: outcomes for unchosen lotteries were not realized.

Table 23: Observed Choices and Policy Benchmarks

Policy and outcome	Estimate	95% interval	p-value
Panel A: Payoff-blind trials relative to 50–50 random choice			
Observed: alignment with own Part 1 choice	2.24 pp	[0.93, 3.56] pp	0.001
Always copy: alignment with own Part 1 choice	5.12 pp	[3.67, 6.57] pp	< 0.001
Observed: expected dollars per paid trial	\$0.0126	[−\$0.0013, \$0.0266]	0.075
Always copy: expected dollars per paid trial	\$0.0160	[−\$0.0009, \$0.0330]	0.064
Panel B: Realized Part 2 payout relative to simulated random choice			
Observed minus random-policy total	−\$0.21	[−\$1.43, \$0.99]	0.735

Notes: Panel A uses 6,328 payoff-blind trials from 226 participants and participant-clustered inference. Alignment is agreement with the subject's Part 1 choice for the same lottery pair. Panel B uses all 1,808 paid Part 2 trials and 100,000 random-policy simulations. Its interval is the 2.5th to 97.5th percentile of observed-minus-simulated mean payouts.

I also ask how outcomes would differ if each subject instead received a random draw

from the 266 eligible Part 1 profiles. The exercise transports each subject's observed copying response to the counterfactual referent pool; it holds behavior fixed and is therefore not a causal estimate of how subjects would respond to random peers. The mean–SD value of option a in lottery pair t for subject i is

$$W_{it}(a) = \mathrm{EV}_t(a) - \widehat{\rho}_{i,-t}\,\mathrm{SD}_t(a),$$

where EV and SD are the lottery's expected payoff and standard deviation in dollars. The weight $\widehat{\rho}_{i,-t}$ comes from a hierarchical logit of Part 1 choices on EV and SD differences, with a participant-specific SD coefficient. It is the negative ratio of that coefficient, including its empirical-Bayes participant effect, to the common EV coefficient. The evaluated pair t is left out when fitting the model; pairs with identical EV and SD are excluded.

Table 24 reports gains relative to choosing each option with probability one-half. The transported policy raises alignment with Part 1 choices by 2.16 percentage points overall and mean–SD value by \$0.0174 per payoff-blind trial. Both gains are concentrated among lower- and middle-risk-taking subjects; estimates for higher-risk-taking subjects include zero. The counterfactual does not produce a positive expected-dollar gain.

Table 24: Transported Outcomes under Random Eligible Referents

Part 1 risk-taking group	Participants	Alignment gain	Mean–SD value gain
Overall	226	2.16 pp	\$0.0174
		[1.56, 2.95] pp	[\$0.0123, \$0.0237]
Lower risk-taking	95	3.02 pp	\$0.0330
Middle risk-taking	67	2.69 pp	\$0.0171
Higher risk-taking	64	0.30 pp	−\$0.0055

Notes: The counterfactual draws from 266 eligible Part 1 profiles and applies the subject's observed copying response. Gains are relative to 50–50 random choice; mean–SD gains subtract $[W_{it}(A) + W_{it}(B)]/2$ from expected policy value and average over payoff-blind trials. Overall intervals use a finite-pool two-way bootstrap. Higher-risk-group intervals include zero for both outcomes; lower- and middle-risk estimates are positive. Mean–SD value is a model-based preference metric, not realized earnings.

The random-profile benchmark also changes which subjects appear to benefit. For lower-, middle-, and higher-risk-taking subjects, the corresponding alignment gains under

the actual selected-extreme assignments are 1.17, 2.19, and 3.91 percentage points. The transported random-profile gains in Table 24 are 3.02, 2.69, and 0.30 points. Top-tail profiles comprise about 36 percent of the eligible random pool but about half of the experimental assignments. The comparison therefore concerns both the composition of information sources and transported behavior; it is not a randomized comparison of referent-selection policies.

J.1 Expected Earnings under Alternative Choice Policies

A different benchmark compares always copying the assigned referent with repeating the subject's own Part 1 choice. Summing expected lottery payments over the eight paid Part 2 blocks gives \$10.525 under always copying and \$10.225 under baseline repetition, a \$0.300 difference (participant-bootstrap 95 percent confidence interval \$0.052 to \$0.547). This comparison includes both information conditions and is not the hidden-information, random-choice benchmark above.

Table 25: Expected Lottery Payments under Alternative Choice Policies

Policy	Expected total (\$)	Payout SD (\$)
Repeat own Part 1 choice	10.225	8.458
Always copy the assigned referent	10.525	8.970
Observed Part 2 choices	10.366	8.848

Notes: Totals cover one paid choice in each of eight Part 2 blocks, excluding fixed completion payments and belief bonuses. Expected totals integrate over lottery outcomes and random selection of a paid trial within each block. The final column averages each participant's standard deviation of total lottery payments under the indicated policy. The always-copy minus baseline-repeat confidence interval uses 50,000 participant-bootstrap draws with seed 20260913.

Always copying increases both expected earnings and payout risk relative to baseline repetition. Expected dollars therefore cannot establish a welfare improvement for risk-averse subjects. The policies hold the assigned referents and available lottery pairs fixed; they are accounting counterfactuals, not randomized behavioral interventions. Nor does an always-copy comparison estimate the earnings effect of the experimentally induced increase in copying.

For scale, independent 50–50 option selection yields an expected Part 2 total of \$10.392. Always choosing the higher-expected-value option yields \$13.337, but requires information unavailable in hidden-payoff blocks. Neither benchmark incorporates the subject's valuation of risk. A second risk summary pools variation across participants as well as lottery outcomes: the standard deviation of this mixture of total payments is \$9.027 under always copying and \$8.835 under baseline repetition. These are different quantities from the mean within-participant standard deviations in Table 25.

K Debrief Evidence and Demand Awareness

The debrief included five mandatory open-response questions and eleven structured items. The open questions asked subjects to describe the study's purpose and whether some choices appeared more appropriate, fair, rational, or expected. The structured items asked whether researchers would view some choices more favorably, whether the information made some choices seem more socially appropriate, whether subjects tried to appear fair, unbiased, or rational, and whether they felt pressure to avoid appearing influenced by identity. All 1,130 open responses and 2,486 active structured responses from the 226 participants were read in full. Reproducible text rules were used only after that review to count conservative subsets. One pasted-template response is omitted from the primary qualitative coding, with a second unusually formatted response omitted in sensitivity checks; neither exclusion changes the 226-person behavioral sample.

Identity-related demand was salient. A conservative coding of the first, unprompted question identifies 115 subjects (50.9 percent) as describing identity or demographic differences as part of the study's purpose. In the structured questions, 127 (56.2 percent) agreed that the identity information made them think the researcher expected a response, 101 (44.7 percent) believed researchers would view some choices more favorably, and 45 (19.9 percent) felt pressure to appear unbiased. Forty-two subjects (18.6 percent) reported both changing choices based on what researchers wanted at least sometimes and being influenced at least somewhat when doing so.

The debrief also clarifies the scope of exact copying. Fourteen subjects who reported little desire to conform nevertheless described relying on the referent; their exact-copying rates were 57.1 percent with visible payoffs and 84.7 percent with hidden payoffs (224/392 and 332/392 trials, respectively). Ten subjects described a directional- adjustment or opposite-choice strategy. They copied on 80.0 percent of trials with hidden payoffs when they rated the referent as similar (84/105), but on 41.7 percent when they rated the referent as more or less risk-taking (73/175). These are small groups defined by post-choice responses, not randomized

subgroups or estimates of a mechanism. Thus, exact copying is one form of referent influence, but some non-copying choices can also reflect deliberate use of the signal.

Table 26: Demand Awareness and Behavioral Contrasts

Difference: demand-aware minus other subjects	Estimate	95% CI
Hidden-minus-visible exact copying	0.0 pp	[−5.7, 5.6] pp
Same-gender exact-copying contrast, hidden payoffs	0.7 pp	[−5.8, 7.3] pp
Extra same-nationality response with hidden payoffs	−4.8 pp	[−12.7, 3.2] pp

Notes: "Demand-aware" denotes the 127 subjects who agreed that the displayed nationality and gender information made them think the researcher expected a particular response. Contrasts are descriptive participant-level differences with HC2 intervals. The debrief was administered after the incentivized choices, so these comparisons do not identify causal effects of demand awareness. They show no concentration of the main payoff- visibility or gender patterns among demand-aware subjects; the nationality contrast is less precise and warrants greater caution.

L Experimental Protocol and Lottery Menu

L.1 Instructions and Comprehension

Subjects completed two asynchronous online sessions. The instructions explained that the referent had completed Part 1 earlier and that the displayed choice came from that person's decision with full lottery information. Only the current subject could have lottery details hidden. The instructions did not make a formal same-population sampling claim. The referent provided no advice or outcome feedback, and there was no two-way interaction. Thus, the displayed action was a record of an earlier choice, not a recommendation tailored to the subject.

Before proceeding, subjects had to answer the comprehension items correctly. Table 27 summarizes their content; answer order was randomized. All 226 completed participants eventually passed the referent-choice and blinding checks. First-attempt success was 196/226 (86.7 percent) and 204/226 (90.3 percent), respectively. Passing these checks establishes comprehension of the tested rules, not continuous attention on every subsequent trial.

Table 27: Comprehension Topics

Topic	Required understanding
Part 1	
Payment	One of the 20 main choices is randomly selected for payment, plus the completion bonus.
Probability	A 0.3 chance corresponds to 30 successes per 100 trials in expectation.
Expected value	Calculate the mean payoff of an example lottery.
Standard deviation	Identify the higher-dispersion lottery at equal expected value.
Riskiness	Identify the riskier lottery at equal expected value.
Part 2	
Payoff independence	The referent's action does not alter the subject's lottery payoffs or probabilities.
Referent changes	The instructions specify a different referent in each block.
Choice labels	The recorded A/B choice refers to the current on-screen labels.
Information asymmetry	The referent saw probabilities, outcomes, and EV/SD information even when the subject's lottery details are hidden.

L.2 Counterbalanced Block Sequences

Table 28 gives the four sequences of nationality match and payoff visibility. Each four-condition sequence is repeated in a rotated order. Within each nationality-by-visibility condition, subjects encounter one female and one male referent; consequently, each of the eight gender-by-nationality-by-visibility cells occurs once per completed participant. These are repeated conditions within the same people, not mutually exclusive participant groups.

Table 28: Nationality-Match and Visibility Sequences

Sequence	Block 1	2	3	4	5	6	7	8
1	SH	SV	DH	DV	SV	DV	SH	DH
2	SV	DV	SH	DH	DV	DH	SV	SH
3	DH	SH	DV	SV	SH	SV	DH	DV
4	DV	DH	SV	SH	DH	SH	DV	SV

Notes: S/D denotes same/different nationality; H/V denotes hidden/visible payoffs. The rows are numbered 1–4 here; the implementation codes them 0–3 in the same order. Referent gender is balanced separately within each nationality-by-visibility condition.

Two dominance trials are inserted in Part 2, with lottery details visible and the referent's choice hidden. In one, both lotteries pay −\$1 with probability 0.50; the remaining outcomes are \$1 and \$6 under A versus \$0.50 and \$4 under B, with probabilities 0.30 and 0.20. In the other, both pay −\$2 with probability 0.40, while the remaining outcomes are \$2 and \$4.40 under A versus \$1.50 and \$3.40 under B, with probabilities 0.35 and 0.25. A is never worse and is sometimes better. These checks are separate from comprehension and excluded from the 56 primary social-information choices.

L.3 Lottery Profiles and Comparisons

Table 29 reports all ten profiles. Profiles 1–2, 3–4, 5–6, 7–8, and 9–10 are constructed to have equal expected values and standard deviations up to rounding, but different distributions of gains and losses. These matched profiles allow comparisons of behavior beyond mean and variance; they do not by themselves identify a loss-aversion parameter. Only 3–4 and 5–6 are compared directly in the 20 Part 1 trials; other comparisons involve a common alternative.

Table 29: Lottery Profiles

Profile	Payoffs in dollars (probabilities)	EV ($)	SD ($)
1	0.60 (.79); −0.40 (.20); −2.00 (.01)	0.374	0.465
2	0.8490 (.395); 0.3105 (.395); −0.40 (.21)	0.374	0.465
3	2.50 (.56); −0.60 (.38); −2.00 (.06)	1.052	1.664
4	3.4071 (.28); 1.2929 (.28); −0.60 (.44)	1.052	1.664
5	3.80 (.55); −1.00 (.40); −2.00 (.05)	1.590	2.452
6	4.6865 (.275); 2.7317 (.275); −1.00 (.45)	1.590	2.452
7	6.3750 (.40); −1.30 (.50); −2.00 (.10)	1.700	3.822
8	7.8668 (.20); 4.5332 (.20); −1.30 (.60)	1.700	3.822
9	12.00 (.21); −0.80 (.78); −2.00 (.01)	1.876	5.221
10	13.1790 (.105); 10.7067 (.105); −0.80 (.79)	1.876	5.221

Notes: EV is expected value and SD is standard deviation. Entries reproduce the implemented payoff distributions; reported moments are rounded to three decimals. Higher dispersion generally comes with higher expected value, so choices reflect risk-return tradeoffs as well as sensitivity to the distribution's shape.

The first ten Part 1 comparisons are 1–4, 1–6, 1–7, 1–9, 4–6, 4–7, 4–9, 6–7, 6–9, and 7–9. The remaining ten are 2–4, 2–7, 3–4, 3–7, 5–6, 5–7, 6–8, 8–9, 4–10, and 6–10. Part 2 uses seven comparisons among profiles 1, 4, 6, 7, and 9: 1–4, 4–6, 6–7, 7–9, 1–7, 4–9, and 1–9. This subset provides pronounced risk-return contrasts within a seven-choice referent block.

L.4 Dispersion in Baseline Risk Taking

The Part 1 risk-taking index averages normalized chosen standard deviations over 20 private choices. Its mean is 0.456, its standard deviation is 0.218, and its interquartile range is 0.30–0.60. Only six of the 226 participants are at either endpoint of the 0–1 index. Table 30 shows the spread rather than treating the sample as two homogeneous risk types.

Table 30: Distribution of the Part 1 Risk-Taking Index

Index range	Participants	Share (%)
0.00–0.20	34	15.0
> 0.20–0.40	78	34.5
> 0.40–0.60	69	30.5
> 0.60–0.80	28	12.4
> 0.80–1.00	17	7.5

L.5 Numeracy and Lottery Summary Statistics

Within payoff-visible trials, the experiment also varies whether expected value and standard deviation are displayed as numerical summaries. Probabilities and possible payoffs remain visible in both conditions. This is a test of access to summary statistics, not the hidden-versus-visible treatment used to identify the informational component of identity influence.

Numeracy is first-submission success on four Part 1 concepts: probability, expected value, standard deviation, and riskiness (the last two are assessed jointly). Low, middle, and high numeracy denote 0–1, 2, and 3–4 correct concepts. Raw rates in Table 31 show slightly more copying and switching toward the referent with the summaries in every group, not more switching without summaries among the least numerate subjects.

Table 31: Visible-Payoff Responses by Numeracy and Summary Availability

	Exact copying (%)		Switch toward (%)	
Numeracy	No summary	EV/SD shown	No summary	EV/SD shown
Low	58.9	60.4	42.3	46.3
Middle	59.8	61.8	38.8	41.4
High	62.2	62.9	42.2	44.8

Notes: Raw means. The low-, middle-, and high-numeracy groups contribute 756, 1,470, and 938 trials per summary condition, respectively. Switching toward the referent is measured only among trials with initial disagreement, so its denominators are smaller. Both conditions show the underlying lottery details.

In participant-fixed-effects models with block-position, lottery-pair, and identity controls, displaying summaries changes copying by 1.31 percentage points (SE 1.55, $p = 0.401$) and switching toward the referent by 2.15 points (SE 1.79, $p = 0.229$). Interacting summary availability with standardized numeracy mistakes gives 0.26 points for copying (SE 1.36, $p = 0.848$) and -0.96 for switching toward the referent (SE 1.58, $p = 0.544$). These estimates do not detect a numeracy gradient in responses to the summaries.

M Additional Field Evidence and Policy Context

The women-in-economics application in Section 1 is one setting in which the identity of an information source may matter. Other field studies establish the stakes of information exposure, experience, and social influence without supplying a common estimate of their relative contributions. The comparisons below preserve each study's intervention and outcome rather than extrapolate the laboratory copying coefficient to unrelated decisions.

M.1 Neighborhood Choice and Experience

Aliprantis et al. (2024) document large differences in neighborhood socioeconomic status that household income and wealth do not explain. For Black households in the sixth income decile, living in a Black rather than non-Black neighborhood is associated with neighborhood poverty of 22.0 rather than 13.5 percent and unemployment of 11.5 rather than 7.0 percent. Their adjusted neighborhood-SES comparison is approximately 33 percentile points. These are sorting patterns, not estimates of what changing an information source would accomplish. The study's homophily interpretation leaves open how experience and information contribute to the preferences expressed in location choices.

Darrah and DeLuca (2014) provide complementary qualitative evidence from 110 Baltimore Mobility Program participants. The program combined counseling with exposure to lower-poverty neighborhoods. Among the suburban-county movers interviewed, 60 percent described changes in how they assessed residential options; another 26 percent evaluated county neighborhoods and schools more favorably without describing such specific changes. This is evidence about evolving evaluations after counseling and experience, not an experimental decomposition of peer learning, identity, and preferences. It concerns the Baltimore Mobility Program, not the separate Moving to Opportunity experiment.

M.2 Learning, Social Information, and Persistence

In Bangladesh, Bryan et al. (2014) report migration rates of 35.9 percent with information alone, 36.0 percent in the control group, and 58.0 percent when cash or credit accompanied the information. One year later, the originally incentivized group migrated 9.2 percentage points more than the nonincentivized group without another incentive. The original experiment found essentially no induced migration among friends and relatives. The result therefore supports learning through one's own subsidized experience, not a claim that similar peers copied migration decisions.

Table 32 collects three studies that manipulate social information more directly. Their outcomes differ from exact copying in the laboratory: financial-asset purchase, water consumption, and continued employment. Their identifying contrasts also differ. Only the investment design separates learning from the additional effect of a peer's possession of the asset.

Table 32: Social-Information Interventions in the Field

Study	Intervention and result	Interpretation
Bursztyn et al. (2014)	Asset purchase: 42% without peer information, 71% with learning alone, and 93% with learning plus peer possession.	Rounded rates imply a 29-point learning contrast and a further 22-point possession contrast. The ratio $29/51 \approx 57\%$ is specific to this design, not a transferable information share.
Ferraro and Price (2013)	Water use: advice plus an appeal reduces use by about 990 gallons per household; adding a peer comparison reduces it by about 1,740 gallons, relative to control.	The added comparison contributes about 750 gallons of reduction. Information, descriptive norms, and social pressure are not separately identified.
Coffman et al. (2017)	Teach For America offers: adding a sentence about prior acceptance increases second-year teaching by 3.2 percentage points (SE 1.2; $N = 6{,}685$), from a control mean of 58.0%.	58.0% plus the adjusted effect is 61.2%, not a reported raw treated mean. The persistence of the effect does not separately identify learning and conformity.

The water comparison can be scaled arithmetically: applying 750 gallons to 106,872

households gives approximately 80.2 million gallons. That is a mechanical illustration, not an observed county-wide treatment total. Likewise, the 57 percent investment ratio is not a parameter to apply to major choice or neighborhood sorting. The relevant policy lesson is to test what an intervention changes: access to facts, the perceived relevance of a source, social incentives, or several of these together.

For women considering economics, the role-model evidence already discussed in Section 1 is the closest application. The additional field cases motivate measurement of information and source relevance, but do not justify replacing that focused application with a numerical projection of field effects from this experiment.