Social Bee Lab

DISCUSSION PAPER SERIES

SBL-DP-00003

When Peer Endorsements Hurt: Source Attribution in Talent Recognition

Manuel Munoz

July 206

When Peer Endorsements Hurt: Source Attribution in Talent Recognition

Manuel Munoz[*] [†]

2026.07.27

Abstract

Peer recognition systems promise to reveal talent that formal evaluators may miss, but peer input is useful only if later evaluators trust the process behind it. I study this credibility problem in an incentivized field study embedded in the final stage of a talent recognition program. The 665 finalists were peer endorsed, authority endorsed, or unendorsed. Evaluators assigned selection probabilities across real finalist categories in two payoff relevant environments: collaboration, where higher partner performance increased payoffs, and tournament, where lower opponent performance increased payoffs. Peer endorsed candidates were avoided as partners but targeted as opponents: relative to unendorsed candidates, they received 10.1 percentage points lower selection probability in collaboration and 6.3 percentage points higher selection probability in tournament. Authority endorsements moved in the opposite direction. Peer endorsed finalists were not observably weaker on GPA or task performance. Open text justifications indicate that the penalty is concentrated among evaluators who attributed peer endorsement to social motives rather than merit. A preregistered complementary experiment shows that disclosing comparable performance partially reduces the penalty. Peer input can reveal talent, but organizations must make its merit basis credible.

Keywords: peer recognition; endorsements; source attribution; talent allocation; organizational design

JEL Classification: C93; D82; D83; D91; M51

[*]New York University Abu Dhabi, Abu Dhabi, United Arab Emirates, and Universidad Autonoma de Bucaramanga, Bucaramanga, Colombia (e-mail: manumunoz@nyu.edu).

[†]I am grateful to Michelle Acampora, Emilio Castilla, Moritz Janas, Mario Molina, and Pedro Rey-Biel for helpful comments. I thank Jhon Diaz from UNAB for invaluable institutional support. The main study reported in this paper obtained ethical approval at LISER and was preregistered at the AEA RCT Registry under number AEARCTR-0015842. The follow-up online experiment obtained ethical approval from NYU Abu Dhabi and was preregistered at the AEA RCT Registry under number AEARCTR-0018263. I used Google's Gemini for light copy editing.

1 Introduction

Organizations increasingly rely on peers to identify talent. Peer recognition programs and awards, lateral referrals, shared leadership systems, and collaborative evaluation tools are designed to surface information that formal authorities may miss: effort in daily work, reliability, helpfulness, and informal contributions to collective outcomes (Gallus 2017; Pearce and Conger 2003; Ensley et al. 2006). The promise of these systems is informational, but that promise depends on whether later evaluators trust the peer signal. A peer endorsement is not only a statement about the candidate; it is also a signal about the process that produced the endorsement. Evaluators may interpret peer input as informed recognition of merit, but they may also infer friendship, reciprocity, favoritism, popularity, or lack of objectivity. This inference is a problem of attribution, because evaluators form beliefs about the causes and motives behind the signal they observe (Kelley 1973; Martinko et al. 2007). It is also a problem of legitimacy, because organizational signals influence evaluation only when they are viewed as appropriate, credible, and tied to accepted criteria of judgment (Suchman 1995; Bitektine 2011).

I study when peer endorsements become credible organizational signals. In standard informational accounts, endorsements reduce uncertainty because an informed insider vouches for a candidate's quality (Spence 1973; Connelly et al. 2011). Prior work shows that referrals, recommendations, and application endorsements shape access to jobs, resources, and elite opportunities because they transmit information, confer credibility, or activate social capital (Fernandez et al. 2000; Fernandez and Fernandez-Mateo 2006; Burt 1992; Castilla and Rissing 2019; Rivera 2012). I argue that this logic depends on source attribution. Vertical endorsements from authorities may be interpreted as institutionally accountable and based on merit. Peer endorsements may be more ambiguous because they are lateral and socially embedded (Burt 1992; McPherson et al. 2001). When evaluators attribute peer endorsement to social motives rather than screening based on merit, peer endorsement can reduce rather than increase demand for the endorsed candidate.

I test this argument in a field study embedded in the final stage of a talent recognition program at a Colombian university. The setting is not an employment relationship, but it is a real talent recognition process in which source labels affected consequential eval-

uation decisions. All eligible students received an invitation to apply. Some invitations informed the student that they had been endorsed by a faculty member or by a peer; the remaining invitations contained no endorsement information. I refer to the three source labels as faculty endorsed, peer endorsed, and unendorsed. The final stage sample includes 665 finalists: 103 faculty endorsed finalists, 99 peer endorsed finalists, and 463 unendorsed finalists. The unendorsed category is the benchmark. These finalists were eligible, invited, and had progressed through the same institutional process, but they did not carry a faculty or peer endorsement label.

I place these real source labels in two environments with opposing incentives. Each finalist first completed an individual task measuring real effort. Finalists then assigned selection probabilities across anonymous but real finalist categories. For clarity, I refer to the first environment as *collaboration*: participants benefited from being matched with a higher performing partner, although the two individuals did not actively work together. I refer to the second as *tournament*: participants benefited from being matched against a lower performing opponent. Evaluators did not choose named individuals and did not observe potential partners' or opponents' GPA, prior scores, or task performance. They observed only the source labels. The probability allocations were then implemented by a matching algorithm that paired participants with actual finalists from the selected category.[1]

The reversal from collaboration to tournament is the key empirical feature. If a source label is interpreted as a positive signal of expected performance, evaluators should assign it more probability in collaboration and less probability in tournament. If it is interpreted as a negative signal, evaluators should avoid it as a partner but target it as an opponent. This reversal helps distinguish an interpretation based on expected performance from simple dislike of the label. A general dislike of peer endorsed candidates would predict avoidance in both environments. A belief that peer endorsement predicts lower performance predicts avoidance in collaboration and targeting in tournament.

The field study does not randomize endorsement status and therefore does not esti-

[1] The design is related to incentivized resume rating, in which respondents evaluate candidate profiles and are then matched with real candidates on the basis of their evaluations (Kessler et al. 2019). It is also related to incentive aligned stated preference and conjoint designs in which stated choices are made behaviorally meaningful through real consequences (Ding et al. 2005; Ding 2007; Hainmueller et al. 2014, 2015). Unlike standard resume rating or conjoint designs, the object of interest here is the source label itself.

mate the causal effect of peer endorsement on candidate ability. Instead, it estimates how evaluators interpret source labels when making incentivized choices among finalist categories. This distinction is central. The source labels shaped who reached the final stage. I use the field study to ask how finalists evaluated those labels once candidates had reached the same final stage pool and individual performance information was not available at the moment of selection.

I find three main results in the field study. First, peer endorsed candidates were avoided in collaboration and targeted in tournament. Participants allocated 25.2% of the collaboration selection probability to peer endorsed candidates, compared with 35.2% to unendorsed candidates and 39.6% to faculty endorsed candidates. Peer endorsed candidates received 10.06 percentage points less than unendorsed candidates as partners. In tournament, the pattern reversed: participants allocated 38.8% to peer endorsed opponents, compared with 32.6% to unendorsed opponents and 28.6% to faculty endorsed opponents. Peer endorsed candidates received 6.29 percentage points more than unendorsed candidates as opponents. The pooled estimate implies a 16.35 percentage point reversal in the peer versus unendorsed gap. Faculty endorsements moved in the opposite direction: faculty endorsed candidates were favored as partners and avoided as opponents.

Second, the peer penalty is not supported by observed performance differences among finalists. Peer endorsed finalists were not observably weaker at the final stage, and equality tests do not reject equal performance across source labels in the individual, collaboration, or tournament tasks. These comparisons are descriptive because endorsement status was not randomly assigned and progression to the final stage was selective. They do show, however, that peer endorsed finalists were not visibly lower performers on the academic and task performance measures observed in the study.

Third, open text justifications provide evidence consistent with source attribution. The strongest peer penalty appears among evaluators who explicitly attributed peer endorsement to friendship, favoritism, reciprocity, bias, personal ties, lack of objectivity, or other motives unrelated to merit. In collaboration, these evaluators assigned peer endorsed candidates an additional 11.49 percentage points less than unendorsed candidates. In tournament, the same attribution was associated with an additional 18.34 percentage points assigned to peer endorsed opponents. This evidence indicates that

the behavioral penalty is concentrated among evaluators whose own explanations articulate the source attribution concern.

I also report a preregistered information experiment with 400 online participants. Participants were assigned to either a control condition that reproduced the baseline information structure or a performance information condition in which they were told that observed objective performance was statistically similar across the three finalist categories. Performance information reduced the collaboration peer penalty by 6.02 percentage points and reduced tournament targeting by 6.17 percentage points. These effects offset 41.4% of the control condition collaboration peer penalty and 51.4% of the control condition tournament targeting premium. The treatment did not eliminate the penalty entirely, but it shows that transparent performance information can partially reduce the attributional penalty attached to peer endorsement.

The paper makes three contributions. First, it identifies a boundary condition for endorsement and referral systems. Prior work shows that referrals, nominations, and recommendations shape access to jobs, educational opportunities, and elite organizational settings because they transmit private information, confer credibility, or activate social capital (Fernandez et al. 2000; Burt 1992; Castilla and Rissing 2019; Rivera 2012). I show that this logic depends on the credibility and legitimacy of the process that produced the source label (Suchman 1995; Bitektine 2011). Peer endorsements may contain useful information, but they can become negative signals when evaluators infer that the endorsement reflects social motives rather than merit. The contribution is not simply that endorsements sometimes fail to help. It is that the same organizational signal can change sign depending on how evaluators attribute the source.

Second, the paper contributes to research on meritocracy and internal labor markets. Prior work shows that formal commitments to meritocracy, evaluation systems, and network based hiring can reproduce inequality through the judgments of organizational gatekeepers (Castilla and Benard 2010; Rivera 2012; Fernandez et al. 2000; Castilla and Rissing 2019). This literature has often focused on centralized evaluators, formal selection criteria, and the ways in which social background or network position shape access to valued opportunities. I identify a different mechanism that can arise when organizations decentralize talent recognition: lateral input may be treated as less legitimate than faculty input even when peer endorsed candidates are not observably

weaker. Peer recognition can expand the information available to the organization, but only if downstream evaluators view the peer signal as evidence of merit.

Third, the paper contributes to organizational design by showing that peer recognition systems require legitimating infrastructure. Research on shared leadership, peer input, and social capital emphasizes that peers may observe effort, reliability, helping behavior, and collaborative contributions that formal authorities miss (Pearce and Conger 2003; Ensley et al. 2006; Burt 1992; McPherson et al. 2001). Research on organizational legitimacy further suggests that signals and practices are more influential when audiences view them as appropriate, credible, and consistent with accepted evaluative criteria (Suchman 1995; Bitektine 2011). The results show that collecting peer information is not enough: organizations must also make the merit basis of peer input visible to later evaluators. Peer endorsements are likely to be more effective when paired with clear criteria, structured evidence, or performance benchmarks that make the source interpretable as evidence of merit.

The rest of the paper proceeds as follows. Section 2 develops the theoretical framework and hypotheses. Section 3 describes the empirical context, protocol, and estimation strategy. Section 4 presents the field study results. Section 5 reports the information experiment. Section 6 concludes. The online appendix reports the experimental materials (Appendix A), the formal model (Appendix B), additional analyses (Appendix C), open text coding (Appendix D), and preregistration details (Appendix E).

2 Theory and Hypotheses

The central theoretical claim is that endorsements are source dependent signals. An endorsement does not only indicate that someone supports a candidate. It also invites evaluators to infer why that support was given. When evaluators attribute an endorsement to credible screening based on merit, the label should increase expected performance. When they attribute it to friendship, reciprocity, favoritism, personal ties, popularity, or lack of objectivity, the same label can become a negative signal.

This argument builds on a mechanism based view of endorsements in selection. Prior work distinguishes whether endorsement advantages arise because decision makers value endorsements differently, because endorsements signal candidate quality, or be-

cause some candidates are more likely to obtain endorsements in the first place (Castilla 2022). I use this distinction to clarify the scope of the paper. I do not study who obtains endorsements or whether receiving an endorsement changes candidate quality. Endorsement status in the field study was not randomly assigned, and it shaped who reached the final stage. The theoretical focus is instead on downstream interpretation: once source labels are visible at a common decision stage, do evaluators treat them as credible evidence of merit or as evidence of the social process that produced the endorsement?

This focus requires separating three explanations for the peer penalty. First, evaluators may avoid peer endorsed candidates because those candidates are objectively weaker. This is a quality signal explanation. Second, evaluators may simply dislike peer endorsed candidates or dislike the peer label. This is a label valuation explanation. Third, evaluators may treat peer endorsement as a negative signal of expected performance because they attribute the source to social motives rather than screening based on merit. This is the source attribution explanation. The empirical design is built to distinguish these possibilities. Observed GPA and task performance speak to the quality signal explanation. The collaboration and tournament environments distinguish label aversion from expected performance beliefs. Open text responses and the information experiment speak to the attribution process and its malleability.

Faculty and peer endorsements may also differ in the type of information evaluators expect them to contain. Peers may observe daily effort, reliability, helpfulness, motivation, and informal collaboration that faculty do not observe. Faculty, by contrast, may be perceived as better positioned to compare academic performance, assess relative ability, and make accountable judgments tied to institutional criteria. The theory therefore does not assume that one source is always more informative than the other. The sign of the endorsement depends on whether evaluators interpret the source as credible evidence of merit or as reflecting motives unrelated to merit

Appendix B formalizes the source attribution logic in a simple belief updating model. The model clarifies that endorsements are comparative. An endorsement is valuable only when evaluators view the source as more informative about merit than the unendorsed benchmark. A noisy endorsement should be discounted, but not necessarily penalized. A penalty arises only when evaluators view the source as less credible or less

tied to merit than the benchmark.

In the empirical setting, the unendorsed category is the benchmark because those finalists were eligible, invited, and reached the same final stage without carrying a faculty or peer endorsement label. Ex ante, peer endorsement can plausibly move choices in either direction. If evaluators believe that peers observe useful information about effort, reliability, or collaborative behavior, peer endorsement should increase demand in collaboration and reduce targeting in tournament. If evaluators instead suspect friendship, reciprocity, favoritism, personal ties, popularity, or lack of objectivity, peer endorsement should reduce demand in collaboration and increase targeting in tournament. The prediction is that evaluators may behave as if the peer label lowers expected performance when they question the credibility of the peer endorsement process.

The collaboration and tournament environments translate these beliefs into opposite behavioral predictions. In collaboration, higher expected performance makes a source more attractive because the evaluator benefits from a stronger partner. In tournament, higher expected performance makes a source less attractive because the evaluator benefits from a weaker opponent. A source treated as a positive signal should therefore receive higher selection probability in collaboration and lower selection probability in tournament. A source treated as a negative signal should receive lower selection probability in collaboration and higher selection probability in tournament.

Hypothesis 1 (Merit attribution and endorsement value) *If evaluators attribute an endorsement source to credible screening based on merit, then candidates from that source should be treated as stronger expected performers than unendorsed candidates. They should receive higher selection probability than unendorsed candidates in collaboration and lower selection probability than unendorsed candidates in tournament.*

Hypothesis 2 (Social attribution and endorsement penalty) *If evaluators attribute an endorsement source to friendship, reciprocity, favoritism, bias, personal ties, lack of objectivity, or other motives unrelated to merit, then candidates from that source should be treated as weaker expected performers than unendorsed candidates. They should receive lower selection probability than unendorsed candidates in collaboration and higher selection probability than unendorsed candidates in tournament.*

Hypotheses 1 and 2 are source generic. They can apply to any endorsement source depending on how evaluators interpret the process that produced the label. In this paper, faculty endorsement serves as a positive signal contrast, while peer endorsement is the focal case for the social attribution penalty. The open text justifications provide supporting evidence by identifying whether evaluators explicitly describe peer endorsement as reflecting social motives rather than merit.

The source attribution account also implies that the penalty should be partly malleable. If evaluators penalize a source because they are uncertain whether it reflects merit, then credible information showing similar observed performance across source labels should weaken the negative inference. The prediction is attenuation rather than full elimination. Performance information may make the merit basis of peer endorsement more visible, but it may not fully override concerns about motives, objectivity, or the social structure behind the endorsement.

Hypothesis 3 (Performance information and penalty mitigation) *If an endorsement source is penalized because evaluators are uncertain whether it reflects merit, then disclosing equivalent objective performance across source categories should reduce the penalty. Candidates from the penalized source should be avoided less as partners in collaboration and targeted less as opponents in tournament.*

In the empirical application, I test Hypothesis 3 for peer endorsement. Performance information should make the merit basis of peer endorsement more credible, reducing both the collaboration peer penalty and the tournament peer targeting premium. Residual gaps would be consistent with the idea that source attributions are not fully overwritten by objective performance information.

3 Methods

This section describes the field setting, sample, protocol, incentive structure, and estimation strategy. I embedded the study in the final stage of a real talent recognition program, where finalists made choices with payoff consequences over three source labels: faculty endorsed, peer endorsed, and unendorsed. Endorsement status was not randomly assigned and shaped progression into the final stage. The field study there-

fore estimates how finalists interpreted source labels under a controlled information structure.

3.1 Setting, Sample, and Estimand

The empirical setting is a university talent recognition program for high performing students. The university first defined an eligible pool using an academic performance threshold: having GPA above the university median. All eligible students received an invitation to apply. For some students, the invitation informed them that they had been endorsed by a faculty member or by a peer. The remaining invitations contained no endorsement information. I refer to these three source labels as faculty endorsed, peer endorsed, and unendorsed.

A total of 3,186 eligible students were invited into the application process: 198 with a faculty endorsement, 285 with a peer endorsement, and 2,703 with no endorsement information. The application process unfolded over three stages. Candidates accumulated points as they completed program activities, and the institution considered the candidates with the highest cumulative scores for a formal award. The award had both professional and monetary value: an institutional certificate designed to enhance recipients' curricula vitae and a monetary prize equivalent to approximately two thirds of a local monthly salary. I embedded the field study in the third and final stage. Appendix Table C-1 reports participant flow by source label.

Endorsement status was strongly associated with progression through the application process. Unendorsed invitees had a final stage completion rate of 17.1 percent. Faculty endorsed invitees were 34.9 percentage points more likely to complete the process and reach the final stage sample ($SE = 3.6$, $p < 0.001$), implying a completion rate of approximately 52.0 percent. Peer endorsed invitees were 17.6 percentage points more likely to complete the process ($SE = 2.9$, $p < 0.001$), implying a completion rate of approximately 34.7 percent. These estimates come from descriptive linear probability models of final stage completion on source label indicators, with unendorsed invitation as the omitted category and robust standard errors. Appendix Table C-1 reports the full participant flow by source label. This flow is why I interpret the field study as evidence on source label evaluation among finalists rather than as evidence on the causal effect

of endorsement status on participation or candidate quality.

The field study sample consists of the 665 applicants who reached the final stage of the program: 463 unendorsed finalists, 103 faculty endorsed finalists, and 99 peer endorsed finalists. All finalists had passed through the same institutional process, had accumulated points toward the same award, and remained eligible to earn additional points in the final stage. Comparisons of GPA and task performance by source label are descriptive. I use them to assess whether peer endorsed finalists were observably weaker on the measures available in the study.

Two institutional features reduced strategic concerns in the endorsement phase. First, awards were allocated within source specific tracks, and peer endorsers were not ranked against the specific students they endorsed. A peer endorsement therefore did not place the endorsed student in direct competition with the peer who endorsed them. Second, each endorser, whether peer or faculty, could endorse only one eligible student, making the endorsement scarce. These features made endorsements meaningful source labels while preserving the fact that endorsement status itself was not randomly assigned.

Throughout the paper, I distinguish between *endorsers*, who generated the faculty and peer endorsement labels before the field study, and *evaluators*, who are the finalists making the selection decisions during the field study. Endorsement status was determined before the collaboration and tournament choices and affected progression into the final stage; the collaboration and tournament choices were made later by finalists evaluating anonymous source labels. Conditional on reaching the final stage, the question is how evaluators use those labels when they affect payoffs.

3.2 Information Environment and Protocol

Appendix A reproduces the final stage instructions. Participants first received information about the program and the three source labels. They were told that all invitees had a cumulative GPA of at least 4.0 (i.e., above the university median), that some invitees had been endorsed by faculty, some had been endorsed by peers, and others had received no endorsement information. They were also told that award competition occurred within source specific tracks.

Before the matching decisions, participants answered brief comprehension questions

and preregistered belief questions about the source labels. These questions verified understanding of the program rules and did not reveal individual scores, GPA, cumulative program points, or performance information by source label. Appendix C.5 reports the belief measures and their relationship with the selection probabilities.

Participants then completed a real effort task across three environments: an individual baseline, a collaboration environment, and a tournament environment. The task displayed 4×4 matrices filled with zeros and ones; participants counted the number of ones under time pressure (Abeler et al. 2011). The task is useful for this setting because performance is objectively scored, requires attention and effort, and does not depend on specialized academic knowledge.

The sequence was as follows. Participants first performed the individual baseline task for one minute and earned one point for each correct answer. They then completed the collaboration and tournament environments in randomized order. In each environment, participants first assigned selection probabilities across the three source labels and then performed the one minute task for that environment. Immediately after each matching decision and task, participants explained the rationale for their probability assignment in an open text response.

The information available at the moment of selection was intentionally limited. Evaluators did not choose named individuals. They did not observe potential partners' or opponents' GPA, baseline score, cumulative program points, or task score. They observed only the three source labels: faculty endorsed finalist, peer endorsed finalist, and unendorsed finalist. Finalists' GPA and task scores are observed by the researcher, but they were not shown to evaluators before the matching decisions.

3.3 Incentivized Source Category Choices

The matching task is an incentivized choice over source categories. The categories were not hypothetical labels created only for the study. They corresponded to actual source labels in the finalist pool, and the probability allocations were implemented by matching evaluators with real finalists from the selected category. This logic parallels incentivized resume rating designs, in which respondents evaluate candidate descriptions and are then matched with real candidates on the basis of their evaluations (Kessler et al. 2019).

Here, the design is simpler and more targeted: evaluators choose among source labels.

In each matching environment, participant i assigned probabilities p_{iF}, p_{iP}, and p_{iU} to the faculty endorsed, peer endorsed, and unendorsed finalist categories, with $p_{iF} + p_{iP} + p_{iU} = 100$. These probabilities determined the source category from which the participant's partner or opponent would be drawn. For payoff implementation, the computer drew one source category according to the participant's stated probability distribution. Conditional on the drawn category, the computer selected a finalist from that category at random, and the selected finalist's realized score in the corresponding task was used to determine participant i's score for that environment. The selected finalist's payoff was not affected by being selected as someone else's partner or opponent.

In the collaboration environment, participant i earned points based on the average output of the dyad multiplied by a premium:

$$\pi_i^C = 1.5 \times \frac{M_i + M_j}{2}, \tag{1}$$

where M_i is participant i's score and M_j is the selected partner's score. Because M_j was not observed before the matching decision, the payoff maximizing strategy was to assign higher probability to the source category believed to have higher expected performance.

In the tournament environment, participant i competed against the selected opponent. If participant i solved more matrices than the opponent, i earned two points per correct answer. If i solved fewer matrices, i earned half a point per correct answer. In case of a tie, the winner was chosen randomly. Equivalently, the expected payoff conditional on the two scores is

$$\pi_i^T = \begin{cases} 2M_i, & \text{if } M_i > M_j, \\ 1.25M_i, & \text{if } M_i = M_j, \\ 0.5M_i, & \text{if } M_i < M_j. \end{cases} \tag{2}$$

Thus, for any given own score, the payoff maximizing strategy was to assign higher probability to the source category believed to have lower expected performance.

Participants evaluated all three categories simultaneously and assigned probabilities from 0 to 100 across the categories. This joint evaluation format required them to

trade off the relative value of the source labels rather than judge each label in isolation (Bohnet 2016). The design therefore identifies relative demand for source labels under opposite incentives.

3.4 Empirical Specification

The main outcome is the selection probability assigned by participant i to source s in environment e, denoted p_{is}^{e}, where $s \in \{F, P, U\}$ indexes faculty endorsed, peer endorsed, and unendorsed finalists, and $e \in \{C, T\}$ indexes the collaboration and tournament environments. The omitted source category is U, the unendorsed finalist type. For each environment, I estimate

$$p_{is}^{e} = \alpha_{i}^{e} + \beta_{P}^{e}\mathbf{1}\{s = P\} + \beta_{F}^{e}\mathbf{1}\{s = F\} + \varepsilon_{is}^{e}. \tag{3}$$

The unit of observation is a participant by source cell. Participant fixed effects, α_{i}^{e}, absorb the fact that each participant allocates exactly 100 percentage points across the three source categories within an environment. Standard errors are clustered at the participant level. In collaboration, positive coefficients indicate greater demand for a source as a partner relative to unendorsed finalists. In tournament, positive coefficients indicate greater targeting of a source as an opponent relative to unendorsed finalists. Because probabilities sum to 100 within participant and environment, the coefficients describe relative source label demand with unendorsed finalists as the benchmark.

To test the reversal directly, I stack the collaboration and tournament source cells and estimate

$$p_{ise} = \alpha_{ie} + \beta_{P}\mathbf{1}\{s = P\} + \beta_{F}\mathbf{1}\{s = F\} + \lambda_{P}\left(\mathbf{1}\{s = P\} \times T_{e}\right) + \lambda_{F}\left(\mathbf{1}\{s = F\} \times T_{e}\right) + \varepsilon_{ise}, \tag{4}$$

where T_{e} equals one for tournament and zero for collaboration. The specification includes participant by environment fixed effects. The interaction coefficient λ_{P} tests whether the peer versus unendorsed gap reverses across environments. A negative peer coefficient in collaboration and a positive peer coefficient in tournament indicate that peer endorsed finalists are avoided as partners but targeted as opponents.

I also report participant level peer gaps:

$$G_{iP}^{C} = p_{iU}^{C} - p_{iP}^{C}, \qquad G_{iP}^{T} = p_{iP}^{T} - p_{iU}^{T}. \tag{5}$$

Positive values of both gaps indicate a penalty against peer endorsed finalists: lower demand in collaboration and greater targeting in tournament. These gap outcomes summarize the peer penalty and correspond directly to the theoretical predictions in Section 2.

The field study was preregistered before final stage data collection. The preregistration specified the main selection probability outcomes, the source categories, and the two incentive environments. Appendix E maps the preregistered outcomes to the manuscript analyses.

3.5 Open Text Measures of Source Attribution

I use the open text justifications to measure the source attributions evaluators articulated when explaining their selection choices. Participants provided one justification after the collaboration allocation and one after the tournament allocation. I classify responses into theory driven categories that capture whether each source is interpreted as conveying merit information or as reflecting social motives. Merit information codes indicate that the respondent described the corresponding source label as informative about ability, effort, performance, reliability, motivation, objectivity, responsibility, or other criteria tied to merit. These codes are defined separately for peer endorsement, faculty endorsement, and the unendorsed category. Social motive codes indicate that the respondent described an endorsement source as reflecting friendship, favoritism, reciprocity, bias, personal ties, popularity, lack of objectivity, or other motives unrelated to merit. These codes are defined for peer endorsement and faculty endorsement. I also code whether the response gives a fairness or diversification rationale, and whether it lacks an interpretable source based rationale.

The coding task is a text as data measurement problem: mapping short natural language explanations into a small set of theory driven attribution categories. This approach follows the broader literature that treats text as a source of economic data (see e.g., Gentzkow et al. 2019; Ash and Hansen 2023), as well as organizational research

using language to measure culture, cognition, and adaptation inside organizations (Srivastava et al. 2018). Recent work also argues that large language models can be useful measurement tools for text analysis when the task is well defined, the coding rule is transparent, and validation or robustness checks are reported (Ash et al. 2024; Ludwig et al. 2025). This is especially relevant for open ended survey responses, which are increasingly used to measure motives, beliefs, mental models, narratives, attention, and information transmission (Haaland et al. 2025).

I use the large language model as a blind, rule based annotator rather than as a discovery device. For each response, the model received only the task context and the respondent's open text justification. The prompt did not include the respondent's selection probabilities, task performance, GPA, source label, evaluator source label, or any outcome variable. Responses were coded in their original language. Appendix D reports the codebook, prompt, response tabulations, and coding details. I use the codes to measure articulated source attributions associated with selection behavior.

4 Results

The field study results are organized around the main behavioral estimand: the selection probability evaluators assigned to each source category when choosing partners and opponents. I first report the source label choices. I then examine whether observed performance differences support a simple alternative explanation. Finally, I use the open text justifications to characterize the source attributions evaluators articulated when explaining their choices.

4.1 Source Labels and the Collaboration–Tournament Reversal

I first test whether the same source label is avoided when evaluators seek a partner but targeted when evaluators seek an opponent. In collaboration, assigning a higher selection probability to a source category indicates greater demand for finalists from that category as partners. In tournament, assigning a higher selection probability indicates greater willingness to face finalists from that category as opponents. Thus, if a source label receives lower probability in collaboration but higher probability in tournament,

evaluators behave as if that label predicts lower expected performance.

Participants used the probability allocation task as an intensive measure rather than as a forced choice. Appendix Table C-5 shows that more than 93% of participants assigned positive probability to all three source categories in each environment, and only about 4% assigned all probability to a single category.

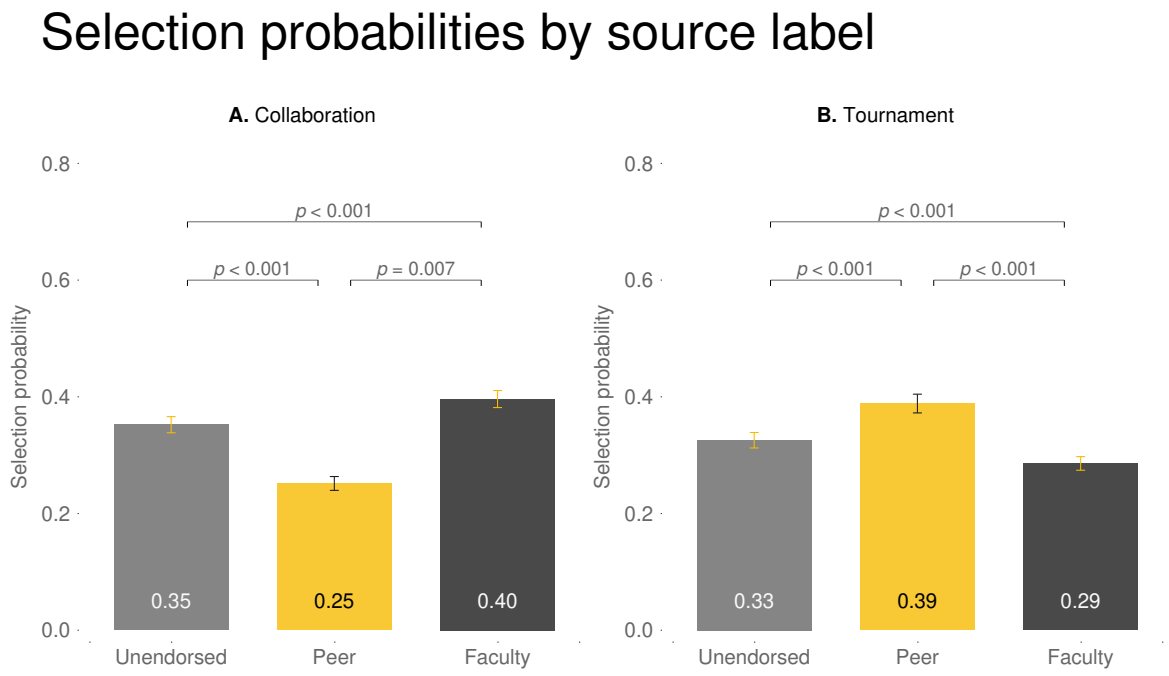

Figure 1 Selection Probabilities by Source Label

The figure reports mean selection probabilities assigned to each source category in the collaboration and tournament environments. Source labels refer to the finalist's endorsement status: unendorsed, peer endorsed, or faculty endorsed. Error bars report 95 percent confidence intervals.

Figure 1 illustrates the main pattern. In collaboration, peer endorsed finalists receive the lowest selection probability: 25.2%, compared with 35.2% for unendorsed finalists and 39.6% for faculty endorsed finalists. In tournament, the ordering reverses. Peer endorsed finalists receive the highest opponent selection probability: 38.8%, compared with 32.6% for unendorsed finalists and 28.6% for faculty endorsed finalists.

Table 1 reports the corresponding regression estimates. Relative to unendorsed finalists, peer endorsed finalists receive 10.06 percentage points less in collaboration and 6.29 percentage points more in tournament. The pooled specification tests the reversal directly: the peer endorsed by tournament interaction is 16.35 percentage points. This pattern is difficult to explain as simple dislike of peer endorsed finalists. Dislike would predict avoidance in both environments. Instead, evaluators behaved as if the

Table 1 Selection Probabilities by Source Label and Environment

This table reports regression estimates for the selection probabilities assigned to each source category. The unit of observation is a participant–source cell. Columns I and II estimate the collaboration (COLLAB) and tournament (TOURN) environments separately and include participant fixed effects. Column III stacks both environments (POOLED) and includes participant by environment fixed effects. The omitted source category is unendorsed finalists. Coefficients are reported in percentage points. Standard errors, clustered at the participant level, are reported in parentheses. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	I COLLAB	II TOURN	III POOLED
Peer endorsed	-10.06^{***}	6.29^{***}	-10.06^{***}
	(1.33)	(1.69)	(1.33)
Faculty endorsed	4.40^{***}	-3.97^{***}	4.40^{***}
	(1.62)	(1.20)	(1.62)
Peer endorsed × Tournament			16.35^{***}
			(2.11)
Faculty endorsed × Tournament			-8.37^{***}
			(1.90)
Observations	1,995	1,995	3,990
Participants	665	665	665
R^2	0.104	0.052	0.078
Participant FE	Yes	Yes	–
Participant × environment FE	–	–	Yes
p-value: Peer = Faculty	0.000	0.000	
p-value: Peer reversal = Faculty reversal			0.000

peer endorsement label predicted lower expected task performance: peer endorsed finalists were avoided when high partner performance was valuable and targeted when low opponent performance was valuable.

Faculty endorsement moves in the opposite direction. Faculty endorsed finalists receive 4.40 percentage points more than unendorsed finalists in collaboration and 3.97 percentage points less in tournament. The contrast between peer and faculty endorsements shows that the result is not a general endorsement effect. It is source specific.

Table 2 expresses the result using gaps computed at the participant level. The collaboration peer penalty is 10.06 percentage points, the tournament peer targeting premium is 6.29 percentage points, and the total peer reversal is 16.35 percentage points. The corresponding faculty gaps are negative in both environments, and the peer and faculty reversals differ by 24.72 percentage points.

Result 1 *Peer endorsed finalists are avoided as partners and targeted as opponents.*

Table 2 Participant Level Source Gaps and Reversals

This table reports participant level source gaps. Columns I, II, and III report the collaboration (COLLAB), tournament (TOURN), and reversal (REV.) gaps, respectively. In COLLAB, the peer gap is $p^C_{iU} - p^C_{iP}$ and the faculty gap is $p^C_{iU} - p^C_{iF}$. In TOURN, the peer gap is $p^T_{iP} - p^T_{iU}$ and the faculty gap is $p^T_{iF} - p^T_{iU}$. Positive values indicate that the source category is treated as lower performing than unendorsed finalists. The REV. column is the sum of the COLLAB and TOURN gaps. Robust standard errors are reported in parentheses. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	I	II	III
	COLLAB	TOURN	REV.
Peer gap	10.06^{***}	6.29^{***}	16.35^{***}
	(1.08)	(1.38)	(1.72)
Faculty gap	-4.40^{***}	-3.97^{***}	-8.37^{***}
	(1.32)	(0.98)	(1.55)
Peer gap − Faculty gap	14.46^{***}	10.26^{***}	24.72^{***}
	(1.16)	(1.26)	(1.86)
Observations	665	665	665

Appendix Tables C-3 and C-4 show that the peer reversal is similar across the randomized order of the two environments and positive across evaluator source labels. Appendix C.5 reports the preregistered belief outcomes. Evaluators expected peer endorsed finalists to have lower participation and lower earlier stage performance, but controlling for these beliefs does not eliminate the peer endorsement pattern: peer endorsed finalists remain less likely to be selected in collaboration and more likely to be selected in tournament.

4.2 Observed Performance as an Alternative Explanation

The selection results show that evaluators behaved as if peer endorsed finalists had lower expected task performance. A natural alternative is that peer endorsed finalists were in fact weaker performers in the realized finalist pool. I examine this possibility descriptively, because endorsement status was not randomly assigned and progression to the final stage was selective.

The observed performance evidence does not support this explanation. Appendix Table C-8 shows that peer endorsed finalists had the highest mean GPA in the finalist pool: 4.43, compared with 4.33 among unendorsed finalists and 4.33 among faculty endorsed finalists.[2] The objective task measures also do not show a consistent peer

[2] Colombian university GPAs are typically reported on a 0.0–5.0 scale, with 5.0 as the maximum grade and 3.3 as the usual minimum passing grade.

endorsed deficit. Equality tests do not reject equal performance across source labels in the individual baseline task, the collaboration task, or the tournament task. Appendix Table C-9 reports the corresponding regression diagnostics and the standardized task performance index.

Result 2 *Peer endorsed finalists are not observably lower performing.*

These comparisons do not imply that endorsement status is as good as randomly assigned, nor do they rule out all unobserved differences across finalist types. They show that peer endorsed finalists were not observably lower performing on the academic and task measures collected in the study. Moreover, evaluators did not observe a potential partner's or opponent's GPA, individual baseline score, cumulative program points, or realized task score when assigning probabilities. The selection results therefore capture how evaluators interpreted source labels under the information available at the time of choice.

4.3 Open Text Evidence on Source Attribution

I next examine whether the peer penalty is associated with the source attributions evaluators articulated in their open text justifications. Participants wrote one justification after the collaboration allocation and one after the tournament allocation. The analysis focuses on whether respondents attributed peer endorsement to social motives: friendship, favoritism, reciprocity, bias, personal ties, lack of objectivity, or other motives unrelated to merit.

These attributions were common enough to matter. Peer social motives appear in 19.7% of collaboration justifications and 24.2% of tournament justifications. By contrast, faculty endorsements were rarely described as socially motivated: 1.5% in collaboration and 0.3% in tournament. Appendix Table C-11 reports all source attribution categories.

Table 3 reports participant level peer gap regressions. Using the collaboration peer penalty, $p^C_{iU} - p^C_{iP}$, as the outcome, respondents who did not mention peer social motives assigned peer endorsed finalists 7.80 percentage points less than unendorsed finalists. Respondents who did mention peer social motives showed an additional 11.49 percent-

age point penalty, implying a total collaboration peer penalty of 19.29 percentage points for this group.

The tournament pattern is even sharper. Among respondents who did not mention peer social motives, the tournament peer targeting premium, $p^T_{iP} - p^T_{iU}$, is small and statistically indistinguishable from zero. Among respondents who did mention peer social motives, the premium increases by 18.34 percentage points, implying a tournament peer targeting premium of 20.19 percentage points.

Table 3 Peer Social Motive Attributions and Selection Gaps

This table reports participant level regressions linking open text source attributions to peer selection gaps. Columns I, II, and III report the collaboration peer penalty (COLLAB), the tournament peer targeting premium (TOURN), and the peer reversal (REV.), respectively. The COLLAB outcome is $p^C_{iU} - p^C_{iP}$, the TOURN outcome is $p^T_{iP} - p^T_{iU}$, and the REV. outcome is $(p^C_{iU} - p^C_{iP}) + (p^T_{iP} - p^T_{iU})$. The collaboration peer social motive indicator equals one when the collaboration justification attributed peer endorsement to friendship, favoritism, reciprocity, bias, personal ties, lack of objectivity, or other motives unrelated to merit. The tournament peer social motive indicator is defined analogously using the tournament justification. Positive coefficients indicate a larger peer penalty. Robust standard errors are reported in parentheses. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	I	II	III
	COLLAB	TOURN	REV.
Peer social motives			
Collaboration	11.49^{***}		9.22^{**}
	(2.34)		(4.67)
Tournament		18.34^{***}	23.00^{***}
		(3.56)	(4.71)
Constant	7.80^{***}	1.85	8.97^{***}
	(1.24)	(1.44)	(1.86)
Observations	665	665	665
R^2	0.027	0.049	0.069

The combined peer reversal shows the same pattern. The reversal is positive even among respondents who did not articulate peer social motive concerns, but it is substantially larger among respondents who did. Mentioning peer social motives in the collaboration justification is associated with a 9.22 percentage point larger peer reversal, and mentioning peer social motives in the tournament justification is associated with a 23.00 percentage point larger peer reversal.

The association is source specific. Appendix Table C-12 reports target cell regressions that interact the peer social motive indicator with the target source category. In collaboration, peer social motive attributions are associated with an additional 11.49 percentage point penalty for peer endorsed finalists, while the corresponding shift for

faculty endorsed finalists is positive. In tournament, peer social motive attributions are associated with an additional 18.34 percentage points assigned to peer endorsed opponents, while the corresponding faculty endorsed interaction is small and statistically insignificant. Appendix Table C-13 shows that the same source specific pattern holds in participant level peer minus faculty gaps and remains similar after controlling for environment order and evaluator source label.

Conversely, Appendix Table C-14 shows that respondents who describe peer endorsement as merit relevant do not show the collaboration peer penalty: the peer versus unendorsed gap changes from -13.23 percentage points among respondents who do not mention peer merit information to $+3.73$ percentage points among those who do. These patterns are consistent with the claim that peer endorsement becomes costly when evaluators attribute it to social motives rather than screening based on merit.

Result 3 *The peer penalty is concentrated among evaluators who attribute peer endorsement to social motives.*

5 Study 2: Performance Information and Peer Signal Credibility

The field study shows that peer endorsed finalists are avoided in collaboration and targeted in tournament. Study 2 tests a design implication of the source attribution account. If evaluators penalize peer endorsement partly because they are uncertain whether it reflects merit, then credible performance information should reduce the peer penalty.

I ran a preregistered information experiment with 400 Prolific participants. Participants evaluated the same three source labels used in the field study: peer endorsed, faculty endorsed, and unendorsed finalists. The CTRL group received the baseline information structure. The INFO group received the same information plus a statement that observed objective performance was statistically similar across the three source labels. Participants then assigned selection probabilities across the source categories in the same collaboration and tournament environments used in the field study. Appendix

Table C-19 reports sample construction and randomization balance.

The estimands mirror the field study. In collaboration, the peer gap is the selection probability assigned to unendorsed finalists minus the selection probability assigned to peer endorsed finalists. In tournament, the peer gap is the selection probability assigned to peer endorsed opponents minus the selection probability assigned to unendorsed opponents. Larger values indicate a larger peer penalty. Hypothesis 3 predicts that performance information should reduce both gaps.

5.1 Performance Information Reduces the Peer Penalty

The CTRL condition reproduces the peer penalty. Participants in the CTRL assign peer endorsed finalists 14.55 percentage points less selection probability than unendorsed finalists in collaboration. In tournament, the pattern reverses: CTRL participants assign peer endorsed finalists 12.00 percentage points more selection probability than unendorsed finalists.

Figure 2 shows that performance information in INFO reduces both peer gaps. In collaboration, the peer penalty falls from 14.55 to 8.53 percentage points, a reduction of 6.02 points. In tournament, the peer targeting premium falls from 12.00 to 5.83 percentage points, a reduction of 6.17 points. The combined peer reversal falls by 12.19 percentage points, from 26.55 to 14.36.

Table 4 reports the treatment effects. INFO reduces the collaboration peer penalty by 6.02 percentage points and tournament peer targeting by 6.17 percentage points. The combined peer reversal falls by 12.19 percentage points. Appendix Table C-20 reports the corresponding target cell regressions.

The INFO treatment does not eliminate the peer penalty. Residual gaps remain in both environments, consistent with the idea that a short performance disclosure does not fully overwrite source attributions. The attenuation is nevertheless informative: the peer endorsement penalty is partly malleable when evaluators see evidence that the peer source tracks merit.

Appendix Tables C-21–C-23 report source specificity checks, order controls, log ratio robustness, and allocation use diagnostics. Performance information has no compara-

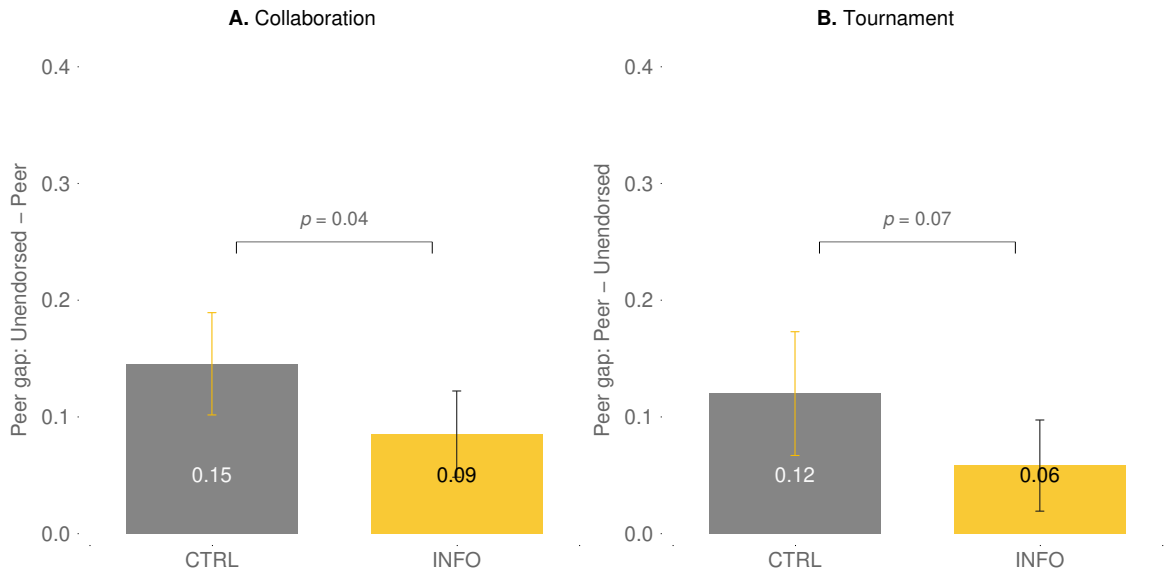

Figure 2 Performance Information and Peer Gap Mitigation

The figure reports the two peer gaps used in Study 2. Panel A reports the collaboration peer penalty, defined as the selection probability assigned to unendorsed finalists minus the selection probability assigned to peer endorsed finalists. Panel B reports the tournament peer targeting premium, defined as the selection probability assigned to peer endorsed opponents minus the selection probability assigned to unendorsed opponents. Larger values indicate a larger penalty against peer endorsed finalists. Error bars report 95 percent confidence intervals.

ble effect on faculty endorsed finalists, reduces the peer minus faculty reversal, and produces similar estimates when controlling for the randomized order of the two environments. The INFO treatment also increases near equal allocations; I interpret this as part of the behavioral response to performance information.

Result 4 *Performance information partially mitigates the peer endorsement penalty.*

Table 4 Performance Information and Peer Gap Mitigation

This table reports participant level treatment effects from Study 2. Columns I, II, and III report the collaboration peer penalty (COLLAB), the tournament peer targeting premium (TOURN), and the peer reversal (REV.), respectively. The COLLAB outcome is the selection probability assigned to unendorsed finalists minus the selection probability assigned to peer endorsed finalists. The TOURN outcome is the selection probability assigned to peer endorsed opponents minus the selection probability assigned to unendorsed opponents. The REV. outcome is the sum of the two gaps. The treatment effect is the coefficient on the performance information condition (INFO) relative to the control condition (CTRL). Larger values indicate a larger penalty against peer endorsed finalists. The offset rows are scaled so that positive values indicate mitigation of the peer gap. Robust standard errors are reported in parentheses. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	I	II	III
	COLLAB	TOURN	REV.
Performance INFO	-6.02^{**}	-6.17^{*}	-12.19^{**}
	(2.92)	(3.36)	(4.81)
CTRL mean	14.55^{***}	12.00^{***}	26.55^{***}
	(2.24)	(2.70)	(3.61)
INFO mean	8.53^{***}	5.83^{***}	14.36^{***}
	(1.88)	(1.99)	(3.19)
Treatment offset	6.02^{**}	6.17^{*}	12.19^{**}
	(2.92)	(3.36)	(4.81)
Offset of CTRL gap (%)	41.4^{***}	51.4^{**}	45.9^{***}
	(15.8)	(19.9)	(14.1)
Observations	400	400	400
R^2	0.010	0.008	0.016

6 Discussion and Conclusion

Organizations often decentralize talent recognition to capture information that formal evaluators miss. This paper shows that collecting peer input is only the first design problem. The second is making peer input credible to the evaluators who later use it. In the field study, peer endorsed finalists were not observably lower performing, yet evaluators avoided them as partners in collaboration and targeted them as opponents in tournament. The result identifies a boundary condition for peer recognition: lateral signals can reveal information and still be penalized when evaluators question the process that produced them.

The collaboration to tournament reversal is central to this conclusion. If evaluators simply disliked peer endorsed finalists, they should have avoided them in both environments. Instead, they behaved as if the peer endorsement label predicted lower expected task performance. Faculty endorsement moved in the opposite direction, which reinforces the source attribution interpretation: the result is not about endorsement in

general, but about how evaluators interpret the source of the endorsement.

The open text evidence points to the attribution behind this pattern. The peer penalty is concentrated among evaluators whose own explanations attribute peer endorsement to social motives unrelated to merit. These justifications were elicited after choices, so I interpret them as articulated attributions associated with selection behavior. They are nevertheless theoretically informative: evaluators do not only observe whether a candidate has been endorsed; they infer what the endorsement says about the process that produced it.

Study 2 provides the design implication. Performance information reduced, but did not eliminate, the peer penalty. This matters because the theory does not imply that a short disclosure should fully overwrite source attributions. Rather, it implies that the penalty should be partly malleable when the organization makes the merit basis of the peer signal visible. The evidence is consistent with that prediction: peer endorsement becomes less costly when evaluators are told that observed performance is similar across source labels.

The main theoretical implication is that endorsements are source dependent signals. Prior research on referrals, recommendations, and nominations emphasizes their informational value: insiders observe candidate quality and transmit that information to evaluators (Fernandez et al. 2000; Castilla and Rissing 2019; Rivera 2012). I show that this logic can reverse when evaluators attribute the source to social motives. A peer endorsement is not only a claim about candidate quality; it is also evidence about a social process. When that process is interpreted as friendship, reciprocity, favoritism, or lack of objectivity, the endorsement can become a liability.

The paper also contributes to research on meritocracy and internal labor markets. Organizations often decentralize evaluation to reduce dependence on formal gatekeepers and to capture information dispersed across social networks. The results identify a risk of decentralization: lateral input may be treated as less legitimate than vertical input even when laterally endorsed candidates are not observably weaker. Peer recognition can expand what the organization knows, but the value of that knowledge depends on whether later evaluators view the peer signal as credible evidence of merit.

6.1 Implications for Organizational Design

The results do not imply that organizations should avoid peer recognition systems. They imply that such systems require a translation layer. Peer input may originate in decentralized observation, but later evaluators need to understand why that input should be interpreted as evidence of merit.

A first design principle is to make the criteria behind peer input visible. A label such as "peer endorsed" may be intended to communicate broad recognition, but evaluators may instead infer social ties. Peer endorsements should therefore be paired with information about the behavior recognized, the threshold met, the rubric used, or the evidence supporting the endorsement.

A second principle is to distinguish using peer input as a search device from displaying peer endorsement as an evaluative label. Peer input may be valuable upstream because it surfaces candidates formal evaluators might miss. Downstream, however, the peer label may need to be contextualized or replaced with standardized evidence before it is used in later evaluation.

A third principle is to align peer input with the dimensions peers are positioned to observe. Peers may be especially informative about reliability, helping, responsiveness, and team contribution. If later evaluators are choosing on individual task performance, organizations need to explain why peer input is relevant to that criterion. Without such alignment, peer endorsement may be interpreted as social recognition rather than evidence of performance.

6.2 Limitations and Future Research

Several scope conditions are important. The field setting is a university talent recognition program rather than an employment relationship involving promotions, wages, or long term careers. The task measures short run individual performance under time pressure rather than complex interdependent work. The source labels are anonymous, whereas workplace endorsements often come from identifiable peers whose status, expertise, and relationship to the candidate may shape credibility. Future research should examine whether the same source attribution penalty appears in hiring, promotion,

internal mobility, peer awards, and team staffing decisions.

These limitations also point to a broader research agenda on how organizations translate decentralized information into credible signals. Future studies could test design features such as anonymizing endorsement sources, adding structured rubrics, requiring behavioral evidence, disclosing endorsement criteria, combining peer and supervisor input, or showing objective performance benchmarks. Such designs would help distinguish when peer input is discounted because it is socially attributed and when it is valued because it reveals information that formal channels miss.

The broader lesson is that decentralized recognition does not become credible simply because it is decentralized. Peer endorsements can reveal talent, but they can also raise doubts about motives. In this setting, those doubts made peer endorsed finalists less attractive as partners and more attractive as opponents. Performance information reduced the penalty but did not eliminate it. Effective peer recognition systems therefore require more than collecting endorsements. They require making the merit basis of those endorsements visible, credible, and usable by the evaluators who make later decisions.

References

Abeler, J., Falk, A., Goette, L., and Huffman, D. (2011). Reference points and effort provision. The American Economic Review, 101(2):470–492.

Ash, E. and Hansen, S. (2023). Text algorithms in economics. Annual Review of Economics, 15:659–688.

Ash, E., Hansen, S., and Muvdi, Y. (2024). Large language models in economics. Technical Report 19479, CEPR Discussion Paper.

Bermejo, V. J., Gago, A., Gálvez, R. H., and Harari, N. (2025). LLMs outperform outsourced human coders on complex textual analysis. Scientific Reports, 15:40122.

Bitektine, A. (2011). Toward a theory of social judgments of organizations: The case of legitimacy, reputation, and status. Academy of Management Review, 36(1):151–179.

Bohnet, I. (2016). What works: Gender equality by design. Harvard University Press.

Burt, R. S. (1992). Structural holes: The social structure of competition. Harvard University Press.

Castilla, E. J. (2022). Gender, race, and network advantage in organizations. Organization Science, 33:2364–2403.

Castilla, E. J. and Benard, S. (2010). The paradox of meritocracy in organizations. Administrative Science Quarterly, 55(4):543–576.

Castilla, E. J. and Rissing, B. A. (2019). Best in class: The returns on application endorsements in higher education. Administrative Science Quarterly, 64(1):230–270.

Connelly, B. L., Certo, S. T., Ireland, R. D., and Reutzel, C. R. (2011). Signaling theory: A review and assessment. Journal of Management, 37(1):39–67.

Ding, M. (2007). An incentive-aligned mechanism for conjoint analysis. Journal of Marketing Research, 44(2):214–223.

Ding, M., Grewal, R., and Liechty, J. (2005). Incentive-aligned conjoint analysis. Journal of Marketing Research, 42(1):67–82.

Ensley, M. D., Hmieleski, K. M., and Pearce, C. L. (2006). The importance of vertical and shared leadership within new venture top management teams: Implications for the performance of startups. The Leadership Quarterly, 17(3):217–231.

Fernandez, R. M., Castilla, E. J., and Moore, P. (2000). Social capital at work: Networks and employment at a phone center. American Journal of Sociology, 105(5):1288–1356.

Fernandez, R. M. and Fernandez-Mateo, I. (2006). Networks, race, and hiring. American Sociological Review, 71(1):42–71.

Gallus, J. (2017). Fostering public good contributions with symbolic awards: A large-scale natural field experiment at wikipedia. Management Science, 63(12):3999–4015.

Gentzkow, M., Kelly, B., and Taddy, M. (2019). Text as data. Journal of Economic Literature, 57(3):535–574.

Gilardi, F., Alizadeh, M., and Kubli, M. (2023). ChatGPT outperforms crowd workers for text-annotation tasks. Proceedings of the National Academy of Sciences, 120(30):e2305016120.

Haaland, I., Roth, C., Stantcheva, S., and Wohlfart, J. (2025). Understanding economic behavior using open-ended survey data. Journal of Economic Literature, 63(4):1244–1280.

Hainmueller, J., Hangartner, D., and Yamamoto, T. (2015). Validating vignette and conjoint survey experiments against real-world behavior. Proceedings of the National Academy of Sciences of the United States of America, 112(8):2395–2400.

Hainmueller, J., Hopkins, D. J., and Yamamoto, T. (2014). Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments. Political Analysis, 22(1):1–30.

Kelley, H. H. (1973). The processes of causal attribution. American Psychologist, 28(2):107–128.

Kessler, J. B., Low, C., and Sullivan, C. D. (2019). Incentivized resume rating: Eliciting employer preferences without deception. American Economic Review, 109(11):3713–3744.

Ludwig, J., Mullainathan, S., and Rambachan, A. (2025). Large language models: An applied econometric framework. Working Paper 33344, National Bureau of Economic Research.

Martinko, M. J., Harvey, P., and Douglas, S. C. (2007). The role, function, and contribution of attribution theory to leadership: A review. The Leadership Quarterly, 18(6):561–585.

McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. Annual Review of Sociology, 27(1):415–444.

Pearce, C. L. and Conger, J. A. (2003). Shared leadership: Reframing the hows and whys of leadership. Sage Publications.

Rivera, L. A. (2012). Hiring as cultural matching: The case of elite professional service firms. American Sociological Review, 77(6):999–1022.

Spence, M. (1973). Job market signaling. The Quarterly Journal of Economics, 87(3):355–374.

Srivastava, S. B., Goldberg, A., Manian, V. G., and Potts, C. (2018). Enculturation trajectories: Language, cultural adaptation, and individual outcomes in organizations. Management Science, 64(3):1348–1364.

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. Academy of Management Review, 20(3):571–610.

Online Appendix:

When Peer Endorsements Hurt: Source Attribution in Talent Recognition

Manuel Munoz

A Experimental Materials and Instructions

A.1 Study 1 Invitation Email

Hello {{name}},

You are about to finish your application to the [Award Name]. Here we invite you to complete the third and final step of your application to win this award {{phrase}}.

[Image Omitted]

Only students who complete the three steps of the application process will be considered. In addition to your GPA, each step gives you a score and we will rank the students with the highest scores to choose the winners.

The award winners will receive a certificate and one million pesos each, and will be announced on [Date]. Remember that there will be several winners for the award, and in your case you only compete with those students who, like you, have been {{source}}.

Finish your application by clicking on the following personal link (do not share it): {{link}}

We hope you participate and we can celebrate your excellence.

Sincerely,

[Signature of University Organizer]

A.2 Study 1 Survey and Task Instructions

WELCOME

This is the third and final step of your application as a candidate for the [Award Name] for University Excellence 2025.

This step consists of 5 parts.

- Part A: Questions about the award (2 minutes)
- Part B: Timed task 1 (2 minutes)
- Part C: Timed task 2 (2 minutes)
- Part D: Timed task 3 (2 minutes)
- Part E: Questions about participation (1 minute)

The entire process takes approximately 9 minutes. All the questions and tasks you complete give you points in the application ranking. Answer them as best as possible.

[Page break]

PARTE A

The people invited to apply to the [Award Name] for University Excellence 2025 are students with a cumulative GPA equal to or greater than 4.0.

All were selected through one of 3 different means:

1. Recommended by professors: They are students recommended by university professors.
2. Directly invited: They are students invited directly (without recommendation).
3. Recommended by peers: They are students recommended by other students who are also applying to the award.

*Remember that students compete for spots only among people in their same group. You do not compete with someone who was invited in a different way than you.

Next, we will ask you to indicate your opinion about the different groups of invited students.

[Page break]

Did you know: The participation rate in the process, meaning the number of students who are applying to the award after being invited, is different among the three groups of students.

Indicate in what order you consider each group is placed: 1, 2, or 3?

The first must be the group with the highest percentage of students who have applied and the third must be the group with the lowest percentage of students who have applied.

- Recommended by professors (1, 2, 3)
- Recommended by peers (1, 2, 3)
- Directly invited (1, 2, 3)

[Page break]

Applicants from all groups complete several tests (tasks, questions, etc.) throughout the 3 steps of the application process.

The tests are the same for everyone and provide a score used to determine the applicants' ranking. Out of 100% of the total points that can be obtained in the 3 application steps, how many points do you think applicants from each group obtain on average?

- Points of those recommended by professors: [slider 0-100]
- Points of those recommended by peers: [slider 0-100]
- Points of those directly invited: [slider 0-100]

[Page break]

DESCRIPTION PARTS B, C & D

In each of the following three parts (Part B, Part C, and Part D) you will have 60 seconds (1 minute), which you can use to solve a series of problems. You will see tables like the one below, containing ones (1) and zeros (0). For each table, you must sum the number of ones. For example, the correct answer for the table below is 9 because the number 1 appears 9 times.

1	0	1	1
0	1	0	0
1	1	0	0
1	1	0	1

[Page break]

PART B

In Part B you will perform the task alone. You have one minute and will obtain one point for each correct answer. Get ready! Your minute will start as soon as you move to the next screen.

[Page break]

[Participants then see up to 30 matrices sequentially like the one abive]

[Page break]

PART C

In Part C you will perform the same task, but this time you will be paired to work as a team and collaborate with another applicant who can come from any of the three groups:

1. Student recommended by a professor
2. Student directly invited
3. Student recommended by a peer

On the next screen, you will have the possibility to indicate with what probability (chance) you will be paired to collaborate with someone from each group.

[Page break]

Score: In this case, we will sum the score of each one (yours and your collaborator's), divide it by two (that is, average it) and multiply it by 1.5 points. For example, if you solve 10 tables correctly and your collaborator solves 8, the average score is 9 which will be multiplied by 1.5 points. In this case, you will earn $9 \times 1.5 = 13.5$ points for this task.

This match affects only your score in this part of the task. Your collaborator will not gain or lose points and their standing in the award will not be affected by being selected by you.

Choice for Part C Please, indicate with what probability you want to be paired to collaborate with someone from the different groups. The total must sum to 100%:

- Student recommended by a peer: [Input %]
- Student directly invited: [Input %]
- Student recommended by a professor: [Input %]
- Total: [Sum %]

[Page break]

[Participants then complete the 60-second real-effort task]

[Page break]

Why did you choose to collaborate with the following probabilities? [Open text response]

[Page break]

PART D

In Part D you will perform the same task, but this time you will be paired to compete with another applicant who can come from any of the three groups:

1. Student recommended by a professor
2. Student directly invited
3. Student recommended by a peer

On the next screen, you will have the possibility to indicate with what probability (chance) you will be paired to compete with someone from each group.

[Page break]

Score: If in the pair, you are the person with the highest number of correct answers you will be the winner and receive 2 points for each correct answer. If not, you receive half a point per correct answer. In case of a tie, the winner is chosen randomly.

This match affects only your score in this part of the task. Your pair will not gain or lose points and their standing in the award will not be affected by being selected by you.

Choice for Part D Please, indicate with what probability you want to be paired to compete with someone from the different groups. The total must sum to 100%:

- Student recommended by a peer: [Input %]
- Student directly invited: [Input %]
- Student recommended by a professor: [Input %]
- Total: [Sum %]

[Page break]

[Participants then complete the 60-second real-effort task]

[Page break]

Why did you choose to compete with the following probabilities? [Open text response]

[Page break]

PART E

This is the last part of the whole application process and consists of a single question. In one of the previous questions, you indicated that the [Group X] had the highest participation rate and the [Group Y] the lowest.

The award application data shows that although all students have high academic performance, the proportion of students who apply after being invited is approximately:

1. Students recommended by professors: 3 out of 5 invitees apply
2. Students recommended by peers: 2 out of 5 invitees apply
3. Invited students (no recommendation): 1 out of 5 invitees apply

Explain the reasons why you think students from the groups have different participation rates. For example, why being recommended by a professor or a peer affects the participation rate in comparison to students who were directly invited (without recommendation). [Open text response]

A.3 Study 2 Survey and Task Instructions

WELCOME

Thank you for participating in this study on economic decision making. The entire process consists of 6 parts and takes about 10 minutes to complete.

Compensation:

- Base Pay: $2.50 for fully completing the study.
- Bonus: Up to $3.00 additional based on the points you earn during the tasks.

Note: You will only be paid if you complete the entire study.

Please read the upcoming instructions carefully, as they explain how your bonus earnings are determined.

[Page break]

PART A

Before you start, please answer these 3 questions.

What is your gender?

- Female
- Male

What is your age?

- 18–24 years old
- 25–34 years old
- 35–44 years old
- 45–54 years old
- 55–64 years old
- 65 years or older

What is the highest degree you have completed?

- Less than a high school diploma
- High school degree or equivalent (e.g., GED)
- Bachelor's degree
- Master's degree
- Doctorate or professional degree (e.g., PhD, MD, JD)

[Page break]

INSTRUCTIONS

You will now read a short description of a Talent Recognition Program launched by a university abroad. Please make sure you understand how the program operates and how applicants are selected, as you will be asked for your opinions on this process next.

[Page break]

THE TALENT RECOGNITION PROGRAM

Leadership at a university in another country launched a Talent Recognition Program.

This program was designed to identify students who showed academic excellence and reward them for it.

[Page break]

REWARDS

The Talent Recognition Program rewards its winners with a recognition certificate, which they can reference in their CV, and monetary prizes equivalent to a month's salary.

[Page break]

ELIGIBILITY

All eligible students being considered for the Talent Recognition Program have a cumulative GPA that is above the median GPA of the university.

[Page break]

INVITATION TYPE #1

All eligible students enter the program through one of three different paths.

Some students are invited to apply because they were recommended by a professor.

[Page break]

INVITATION TYPE #2

Some students are invited to apply because they were recommended by a peer, another student who is also an applicant for the award.

[Page break]

INVITATION TYPE #3

Some others are invited directly based on their records, without any recommendation.

[Page break]

APPLICATION FOR THE TALENT RECOGNITION PROGRAM

The application process goes over a 3-week period, where applicants complete a series of tests.

At the end of the three weeks everyone is scored based on their performance during the entire process.

[Page break]

LAST STAGE OF THE PROCESS

On the last stage, week 3, applicants had to perform a task where they had to identify the number of "ones" in a series of tables, for one minute.

Example: this table has 8 "ones."

[Page break]

PART B

Remember: All applicants invited to the Talent Recognition Program had a cumulative GPA above the university median. They were selected through one of three channels:

1. Recommended by a professor
2. Recommended by a peer, another student who is also applying

3. Invited directly, without any recommendation

[Page break]

The following questions ask for your opinions about these three groups.

The application process had three stages. We want you to focus only on the points applicants scored in the first two stages, excluding the final counting task. Please rank the applicant groups based on who you think scored the most points during these initial stages:

- Rank 1: Highest scoring group
- Rank 2: Second highest scoring group
- Rank 3: Lowest scoring group

- Students invited directly: [Rank 1, 2, or 3]
- Students recommended by a peer: [Rank 1, 2, or 3]
- Students recommended by a professor: [Rank 1, 2, or 3]

[Page break]

You ranked the groups' scores as follows:

- Recommended by a professor: [Displayed rank]
- Recommended by a peer: [Displayed rank]
- Invited directly: [Displayed rank]

Please explain why you ranked the groups in this specific order. [Open text response]

[Page break]

INFO CONDITION ONLY

INFORMATION UPDATE

Program data reveals a key fact: At the end of Stage 2, all three groups of applicants had equally high scores.

There were no meaningful differences in performance between students recommended by a professor, recommended by a peer, or invited directly.

In your opinion, why did these three groups perform equally well during the first two stages? [Open text response]

[Page break]

DESCRIPTION OF PARTS C, D & E

In the next three parts, you will solve a series of counting problems.

- The Task: You will see tables filled with zeros (0) and ones (1). You must count the total number of ones and type your answer.
- The Time Limit: You have exactly 60 seconds for each part.

Example: The correct answer for the table below is 9, because the number 1 appears nine times.

[Example matrix omitted.]

[Page break]

PART C

In this part, you will complete the counting task individually.

- Time: You have exactly 1 minute.
- Scoring: You will earn 1 point for each correct answer.

Get ready! Your timer will begin as soon as you move to the next screen.

[Page break]

[Participants then saw up to 15 matrices sequentially and entered the number of ones in each table.]

[Page break]

Great job! You have completed Part C. Please click to proceed to the next screen.

[Page break]

PART D: FIRST MATCHING ENVIRONMENT

[The order of the two matching environments was randomized. Depending on the assigned order, participants first saw either the tournament version or the collaboration version of Part D.]

Tournament version

In this part, you will perform the same counting task, but this time you will compete against an applicant to the Talent Recognition Program, who has already completed this task. Your opponent may be from any of the three applicant groups.

Scoring Rules:

- Win: If you answer more tables correctly than your opponent, you receive 2 points for every correct answer.
- Lose: If your opponent answers more correctly, you receive 0 points.
- Tie: Ties are resolved randomly by the computer.

Note: Points are individual and only affect YOU. Your performance will not affect the applicant; their past score is only used as a reference to calculate your points.

[Page break]

Collaboration version

In this part, you will perform the same counting task, but this time you will work as a team with an applicant to the Talent Recognition Program, who has already completed this task. Your teammate may be from any of the three applicant groups.

Scoring Rules:

- Your score and your teammate's score will be added together and averaged.
- That average is then multiplied by 1.5 to determine your final points.
- Example: If you solve 10 tables correctly and your teammate solves 8, the average is 9. You will earn $9 \times 1.5 = 13.5$ points for this task.

Note: Points are individual and only affect YOU. Your performance will not affect the applicant; their past score is only used as a reference to calculate your points for being in a team.

[Page break]

Choice for Part D

Please choose the probability of being matched to [compete against / work as a team with] someone from each of the three groups.

Assign a percentage to each group below. A higher percentage means a higher chance of the computer pairing you with someone from that group.

Your total must sum to 100%.

- Student recommended by a peer: [Input %]
- Student recommended by a professor: [Input %]
- Student invited directly: [Input %]
- Total: [Sum %]

Get ready! Your timer will begin as soon as you move to the next screen.

[Page break]

[Participants then saw up to 15 matrices sequentially and entered the number of ones in each table.]

[Page break]

Great job! You have completed Part D. Please click to proceed to the next screen.

[Page break]

You chose to [compete / work as a team] with the groups using the following probabilities:

- Professor recommendation: [Displayed %]
- Peer recommendation: [Displayed %]
- Direct invite: [Displayed %]

Please explain why you allocated your percentages this way. [Open text response]

[Page break]

PART E: SECOND MATCHING ENVIRONMENT

[Part E used the remaining matching environment. Participants who first completed the tournament version in Part D completed the collaboration version in Part E, and participants who first completed the collaboration version in Part D completed the tournament version in Part E.]

Collaboration version

In this part, you will perform the same counting task, but this time you will work as a team with an applicant to the Talent Recognition Program, who has already completed this task. Your teammate may be from any of the three applicant groups.

Scoring Rules:

- Your score and your teammate's score will be added together and averaged.

- That average is then multiplied by 1.5 to determine your final points.
- Example: If you solve 10 tables correctly and your teammate solves 8, the average is 9. You will earn $9 \times 1.5 = 13.5$ points for this task.

Note: Points are individual and only affect YOU. Your performance will not affect the applicant; their past score is only used as a reference to calculate your points for being in a team.

[Page break]

Tournament version

In this part, you will perform the same counting task, but this time you will compete against an applicant to the Talent Recognition Program, who has already completed this task. Your opponent may be from any of the three applicant groups.

Scoring Rules:

- Win: If you answer more tables correctly than your opponent, you receive 2 points for every correct answer.
- Lose: If your opponent answers more correctly, you receive 0 points.
- Tie: Ties are resolved randomly by the computer.

Note: Points are individual and only affect YOU. Your performance will not affect the applicant; their past score is only used as a reference to calculate your points.

[Page break]

Choice for Part E

Please choose the probability of being matched to [compete against / work as a team with] someone from each of the three groups.

Assign a percentage to each group below. A higher percentage means a higher chance of the computer pairing you with someone from that group.

Your total must sum to 100%.

- Student recommended by a professor: [Input %]
- Student recommended by a peer: [Input %]
- Student invited directly: [Input %]
- Total: [Sum %]

[Page break]

Get ready! Your timer will begin as soon as you move to the next screen.

[Page break]

[Participants then saw up to 15 matrices sequentially and entered the number of ones in each table.]

[Page break]

Great job! You have completed Part E. Please click to proceed to the next screen.

[Page break]

You chose to [compete / work as a team] with the groups using the following probabilities:

- Professor recommendation: [Displayed %]
- Peer recommendation: [Displayed %]
- Direct invite: [Displayed %]

Please explain why you allocated your percentages this way. [Open text response]

[Page break]

CTRL CONDITION ONLY: POST-CHOICE INFORMATION SCREEN

At the beginning you ranked the groups' scores as follows:

- Recommended by a professor: [Displayed rank]
- Recommended by a peer: [Displayed rank]
- Invited directly: [Displayed rank]

Program data reveals a key fact: At the end of Stage 2, all three groups of applicants had equally high scores.

There were no meaningful differences in performance between students recommended by a professor, recommended by a peer, or invited directly.

In your opinion, why did these three groups perform equally well during the first two stages? [Open text response]

B A Simple Source Attribution Model

This appendix formalizes the source attribution argument in the main text. The purpose of the model is not to describe the institution's actual endorsement or invitation rule. Endorsement status was not randomly assigned, and endorsers may have used many kinds of information. Instead, the model describes how downstream evaluators interpret source labels when they do not observe a potential partner's or opponent's individual task performance at the moment of selection.

The model makes three points. First, source labels are comparative: an endorsement is positive only if evaluators perceive it as more informative about high expected performance than the unendorsed benchmark. Second, an endorsement is not harmful merely because it is noisy. A noisy source should be discounted. A penalty requires evaluators to view the source as less credible or less tied to merit than the benchmark. Third, performance information should attenuate, but need not eliminate, a peer endorsement penalty.

B.1 Source Labels and Posterior Beliefs

Each candidate i has latent task ability. The program first imposes a common eligibility threshold. Let $Q_i = 1$ denote that candidate i satisfies this threshold and belongs to the relevant finalist pool. Eligibility is a coarse signal: all finalists satisfy the threshold, but they may still differ in where they fall within the above threshold performance distribution.

For parsimony, represent residual ability within the eligible pool using two types,

$$\theta_i \in \{H, L\},$$

where $H > L$. Both types satisfy the eligibility requirement. Thus, L should not be interpreted as ineligible or low absolute ability. Rather, H and L denote higher and lower expected task ability within the pool of candidates who already satisfy the program's threshold.

Evaluators do not observe θ_i. Conditional on eligibility, they begin with the prior belief

$$\pi = \Pr(\theta_i = H \mid Q_i = 1),$$

where $0 < \pi < 1$.

The evaluator observes the candidate's source label,

$$S_i \in \mathcal{S} = \{F, P, U\},$$

where F denotes a faculty endorsed finalist, P denotes a peer endorsed finalist, and U denotes an unendorsed finalist. The unendorsed label is the benchmark. It refers to finalists who were eligible, invited, and advanced through the same institutional process, but whose invitation did not communicate a faculty or peer endorsement.

Evaluators hold subjective beliefs about how likely each source label is for higher and lower ability candidates within the eligible pool. For $s \in \mathcal{S}$ and $\theta \in \{H, L\}$, let

$$\ell_{s\theta} = \Pr(S_i = s \mid \theta_i = \theta, Q_i = 1),$$

with $\ell_{s\theta} > 0$. These are subjective likelihoods. They capture how evaluators believe the source label process works, not necessarily how the institution actually generated endorsements or invitations.

After observing source label s, the evaluator's posterior belief that the candidate is high ability is

$$\mu_s \equiv \Pr(\theta_i = H \mid S_i = s, Q_i = 1) = \frac{\ell_{sH}\pi}{\ell_{sH}\pi + \ell_{sL}(1-\pi)}.$$

Define the source likelihood ratio

$$R_s = \frac{\ell_{sH}}{\ell_{sL}}.$$

Bayes' rule implies

$$\frac{\mu_s}{1-\mu_s} = \frac{\pi}{1-\pi} R_s.$$

Thus, posterior expected ability is increasing in R_s.

The empirical comparison is relative to the unendorsed label. Define the diagnosticity

of source s relative to the unendorsed benchmark as

$$D_s = \frac{R_s}{R_U},$$

and its log form as

$$d_s = \log D_s.$$

Then, for any endorsed source $s \in \{F, P\}$,

$$\mu_s > \mu_U \quad \Longleftrightarrow \quad D_s > 1 \quad \Longleftrightarrow \quad d_s > 0,$$

and

$$\mu_s < \mu_U \quad \Longleftrightarrow \quad D_s < 1 \quad \Longleftrightarrow \quad d_s < 0.$$

This is the central theoretical object. A source label is interpreted positively when evaluators perceive it as more diagnostic of high performance than the unendorsed benchmark. It is interpreted negatively when evaluators perceive it as less diagnostic of high performance than the unendorsed benchmark. A peer endorsement therefore need not be objectively harmful, or even objectively uninformative, to be penalized. It is enough that evaluators assign the peer source lower diagnosticity than the unendorsed benchmark.

B.2 Source Attribution and Credibility

The diagnosticity parameter d_s is not a primitive fact about candidates. It reflects how evaluators interpret the process that generated source label s. I model source attribution as perceived credibility.

For each endorsed source $s \in \{F, P\}$, suppose evaluators consider two possible interpretations of the endorsement process. A merit interpretation treats the label as evidence that the candidate was endorsed because of expected performance. A social interpretation treats the label as partly reflecting friendship, reciprocity, favoritism, personal ties, popularity, or lack of objectivity.

Let $c_s \in [0, 1]$ denote the perceived credibility of source s, or equivalently the weight evaluators place on the merit interpretation. Let d_s^M denote the relative log diagnosticity

of source s under the merit interpretation, and let d_s^S denote the relative log diagnosticity of source s under the social interpretation. The evaluator's perceived diagnosticity is

$$d_s(c_s) = c_s d_s^M + (1 - c_s) d_s^S.$$

Assume that the merit interpretation is more favorable to expected performance than the social interpretation:

$$d_s^M > d_s^S.$$

Then

$$\frac{\partial d_s(c_s)}{\partial c_s} = d_s^M - d_s^S > 0.$$

Higher perceived credibility makes the source label more positive.

The case of theoretical interest is

$$d_s^S < 0 < d_s^M.$$

Under this condition, there exists a credibility threshold

$$c_s^* = \frac{-d_s^S}{d_s^M - d_s^S} \in (0, 1),$$

such that

$$d_s(c_s) > 0 \quad \Longleftrightarrow \quad c_s > c_s^*,$$

and

$$d_s(c_s) < 0 \quad \Longleftrightarrow \quad c_s < c_s^*.$$

Thus, the same source label can be interpreted as either positive or negative depending on how evaluators attribute the process that generated it. Faculty endorsements may be perceived as institutionally accountable, informed, and tied to accepted criteria of merit, leading to

$$d_F(c_F) > 0.$$

Peer endorsements may contain valuable information because peers observe one another in daily academic and social environments. But their lateral and socially embedded nature can also create attributional ambiguity. If evaluators suspect friendship,

reciprocity, favoritism, popularity, or lack of objectivity, perceived credibility falls. If it falls below the threshold c_P^*, then

$$d_P(c_P) < 0.$$

This inequality does not imply that peer endorsed candidates are objectively weaker. It means that evaluators may infer lower expected task ability from the peer endorsement label because they attribute the source process to social motives rather than screening based on merit.

B.3 Selection Gaps and the Collaboration–Tournament Reversal

The empirical design uses two incentive environments. In the collaboration environment, evaluators benefit from being matched with a higher performing partner, although the two individuals do not actively work together. In the tournament environment, the incentive is reversed: evaluators benefit from being matched against a lower performing opponent. This reversal allows the same source label to be tested under opposite payoff incentives.

Let M_j denote the task score of a potential partner or opponent. Assume that higher type implies a better score distribution: the score distribution for type H first order stochastically dominates the score distribution for type L, and

$$\mathbb{E}[M_j \mid H] > \mathbb{E}[M_j \mid L].$$

The expected score of a candidate from source s is

$$\mathbb{E}[M_j \mid S_j = s, Q_j = 1] = \mu_s \mathbb{E}[M_j \mid H] + (1 - \mu_s)\mathbb{E}[M_j \mid L],$$

which is increasing in μ_s, and therefore increasing in d_s.

In the collaboration environment, higher μ_s increases the expected value of choosing source s as a partner. In the tournament environment, higher μ_s decreases the expected value of choosing source s as an opponent, because a higher ability opponent is more likely to obtain a high score.

To map subjective values into selection probabilities, assume assigned probabilities are monotone in subjective expected value. One convenient representation is the logit allocation rule

$$p_{is}^{e} = \frac{\exp(\lambda_i^e V_{is}^e)}{\sum_{k\in\mathcal{S}} \exp(\lambda_i^e V_{ik}^e)}, \qquad \lambda_i^e > 0,$$

where $e \in \{C, T\}$ indexes the collaboration and tournament environments, V_{is}^e is evaluator i's subjective value of source s in environment e, and the empirical probabilities can be read as 100 times these shares. The logit form is not essential; it is used only to capture interior probability allocations while preserving monotonicity.

For each endorsed source $s \in \{F, P\}$, define the collaboration source gap as

$$G_s^C = \bar{p}_U^C - \bar{p}_s^C,$$

and define the tournament source gap as

$$G_s^T = \bar{p}_s^T - \bar{p}_U^T.$$

Both gaps are scaled so that positive values indicate that source s is treated as lower performing than the unendorsed benchmark. In collaboration, $G_s^C > 0$ means that source s receives lower selection probability as a partner than the unendorsed source. In the tournament, $G_s^T > 0$ means that source s receives higher selection probability as an opponent than the unendorsed source.

Proposition 1 (Source beliefs and incentive reversal) *Suppose higher type first order stochastically dominates lower type in task scores and assigned probabilities are monotone in subjective expected value. For any endorsed source $s \in \{F, P\}$,*

$$d_s > 0 \quad \Longrightarrow \quad G_s^C < 0 \quad \textit{and} \quad G_s^T < 0,$$

whereas

$$d_s < 0 \quad \Longrightarrow \quad G_s^C > 0 \quad \textit{and} \quad G_s^T > 0.$$

Proof. Bayes' rule implies that $d_s > 0$ if and only if $\mu_s > \mu_U$, and $d_s < 0$ if and only if $\mu_s < \mu_U$. First order stochastic dominance implies that higher μ_s raises expected partner value in collaboration and lowers expected opponent value in the tournament.

By monotonicity of the allocation rule, a source with higher subjective value receives higher selection probability. Therefore, if $d_s > 0$, source s receives higher probability than the unendorsed source in collaboration and lower probability than the unendorsed source in the tournament, implying $G_s^C < 0$ and $G_s^T < 0$. If $d_s < 0$, the inequalities reverse, implying $G_s^C > 0$ and $G_s^T > 0$. □

The collaboration–tournament reversal helps distinguish a performance inference from simple dislike of a source label. A pure dislike of a label would reduce selection of that label in both environments. By contrast, a belief that the label predicts lower performance implies avoidance in collaboration and targeting in the tournament.

B.4 Model Implications

The standard informational account of endorsements corresponds to $d_s > 0$: the endorsed source is perceived as more diagnostic of high performance than the unendorsed benchmark. This prediction is most natural for faculty endorsements, which may be perceived as institutionally accountable, informed, and tied to accepted criteria of merit. The model therefore implies

$$d_F > 0 \quad \Longrightarrow \quad \bar{p}_F^C > \bar{p}_U^C \quad \text{and} \quad \bar{p}_F^T < \bar{p}_U^T.$$

Equivalently,

$$G_F^C < 0 \quad \text{and} \quad G_F^T < 0.$$

The source attribution account allows the sign of an endorsement to depend on the perceived motives and credibility of the source. Peer endorsement is the key case. Peers may observe useful information about effort, reliability, helping, and daily performance. At the same time, peer endorsement is lateral and socially embedded. If evaluators attribute peer endorsement to friendship, reciprocity, favoritism, bias, personal ties, popularity, or lack of objectivity, then perceived credibility c_P falls. If $c_P < c_P^*$, then

$$d_P < 0.$$

The model therefore implies

$$d_P < 0 \quad \Longrightarrow \quad \bar{p}_P^C < \bar{p}_U^C \quad \text{and} \quad \bar{p}_P^T > \bar{p}_U^T.$$

Equivalently,

$$G_P^C > 0 \quad \text{and} \quad G_P^T > 0.$$

This implication is not a claim that peer endorsed candidates are objectively weaker. It is a claim about interpretation. The empirical test asks whether evaluators behave as if the peer endorsement label lowers expected task performance relative to the unendorsed benchmark.

The same logic yields an implication for open text justifications. Evaluators who explicitly describe peer endorsements as reflecting friendship, favoritism, reciprocity, bias, personal ties, popularity, lack of objectivity, or other motives unrelated to merit are articulating a lower perceived credibility weight c_P. The peer penalty should therefore be larger among evaluators who articulate these source attribution concerns. This implication is mechanism consistent evidence rather than the main behavioral test, because the main test comes from incentivized selection probabilities.

B.5 Performance Information

The source attribution account also generates a prediction about performance transparency. The eligibility threshold tells evaluators that all finalists satisfy a minimum standard. However, it does not tell them where finalists from different source labels fall within the above threshold performance distribution. Evaluators may still believe that source labels occupy different regions of that distribution.

Performance information refines this belief. In the empirical application, the information intervention tells evaluators that observed objective performance is statistically similar across source labels. This information is a public signal about one component of payoff relevant ability. It need not fully reveal all traits relevant for collaboration or tournament performance, and it need not fully eliminate concerns about the motives behind the endorsement process. The prediction is therefore attenuation, not necessarily full elimination.

Let d_{iP}^{0} denote evaluator i's baseline perceived diagnosticity of peer endorsement relative to the unendorsed benchmark. Let d_{iP}^{Info} denote the corresponding diagnosticity after receiving equivalent performance information. A simple reduced form representation of partial information is

$$d_{iP}^{Info} = (1 - \omega_i) d_{iP}^{0}, \qquad 0 < \omega_i < 1.$$

If peer endorsement is initially penalized, $d_{iP}^{0} < 0$, then

$$d_{iP}^{0} < d_{iP}^{Info} < 0.$$

Thus, performance information makes the peer label less negative, but it does not require the peer label to become neutral or positive.

Under the monotone selection logic in Proposition 1, attenuation of negative peer diagnosticity should reduce both the collaboration peer penalty and the tournament peer targeting premium:

$$G_{P,\text{Info}}^{C} < G_{P,\text{Control}}^{C},$$

and

$$G_{P,\text{Info}}^{T} < G_{P,\text{Control}}^{T}.$$

This prediction allows source labels to continue mattering after performance information is disclosed. Equivalent observed performance weakens the inference that peer endorsed finalists are lower performing, but it may not fully override beliefs about friendship, reciprocity, favoritism, personal ties, popularity, or lack of objectivity in the peer endorsement process.

C Additional Analyses and Robustness Checks

C.1 Participant Flow and Observable Characteristics

This appendix subsection reports descriptive information about the field study sample. Table C-1 shows how invited students progressed through the talent recognition program by source label. Table C-2 compares observable characteristics and objective task performance among finalists. These analyses are descriptive because endorsement status was not randomly assigned and progression through the application process was selective.

Table C-1 Participant Flow by Source Label

This table reports the flow of invited students through the talent recognition program by source label. Columns report unendorsed invitees (UNEND.), peer endorsed invitees (PEER), faculty endorsed invitees (FACULTY), and the pooled sample (TOTAL). Entries are counts, with cumulative retention rates relative to invited students in each source label reported below in parentheses. The invited row is the denominator for each source column. The final column reports robust equality test p-values from linear probability models of each stage completion indicator on source label indicators, with unendorsed invitees as the omitted category. Endorsement status was not randomly assigned, so these comparisons are descriptive.

	UNEND.	PEER	FACULTY	TOTAL	p-value
Invited / eligible	2,703	285	198	3,186	
	(100.0)	(100.0)	(100.0)	(100.0)	
Completed Stage 1	816	181	127	1,124	< 0.001
	(30.2)	(63.5)	(64.1)	(35.3)	
Completed Stage 2	587	131	116	834	< 0.001
	(21.7)	(46.0)	(58.6)	(26.2)	
Completed Stage 3	463	99	103	665	< 0.001
	(17.1)	(34.7)	(52.0)	(20.9)	
# Invited observations	2,703	285	198	3,186	

Table C-1 shows that endorsement status was strongly associated with progression through the application process. Unendorsed invitees had the lowest final stage completion rate, while peer endorsed and faculty endorsed invitees were more likely to complete all three stages. This pattern is important for interpreting the field study: source labels shaped selection into the final stage sample, so the field study estimates how evaluators interpreted source labels among finalists rather than the causal effect of endorsement status on candidate quality or progression.

Table C-2 provides descriptive evidence on observable finalist characteristics and performance. Peer endorsed finalists were not observably lower performing on the measures

Table C-2 Source Labels, Observable Characteristics, and Objective Performance

This table compares finalists across source labels in the talent recognition program. Columns report means for unendorsed finalists (UNEND.), peer endorsed finalists (PEER), and faculty endorsed finalists (FACULTY). Binary variables are reported as percentages. Standard deviations are reported below means in parentheses. The final column reports robust equality test p-values from regressions of each variable on source label indicators. Endorsement status was not randomly assigned and progression to the final stage was selective, so these comparisons are descriptive.

	UNEND.	PEER	FACULTY	p-value
Panel A. Observable characteristics				
Female (%)	64.58	58.59	63.11	0.543
	(47.88)	(49.51)	(48.49)	
Low SES (%)	37.37	35.35	31.07	0.461
	(48.43)	(48.05)	(46.50)	
Ethnic minority (%)	2.59	4.04	0.97	0.254
	(15.91)	(19.79)	(9.85)	
Rural (%)	33.91	30.30	29.13	0.550
	(47.39)	(46.19)	(45.66)	
Age	24.23	22.09	22.21	< 0.001
	(7.48)	(5.01)	(5.49)	
GPA	4.33	4.43	4.33	< 0.001
	(0.21)	(0.21)	(0.28)	
Panel B. Objective task performance				
Individual score	8.34	8.67	8.66	0.787
	(5.23)	(5.73)	(5.79)	
Collaboration score	11.77	11.09	11.66	0.486
	(5.63)	(4.99)	(5.83)	
Tournament score	11.50	12.29	11.51	0.537
	(5.37)	(6.72)	(5.83)	
# Observations	463	99	103	

collected in the study. They had the highest mean GPA, and equality tests do not reject equal performance across source labels in the individual, collaboration, or tournament tasks. These comparisons do not establish balance on unobserved dimensions, but they show that the main peer endorsement penalty is not supported by visible differences in GPA or task performance among finalists.

C.2 Selection Robustness

This appendix subsection reports additional analyses for the source based selection results in Section 4.1. The analyses examine whether the peer reversal differs by the randomized order of the collaboration and tournament environments, whether it dif-

fers by evaluator source label, whether participants used the probability allocation task intensively, and whether the main result is robust to alternative outcome definitions.

Table C-3 Order of Environments and Source Label Reversal

This table reports source label reversals by the randomized order in which participants completed the collaboration and tournament environments. Columns I, II, and III report reversals for participants who completed the tournament first (TOURN FIRST), participants who completed collaboration first (COLLAB FIRST), and the difference between the two orders (DIFF.), respectively. The peer reversal is the change in the peer versus unendorsed gap from collaboration to tournament. The faculty reversal is defined analogously. Standard errors, clustered at the participant level, are reported in parentheses. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	I TOURN FIRST	**II** COLLAB FIRST	**III** DIFF.
Peer reversal	15.50^{***}	17.27^{***}	1.77
	(2.99)	(2.98)	(4.23)
Faculty reversal	-11.19^{***}	-5.35^{**}	5.84
	(2.65)	(2.71)	(3.79)
Participants	344	321	665
p-value: peer difference by order			0.676
p-value: faculty difference by order			0.124
p-value: joint test of order interactions			0.303

Table C-3 shows that the peer reversal is positive and statistically significant regardless of whether the tournament environment or the collaboration environment appeared first. The peer reversal is 15.50 percentage points when the tournament appears first and 17.27 percentage points when collaboration appears first; the difference is not statistically significant.

Table C-4 shows that the components of the peer penalty vary across evaluator source labels. Unendorsed evaluators show a larger collaboration penalty, while peer endorsed and faculty endorsed evaluators show larger tournament targeting. However, the total peer reversal is positive in all three groups, and equality tests do not reject equal reversals across evaluator source labels.

Table C-5 shows that participants generally used the probability allocation task as an intensive measure. More than 93% assigned positive probability to all three source categories in each environment, and only about 4% assigned all probability to a single category.

Table C-6 shows that the directional evidence is consistent with the mean allocation results. The strict pairwise tournament indicator is not above 0.5 when ties are included, but it becomes significant after excluding exact ties, and the nonparametric signed rank

Table C-4 Peer Gaps by Evaluator Source Label

This table reports participant level peer gaps by the evaluator's own source label. Columns I–V report estimates for unendorsed evaluators (UNEND.), peer endorsed evaluators (PEER), faculty endorsed evaluators (FACULTY), all evaluators (ALL), and equality tests across evaluator source labels (P-VAL.), respectively. The collaboration peer gap is $p^C_{iU} - p^C_{iP}$. The tournament peer gap is $p^T_{iP} - p^T_{iU}$. The peer reversal is the sum of these two gaps. Positive values indicate that peer endorsed finalists are treated as lower expected performers than unendorsed finalists. Standard errors are robust and reported in parentheses. Evaluator source label was not randomly assigned, so the table is descriptive heterogeneity evidence. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	I	II	III	IV	V
	UNEND.	PEER	FACULTY	ALL	P-VAL.
Collaboration peer gap	13.55^{***}	-0.73	4.79^{**}	10.06^{***}	0.000
	(1.32)	(2.83)	(2.15)	(1.08)	
Tournament peer gap	2.58	13.92^{***}	15.60^{***}	6.29^{***}	0.000
	(1.69)	(3.02)	(3.35)	(1.38)	
Peer reversal	16.13^{***}	13.19^{***}	20.39^{***}	16.35^{***}	0.500
	(2.08)	(4.28)	(4.43)	(1.72)	
Participants	463	99	103	665	

Table C-5 Probability Allocation Task Usage

This table describes how participants used the probability allocation task in the collaboration (COLLAB) and tournament (TOURN) environments. The first three rows report the share of participants assigning positive probability to one, two, or three source categories. The fourth row reports the share assigning all probability to one category. The fifth row reports the share making a near equal allocation, defined as a maximum minus minimum probability difference of at most one percentage point across the three source categories.

	COLLAB	TOURN
Positive probability to one category only	3.91%	4.06%
Positive probability to two categories	2.56%	1.65%
Positive probability to all three categories	93.53%	94.29%
All probability to one category	3.91%	4.06%
Near equal allocation	11.88%	12.78%
Participants	665	665

and sign tests reject equality between peer endorsed and unendorsed tournament probabilities.

Table C-7 shows that the peer reversal is robust to excluding near equal allocators and to using log ratio transformations that address the compositional nature of the probability allocations. The peer/unendorsed log ratio is negative in collaboration, positive in the tournament, and strongly positive in the reversal. Faculty endorsed finalists move in the opposite direction.

Table C-6 Directional Selection Checks

This table reports binary and nonparametric checks for the direction of source label choices. Columns report the estimate (EST.), benchmark value (BENCH.), and p-value (P-VAL.). Rows 1–2 report pairwise comparisons between peer endorsed and unendorsed finalists in collaboration (COLLAB) and tournament (TOURN). Rows 3–4 report whether peer endorsed finalists receive the lowest COLLAB probability or highest TOURN probability among all three source categories. Rows 5–6 repeat the pairwise comparisons after excluding exact ties between peer endorsed and unendorsed finalists. The final rows report Wilcoxon signed rank and sign test p-values for the pairwise peer versus unendorsed comparisons.

	EST.	BENCH.	P-VAL.
Peer < unendorsed in			
COLLAB	0.580	0.500	0.000
TOURN	0.481	0.500	0.332
Peer lowest among all three in			
COLLAB	0.402	0.333	0.000
TOURN	0.424	0.333	0.000
Peer < unendorsed in			
COLLAB, excluding ties	0.731	0.500	0.000
TOURN, excluding ties	0.559	0.500	0.005
Wilcoxon signed rank test: COLLAB			0.000
Wilcoxon signed rank test: TOURN			0.000
Sign test: COLLAB			0.000
Sign test: TOURN			0.005
Participants	665		

Table C-7 Near Equal Allocation and Log Ratio Robustness

This table reports robustness checks for the main source label reversal. Columns report estimates (EST.) and observations (OBS.). Panel A excludes near equal allocators. In collaboration (COLLAB) and tournament (TOURN), near equal allocation is defined as a maximum minus minimum difference of at most one percentage point across the three source categories. For the total reversal (REV.), the exclusion removes participants who made near equal allocations in both environments. Panel B reports log ratio transformations using $\log((p_{is} + 0.5)/(p_{iU} + 0.5))$, where s is either peer endorsed or faculty endorsed and U is unendorsed. Standard errors are reported in parentheses. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	EST.	OBS.
Panel A. Excluding near equal allocators		
Peer COLLAB gap	11.39^{***}	586
	(1.22)	
Peer TOURN gap	7.19^{***}	580
	(1.58)	
Peer REV.	17.85^{***}	608
	(1.88)	
Panel B. Log ratio outcomes		
Peer/unendorsed log ratio, COLLAB	-0.418^{***}	665
	(0.049)	
Peer/unendorsed log ratio, TOURN	0.221^{***}	665
	(0.056)	
Peer log ratio REV.	0.639^{***}	665
	(0.077)	
Faculty/unendorsed log ratio, COLLAB	0.167^{***}	665
	(0.049)	
Faculty/unendorsed log ratio, TOURN	-0.155^{***}	665
	(0.042)	
Faculty log ratio REV.	-0.322^{***}	665
	(0.065)	
Peer REV. − faculty REV.	0.960^{***}	665
	(0.084)	

C.3 Observed Performance Diagnostics

This appendix subsection reports descriptive diagnostics for the observable performance alternative discussed in Section 4.2. Endorsement status was not randomly assigned, and progression to the finalist stage was selective. The analyses should therefore be interpreted as descriptive comparisons among finalists, not as balance tests.

Table C-8 Observed Finalist Performance by Source Label

This table reports mean GPA and objective task performance by source label among finalists. Columns report unendorsed finalists (UNEND.), peer endorsed finalists (PEER), faculty endorsed finalists (FACULTY), and the p-value from an omnibus equality test across the three source labels (P-VAL.). GPA is measured on the university's 0–5 scale. Task scores report the number of correctly solved matrices in the individual baseline, collaboration, and tournament tasks. These comparisons are descriptive because endorsement status was not randomly assigned and progression to the finalist stage was selective.

	UNEND.	PEER	FACULTY	P-VAL.
GPA	4.33	4.43	4.33	0.000
Individual baseline score	8.34	8.67	8.66	0.787
Collaboration score	11.77	11.09	11.66	0.486
Tournament score	11.50	12.29	11.51	0.537
Finalists	463	99	103	
GPA observations	463	99	100	

Table C-8 shows that peer endorsed finalists were not lower performing on the observed measures. Peer endorsed finalists had the highest mean GPA and did not score significantly worse in any of the three task environments.

Table C-9 shows no consistent peer endorsed task performance deficit. The peer endorsed coefficient is positive and statistically insignificant in the individual baseline and tournament tasks. It is negative in the collaboration task, but statistically insignificant without controls and only marginally significant after adding controls. The standardized task performance index is close to zero in both specifications.

Table C-10 shows that peer endorsed finalists are not disproportionately represented in the lower tail of the task performance distribution. The peer endorsed coefficients are small and statistically insignificant across all three task environments.

Table C-9 Task Performance by Source Label: Regression Diagnostics

This table reports regressions of objective task performance on source label indicators. Columns I–IV report the individual baseline score (INDIV), collaboration score (COLLAB), tournament score (TOURN), and standardized task index (INDEX), respectively. Unendorsed finalists are the omitted category. The INDEX outcome is the average of standardized INDIV, COLLAB, and TOURN scores. Panel A reports specifications without controls. Panel B adds GPA and demographic controls: gender, low socioeconomic status, first generation status, ethnic minority status, rural background, and age. Robust standard errors are reported in parentheses. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	I	II	III	IV
	INDIV	COLLAB	TOURN	INDEX
Panel A. No controls				
Peer endorsed	0.330	−0.678	0.796	0.027
	(0.624)	(0.565)	(0.718)	(0.078)
Faculty endorsed	0.323	−0.109	0.018	0.015
	(0.619)	(0.630)	(0.626)	(0.079)
Omnibus p-value	0.787	0.486	0.537	0.935
Observations	665	665	665	665
Panel B. With controls				
Peer endorsed	0.123	−0.999*	0.556	−0.019
	(0.615)	(0.598)	(0.729)	(0.080)
Faculty endorsed	0.258	−0.093	0.018	0.011
	(0.638)	(0.641)	(0.633)	(0.080)
Controls	Yes	Yes	Yes	Yes
Observations	662	662	662	662

Table C-10 Lower Tail Task Performance by Source Label

This table reports lower tail performance checks. Columns I, II, and III report indicators for scoring at or below the sample 25th percentile cutoff in the individual baseline task (INDIV), collaboration task (COLLAB), and tournament task (TOURN), respectively. Because of ties in task scores, these indicators include more than exactly 25 percent of the sample. Unendorsed finalists are the omitted category. Robust standard errors are reported in parentheses. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	I	II	III
	INDIV	COLLAB	TOURN
Peer endorsed	−0.004	0.040	0.038
	(0.053)	(0.054)	(0.054)
Faculty endorsed	0.011	0.034	0.022
	(0.052)	(0.053)	(0.053)
Cutoff	6	9	9
Unendorsed mean	0.348	0.354	0.356
Omnibus p-value	0.969	0.665	0.750
Observations	665	665	665

C.4 Open Text Source Attribution Analyses

This appendix subsection reports additional analyses for the open text source attribution results in Section 4.3. The text measures are based on open text justifications elicited

after the collaboration and tournament probability allocations. Responses were coded into theory driven, nonmutually exclusive categories.

Table C-11 Prevalence of Source Attribution Codes

This table reports respondent level prevalence of source attribution codes in the collaboration (COLLAB) and tournament (TOURN) justifications. Columns report counts (N) and percentages (PCT.). Categories are not mutually exclusive. The coding variables are repeated across source rows in the long format data, so the table reports prevalence using one observation per respondent. Percentages are based on 665 respondents.

	COLLAB		TOURN	
	N	PCT.	N	PCT.
Peer merit information	124	18.6%	116	17.4%
Peer social motives	131	19.7%	161	24.2%
Faculty merit information	335	50.4%	271	40.8%
Faculty social motives	10	1.5%	2	0.3%
Unendorsed merit information	242	36.4%	212	31.9%
Fairness or diversification rationale	171	25.7%	221	33.2%
No source based rationale	103	15.5%	140	21.1%
Respondents	665		665	

Table C-11 shows that peer social motive attributions are common in both environments, while faculty social motive attributions are rare. Faculty endorsements are much more often described as merit relevant than as socially motivated.

Table C-12 shows that peer social motive language is associated with a larger penalty for peer endorsed finalists in collaboration and larger targeting of peer endorsed finalists in the tournament. The corresponding faculty endorsed interactions do not move in the same direction. This source specificity is consistent with the theory: the attribution concern is attached to the peer source rather than to endorsement in general.

Table C-13 shows that the peer social motive association is robust to adding controls for environment order and evaluator source label. Panel B further shows that the association is source specific: peer social motive language predicts a larger peer penalty relative to faculty endorsement, not simply a general change in the use of all source labels.

Table C-14 provides a contrast to the peer social motive results. Respondents who describe peer endorsement as merit relevant do not show the collaboration peer penalty. Among respondents who do not mention peer merit information, peer endorsed finalists receive 13.23 percentage points less than unendorsed finalists. Among respondents who do mention peer merit information, the implied peer versus unendorsed gap is 3.73 percentage points and statistically indistinguishable from zero. This pattern is consistent

Table C-12 Target Cell Selection and Peer Social Motive Attributions

This table reports target cell regressions in which the dependent variable is the selection probability assigned to a source category. Columns I and II report the collaboration (COLLAB) and tournament (TOURN) environments, respectively. The omitted source category is unendorsed finalists. The omitted attribution category is respondents who do not mention peer social motives. The COLLAB column uses the collaboration peer social motive code, and the TOURN column uses the tournament peer social motive code. Participant fixed effects absorb respondent level main effects. Standard errors are clustered at the participant level. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	I	**II**
	COLLAB	TOURN
Peer endorsed	-7.80^{***}	1.85
	(1.52)	(1.76)
Faculty endorsed	2.91	-3.61^{**}
	(1.79)	(1.43)
Peer endorsed × peer social motives	-11.49^{***}	18.34^{***}
	(2.87)	(4.37)
Faculty endorsed × peer social motives	7.56^{*}	-1.49
	(4.12)	(2.58)
Constant	35.22^{***}	32.56^{***}
	(0.86)	(0.82)
Observations	1,995	1,995
Participants	665	665
R^2	0.134	0.095
Participant FE	Yes	Yes
p-value: peer interaction = faculty interaction	0.000	0.000

with the idea that the peer penalty depends on how evaluators attribute the source process.

Table C-15 validates the coding logic. In collaboration, respondents who describe faculty endorsement as merit relevant strongly favor faculty endorsed finalists. The interaction is large and source specific: faculty merit language shifts the evaluation of faculty endorsed finalists, not peer endorsed finalists. The tournament validation is weaker, but the collaboration pattern shows that the coded attribution categories correspond to meaningful differences in source label evaluation.

Table C-13 Peer Social Motive Attributions and Source Specific Gaps

This table reports participant level gap regressions. Columns I–III report specifications without controls; columns IV–VI add controls for the randomized order of the collaboration and tournament environments and the evaluator's own source label. Within each set, columns report the collaboration gap (COLLAB), tournament gap (TOURN), and reversal gap (REV.). Panel A reports peer gaps. Panel B reports peer minus faculty gap differences. Positive coefficients indicate a larger penalty against peer endorsed finalists. Robust standard errors are reported in parentheses. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	NO CTRLS.			CTRLS.		
	I	**II**	**III**	**IV**	**V**	**VI**
	COLLAB	TOURN	REV.	COLLAB	TOURN	REV.
Panel A. Peer gaps						
Collaboration peer social motives	11.49^{***}		9.22^{**}	10.29^{***}		9.17^{*}
	(2.34)		(4.67)	(2.43)		(4.72)
Tournament peer social motives		18.34^{***}	23.00^{***}		18.14^{***}	22.87^{***}
		(3.56)	(4.71)		(3.50)	(4.72)
Panel B. Peer minus faculty gap differences						
Collaboration peer social motives	19.05^{***}		11.16^{**}	18.28^{***}		12.11^{**}
	(2.60)		(5.11)	(2.73)		(5.13)
Tournament peer social motives		19.83^{***}	33.53^{***}		20.10^{***}	32.65^{***}
		(3.38)	(5.06)		(3.39)	(5.02)
Order controls	No	No	No	Yes	Yes	Yes
Evaluator source controls	No	No	No	Yes	Yes	Yes
Observations	665	665	665	665	665	665

Table C-14 Peer Merit Information as a Contrast

This table reports target cell regressions interacting source category with an indicator for whether the respondent described peer endorsement as merit relevant. Columns I and II report the collaboration (COLLAB) and tournament (TOURN) environments, respectively. The omitted source category is unendorsed finalists, and the omitted attribution category is respondents who do not mention peer merit information. Participant fixed effects absorb respondent level main effects. Standard errors are clustered at the participant level. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	I COLLAB	**II** TOURN
Peer endorsed	-13.23^{***}	6.38^{***}
	(1.41)	(1.86)
Faculty endorsed	4.22^{**}	-3.93^{***}
	(1.86)	(1.33)
Peer endorsed × peer merit information	16.96^{***}	-0.52
	(3.48)	(4.41)
Faculty endorsed × peer merit information	0.97	-0.21
	(3.62)	(3.09)
Constant	35.22^{***}	32.56^{***}
	(0.86)	(0.83)
Observations	1,995	1,995
Participants	665	665
R^2	0.130	0.052
Participant FE	Yes	Yes
p-value: peer interaction = faculty interaction	0.000	

Table C-15 Faculty Merit Information and Faculty Endorsement Choices

This table reports validation checks for the text coding. Columns I and II report the collaboration (COLLAB) and tournament (TOURN) environments, respectively. The table focuses on whether respondents who describe faculty endorsement as merit relevant also behave as if faculty endorsement is a positive signal. The omitted source category is unendorsed finalists. Participant fixed effects absorb respondent level main effects. Standard errors are clustered at the participant level. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	I COLLAB	**II** TOURN
Faculty endorsed, no faculty merit attribution	-9.22^{***}	-4.52^{***}
	(1.84)	(1.61)
Faculty endorsed, faculty merit attribution	17.81^{***}	-3.17
	(2.32)	
Faculty endorsed × faculty merit attribution	27.02^{***}	1.35
	(2.97)	(2.40)
Peer endorsed × faculty merit attribution	-0.00	3.58
	(2.67)	(3.41)
Observations	1,995	1,995
Participants	665	665
R^2		0.053
Participant FE	Yes	Yes
p-value: faculty interaction = peer interaction	0.000	

C.5 Preregistered Beliefs and Choice Diagnostics

This appendix subsection reports the belief outcomes that were specified in the field study preregistration. Before the collaboration and tournament decisions, participants reported beliefs about the three source labels. First, they ranked which source label they expected to have the largest number of applicants. Second, they reported the average percentage of possible points they expected finalists from each source label to have earned through the first two stages of the application process. I report these beliefs descriptively and then examine whether the main source label results are robust to controlling for the belief measures. I also report a same source allocation diagnostic, because the preregistration anticipated that evaluators might prefer finalists from their own source label.

Table C-16 Preregistered Beliefs and Belief Choice Consistency

This table reports preregistered belief outcomes and belief choice diagnostics for the field study. Panel A reports respondent beliefs by source label: unendorsed finalists (UNEND.), peer endorsed finalists (PEER), and faculty endorsed finalists (FACULTY). The last two columns report p-values for equality between peer endorsed and unendorsed finalists (PEER = UNEND.) and between peer endorsed and faculty endorsed finalists (PEER = FACULTY). The participation index is coded so that 3 means the source label was ranked as having the most applicants and 1 means it was ranked as having the fewest applicants. The expected score is the respondent's belief about the average percentage of possible points earned by finalists from that source label through the first two stages of the application process. Panel B reports participant fixed effects regressions of selection probability on expected score in collaboration (COLLAB) and tournament (TOURN). Standard errors are clustered at the participant level. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	UNEND.	PEER	FACULTY	PEER = UNEND.	PEER = FACULTY
Panel A. Beliefs by source label					
Expected participation index	2.23	1.73	2.04	0.000	0.000
Ranked as most applicants	0.465	0.230	0.305	0.000	0.030
Expected score through stages 1–2	74.94	62.71	74.50	0.000	0.000
Panel B. Belief choice consistency					
	COLLAB		TOURN		
Expected score	0.134^{**}		-0.252^{***}		
	(0.052)		(0.060)		
Observations	1,995		1,995		
Participants	665		665		
Participant FE	Yes		Yes		

Table C-16 shows that evaluators held less favorable beliefs about peer endorsed finalists before making their incentivized selection decisions. Peer endorsed finalists were least likely to be ranked as the source label with the most applicants, and their expected

score through the first two stages was 62.71 percent, compared with 74.94 percent for unendorsed finalists and 74.50 percent for faculty endorsed finalists. The belief choice consistency regressions show that source labels expected to perform better receive higher selection probability in collaboration and lower selection probability in the tournament.

Table C-17 Selection Probabilities with Belief Controls

This table reports participant–source cell regressions of selection probabilities on source label indicators and preregistered belief measures. Columns I and II report the collaboration (COLLAB) and tournament (TOURN) environments, respectively. The omitted source category is unendorsed finalists. Expected score is the respondent's belief about the average percentage of possible points earned by finalists from the corresponding source label through the first two stages. The expected participation index is coded so that higher values indicate that the respondent expected more applicants from that source label. All specifications include participant fixed effects. Standard errors are clustered at the participant level. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	I	II
	COLLAB	TOURN
Peer endorsed	-11.19^{***}	5.18^{***}
	(1.76)	(1.85)
Faculty endorsed	4.86^{***}	-4.05^{***}
	(1.57)	(1.20)
Expected score	-0.208^{***}	-0.082
	(0.070)	(0.062)
Expected participation index	2.840^{***}	-0.198
	(0.791)	(0.786)
Constant	44.44^{***}	
	(5.63)	
Observations	1,995	1,995
Participants	665	665
R^2	0.129	0.053
Participant FE	Yes	Yes

Table C-17 shows that the main source label pattern is not eliminated by the preregistered belief measures. Peer endorsed finalists remain less likely to be selected in collaboration and more likely to be selected in the tournament after controlling for expected participation and expected performance by source label. The belief measures therefore help characterize evaluators' priors, but they do not replace the main behavioral estimand based on incentivized selection probabilities.

Table C-18 shows some same source allocation. Evaluators assign more selection probability to finalists who share their own source label, and this pattern remains after controlling for target source labels and preregistered beliefs. However, same source allocation does not account for the main peer endorsement pattern: peer endorsed fi-

Table C-18 Same Source Allocation Diagnostic

This table reports the preregistered same source allocation diagnostic. The dependent variable is the selection probability assigned to each source category. Columns I–III report collaboration choices (COLLAB); columns IV–VI report tournament choices (TOURN). Same source equals one when the evaluator's own source label matches the target source label. Columns I and IV include only participant fixed effects. Columns II and V add source label controls. Columns III and VI add source label controls and preregistered belief measures. Standard errors are clustered at the participant level and reported in parentheses. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	COLLAB			TOURN		
	I	**II**	**III**	**IV**	**V**	**VI**
Same source	7.56^{***}	8.46^{***}	7.78^{***}	2.29^{*}	4.22^{***}	4.54^{***}
	(1.37)	(1.56)	(1.56)	(1.35)	(1.45)	(1.51)
Target source controls	No	Yes	Yes	No	Yes	Yes
Belief controls	No	No	Yes	No	No	Yes
Participant FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,995	1,995	1,995	1,995	1,995	1,995
Participants	665	665	665	665	665	665

nalists remain less likely to be selected in collaboration and more likely to be selected in the tournament after belief controls are included.

C.6 Study 2 Additional Analyses

This appendix subsection reports additional analyses for the performance information experiment in Section 5. The analyses report sample construction and randomization balance, target cell regressions corresponding to the participant level peer gap estimates, source specificity checks, order controls, compositional robustness, and probability allocation usage. The treatment is the performance information condition.

Table C-19 shows that treatment assignment is balanced on the measured prechoice variables. The order of the collaboration and tournament environments is also balanced across treatment conditions.

Table C-20 shows the same treatment pattern using participant–target source cells. Performance information raises the relative selection probability assigned to peer endorsed finalists in collaboration and lowers the relative selection probability assigned to peer endorsed finalists in the tournament.

Table C-21 shows that performance information does not meaningfully shift faculty endorsed finalists on the same gap scale. The treatment instead reduces the peer minus faculty gap, especially in the combined reversal.

Table C-19 Information Treatments, Observable Characteristics, and Objective Performance

This table compares observable characteristics and task performance across the control condition (CTRL) and the performance information condition (INFO) in Study 2. Columns report means and standard deviations in parentheses. Binary variables are reported as percentages. The final column reports robust equality test p-values from regressions on the treatment indicator.

	CTRL		INFO		P-VAL.
	MEAN	SD	MEAN	SD	
Panel A. Observable characteristics					
Female (%)	54.76	(49.89)	46.32	(50.00)	0.09
College degree (%)	60.95	(48.90)	67.89	(46.81)	0.15
Age 18–35	38.095	(48.68)	40.526	(49.22)	0.620
Age 35–55	48.095	(50.08)	42.105	(49.50)	0.230
Age 55+	13.810	(34.58)	17.368	(37.98)	0.330
Panel B. Objective task performance					
Individual score	9.371	(3.00)	9.311	(2.97)	0.839
Collaboration score	11.524	(2.90)	11.205	(2.56)	0.244
Tournament score	11.533	(2.64)	11.274	(2.83)	0.345
Observations	210		190		

Table C-22 shows that the treatment effects are similar after controlling for the randomized order of the two environments. The log ratio specifications also support the main conclusion: the peer/unendorsed log ratio increases in collaboration, decreases in the tournament, and the peer log ratio reversal is significantly attenuated.

Table C-23 shows that performance information changed how participants used the probability allocation task. Participants in the information condition were more likely to assign positive probability to all three categories and more likely to make near equal allocations. I do not condition the primary treatment estimates on near equal allocation because allocation style is itself affected by the treatment. Instead, I interpret this pattern as part of the behavioral response to performance information: the disclosure made the three source categories appear less differentiated, which is consistent with partial mitigation of source based penalties.

Table C-20 Study 2 Target Cell Regressions

This table reports target cell regressions from the performance information experiment. Columns I and II report the collaboration (COLLAB) and tournament (TOURN) environments, respectively. The dependent variable is the selection probability assigned to a source category. The omitted source category is unendorsed finalists. The performance information treatment coefficient is the treatment effect for unendorsed finalists. The interaction terms report how the treatment changes the corresponding source gap relative to unendorsed finalists. The CTRL peer gap equals $p^C_{iU} - p^C_{iP}$ in COLLAB and $p^T_{iP} - p^T_{iU}$ in TOURN. The treatment offset rows are scaled so that positive values indicate mitigation of the peer gap. Standard errors are clustered at the participant level and reported in parentheses. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	I	**II**
	COLLAB	TOURN
Peer endorsed	-14.55^{***}	12.00^{***}
	(2.24)	(2.71)
Faculty endorsed	-4.61^{*}	-0.17
	(2.48)	(1.64)
Performance INFO	-1.63	2.01
	(1.86)	(1.46)
Peer endorsed × INFO	6.02^{**}	-6.17^{*}
	(2.93)	(3.36)
Faculty endorsed × INFO	-1.13	0.12
	(3.10)	(1.93)
Observations	1,200	1,200
Participants	400	400
R^2	0.087	0.066
CTRL peer gap	14.55	12.00
Treatment offset	6.02	6.17
Remaining peer gap under INFO	8.53	5.83
Offset of CTRL peer gap (%)	41.4	51.4
p-value: peer interaction = faculty interaction	0.002	0.056

Table C-21 Study 2 Source Specific Gap Checks

This table reports participant level treatment effects for faculty gaps and peer minus faculty gap differences. Columns I, II, and III report the collaboration gap (COLLAB), tournament gap (TOURN), and reversal gap (REV.), respectively. Faculty gaps are defined on the same penalty scale as peer gaps: $p^C_{iU} - p^C_{iF}$ in COLLAB and $p^T_{iF} - p^T_{iU}$ in TOURN. The peer minus faculty gap subtracts the faculty gap from the peer gap. Positive values indicate that peer endorsed finalists are penalized more than faculty endorsed finalists. Robust standard errors are reported in parentheses. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	I	II	III
	COLLAB	TOURN	REV.
Panel A. Faculty gaps			
INFO treatment	1.13	0.12	1.25
	(3.09)	(1.93)	(3.36)
CTRL mean	4.61^{*}	−0.17	4.45^{*}
	(2.48)	(1.64)	(2.67)
Panel B. Peer minus faculty gaps			
INFO treatment	-7.15^{***}	-6.29^{*}	-13.44^{***}
	(2.30)	(3.28)	(4.41)
CTRL mean	9.94^{***}	12.17^{***}	22.10^{***}
	(1.93)	(2.61)	(3.68)
Observations	400	400	400

Table C-22 Study 2 Robustness: Order Controls and Log Ratio Outcomes

This table reports robustness checks for Study 2. Columns I, II, and III report the collaboration gap (COLLAB), tournament gap (TOURN), and reversal gap (REV.), respectively. Panel A reports treatment effects on participant level peer gaps after controlling for the randomized order of the COLLAB and TOURN environments. Panel B reports log ratio outcomes using $\log((p_{is} + 0.5)/(p_{iU} + 0.5))$, where s is peer endorsed or faculty endorsed and U is unendorsed. In the COLLAB log ratio, a positive treatment effect means that peer endorsed finalists become more attractive relative to unendorsed finalists. In the TOURN log ratio and REV. outcomes, negative treatment effects indicate mitigation. Robust standard errors are reported in parentheses. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	I	II	III
	COLLAB	TOURN	REV.
Panel A. Peer gaps with order controls			
INFO treatment	-6.11^{**}	-5.82^{*}	-11.93^{**}
	(2.93)	(3.29)	(4.79)
Collaboration first	−3.15	12.84^{***}	9.69^{**}
	(2.95)	(3.39)	(4.86)
Panel B. Log ratio outcomes			
Peer/unendorsed log ratio	0.351^{**}	-0.300^{*}	-0.651^{***}
	(0.152)	(0.154)	(0.246)
Faculty/unendorsed log ratio	0.026	0.000	−0.025
	(0.144)	(0.089)	(0.164)
Peer minus faculty log ratio REV.			-0.626^{***}
			(0.222)
Observations	400	400	400

Table C-23 Study 2 Probability Allocation Usage

This table reports how participants used the probability allocation task in Study 2. Columns report the control condition (CTRL), the performance information condition (INFO), and the treatment effect (EFFECT). "All three categories" equals one if the participant assigned positive probability to all three source categories. "All probability to one category" equals one if the participant assigned 100 percent to a single source category. "Near equal allocation" equals one if the maximum minus minimum difference across the three source probabilities is at most one percentage point. The treatment effect is the coefficient from a regression of each indicator on the performance information treatment. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	CTRL	INFO	EFFECT
Panel A. COLLAB			
All three categories	0.862	0.926	0.064^{**}
All probability to one category	0.086	0.037	-0.049^{**}
Near equal allocation	0.262	0.500	0.238^{***}
Panel B. TOURN			
All three categories	0.871	0.958	0.086^{***}
All probability to one category	0.100	0.037	-0.063^{**}
Near equal allocation	0.305	0.516	0.211^{***}
Observations	210	190	400

D Open Text Coding

This appendix describes the coding of the open text justifications used in Section 4.3. Participants provided one justification after assigning selection probabilities in the collaboration environment and one justification after assigning selection probabilities in the tournament environment. These responses are useful because the behavioral choices identify how evaluators acted on source labels, whereas the text helps characterize the attributions evaluators articulated when explaining those choices. I therefore use the open text responses as mechanism consistent evidence. They help assess whether the peer endorsement penalty is associated with explicit concerns about friendship, favoritism, reciprocity, bias, personal ties, lack of objectivity, or other motives unrelated to merit.

The coding task is a text as data measurement problem. The goal is to map short natural language explanations into a small set of theory driven attribution categories. This approach follows the broader economics literature that treats text as a source of economic data (see e.g., Gentzkow et al. 2019; Ash and Hansen 2023). Recent work in economics also argues that large language models can be useful measurement tools for text analysis when the coding rule is transparent, the task is well defined, and validation or robustness checks are reported (Ash et al. 2024; Ludwig et al. 2025). This is especially relevant for open ended survey responses, which are increasingly used to measure motives, beliefs, mental models, narratives, attention, and information transmission (Haaland et al. 2025). Related benchmarking evidence finds that large language models can reproduce expert text annotations and, in some settings, outperform outsourced human coders, including in Spanish language corpora (Bermejo et al. 2025); related evidence in political science shows that ChatGPT can outperform crowd workers on several annotation tasks (Gilardi et al. 2023).

I use the large language model as a blind, rule based annotator rather than as a discovery device. For each response, the model received only two inputs: the task context, either collaboration or tournament, and the respondent's open text justification. The prompt did not include the respondent's selection probabilities, task performance, GPA, source label, evaluator source label, or any outcome variable. This restriction is important because the source attribution variables should be based on what the respondent

wrote, not on the choices the variables are later used to explain. The model was instructed to code only content present in the text and not to infer motives from the probability allocation.

The coding scheme is source specific and organized around four types of rationale. First, *merit information* codes indicate that the respondent described a source label as informative about ability, effort, performance, reliability, motivation, objectivity, responsibility, or other criteria tied to merit. These codes are defined separately for peer endorsement, faculty endorsement, and the unendorsed category. Second, *social motive* codes indicate that the respondent described an endorsement source as reflecting friendship, favoritism, reciprocity, bias, personal ties, popularity, lack of objectivity, or other motives unrelated to merit. These codes are defined for peer endorsement and faculty endorsement. There is no corresponding social motive code for the unendorsed category because the unendorsed label does not involve an endorser. Third, *fairness or diversification* captures rationales based on equalizing chances, spreading probabilities across categories, or avoiding concentration in one source label. Fourth, *no source based rationale* captures responses that do not provide an interpretable source based explanation.

The categories are theory driven and not mutually exclusive. For example, a respondent could state that peers observe daily effort while also expressing concern that peer endorsements may reflect friendship. Such a response would receive both the peer merit information code and the peer social motive code. Similarly, a respondent could describe faculty endorsement as evidence of merit while also mentioning fairness as a reason to spread probabilities across categories.

The variable names retain legacy abbreviations from the coding files. The variables `pmi`, `fmi`, and `dmi` all refer to merit information, with the source varying across peer endorsement, faculty endorsement, and the unendorsed category. The variables `psm` and `fsm` refer to social motives for peer endorsement and faculty endorsement, respectively. The variables `fd` and `ns` refer to fairness or diversification and no source based rationale. The environment prefixes `collab_` and `tour_` identify whether the code comes from the collaboration or tournament justification.

The production coding was conducted in the ChatGPT Pro web interface in April 2026. I used the model available in the ChatGPT model picker at the time of coding; the web

interface did not provide a stable API model identifier for this interaction. For transparency, the replication materials include the exact prompt, the date range of coding, the input file, the raw model outputs, and the final coding file used in the analysis. Responses were coded in their original language. The prompt returned binary indicators for each category listed in Table D-1.

D.1 Coding Categories

Table D-1 standardizes the interpretation of the coding variables. The merit information codes have the same meaning across sources: the respondent treats the corresponding source label as conveying evidence of merit. The social motive codes also have the same meaning across endorsed sources: the respondent treats the corresponding endorsement as potentially shaped by motives unrelated to merit.

D.2 Coding Prompt

The coding prompt instructed the model to classify each response using the categories in Table D-1. The prompt emphasized three rules. First, classify the respondent's stated rationale, not the probability allocation. Second, assign source specific codes only when the response refers to the corresponding source label. Third, allow multiple codes when a response contains multiple rationales. The prompt did not include any information about the respondent's actual selection probabilities or observed performance.

> You will classify an open text justification written by a participant after assigning probabilities across three source labels in a talent recognition task. The source labels are: peer endorsed finalists, faculty endorsed finalists, and unendorsed finalists. The participant may be explaining choices in either collaboration, where a better partner increases the participant's payoff, or tournament, where a weaker opponent increases the participant's payoff.
>
> For each response, return binary indicators for the following categories: peer merit information, peer social motives, faculty merit information, faculty social motives, unendorsed merit information, fairness or diversification, and no source based rationale. Categories are not mutually exclusive.

Use the following definitions. Merit information means the response treats the corresponding source label as evidence about ability, effort, performance, reliability, motivation, objectivity, responsibility, or other criteria tied to merit. Social motives means the response treats the endorsement as potentially reflecting friendship, favoritism, reciprocity, bias, personal ties, popularity, lack of objectivity, or other motives unrelated to merit. Fairness or diversification means the response justifies the allocation by giving equal opportunities, diversifying, balancing probabilities, or avoiding concentration in one category. No source based rationale means the response provides no interpretable rationale about source labels.

Return only the binary indicators and do not infer information from variables not provided.

Table D-1 Open Text Coding Categories

This table reports the coding categories used to classify open text justifications. The same categories were applied separately to the collaboration and tournament justifications. The variable prefixes `collab_` and `tour_` identify the environment. Categories are nonmutually exclusive.

Category	Variables	Coding rule
Peer merit information	`pmi`	Equals one when the response describes peer endorsement as informative about merit. This includes claims that peers know the candidate's effort, responsibility, reliability, motivation, daily performance, collaboration, or other qualities that could predict task performance.
Peer social motives	`psm`	Equals one when the response describes peer endorsement as reflecting friendship, favoritism, reciprocity, personal ties, popularity, bias, lack of objectivity, or other motives unrelated to merit. This is the focal text measure for the source attribution analysis.
Faculty merit information	`fmi`	Equals one when the response describes faculty endorsement as informative about merit. This includes claims that faculty are knowledgeable, objective, institutionally accountable, familiar with student performance, or better positioned to assess ability, effort, responsibility, or academic quality. The variable name retains the legacy abbreviation `ami`, but the manuscript labels this category as faculty merit information.
Faculty social motives	`fsm`	Equals one when the response describes faculty endorsement as reflecting bias, favoritism, personal ties, lack of objectivity, or other motives unrelated to merit. The variable name retains the legacy abbreviation `asm`, but the manuscript labels this category as faculty social motives.
Unendorsed merit information	`dmi`	Equals one when the response describes the unendorsed category as informative about merit. This includes claims that unendorsed finalists reached the same stage without an endorsement, applied on their own initiative, satisfied the same eligibility criteria, or may be especially motivated, independent, or deserving. The variable name retains the legacy abbreviation `dmi`, but the manuscript labels this category as unendorsed merit information.
Fairness or diversification	`fd`	Equals one when the response justifies the probability allocation using fairness, equal chances, diversification, balance across categories, or avoiding concentration in a single source label, rather than a performance inference about a specific source.
No source based rationale	`ns`	Equals one when the response does not provide an interpretable rationale about source labels. This includes vague responses, random choice explanations, statements unrelated to the source labels, or responses that do not explain why a particular source category received more or less probability.

E Preregistration Details and Deviations

E.1 Study 1 Preregistration

The field study was preregistered at the AEA RCT Registry before the final stage of data collection. The preregistration described a study of how referral source shapes selection decisions in cooperation and competition. It specified three source categories, which the registry described as faculty referral, peer referral, and direct invitation; the manuscript refers to the same empirical categories as faculty endorsed, peer endorsed, and unendorsed. It also specified two payoff environments, which the registry described as cooperation and competition; the manuscript refers to them as collaboration and tournament. These terminology changes do not change the empirical design or the source categories being compared.

The core elements of the preregistered design are retained. Participants evaluated the same three source categories, assigned probabilities over source categories for collaboration and tournament matching, completed the real effort task, and were randomly assigned to the order of the collaboration and tournament environments. The preregistration listed partner and competitor selection probabilities and beliefs as primary outcomes, and task performance, consistency between beliefs and choices, and qualitative justifications as secondary outcomes. The main text focuses on the incentivized selection probabilities because they are the central behavioral outcome for the paper's source attribution argument. Belief outcomes, task performance, order checks, and text-based source attribution analyses are reported as diagnostic and mechanism-consistent evidence.

There are four substantive differences between the preregistration and the manuscript. First, the realized final stage sample is smaller than the expected sample in the registry. The preregistration anticipated approximately 1,020 final stage participants, while 665 finalists completed the final stage and enter the analysis. This difference reflects progression through the institutional application process rather than an analysis exclusion. Appendix Table C-1 reports participant flow by source label.

Second, the theoretical framing changed. The preregistration stated source-specific predictions comparing faculty referred, peer referred, and directly invited applicants. The

manuscript develops a more general source attribution theory: an endorsement should be valued when evaluators attribute it to credible screening based on merit, and penalized when evaluators attribute it to friendship, reciprocity, favoritism, lack of objectivity, or other motives unrelated to merit. The empirical tests still use the same source categories and the same selection probability outcomes. The unendorsed category remains the omitted benchmark.

Third, several preregistered outcomes are used as supporting analyses rather than as main-text outcomes. Objective task performance is reported in Section 4.2 and Appendix Tables C-8–C-10 to assess whether peer endorsed finalists were observably weaker. Order effects are reported in Appendix Table C-3. Belief outcomes and the consistency between beliefs and choices are reported in Appendix Table C-16. Same source allocation is reported in Appendix Table C-18. These analyses are not the main estimand because the paper focuses on how evaluators use source labels in payoff-relevant selection decisions.

Fourth, the preregistration listed the quality of justifications as a secondary outcome. The manuscript analyzes the open text responses using theory-driven source attribution categories rather than a general quality measure. This coding is used to assess whether the peer penalty is concentrated among evaluators whose own explanations attribute peer endorsement to social motives. Because the justifications were elicited after the probability allocations, these analyses are interpreted as mechanism-consistent rather than causal. The coding protocol and additional text analyses are reported in Appendix C.4 and Appendix D.

Table E-1 Mapping Preregistered Outcomes to Manuscript Analyses

This table summarizes how preregistered field study outcomes are used in the manuscript. The registry used the terms faculty referral, peer referral, and direct invitation; the manuscript uses faculty endorsed, peer endorsed, and unendorsed for the same empirical source categories. The registry used cooperation and competition; the manuscript uses collaboration and tournament.

Preregistered item	**Use in manuscript**	**Location**
Partner/competitor selection probabilities	Main behavioral outcome; reported as source-specific selection probabilities and participant-level peer gaps	Section 4.1; Tables 1–2
Beliefs about participation and performance	Supporting diagnostic; used to characterize beliefs and belief-choice consistency	Appendix Table C-16
Task performance	Supporting diagnostic for observable performance differences across source labels	Section 4.2; Appendix Tables C-8–C-10
Consistency between beliefs and choices	Supporting diagnostic for whether selection probabilities are aligned with stated performance beliefs	Appendix Table C-16
Qualitative justification coding	Mechanism-consistent source attribution evidence; not interpreted causally because text was elicited after choices	Section 4.3; Appendix C.4; Appendix D
Order effects	Robustness check for randomized order of collaboration and tournament environments	Appendix Table C-3
Same-source selection / homophily	Exploratory diagnostic using evaluator source label	Appendix Table C-18

E.2 Study 2 Preregistration

Study 2 was preregistered separately at the AEA RCT Registry before the online data collection. The preregistration described an online experiment on whether an information treatment reduces the penalty attached to peer endorsed candidates in what the registry called cooperation and competition. The registered intervention randomized participants into a control condition and a debiasing condition in which participants were told that candidate performance was statistically the same across endorsement groups. The registered primary outcomes were the probabilities assigned to peer endorsed candidates as partners and as opponents. The preregistration also specified the two main hypotheses: the information treatment should increase the probability as-

signed to peer endorsed candidates as partners and reduce the probability assigned to peer endorsed candidates as opponents.

The manuscript retains the registered treatment, randomization, sample size, and primary outcome logic. I recruited 400 Prolific participants, close to the registered allocation of approximately 200 participants per condition. The realized treatment assignment was 210 participants in control and 190 participants in the information condition. Randomization occurred at the individual level inside the online experiment.

There are three presentation changes relative to the preregistration. First, the preregistration uses the term "debiasing treatment," while the manuscript uses "performance information treatment." This change is only terminological: it refers to the same intervention informing participants that observed performance was statistically similar across source labels. Second, the preregistration uses "cooperation" and "competition," while the manuscript uses "collaboration" and "tournament" to match the terminology used throughout the paper. Third, the manuscript reports the primary outcomes as peer gaps relative to the unendorsed benchmark: $p^C_{iU} - p^C_{iP}$ in collaboration and $p^T_{iP} - p^T_{iU}$ in tournament. This scaling follows the field study and makes both outcomes increasing in the peer penalty. It is equivalent to testing the registered directional hypotheses because a reduction in the collaboration peer gap corresponds to increasing the relative selection probability of peer endorsed partners, and a reduction in the tournament peer gap corresponds to reducing the relative targeting of peer endorsed opponents.

The manuscript also reports supporting analyses that were not separately specified in the preregistration: the combined peer reversal, target cell regressions, source specificity checks comparing peer and faculty endorsement, order controls, log ratio robustness, and allocation use diagnostics. These analyses are reported as robustness and interpretation checks. The main Study 2 conclusion is based on the registered comparison between the control and information conditions for peer endorsed partner and opponent selection.

Table E-2 Mapping Study 2 Preregistered Outcomes to Manuscript Analyses

This table summarizes how the Study 2 preregistration maps to the manuscript. The preregistration used the terms debiasing treatment, cooperation, and competition. The manuscript uses performance information treatment, collaboration, and tournament for the same intervention and environments.

Preregistered item	**Use in manuscript**	**Location**
Control versus debiasing treatment	Same randomized comparison; manuscript labels the treatment as performance information	Section 5; Table 4
Partner selection of peer endorsed candidates	Reported as the collaboration peer gap, $p^{C}_{iU} - p^{C}_{iP}$, so lower values indicate mitigation	Section 5; Figure 2; Table 4
Opponent selection of peer endorsed candidates	Reported as the tournament peer targeting premium, $p^{T}_{iP} - p^{T}_{iU}$, so lower values indicate mitigation	Section 5; Figure 2; Table 4
Hypothesis: increase peer partner selection	Tested by whether performance information reduces the collaboration peer gap	Table 4; Appendix Table C-20
Hypothesis: reduce peer opponent targeting	Tested by whether performance information reduces the tournament peer targeting premium	Table 4; Appendix Table C-20
Expected sample of 400 Prolific participants	Realized analysis sample includes 400 participants: 210 control and 190 performance information	Appendix Table C-19
Additional robustness analyses	Supporting checks: target cell regressions, source specificity, order controls, log ratio outcomes, and allocation use diagnostics	Appendix Tables C-20–C-23