Identity Uncertainty*

Anujit Chakraborty† Arkadev Ghosh‡ Matt Lowe§ Gareth Nellis¶

June 2026

Abstract

Many group identities that influence economic behavior are imperfectly observed. Individuals and institutions often conceal identity markers to limit discrimination. Yet concealment also creates uncertainty about group membership, hampering coordination in social interaction. To study this tradeoff, we paired high- and low-caste men for collaborative data entry work in North India. We randomly assigned each mixed-caste pair to either be: (i) introduced by full names, making caste common knowledge; (ii) introduced by first names only, making caste disclosure a choice; or (iii) instructed not to disclose caste. The two concealment conditions substantially reduce the accuracy of beliefs about a partner's caste and confidence in those beliefs. They also weaken workplace relations, lowering trust, willingness to interact, and perceived productivity—consistent with identity helping structure social coordination. Evidence on mechanisms shows that identity concealment inhibits authentic interaction, making workers less able to express their "true selves," while certainty about a partner's identity is associated with stronger workplace ties. Concealment leaves a sizable caste disparity in higher-status role assignment intact, suggesting minimal impacts on discrimination. We conclude that where group identities are socially entrenched, reducing their legibility may undermine intergroup relations.

Keywords: Social identity; caste; team production; cooperation; field experiment.

JEL codes: C93, D83, J71, O12, Z13

*This project was pre-registered in the AEA RCT Registry with ID AEARCTR-0015151. We thank Suryashis Ghosh, Ujjwal Kumar, Ved Sharma, and the team of enumerators for excellent research assistance. We are grateful for funding from the Weiss Fund for Research in Development Economics, UCSD (Senate General Research), CEGA (Development Economics Challenge), and the Institute for Humane Studies. We thank seminar participants at Duke University, University of Chicago, University of Arkansas, UC Davis, UC San Diego, UC Louvain, and the London School of Economics for useful feedback and comments.

†University of California, Davis. Email: chakraborty@ucdavis.edu.

‡Duke University and NBER. Email: arkadev.ghosh@duke.edu.

§University of British Columbia. Email: matt.lowe@ubc.ca.

¶University of Pennsylvania. Email: gnellis@upenn.sas.edu.

Research increasingly recognizes the importance of social identity for economic outcomes (Akerlof and Kranton, 2000; Alesina and La Ferrara, 2005; Horowitz, 2000; Hjort, 2014; Kline et al., 2022; Oh, 2023). An implicit assumption running through much of this work is that individuals' group identities are readily observable.[1] In practice, however, group membership is often hard to discern, whether because individuals actively conceal their identities, or because social group categories are inherently blurry and indeterminate (Southall, 1970). As such, people must rely on noisy cues—such as names, patterns of speech, dress, or skin tone—to make guesses about the group membership of strangers and acquaintances, with no guarantee of accuracy. For example, Moerman (1965, p. 1218) notes that "it is difficult to delimit Thai peoples by means of traits ... the green sarong which sometimes distinguishes the Khyn from their Lue neighbors may elsewhere characterize the Lue." In his ethnography of Belfast, Northern Ireland, Burton (1978, p. 63) describes the fallible everyday practice of "telling" who is Catholic or Protestant, with one Catholic noting that it was sensible to "treat everyone as an Orangee [Protestant] until you knew different."[2] In short, correctly assigning someone to an ethnic, racial, or religious group is fraught with challenges—particularly on first encounter, and sometimes long after. What are the consequences of the resulting identity uncertainty for intergroup relations? Do members of different groups get along better or worse when they are not sure about one another's social backgrounds?

Theory offers competing predictions. In societies riven by group divides, a natural expectation is that greater identity uncertainty will ease cross-group interactions. Under what we term the *discrimination hypothesis*, an audience that cannot identify someone's group membership cannot discriminate against them on that basis, and will fall back instead on meritocratic, individual-specific evaluation criteria—judging people by "the content of their character." This logic underlies masking interventions such as blind auditions and CV anonymization schemes (Goldin and Rouse, 2000; Behaghel et al., 2015), and sometimes motivates members of stigmatized groups to distort identity signals so as to "pass" as members of advantaged groups (Dahis et al., 2026; Hobbs, 2014).

Still, there are reasons to expect the opposite. Under the *coordination hypothesis*, in settings where social identity is highly consequential, legible identities are a prerequisite for harmonious social encounters, providing each party to the encounter with shared norms about how it should proceed. In Goffman's (1959) "dramaturgical"

[1] In Akerlof and Kranton's (2000) model, "each person is endowed with a conception of her own categories and that of all other people" (718); that is, individuals can observe and categorize those they interact with. The paper acknowledges that people "with nondistinguishing physical features may be able to 'pass' as a member of another group" but suggests that identity choice "is often very limited" (726).

[2] The full quotation from Burton (1978, p. 63) richly conveys the challenges involved in making inferences about group identification: "One Sunday I walked into the house I was living in and a group of men and boys were discussing Protestant characteristics. Gerard remarked it was easy to tell some 'Prods' but harder to do so with others, particularly the young. He knew because he had been wrong a couple of times at work and had 'only just sorted it out.' Two young lads said it was their voices but were laughed out of it. By contrast it was claimed if anything it was their looks. Raymond thought the whole discussion a joke and recalled the 'buck eejit' (a prize idiot) who after the 1969 riots proclaimed he would always now be able to tell an Orangeman. Kieran, ever reflective, said that it was sensible to 'treat everyone as an Orangee until you knew different.' Paul challenged one of the lads who was adamant he could tell with the question: 'What about the wee girls, then, can you tell with them?' There followed a confused admission that that was a little harder, though from another room a young girl shouted she could."

framework, people are strongly averse to awkwardness and loss of face in intimate social settings. Interactions are akin to theatrical performances that require a shared script or "working consensus" of the situation to be successful. Knowing one another's group identities is a critical input into that consensus, helping interlocutors calibrate tone and avoid offense. In urban Zambia, conspicuous ethnic markers enable members of different groups to immediately "fit their neighbours and acquaintances into categories which determine the mode of behaviour towards them" (Mitchell, 1956, 32). Consider—more prosaically—the conversational missteps that can arise from being unaware whether a new acquaintance at an academic conference is a tenured professor or a first-year PhD student. Under this framework, where identities cannot be read, key social guideposts are missing, and friction and discomfort follow.

The effects of concealing identity are therefore an open empirical question. We study the case of caste in Uttar Pradesh, India, a setting where social identity remains deeply embedded in economic, political, and everyday life (Munshi, 2019; Oh, 2023; Cassan et al., 2021; Shukla, 2025; Banerjee et al., 2013; Lowe, 2021; Anderson, 2011; Asher et al., 2026; Chandra, 2007). However, strangers cannot straightforwardly pinpoint each other's caste when first meeting, because the observable markers associated with caste are frequently ambiguous or misleading (Valmiki, 2008; Satyanarayanan and Lee, 2023). For our experiment, we recruited young men from General Castes (GCs, or "upper castes") and Scheduled Castes (SCs, or "lower castes") with caste-distinctive surnames, and randomly paired them for a day of paid collaborative data-entry work, approximating short-term spot labor arrangements common in the region. We randomly assigned cross-caste pairs to one of three information environments: *Revealed*—enumerators announced both partners' full names, making caste common knowledge; *First Name Only*—enumerators introduced partners by first names only, permitting participants to choose whether to reveal caste or not; or *Hidden*—identical to *First Name Only*, other than the addition of instructions explicitly urging participants not to disclose caste-revealing information. The treatments are designed to mirror realistic workplace information environments: from village labor markets in India, where workers grew up together, and thus tend to know each other's caste (as in *Revealed*), to French public institutions, where policies of laïcité discourage displays of religious identity (as in *Hidden*).

Worker pairs collaborated on data-entry tasks for one day, with breaks in between tasks permitting social interaction. We paid pairs a piece rate based on joint productivity, and for all but one task the pair selected a "senior partner" who received a bonus payment. At the conclusion of the work period, each participant privately completed measures of trust and social relations, along with beliefs about their partner's caste.

Our first set of results considers beliefs about caste identity. Our rich measurement of beliefs allows us to contribute new descriptive facts to a nascent literature on "passing" that relies chiefly on administrative data (Adukia et al. 2025; Mitrut et al. 2025; Dahis et al. 2026). As intended, caste identity is effectively common knowledge in the

Revealed condition: 97% of participants correctly guess their partner's caste category (from among five categories), and 97% believe that their partner will guess their own caste category correctly. Participants in *First Name Only* are significantly less accurate and less confident when guessing the caste of their partners, with further losses of accuracy and confidence in *Hidden* ($p < 0.01$ for all pairwise comparisons). Caste category guesses are 14 percentage points less likely to be correct in *First Name Only*, and 22 percentage points less likely to be correct in *Hidden*. These effects should be considered a lower bound of the effective first stage on identity beliefs during the interaction, since they rely on data measured at the end of the workday, when partners have maximal information about each other. The results nevertheless show that a large fraction of participants correctly work out the caste of their partners even in the concealment conditions.

How do participants arrive at accurate guesses without knowledge of last names? First, caste often comes up explicitly in conversation: in *First Name Only*, 42% of participants report that their partner told them their caste. Second, many participants report making a guess based on how their partner spoke (e.g., how formal their Hindi was) or how they appeared (e.g., cleanliness). Combining ratings of clothing and speech with other potentially observable participant characteristics, we reach an out-of-sample predictive accuracy of 66%, shedding light on why even noisy signals of caste often yield accurate beliefs.

Surprisingly, low-caste and high-caste participants guess their partner's caste with similar accuracy in the *First Name Only* treatment, despite the expectation that marginalized groups have stronger incentives to keep their identity concealed. The evidence points to two explanations. First, identity disclosure is highly correlated across pairs. Most partnerships either have both members disclose (and learn) caste, or neither, rather than just one of the two. This correlation is a natural consequence of conversational norms, whereby one party's disclosure of personal information tends to elicit reciprocal disclosure from the other. Second, we do find the hypothesized asymmetry in *Hidden*: when instructed to hide caste identity, low-caste participants do so more successfully than high-caste participants. Thus, low-caste identities are concealed more effectively when concealment is explicitly encouraged, perhaps due to practice.

Our second set of results considers downstream impacts on social and economic relations. Here, the evidence we amass strongly favors the coordination hypothesis over the discrimination hypothesis. Participants in the *First Name Only* and *Hidden* conditions report significantly weaker partner bonds compared to those in *Revealed*: identity fusion, measured on a 0–4 scale, declines by 0.29 and 0.47 points, respectively ($p < 0.01$ and $p < 0.001$), and the perceived likelihood of meeting one's partner again falls by 16 and 20 percentage points ($p < 0.01$ and $p < 0.001$). Trust game behavior likewise deteriorates when last names are not revealed. Participants in the *First Name Only* and *Hidden* conditions send roughly 15% less to their partners ($p < 0.01$ for both comparisons), expect their partners to

send 9 to 13% less ($p = 0.05$ and $p = 0.01$), and, in the *Hidden* condition, return 9% less of the final pool of money ($p = 0.01$). Finally, in an incentivized exercise, participants in the *Hidden* condition require roughly Rs. 37 more per day to choose to work again with their current cross-caste partner (rather than switch to an otherwise unknown same-caste worker), relative to participants in the *Revealed* condition ($p < 0.1$). Importantly, the negative effects of concealment are present for both lower and upper castes, ruling out the possibility that negative effects are driven by upper castes' loss of social dominance when caste is hidden.

Concealment minimally affects team productivity; it appears that the task structure and piece-rate financial incentives discipline short-run production. But, strikingly, concealment substantially lowers *perceptions* of both own and partner performance (by 0.35 to 0.52σ, all $p < 0.01$). Consequently, our treatments leave actual output mostly unchanged while distorting perceptions of competence—a margin likely to matter in this setting for future job referrals and collaboration.

Why does the removal of identity cues undermine social relations? Our central claim is that uncertainty over partner identity inhibits authentic interaction in settings where the social costs of misreading identity are high. Two pieces of evidence lend credence to this interpretation. *(i) Loss of self:* Participants in *First Name Only* and *Hidden* report feeling substantially less able to show their "true selves" to their partners, with parallel declines of roughly one-third of a standard deviation relative to *Revealed* ($p < 0.001$). These effects are most pronounced among participants who rank caste as more important for their personal identity at baseline. *(ii) Confidence:* Confidence in one's guess about a partner's caste identity descriptively explains much of the observed treatment effects. Once guess confidence is controlled for, the *First Name Only* and *Hidden* coefficients shrink in magnitude and most become statistically insignificant across the suite of primary outcomes. More than that, however, participants who report being fully (i.e., 100%) certain of their partner's caste demonstrate discrete improvements in trust, identity fusion, and the likelihood of meeting their partner again (compared to those who are close to certain). In places where the costs of misreading identity might be substantial, knowing a counterpart's identity is qualitatively different from—and more desirable than—merely suspecting it.

If certainty about identity is valuable, why don't more participants in *First Name Only* reveal their identity? Additional evidence on selection into revelation speaks to this. Participants do not reveal indiscriminately; they are more likely to do so when (non-caste) socio-economic distance with their partner is lower, making disclosure appear safer. Identity uncertainty can therefore persist, likely due to the uncertain social costs that follow self-disclosure.

While our reduced form findings speak against the discrimination channel, we also explore the channel more directly using data on the allocation of the better-paid senior partner role. Despite similar ability by caste, the upper-caste worker takes the senior role 60% of the time in the *Revealed* condition ($p < 0.001$ for the null hypothesis

of a 50:50 allocation). Neither concealment condition affects this caste disparity. The disparity then appears to be a feature of caste differences in *self*-perception, e.g., the greater confidence of upper castes, rather than social perception.

We rule out several other explanations for the adverse impacts of identity concealment. First, the results are unlikely to reflect demand effects: our conclusions are similar if we ignore the more bundled *Revealed* condition that directly mentions caste, and participants rarely guess the study purpose to concern caste, nor are guesses affected by treatment. Second, while our concealment conditions also increase beliefs that partners share the same caste, we find that same-caste partnerships have *stronger* social relations. Third, the negative effects are not a consequence of thinner conversations or having less "individuating" information about one's partner. In terms of topics, conversations remain broad-reaching in *First Name Only* and *Revealed*, even if the ease of those conversations differs substantially. Finally, the results do not reflect a "worst-case" statistical discrimination explanation, whereby participants treat an unidentified partner as belonging to an even more socially distant group than the reality: Muslim guesses are extremely rare in the concealment conditions.

Related literature Our paper contributes to a small but growing literature in economics on the social construction of group identity. Building on the identity-economics tradition, this work treats identities not simply as fixed traits, but as social categories that are endogenously concealed and inferred (Akerlof and Kranton, 2000; Rose, 2023). Recent work shows that the classification of group identity is shaped by observer perceptions and economic incentives: wealthier and literate individuals were more likely to be classified by census enumerators as White or Mulatto in the U.S. census (Adukia et al., 2025); more-educated Roma were more likely to "pass" as non-Roma in Romanian census data (Mitrut et al., 2025); and Indians manipulated caste identity to secure eligibility for land acquisition (Cassan, 2015). Most closely related, Dahis et al. (2026) show that Black men in the U.S. who passed as White earned more but experienced weaker family ties. We bring an experimental approach to this literature, enabling us to (i) estimate the first *causal* effects of passing versus identity revelation,[3] and (ii) characterize the effectiveness of concealment using rich measures of beliefs—measures that are missing from the administrative data used in existing studies.[4]

Second, we contribute to the experimental literature on discrimination (Bertrand and Duflo, 2017). Audit studies observe real interactions after identity is perceived (Fix and Struyk, 1993), but auditors may differ along many dimensions other than group identity. Correspondence and related online experiments solve this problem by random-

[3] Or more precisely, the effects of opportunities to pass enabled by our *First Name Only* condition, and the effects of instructions to conceal identity using our *Hidden* condition.

[4] A small literature in political science likewise examines the visibility and identifiability of group identities, showing that ethnic markers are often less legible and fixed than commonly assumed (Robinson, 2024; Harris et al., 2024).

izing identity cues, such as names, but usually observe selection or evaluation margins such as callbacks or purchases (Bertrand and Mullainathan, 2004; Kline et al., 2022; Kelley et al., 2023). Our experiment design combines features of both approaches, enabling us to cleanly identify the effects of perceived identity on in-person interactions. In doing so, we uncover new insights relevant for policies around blinding (Goldin and Rouse, 2000; Behaghel et al., 2015; Boring et al., 2025). While blinding may be a useful tool at the application stage, by reducing discrimination, it may backfire when applied to in-person interactions, given the social costs of failed coordination.[5]

Finally, we contribute to the literature on intergroup contact. This work shows that contact across group lines can reduce prejudice, but that effects depend on how the contact is structured (Rao, 2019; Mousa, 2020; Lowe, 2021; Chakraborty et al., 2024; Ghosh, 2025; Ghosh et al., 2026; Greene et al., 2025; Lowe, 2025). A central tension concerns the role of group identity itself. Economic models emphasize that social categories shape norms and expectations (Akerlof and Kranton, 2010), and laboratory evidence from minimal-group settings shows that uncertainty about group identity can reduce coordination efficiency (Bronchal, 2023). Social psychology similarly argues that group identities provide interpretive frames through which individuals understand one another (Tajfel et al., 1971; Pettigrew and Tropp, 2006). Nevertheless, psychologists hold polar opposite views on whether contact works better when group identities are salient, or when interactions follow an approach of personalization (Brewer and Miller, 1984; Hewstone and Brown, 1986; Miller, 2002; Dovidio et al., 2003). For our core experimental comparisons we hold intergroup contact fixed while varying knowledge of group identity. Contrary to the personalization approach, we find that intergroup interactions benefit from clear cues to group identity.

Setting, experiment design, and data

Setting We ran the experiment in the rural districts of Barabanki and Ayodhya in Uttar Pradesh, India's largest state, with a population of nearly 250 million people. Caste is a salient social cleavage in our study districts. It is usually inferable from surnames, neighborhood of residence (given physical segregation within villages), and familiarity with local family histories and lineages; it might also be partially gauged through the noisier signals emitted by speech patterns, conversational content, appearance, and occupation.[6] Many economic interactions involve people

[5] A related literature considers how individuals respond to the prospect of expected discrimination during an application process (Kang et al., 2016; Charness et al., 2020; Lepage et al., 2022; Ruebeck, 2023; Avery et al., 2024; Del Carpio and Fujiwara, 2026; Angeli et al., 2023). For example, Angeli et al. (2023) show that favela residents in Brazil perform worse at interview when they expect the interviewer to know their favela status. In addition, work on stereotype threat finds that the revelation of group identity can harm individual performance. Most close to us, Hoff and Pandey (2006) find that the public revelation of caste led to a decline in task performance—solving mazes—by low-caste subjects, when grouped with high-caste subjects. Our approach differs from these papers in that we study whether and how group identity becomes revealed to others (as opposed to just randomizing whether it is revealed or not), and we study implications for social interactions (rather than individual decision-making, like whether to apply for a job).

[6] Little quantitative evidence exists on the accuracy of caste guesses made via these noisier signals, though we note that in another Indian state, Odisha, Oh (2023, p. 2077) finds that 71% of workers struggle to form any guess about the caste of surveyors based on appearance alone.

who do not know one another, including short-term work arrangements, hiring through informal networks, or market transactions in nearby towns. This makes the setting well-suited for studying the feasibility and consequences of identity uncertainty.

Caste bias in the study region Caste is a system of hereditary social stratification that permeates social, economic, and political life in Uttar Pradesh. Thousands of *jatis* are grouped into several broad administrative categories: General Castes, Other Backward Classes, Scheduled Castes, and Scheduled Tribes. These categories determine eligibility for affirmative action quotas in government jobs and higher education. Our study focuses on interactions between General Caste and Scheduled Caste participants, the latter being a marginalized group routinely subject to stigma and "untouchability."

To avoid priming respondents, we deliberately opted not to measure caste-related bias in the experimental baseline survey. Instead, we use two sources of evidence to assess bias in the study population. First, as we show in the endline results, participants exhibit ingroup bias in trust giving and social relations; role allocation in tasks also strongly favors GC participants. Second, we document strong caste bias in a scoping survey that we carried out in the same region. Appendix Figure A2 shows that both GC and SC caste groups strongly privilege their in-group during a hypothetical dictator game, over 70% of GC respondents reported that no one in their village would accept food from an SC individual, and almost all GC respondents reported that no one in their family would share a meal with an SC person. Notably, the share of GCs saying no one in their village would shake hands with SCs is considerably lower, at under 20%, indicating that untouchability manifests primarily around food, which triggers purity-pollution concerns, rather than everyday working interactions.[7]

Naming and the concealment of caste Social reform movements have repeatedly sought to reduce the salience of caste identity through changes to naming practices. In Sikhism, the collective surnames Singh (for men) and Kaur (for women) were adopted explicitly to help efface caste distinctions.[8] Similar impulses have appeared in modern India. To subvert caste hierarchies, the Dravidian and Self-Respect movements in early twentieth-century Tamil Nadu—led by E. V. Ramasamy, known as Periyar—urged people to replace caste surnames with initials. Srinivas (1966) coined the term "Sanskritization" to capture efforts by lower-caste groups to emulate upper-caste practices, rituals, and surnames in pursuit of social acceptance and upward social mobility. More recently, urbanization and

[7] This fact helps alleviate ex ante concerns that participation would only attract more progressive upper-caste men willing to work with lower-caste partners.

[8] Sikh historical tradition holds that, at the founding of the Khalsa in 1699, Guru Gobind Singh initiated the *Panj Piare* ("Beloved Five")—who hailed from both upper- and lower-caste backgrounds—into a shared "casteless fraternity" by having them drink from a common cup and abandon their Hindu caste names (Singh, 2004, 80). In practice, many Sikhs have retained additional family names indicative of clan or caste origins.

migration have made such strategic identity management increasingly feasible at the individual level, also: "In rural settlements, people can easily identify the caste identity of others ... But city life is characterized by a high degree of anonymity ... [which] provides the opportunity ... to transcend or hide [SCs'/Dalits'] 'Untouchable-hood' by changing surnames" (Parmar, 2020, 221).

Casual daily labor markets Casual daily labor markets are ubiquitous in developing countries and employ a large share of the world's poor (Kaur, 2019; Cefala et al., 2024). In India, where youth unemployment is high, jobseekers frequently gather at labor stands for temporary work, especially in construction. Many manufacturing firms rely on daily or short-term labor to meet fluctuating manpower needs. Such short-term labor arrangements were also widespread near our study sites. Since employment in these settings is highly transitory, often with one-day contracts, our design closely approximates an important real-world labor-market environment. At the same time, even brief jobs can generate valuable social ties. They allow workers to exchange information, observe ability and reliability, and build weak ties that may facilitate future work, referrals, or access to opportunities elsewhere. The quality of ties formed in short interactions can thus matter economically, even when the employment relationship itself is brief.

Recruitment and baseline We recruited subjects through door-to-door campaigns in villages located near the study sites, using three eligibility criteria. First, respondents had to be men between the ages of 18 and 40. Second, they were required to have a last name drawn from a pre-specified list of distinctive caste surnames—names that a prior scoping survey had shown to be almost exclusively associated with a single caste category (see Appendix Section D.1).[9] We primarily recruited participants from SC or GC backgrounds, though we included a small number of OBCs for obfuscation purposes, as we explain below. Third, respondents completed short, timed exercises testing basic competencies in typing and document reading; only those who demonstrated at least a minimal proficiency in both skills could continue. Individuals who met these three criteria, who provided informed consent, and who then completed a short baseline survey, were invited to participate in the study. They were given a unique identification code, and assigned a date on which they would work.

Flow of the experiment Each iteration of the experiment lasted one full day and ran as follows:

1. *Arrival.* Participants ($4 \leq N \leq 32$) arrived at the study site and checked in with facilitators using their unique identification code. Upon arrival, participants were seated in widely-spaced chairs and instructed not to interact with one another, with enumerators monitoring compliance.

[9] This was, of course, not made salient to participants. The recruitment survey further confirmed that the respondent's reported caste category matched the caste typically linked to their surname.

2. *Random assignment.* We randomly assigned individuals to pairs, and pairs to one of the three treatment conditions.

3. *Professionalism instructions and survey.* Each participant was individually called to complete a short "professionalism" survey, which presented a set of workplace rules and checked for comprehension of those rules. The rules were repeated until they were fully understood.

4. *Tasks in pairs.* Participants were escorted to private booths, where they worked in pairs on a sequence of five collaborative tasks lasting approximately four hours.

5. *Endline survey.* After completing the tasks, participants individually responded to an endline survey, led by an enumerator.

6. *Exit.* Immediately upon finishing the endline survey, participants were instructed to leave the study site. Enumerators ensured that participants did not interact across booths before departure.

7. *Incentive payments.* Participants were informed about their team earnings and bonus payments (see below), with base payments made on the same day. Bonus payments were disbursed within three days.

Randomization At the start of each workday, participants provided enumerators with their unique identification codes. We used these codes to retrieve the participant's caste from a pre-loaded database, minimizing the salience of caste during the session itself. A research assistant entered the codes of attending participants into a randomization platform we built (an `R` Shiny app, see Appendix Figure A1), which used the preloaded baseline data to conduct that day's randomization. The initial step of the randomization involved assigning participants to pairs:

- *Mixed GC-SC pairs (primary sample, $N = 532$ pairs).* Most participants were allocated to high-caste-with-low-caste (GC-SC) dyads.

- *Homogeneous GC–GC or SC–SC pair (obfuscation sample, $N = 49$ pairs).* For each daily session, one caste group (either GC or SC) was assigned to form that day's homogeneous pair. On even-numbered days (e.g., July 2), two GC participants were randomly selected to constitute a homogeneous dyad, while on odd-numbered days, two SC participants were selected instead.[10]

- *Mixed GC–OBC or SC–OBC pair (obfuscation sample, $N = 18$ pairs).* If at least one OBC participant was present at a given session, exactly one was randomly selected and paired with a participant from the caste

[10] This dyad type and the next were introduced partway through the experiment to enhance obfuscation.

opposite to that day's designated homogeneous group.[11]

We included homogeneous and OBC-containing dyads for obfuscation. If, following the experiment, participants discussed their assignments outside the experimental setting, the presence of these additional pair types helped ensure that the study design did not appear to be focused exclusively on GC-SC pairings. This reduced the likelihood that the intergroup relations angle of the experiment would be evident to future participants.

After the randomization was implemented, any unmatched individuals were placed on a waitlist and asked to return on another day at their convenience. In practice, enumerators sought to invite participants in the appropriate proportions each day so as to limit the number of those waitlisted.

Treatment conditions We assigned all GC-SC pairs, with equal probability, to one of three conditions—*Revealed*, *First Name Only*, or *Hidden*. Since the purpose of the remaining homogeneous-caste and OBC-containing dyads was for participants to learn that non-GC-SC pairings were possible, we always assigned these dyads to *Revealed*. Our information treatments are as follows (see Table 1 for the exact wording):

- *Revealed ($N = 176$ GC-SC pairs).* Prior to the pairs of work partners being introduced to one another, participants privately completed the professionalism survey, which lists some neutral workplace guidelines (e.g., not communicating across booths, limiting phone use, coordinating short breaks). At the start of the task survey, with both partners present around a table in their assigned booth, an enumerator read aloud each participant's full (i.e., first and last) name twice, purportedly to confirm that the correct individuals were in the correct booth.[12] Full names—preloaded by the enumerator—were then also displayed on the tablet screen throughout the tasks and were written on slips of paper placed on the desk. We asked participants to use these slips to record who served as the senior partner for each task. Given that we selected participants with caste-distinctive last names, the intent of the *Revealed* treatment was to ensure that caste identity is common knowledge among partners.

- *First Name Only ($N = 179$ GC-SC pairs).* Identical to *Revealed*, but with first names used throughout, rather than full names. Since first names are not caste-revealing in this context, this treatment left participants free to conceal or disclose caste to their partners.

[11] This brings the total number of participants in the experiment to 1,198. Out of these, 1,144 completed both the work tasks and the endline survey. Some subjects left during the day because of personal or family obligations. We keep participants' data for as long as they participated; that is, even if they did not complete all parts of the survey. All participants were compensated based on the parts of the study they completed.

[12] During the informed consent process, all potential participants were explicitly informed that their first and last names might be shared with a work partner who "may come from a similar background to you, or from a very different background." Participants were further instructed: "You should only participate in this study if you are comfortable with your first name, last name, and age being shared with your work partner, who is someone you do not know."

- *Hidden ($N = 177$ GC-SC pairs).* Identical to *First Name Only*, except that the professionalism survey additionally instructed participants not to disclose their last names or any other information that could reveal their caste. We justified the instruction on the grounds of respecting one another's privacy. Participants answered one additional comprehension question concerning the extra instruction. Thus, this condition attempted to make caste identity less observable by discouraging the exchange of clear caste-revealing information during the interaction.[13]

Data entry tasks and piece-rate payments We randomly assigned pairs to booths to prevent any perceived association between treatment status and physical work location. We instructed pairs to perform a sequence of five paid collaborative tasks. Each task lasted up to 45 minutes, with 15-minute breaks in between, and the entire exercise took roughly four hours to complete. The tasks were designed to reflect a clear status hierarchy: one partner, dubbed the *senior partner*, operated the computer to type in the information, while the other, dubbed the *junior partner*, read aloud data from paper forms contained in labeled envelopes. Worker pairs chose which worker would be the senior partner for each task, with the exception of one randomly-chosen task, in which the senior partner role was assigned randomly. The tasks varied in content and difficulty; all, however, demanded close coordination under time pressure. To build familiarity with the technology and procedures, pairs started with the two simplest tasks, in this order:

1. *Ordered data entry.* Participants are given a long list of names, each accompanied by an exam grade. The same names appear on their computer screen in the same order, but without the grades. The task was to enter the correct grade for each name.

2. *Unordered data entry.* Participants again receive a list of names, this time accompanied by years of birth. Unlike the previous task, however, the names appear in a different order on the computer screen. Participants' task was to locate each name in the paper list and enter the corresponding year of birth in a four-digit format (e.g., "1999"). Because completing the task required searching and cross-referencing across the two lists, it was more demanding.

Next, pairs carried out three more challenging tasks, in random order, allowing us to purge order effects in the analysis:

[13] Our treatments map onto a documented institutional response to caste-linked identification in Indian workplaces: suppressing surnames. For example, the Himachal Pradesh police adopted a policy of removing surnames from official communication, to explicitly reduce caste and regional divisions within the force (see *Indian Express*, "We don't have surnames, we are the Himachal Pradesh Police", Jun 24, 2011 and *Times of India*, "Himachal Pradesh police no-surname policy attracts Kerala", Nov 6, 2011). Our *First Name Only* condition captures this kind of partial masking; surnames are withheld, but caste may still be inferred or disclosed during interaction. The *Hidden* condition goes a step further by removing name-based cues altogether.

Table 1: Experimental manipulations, by treatment condition.

	Professionalism survey	Tasks survey		
		Enumerator introductory checks	Tablet display, appeared after every task	Record sheet, sitting on desk throughout
Revealed	We wish to create a positive and productive work environment. To contribute to this environment, you must follow these factory guidelines ... You can take breaks when you need to, but do not make them too long. Ideally, coordinate short breaks together with your partner to maintain a balanced workflow and avoid interruptions for your partner. Do not communicate with workers in the other booths at all over the course of the work day. Please limit the use of your mobile phone while you are working.	Before you begin the work tasks, I must confirm that the correct partners are in the booth. Is ***Pramod Gautam*** in this booth? To repeat, is ***Pramod Gautam*** in this booth? [CONFIRM YES] Thank you. Is ***Anup Pandey*** in this booth? To repeat, is ***Anup Pandey*** in this booth? [CONFIRM YES]	Who do you select to be the senior partner for this task (enters information on the tablet)? Remember, the person chosen as senior partner will be paid 50 INR extra for this task. PARTICIPANTS SELECT: ***Pramod Gautam*** OR ***Anup Pandey***	Number of times ***Pramod Gautam*** was senior partner [PARTICIPANTS WRITE NUMBER]; Number of times ***Anup Pandey*** was senior partner [ENTER NUMBER]
First Name Only	« Same as *Revealed* »	Before you begin the work tasks, I must confirm that the correct partners are in the booth. Is ***Pramod*** in this booth? To repeat, is ***Pramod*** in this booth? [CONFIRM YES] Thank you. Is ***Anup*** in this booth? To repeat, is ***Anup*** in this booth? [CONFIRM YES]	Who do you select to be the senior partner for this task (enters information on the tablet)? Remember, the person chosen as senior partner will be paid 50 INR extra for this task. PARTICIPANTS SELECT: ***Pramod*** OR ***Anup***	Number of times ***Pramod*** was senior partner [PARTICIPANTS WRITE NUMBER]; Number of times ***Anup*** was senior partner [ENTER NUMBER]
Hidden	We wish to create a positive and productive work environment. To contribute to this environment, you must follow these factory guidelines ... You can take breaks when you need to, but do not make them too long. Ideally, coordinate short breaks together with your partner to maintain a balanced workflow and avoid interruptions for your partner. Do not communicate with workers in the other booths at all over the course of the work day. ***Do not tell your last name to your work partner, or provide information that could reveal your caste. It is important to respect each other's privacy.*** Please limit the use of your mobile phone while you are working.	« Same as *First Name Only* »	« Same as *First Name Only* »	« Same as *First Name Only* »

Notes: This table presents relevant excerpts from the professionalism and tasks surveys, where the key experimental manipulations were delivered.

3. *Cross-validation.* Here, the envelope contains two separate paper lists of names with grades. The screen displays names from List 1 in the same order. Participants had to check whether the grade in List 1 matched the grade in List 2, and then enter "0" if the name appeared in both lists but the grades differed, "1" if the name and grade matched across both lists, and "2" if the name from List 1 was missing from List 2.

4. *Receipts.* For this task, participants are given mock shop receipts showing purchased items and prices. On screen, they had to enter totals or check whether the items and prices matched across versions. The job mimics everyday record-keeping in workplace or market contexts, involving checks of discrepancies in numerical information.

5. *Ravens.* This task consisted of a Raven's Progressive Matrices–style cognitive test, to measure abstract reason-

ing ability, rather than a data-entry exercise. Participants were shown visual patterns with one piece missing and had to choose the correct option to complete the pattern.

We compensated participants based on the joint accuracy of their performance across tasks. Team earnings ranged from 150 to 450 INR (roughly $1.50 to $4.50) per person, with compensation increasing in the number of correctly completed task items. In addition, the partner designated as the *senior partner* for a given task received a personal bonus of 50 INR. In total, a participant who served as senior for all five tasks could earn up to 250 INR in bonuses, on top of the team-based payment.

Outcomes We pre-specified five families of primary outcomes (AEA RCT Registry #15151), covering beliefs about group identity, social relations, trust, productivity and task allocation, and well-being at work. The paragraphs that follow denote the exhaustive set of pre-specified primary outcomes. We also refer to some pre-specified secondary outcomes, explicitly flagging those outcomes as such. We detail minor deviations from the pre-registration in Appendix Section D.2.

Beliefs about group identity. At the end of the work session, we asked each worker to guess both the caste category (SC, ST, OBC, GC, or Muslim) and the specific *jati* (e.g., Brahmin, Chamar) of their partner. We then classify their answer as correct or not by matching their guess with the baseline data for their partner. Participants reported their level of confidence in the caste category guess on a 0 to 100% scale—with the answer reflecting the participant's perceived probability that they guessed correctly—and we also asked participants whether they thought their partner would be able to correctly identify their own caste (a second-order belief). This allows us to characterize not only raw accuracy, but also workers' meta-beliefs about their ability to conceal or reveal caste identity. To guard against strategic behavior, we did not incentivize guesses.[14] For obfuscation, we also asked participants to make guesses about their partners' marital status, age, asset ownership, and father's occupation.

Social relations. We measured how participants valued future interaction with their assigned partner. First, in a series of incentivized willingness-to-pay (WTP) and willingness-to-accept (WTA) choices, participants decided whether they would forgo higher wages in order to work again with their current partner rather than a randomly selected co-worker of the same caste (signaled by full name). Participants also reported whether they expected to meet up with their partner in the future outside the study context. Finally, we used an overlapping circles (inclusion-of-other-in-self) measure to assess subjective identity fusion with the partner on a scale of 0 to 4.

[14]For example, if we had incentivized guesses, and new participants had learned from previous participants that this was the case, the new participants might have exerted more effort to learn the caste of their partner in the *First Name Only* and *Hidden* conditions, reducing the external validity of our measured beliefs about identity.

Trust. Pairs played a standard incentivized trust game (Berg et al., 1995). In this game, Player A (the sender) received Rs. 50 and decided how much to send to their work partner. Transfers were tripled, and Player B (the receiver) then chose how much to return. Participants stated their intended transfer amounts if chosen as Player A, their expectations about how much their partner would transfer if assigned as Player A, and their conditional return amounts for each possible transfer if chosen as Player B (using the strategy method). These data allow us to capture multiple dimensions of trust and trustworthiness: willingness to place resources at risk, expectations of reciprocity, and actual reciprocity when endowed with power. We randomly assigned one partner to have their sender choice implemented, with the other having their return choice implemented.[15]

Productivity and task (mis)allocation. We measure productivity on the data-entry tasks as the number of fields correctly entered, a joint measure of quantity and quality. Because pairs were free to decide which worker would serve as the senior partner (i.e., operate the tablet) on all tasks but one, we also track the identity of the partner who assumed this higher-status role, task-wise. We complement these objective performance outcomes with a secondary one: perceptions of own and partner's performance on a scale of 1 to 10.

Well-being at work. Finally, we measured workers' affective experience during the session. Respondents rated the extent to which they felt happy, self-conscious, confident, relaxed, and anxious, each on a scale from 0 = not at all to 3 = very. As secondary outcomes, we also measured workers' perceptions of the interpersonal aspects of the interaction, asking them to report the extent to which they were able to show their "true self" to their partner, as well as their partner's friendliness, rudeness, bossiness, and the degree of tension experienced during the interaction, on scales ranging from 1 = not at all to 10 = very.

Estimation

We estimate our pre-specified intent-to-treat specification:

$$Y_i = \alpha + \beta_1 \textit{First Name Only}_{p(i)} + \beta_2 \textit{Hidden}_{p(i)} + \gamma X_i + \delta Y_{i0} + \lambda_s + \varepsilon_i, \tag{1}$$

where Y_i denotes the outcome of interest for individual i. Our analysis focuses on the individuals assigned to GC-SC pairs unless stated otherwise. The indicators $\textit{First Name Only}_{p(i)}$ and $\textit{Hidden}_{p(i)}$ denote assignment of pair p to the *First Name Only* or *Hidden* conditions, with the *Revealed* condition serving as the omitted reference category. X_i is a vector of individual-level controls (age and fixed effects for the participant's level of education), Y_{i0} is the baseline

[15] Enumerators guided participants through a series of comprehension questions and worked examples to ensure that they understood the game's setup.

measure of the outcome to improve precision (if available), and λ_s are randomization strata fixed effects (day-by-caste, where caste is GC or SC). We cluster standard errors at the pair level for outcomes measured at the individual level, and we estimate heteroskedastic-robust standard errors in specifications that include only one of the two caste categories.

Summary statistics Appendix Table A1 presents summary statistics for the experimental sample. The average participant is 23 years old. Additionally, 73% of the sample report having completed 12th grade. College completion is substantially lower at around 20%, partly because many individuals have not yet reached college completion age. The average monthly income is Rs. 10,700 (roughly 120 USD), indicating that the participants are relatively low-income. Finally, in terms of identity rankings, participants most commonly identify as "Indian" first, followed by personal relationships (e.g., being a father, son, or brother), and then religion, caste, and state (Uttar Pradesh) identity, all of which receive similar rankings.

As expected, GC participants are, on average, more educated, come from wealthier families, and report having fewer siblings, consistent with lower fertility in this group (Appendix Table A2). GCs rate their Indian identity as more important than SCs do, and their caste identity as less important. At first glance, this might seem counterintuitive given the historical dominance and better socioeconomic outcomes of high castes. But recall that participants were asked to provide relative rankings of identities. One plausible explanation is that high castes may feel a stronger sense of pride in the nation, perceiving themselves as more central to the national identity and its historical narrative, whereas low castes may place greater emphasis on caste identity as a counter-response to marginalization.

Sample selection We approached 1,677 potential subjects in Barabanki and Ayodhya through door-to-door campaigns to arrive at our final experimental sample of 1,198 men. We do not find large or systematic differences between those who participated in the experiment and those who ultimately did not (Appendix Table A3). Participants are slightly younger and, consequently, somewhat less likely to have completed college. They also rank their religious identity marginally lower and personal identities (e.g., father, son, brother) marginally higher, while ranking their caste identity similarly. This likely reflects the fact that younger men had greater scheduling flexibility to participate. Crucially, however, we do not observe selection on the importance of caste identity. Participants are also positively selected on baseline productivity for both castes (Appendix Figure A3);[16] for General caste workers in particular, this suggests that higher-ability men were not deterred from participating—whether by the nature of the task or by the prospect of contact with other castes.

[16]This is among men who met the selection criteria based on productivity; in practice very few men we approached did not.

Balance tests Appendix Table A4 indicates that the randomization was successful. We estimate specification 1 with participant characteristics as the dependent variable and report p-values for the full set of pairwise comparisons between treatment groups. Across all comparisons, the estimated coefficients are small in magnitude and not statistically significant.

Attrition Out of the 1,198 men who were randomized into a treatment condition, 1,144 (95%) completed both the work tasks and the endline survey. Some subjects left during the day because of personal or family obligations. We keep participants' data for as long as they participated, that is, even if they did not complete all parts of the survey.

Results

Beliefs about group identity Our *Revealed* treatment condition works as intended: 97% of participants correctly guess their partner's caste category at the end of the workday, 88% correctly guess their partner's specific subcaste (*jati*), participants report that they are 89% confident in their caste category guess on average, 53% report being completely (100%) confident in their caste-category guess, and 97% predict that their partner will guess their caste category accurately (see the omitted category means for columns 1 to 5, Panel A, Table 2). These numbers confirm that caste is approximately common knowledge in the *Revealed* condition.

The *First Name Only* and *Hidden* conditions sharply reduce caste knowledge. Relative to the *Revealed* condition, correct caste category guesses fall by 14 and 22 percentage points (Table 2, Panel A, column 1), and correct *jati* guesses by 20 and 34 points (column 2). Participants are more uncertain about their guesses, with negative effects on confidence of 13 and 22 points (column 3), and even larger effects on the dummy variable for complete confidence (column 4). Meta-perceptions change, too: participants are less likely to believe that their partners can correctly guess their own caste category, and less likely to have accurate beliefs about whether the partner guesses correctly (columns 5 and 6). All pairwise differences are statistically significant at the 1% level, implying that *First Name Only* weakens common knowledge relative to *Revealed*, and *Hidden* reduces it still further. We further note that these estimates likely understate the effective first stage. By design, participants in *Revealed* learn their partner's caste immediately, generating four full hours of common knowledge. In the other conditions, by contrast, information likely filters through gradually over the course of the workday. Thus, although end-of-day guess rates are quite high, participants likely spend much of the interaction under conditions of uncertainty, with only incomplete and evolving information about their partner's identity.

When participants guess incorrectly, errors are not confined to adjacent caste categories. Among GC respondents (whose partner is SC), 54% of wrong guesses place the partner in the (intermediate) OBC category, but 41% classify

Table 2: Guessing partner's caste identity by treatment and caste group.

	Correctly guessed caste category (1)	Correctly guessed jati (2)	Confidence in caste category guess (3)	100% confidence in caste category guess (4)	Partner can guess my caste category (5)	Second-order belief accuracy (6)
A. Full sample						
First Name Only	-0.14***	-0.20***	-12.67***	-0.28***	-0.15***	-0.12***
	(0.03)	(0.04)	(1.90)	(0.05)	(0.03)	(0.03)
Hidden	-0.22***	-0.34***	-22.06***	-0.41***	-0.35***	-0.23***
	(0.03)	(0.03)	(1.86)	(0.05)	(0.04)	(0.03)
First Name Only vs. Hidden *p*-value	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Revealed mean	0.97	0.88	89.10	0.53	0.97	0.94
N	1,013	1,013	1,013	1,013	1,013	1,013
B. SC sample						
First Name Only	-0.14***	-0.20***	-12.27***	-0.24***	-0.17***	-0.12***
	(0.03)	(0.05)	(2.05)	(0.05)	(0.04)	(0.04)
Hidden	-0.16***	-0.28***	-22.31***	-0.37***	-0.34***	-0.24***
	(0.03)	(0.04)	(2.08)	(0.05)	(0.04)	(0.04)
First Name Only vs. Hidden *p*-value	0.66	0.13	<0.01	<0.01	<0.01	0.01
Revealed mean	0.98	0.90	89.06	0.51	0.96	0.93
N	506	506	506	506	506	506
C. GC sample						
First Name Only	-0.13***	-0.20***	-13.31***	-0.31***	-0.15***	-0.12***
	(0.03)	(0.04)	(2.09)	(0.05)	(0.03)	(0.03)
Hidden	-0.28***	-0.40***	-22.00***	-0.46***	-0.36***	-0.23***
	(0.04)	(0.05)	(2.01)	(0.05)	(0.04)	(0.04)
First Name Only vs. Hidden *p*-value	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Revealed mean	0.95	0.86	89.13	0.55	0.98	0.96
N	507	507	507	507	507	507
D. Difference in differences						
First Name Only × GC (vs. SC) *p*-value	0.81	0.92	0.52	0.03	0.51	0.90
Hidden × GC (vs. SC) *p*-value	<0.01	0.03	0.81	<0.01	0.64	0.89

Notes: Each column reports estimates from an OLS regression of the outcome on indicators for assignment to the *First Name Only* and *Hidden* treatments, with *Revealed* as the omitted category. Panel D reports difference-in-differences *p*-values from regressions that interact each treatment indicator with an indicator for GC participant (SC is the omitted group); the reported *p*-values correspond to the interaction terms. Outcomes are defined as follows: (1) equals 1 if the participant correctly guessed their partner's caste category; (2) equals 1 if the participant correctly guessed their partner's jati; (3) is confidence in the caste-category guess on a 0–100% scale; (4) equals 1 if the participant reported they were 100% confident in their caste-category guess; (5) equals 1 if the participant believes their partner can correctly guess their caste category; (6) equals 1 if the participant correctly predicted whether their partner would guess their caste category correctly. Standard errors are reported in parentheses. In Panels A and D, standard errors are clustered at the pair level. In Panels B and C (the SC and GC subsamples), standard errors are heteroskedasticity-robust. All specifications include randomization strata fixed effects and pre-specified controls for participants' education level and age. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

the SC partner as GC—the same caste category as the respondent. Similarly, among SC respondents (whose partner is GC), 63% of wrong guesses place the partner in the OBC category, while 32% classify the GC partner as SC. In sum, while many mistakes reflect movement toward the middle, a large share also involve perceived "passing" across the full caste-status divide. This passing is even more remarkable when put in the context of caste population shares in Uttar Pradesh: OBCs are roughly twice as common as either SCs or GCs; but we see frequent incorrect guesses of SC and GC despite their relatively smaller population shares.[17]

Effects on caste knowledge remain strong when considering GCs and SCs separately (Table 2, Panels B and C). Surprisingly, the effects of *First Name Only* are similar for both high and low-caste participants: we do not see evidence that SCs are hiding their identity more often than GCs in this "choice" condition. As we elaborate below, one explanation is that disclosure tends to be correlated—typically when one partner discloses (or discovers) identity, so does the other.

With that said, we observe substantial heterogeneity by caste when comparing the *Hidden* condition to *First Name Only*. SC participants are able to guess their partner's caste category with similar accuracy in these two conditions (Table 2, Panel B, column 1). In contrast, for GC participants, there is a large gap: their accuracy in *Hidden* is 15 percentage points lower than in *First Name Only* ($p < 0.01$), amounting to a nearly 30 percentage point loss in accuracy relative to the *Revealed* condition. This disparity suggests that SC individuals are more willing—or more effective—at concealing their identity when prompted to do so, perhaps because they are more practiced at concealing, or because high-caste individuals reveal their identity more readily (deliberately or inadvertently) through other social cues. A comparison of column 1 with column 5 speaks to SCs' greater sophistication when hiding caste: while both GCs (Panel C) and SCs (Panel B) expect their partners to make worse caste guesses in *Hidden* than in *First Name Only* (column 5), only SCs are correct in this belief (column 1).

How caste is revealed Our treatments affect the strategies by which participants infer the caste of their partner, as demonstrated in Table 3. In *Revealed*, 74% report having used the partner's last name to guess their caste, versus only 17% in *First Name Only* and 8% in *Hidden* (Panel A, column 2). In *First Name Only*, the most common explanation for how the participant guessed caste is that the partner told them their caste directly (42%). In these cases, participants rarely appear to misrepresent their caste: among those who reported that their partner had told them their caste, 97% correctly identified their partner's caste at endline. In addition, GCs and SCs are similarly likely to tell their partner their caste, ruling out an equilibrium with unraveling, in which only GCs (as the members of the more advantaged social group) reveal, making the signal from SCs' silence informative of their lower-caste

[17] More specifically, the wrong guesses of SCs are roughly consistent with a Bayesian guesser with information only on population shares. The wrong guesses of GCs are consistent with a Bayesian guesser with information on population shares in addition to a signal that the partner might be GC.

status.

After names, the way the partner spoke (36%) and where the partner lives (35%) were the next most commonly cited reasons given for guesses (excepting the "other" category). In *Hidden*, we see evidence that participants followed our instructions, as far fewer report that their partner told them their caste directly (15%). The most common answers are that they guessed based on the way the partner spoke (49%) or based on where the partner lives (32%). The frequency of participants organically revealing their caste in *First Name Only* might suggest that the *Hidden* treatment is costly in its paternalism: it prevents a substantial fraction of participants from playing their "best response," by revealing their caste. We explore this by estimating effects on downstream outcomes in the subsequent sub-sections.

Figure 1 breaks down the reported reasons for caste guesses according to the ultimate accuracy of those guesses. The overall finding is that accurate caste guesses rely disproportionately on direct or high-quality signals, whereas incorrect guesses lean more heavily on noisier, indirect ones. Among correct guesses (Panel A), participants most often mention knowing their partner's last name (38%), and being told caste directly (32%). By contrast, incorrect guesses (Panel B) are more likely to be based on speech and physical appearance, which are more ambiguous proxies for caste. This suggests that when clear identifiers are unavailable, participants substitute toward weaker heuristics.

If participants mentioned that physical appearance or speech cues as the basis for identifying their partner's caste, enumerators asked what specific aspects of these cues informed their judgment. Most responses referred to what their partner said or wore (Appendix Figure A4). Examples include: "He was talking about dairy work on the phone" (a Yadav (OBC)-associated occupation), "he said *Jai Bhim* on the phone" (an SC/Ambedkarite salutation), "he had a *tilak* and *kalawa*" (Brahmin (GC) religious markers), or "he spoke like an educated person." Although these heuristics are less accurate than explicit identifiers, participants tend to rely on them, and they appear to provide some signal. This may help explain why caste-identification rates remain high even in the concealed arms.

To better understand the high rates of correct caste identification in *First Name Only* and *Hidden*, we estimate a logit model predicting SC versus GC status using characteristics that could plausibly be observed, or quickly learned, in the workplace environment. These include demographics; family characteristics that are commonly discussed in Indian workplaces (such as marital status and household size); income; self-assessed ability; self-consciousness; and independent ratings of clothing, speech fluency, and whether the participant's Hindi sounded *shudh* ("pure") or not.[18] The model achieves 66.3% out-of-sample accuracy in a 50/50 train–test split, helping explain the high accuracy of caste guesses even in the concealment conditions (Appendix Tables A7 and A8). Importantly, clothing quality and voice ratings add little predictive power over and above the other observables. The predictive signal appears to be

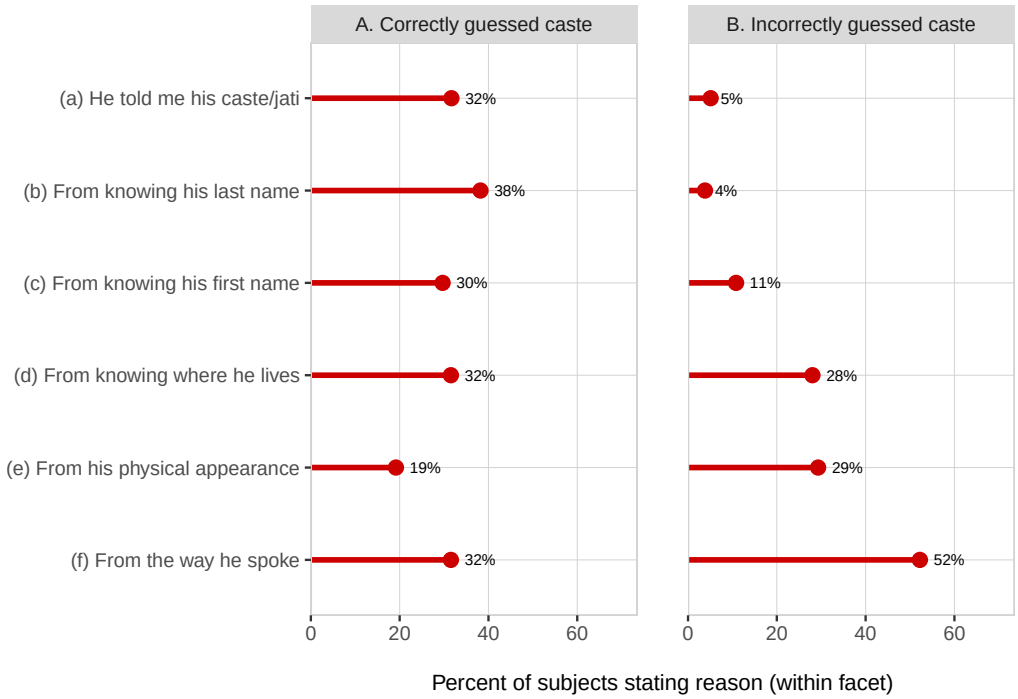
[18]The latter ratings were provided by three Indian research assistants, based on a photograph taken on the work day and two short voice recordings: one from a common prompt administered to all participants, and one from a participant-specific response describing future plans.

Table 3: Reported cues used to infer partner's caste identity, by treatment and caste group.

	He told me (1)	Used last name (2)	Used first name (3)	Used speaking (4)	Used appearance (5)	Used where lives (6)	Used other (7)
A. Full sample							
First Name Only	0.17***	-0.57***	-0.17***	0.18***	0.13***	0.09**	0.33***
	(0.05)	(0.04)	(0.05)	(0.04)	(0.03)	(0.05)	(0.04)
Hidden	-0.10**	-0.66***	-0.28***	0.29***	0.20***	0.07	0.48***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)
First Name Only vs. Hidden p-value	<0.01	<0.01	<0.01	0.01	0.08	0.60	<0.01
Revealed mean	0.25	0.74	0.42	0.19	0.10	0.26	0.45
N	1,013	1,013	1,013	1,013	1,013	1,013	1,013
B. SC sample							
First Name Only	0.14***	-0.58***	-0.15***	0.15***	0.14***	0.09*	0.32***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)
Hidden	-0.12***	-0.66***	-0.27***	0.29***	0.22***	0.04	0.46***
	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)
First Name Only vs. Hidden p-value	<0.01	0.04	<0.01	0.01	0.12	0.28	<0.01
Revealed mean	0.27	0.76	0.41	0.19	0.13	0.27	0.46
N	506	506	506	506	506	506	506
C. GC sample							
First Name Only	0.20***	-0.56***	-0.18***	0.20***	0.13***	0.09*	0.35***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)
Hidden	-0.08*	-0.66***	-0.29***	0.29***	0.18***	0.11**	0.50***
	(0.05)	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.04)
First Name Only vs. Hidden p-value	<0.01	<0.01	0.02	0.09	0.25	0.82	<0.01
Revealed mean	0.23	0.73	0.43	0.20	0.06	0.25	0.45
N	507	507	507	507	507	507	507
D. Difference in differences							
First Name Only × GC (vs. SC) p-value	0.14	0.56	0.37	0.36	0.77	0.93	0.62
Hidden × GC (vs. SC) p-value	0.25	0.91	0.50	0.94	0.52	0.13	0.34

Notes: Each column reports estimates from an OLS regression of the outcome on indicators for assignment to the *First Name Only* and *Hidden* treatments, with *Revealed* as the omitted category. Panel D reports difference-in-differences p-values from regressions that interact each treatment indicator with an indicator for GC participant (SC is the omitted group); the reported p-values correspond to the interaction terms. The outcomes are binary indicators for whether respondents selected each stated reason in response to the question: "How were you able to guess your partner's caste category or jati? Even if you are not completely sure about your guess, please tell us what made you guess the way you did. Select all that apply": (1) being told directly ("he told me his caste/jati"); (2) from knowing his last name; (3) from knowing his first name; (4) from the way he spoke; (5) from his physical appearance; (6) from knowing where he lives; and (7) other reasons. Standard errors are reported in parentheses. In Panels A and D, standard errors are clustered at the pair level. In Panels B and C (the SC and GC subsamples), standard errors are heteroskedasticity-robust. All specifications include randomization strata fixed effects and pre-specified controls for participants' education level and age. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Figure 1: Reported reasons for caste guesses, by accuracy of guess.



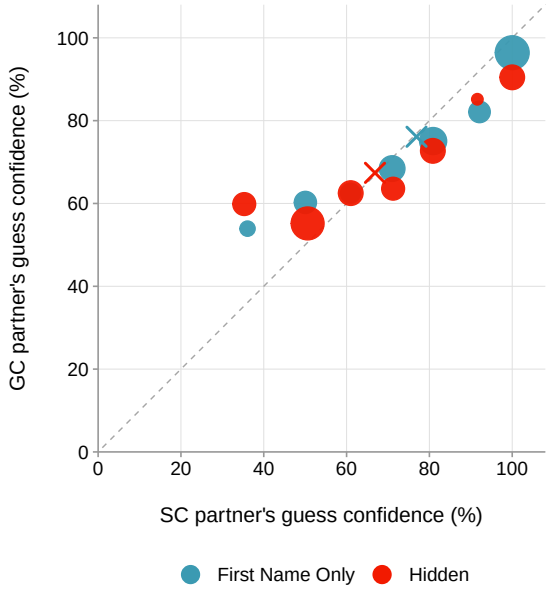
Notes: The unit of analysis is the participant. The outcomes are binary indicators for whether respondents selected each stated reason in response to the question: "How were you able to guess your partner's caste category or jati? Even if you are not completely sure about your guess, please tell us what made you guess the way you did. Select all that apply." Within each panel, the figure shows the percentage of participants citing each stated reason, conditional on whether the participant correctly or incorrectly guessed their partner's caste.

carried not primarily by how participants look or sound, therefore, but other observable features.

Correlated identity disclosure Worker pairs tend to either reveal their caste identities together or not at all. Pooling across the *First-Name Only* and *Hidden* arms, 231 pairs had neither person tell the partner their caste identity, 86 pairs had both tell it, and only 24 pairs involved one-sided disclosure. Consistent with this process of correlated disclosure, subjects' confidence in their caste guesses is strongly associated within pairs, for each concealment condition; that is, when the SC partner is confident in his caste guess, the GC partner tends to be confident in his guess also (Figure 2). It appears that the social script governing identity disclosure is reciprocal: if you tell me your caste, I will tell you mine. Such reciprocity norms inherent in conversations may in turn help explain why caste knowledge is symmetric across SCs and GCs in *First Name Only*, where caste can be learned about through talking, but asymmetric in *Hidden*, where conversation about caste is proscribed (a rule that was largely followed).

Social relations We now consider the first family of downstream outcomes. Recall the competing predictions of the discrimination and coordination hypotheses: the former predicts identity concealment to improve intergroup relations, while the latter predicts negative effects.

Figure 2: Correlation between partners' confidence in guesses about each other's caste.



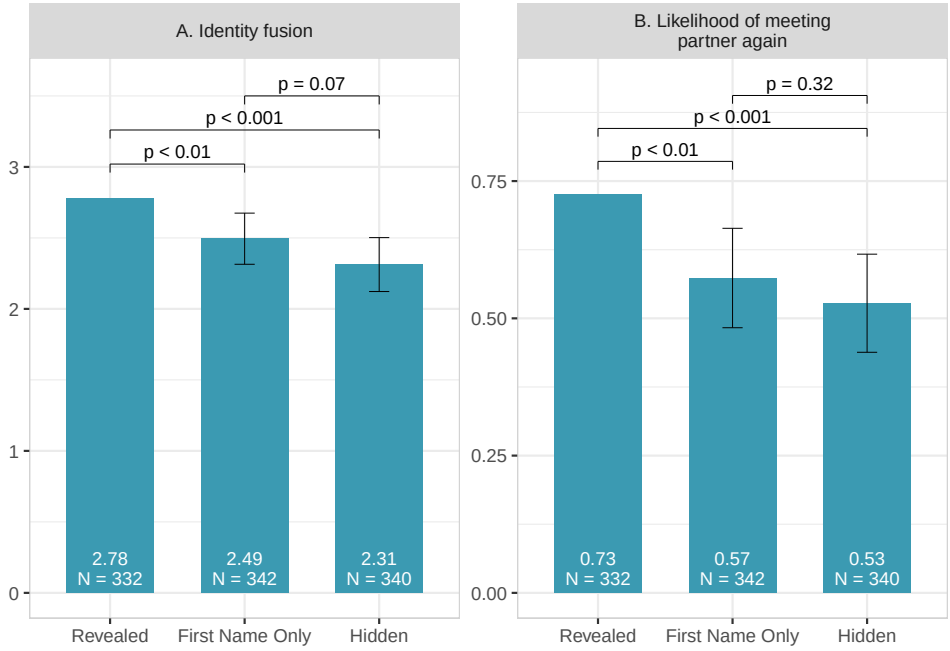
Notes: This figure plots, for each worker pair, each partner's confidence in their guess of the other's caste identity. The GC partner's confidence in their guess about their SC partner is given on the vertical axis, and the SC partner's confidence in their guess about their GC partner is given on the horizontal axis. Nearby values are grouped into bins, with point size proportional to the number of pairs in each bin. The crosses (×) denote arm means for the *First-Name Only* and *Hidden* treatments.

Our two identity-hiding conditions substantially worsen worker relations (Figure 3). Compared to *Revealed*, participants in *First Name Only* are less likely to report feeling close with their partner, using the 0 to 4 scale of identity fusion (Panel A, $\hat{\beta} = -0.29$ or 0.32σ, $p < 0.01$), and 15 percentage points less likely to report expecting to meet up with their partner again ($p < 0.01$), off a high base of 73% (Panel B).[19] These are large effects, especially given that (i) the *First Name Only* treatment is subtle—merely the use of first names rather than full names—and (ii) *First Name Only* participants can in principle replicate the information content of the *Revealed* condition by simply telling their full name to their partner. Social relations worsen still further in the *Hidden* condition, consistent with the stronger effects of *Hidden* on caste knowledge, though these declines are smaller and less precisely estimated than the effects of moving from *Revealed* to *First Name Only* ($p = 0.07$ for identity fusion and $p = 0.32$ for expectation of future interaction).

For an incentivized measure of worker relations, we offered subjects the choice between working another day with their partner, versus a same-caste stranger who had participated previously in the research, and varied the wage offered across these two choices. We started with a wage of Rs. 400 ($\approx$ \$4) for both choices and then gradually increased the offered wage for the less-preferred choice, to ascertain their switching point. Table 4 presents Tobit results where the outcome variable is the wage premium required to switch from working with the partner to working

[19] This high base cannot be attributed to pre-existing relationships: only 3% of participants in the sample reported encountering or having any knowledge of their work partner in the past.

Figure 3: Estimated treatment effects on social relations.



Notes: The unit of analysis is the participant. The outcomes are: (A) Which picture best describes your relationship with your work partner today? (0 = self and partner circles fully apart to 4 = self and partner circles fully overlapping); (B) Do you think you will meet up with your partner again in future? (0 = No, 1 = Yes). The *Revealed* bar in each panel shows the mean of the outcome variable for *Revealed* participants. The *First Name Only* and *Hidden* bars add the corresponding treatment effects, estimated from regressions that control linearly for participants' age, and include randomization-strata and education-bin fixed effects. The 95% confidence intervals are based on standard errors clustered at the pair level. Brackets report two-sided p-values for pairwise differences across arms.

with the stranger, with a larger wage premium reflecting stronger partner relations.[20]

Subjects in the *Revealed* condition would prefer to work with their partner rather than with a same-caste stranger on average, requiring a Rs. 173 wage premium to switch (omitted mean, Table 4, column 1). The *Hidden* treatment harms partner relations, reducing the wage premium by 21%, or Rs. 37 ($p = 0.08$). This is driven primarily by GC participants (column 2, Rs. 43, $p = 0.08$) and less so by SCs (column 3, Rs. 30, $p = 0.18$). The effects for *First Name Only* are relatively small and not statistically significant, perhaps because subjects retain the possibility of learning about their partner's identity—much as they might with a stranger—leaving them more willing to continue working with their partner relative to the *Hidden* condition, where this opportunity is foreclosed.

[20] If a subject reported indifference at some step, then this variable is equal to the midpoint between that wage and the next higher wage premium.

Table 4: Estimated treatment effects on the daily wage participants are willing to sacrifice to work again with partner rather than same-caste stranger.

	Full sample (1)	GC sample (2)	SC sample (3)
First Name Only	-4.07	-7.21	-0.56
	(21.59)	(25.16)	(22.76)
Hidden	-36.88*	-43.38*	-30.48
	(21.33)	(24.40)	(22.64)
First Name Only vs. Hidden p-value	0.12	0.14	0.19
Revealed mean	173	171	175
N	1,012	507	505

Notes: The unit of analysis is the participant. The outcome is the wage (in Rs.) that a participant would be willing to sacrifice to work with their partner again, rather than with a same-caste stranger. Tobit regressions control linearly for participants' age, and include randomization strata and education-bin fixed effects. Standard errors are clustered at the pair level. $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

Trust Caste concealment also harms trust. Participants in the *Revealed* condition share Rs. 27 (out of 50) of their original endowment on average (Figure 4, Panel A).[21] Both the *First Name Only* and *Hidden* conditions reduce trust giving by roughly 15% (both $p < 0.01$). Meanwhile, the *First Name Only* and *Hidden* arms are statistically indistinguishable from one another ($p = 0.84$).

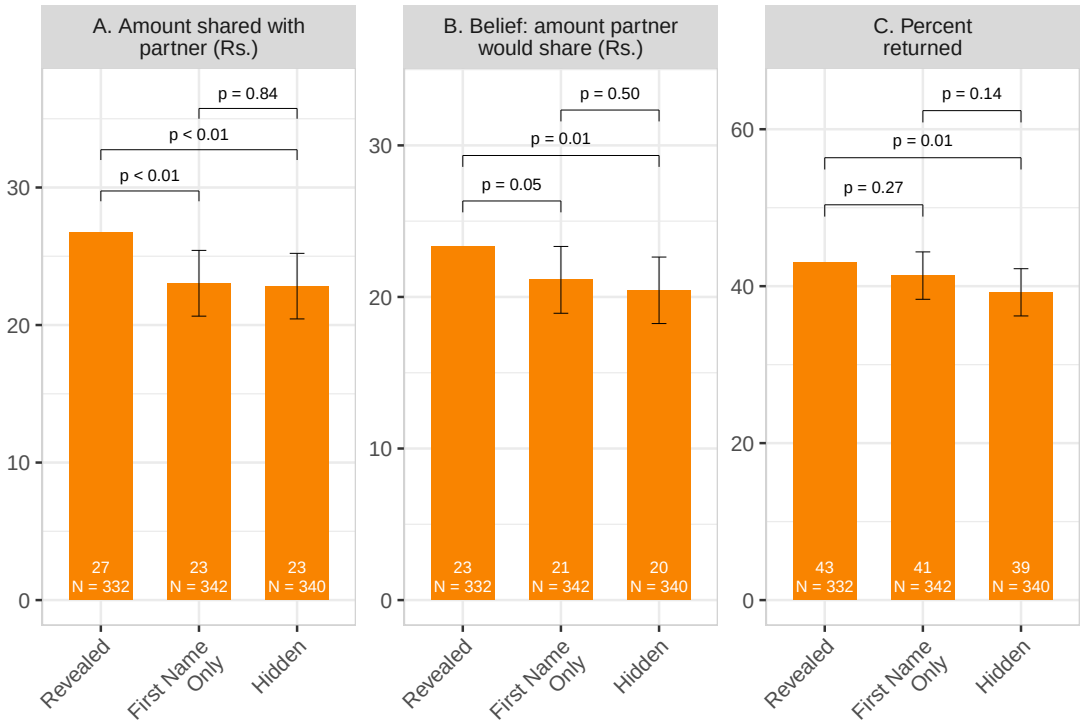
The similarity of the effects produced by the *First Name Only* and *Hidden* manipulations is telling. One concern is that directing participants to actively conceal their identity—as in *Hidden*—creates an artificially cagey setting, and that this is what dampens trust. But the *First Name Only* treatment involves no such instruction; rather, it withholds full caste-identifying information in a natural, unobtrusive way, yet the reduction in trust is near-identical. This suggests that what matters for trust is caste concealment, not concealment more generally.[22]

Negative effects on Sender behavior could be driven by beliefs (expecting the partner to return less) or preferences (feeling less altruistic toward the partner). To unpack these channels, we turn to our trust-game measures of beliefs and Receiver behavior. Participants in the *Revealed* condition expect their partner to share Rs. 23 of their endowment if selected as the Sender, compared to Rs. 21 and Rs. 20 in the *First Name Only* and *Hidden* conditions, respectively ($p = 0.05$ and $p = 0.01$ for the comparisons with *Revealed*). Once again, there is no significant difference between the *First Name Only* and *Hidden* treatment arms. The decline in transfers when identities are more opaque seems at least partly due to diminished expectations of reciprocity, therefore. Nevertheless, we also see a

[21] There is no statistically significant difference in the amounts sent by SC and GC participants.

[22] A separate concern is that higher trust in the *Revealed* treatment reflects a fear of being tracked down rather than genuine liking, since participants' last names are known to their partner. We think this is unlikely: tracking someone down requires far more cost and effort than the few rupees at stake in the trust game. But in any case, if the concern were valid, trust in *Revealed* should be less positively correlated with identity fusion (a direct measure of social connection) than in other arms. We find no evidence of this: the correlation between trust sent and identity fusion is 0.27 in *Revealed* and 0.23 in *First Name Only* (both $p < 0.001$), and the difference is not statistically significant ($p = 0.48$)

Figure 4: Estimated effects on trust game outcomes.



Notes: The unit of analysis is the participant. The outcomes are: (A) the amount (Rs. 0–50) sent to the partner when the participant is Player A; (B) the participant's belief about the amount (Rs. 0–50) the partner would send if selected as Player A; and (C) the percentage of the tripled amount the participant (as Player B) would return to their partner, averaged across all possible transfer levels. The *Revealed* bar in each panel shows the mean of the outcome variable for *Revealed* participants. The *First Name Only* and *Hidden* bars add the corresponding treatment effects, estimated from regressions that control linearly for participants' age and include randomization-strata and education-bin fixed effects. The 95% confidence intervals are based on standard errors clustered at the pair level. Brackets report two-sided p-values for pairwise differences across arms.

role for preferences: Panel C of Figure 4 examines treatment effects on actual reciprocity, measured as the percentage of the tripled amount that participants would return, averaged across all possible transfer levels. Participants in *Revealed* return 43% on average, compared to 41% in *First Name Only* and 39% in *Hidden*. The reduction in reciprocity is statistically significant for the *Revealed* versus *Hidden* comparison ($p = 0.01$) but not for the other two possible comparisons. For the case of reciprocal behavior, only the explicitly instructed hiding condition produces a detectable decline.

Taken as a whole, we see that treatments masking caste identity have measurable efficiency costs, cutting the total surplus generated in the trust game by around 15%. Reciprocity also declines, with participants returning a smaller share of a smaller pie.

Heterogeneity by caste According to the pooled results presented so far, caste concealment harms relationships and trust. At first glance, this appears to support the coordination hypothesis, since identity uncertainty makes interactions harder rather than easier. But pooled effects could obscure heterogeneity. The discrimination hypothesis

predicts that SC-reported relations should benefit from concealment because it shields them from discrimination; while if GCs have a preference for the hierarchy enabled by their caste status, they may be hurt more by concealment.

Appendix Table A5 finds that *First Name Only* and *Hidden* reduce identity fusion, willingness to meet the partner again, and trust for *both* GC and SC participants, firmly corroborating the coordination hypothesis. At the same time, the effects are more negative for GCs, with the difference statistically significant at at least the 10% level for two of the six tests. In addition, Table 4 showed previously that the negative effects on willingness to work together are also larger for GCs. These results suggest that while identity concealment undermines relations for both groups, the costs may be somewhat greater for the higher-status group in this context (GCs), consistent with a GC preference for hierarchy—albeit a limited one—alongside the more powerful coordination effects.

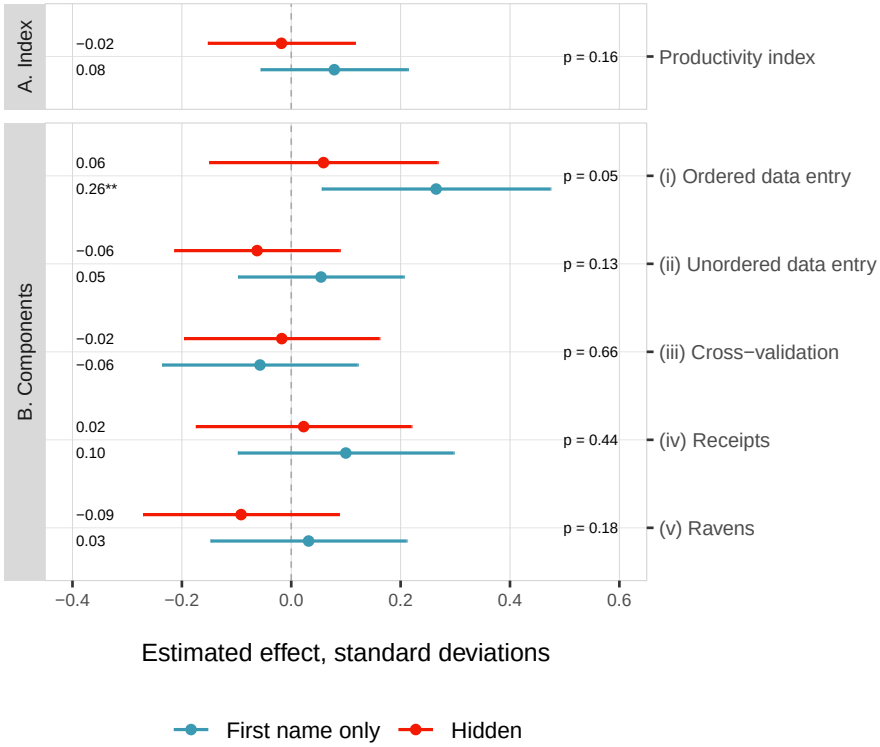
Productivity The harms to worker relations may be less concerning if accompanied by productivity and income gains. Do weaker social relations bring about a greater focus on work, as shown in Park (2019)? Figure 5 reports treatment effects on team productivity across the five tasks, and on an index measure summarizing those five outcomes. Neither the *First Name Only* nor *Hidden* condition produces a statistically significant change in overall productivity, and the difference between the two treatments is small (Panel A). Disaggregating into task components (Panel B) reveals only limited effects. The *First Name Only* condition increases productivity in the ordered data entry task, the simplest task, and the one performed first,[23] relative to both *Revealed* and *Hidden*. For all other tasks, however, the estimated treatment effects are small. The social harms of caste-hiding are not accompanied by economic benefits.

While productivity is not affected by caste-hiding, we do see some evidence of the elasticity of productivity with respect to social identity. Using our more limited data on same-caste teams, we find that same-caste teams have 0.16 standard deviations higher productivity on average ($p = 0.11$, Appendix Figure A5), consistent with other evidence from India that socially proximate teams perform better in collaborative tasks (Ghosh, 2025; Afridi et al., 2024). Strikingly, the lack of effects on *actual* productivity (Figure 5) contrasts sharply with Figure 6, which shows large negative effects on *perceived* productivity. Participants in both the *First Name Only* and *Hidden* conditions rate both their own and their partner's productivity substantially lower, by roughly one-third to one-half of a standard deviation relative to *Revealed*, with especially large effects among GC participants. The cost of concealment thus also operates through misperception: when caste identity is uncertain, participants appear to interpret the same work interaction less favorably and infer lower competence from it. Task structure and piece-rate incentives may discipline realized output. However, they do not prevent social friction from degrading how ability and performance are perceived,

[23] One interpretation is that the "first stage" on beliefs about caste identity is largest at the start of the day, explaining why we only see a productivity effect on the first task. But multiple hypothesis testing concerns are also plausible: of 12 tests in Figure 5, only one is statistically significant.

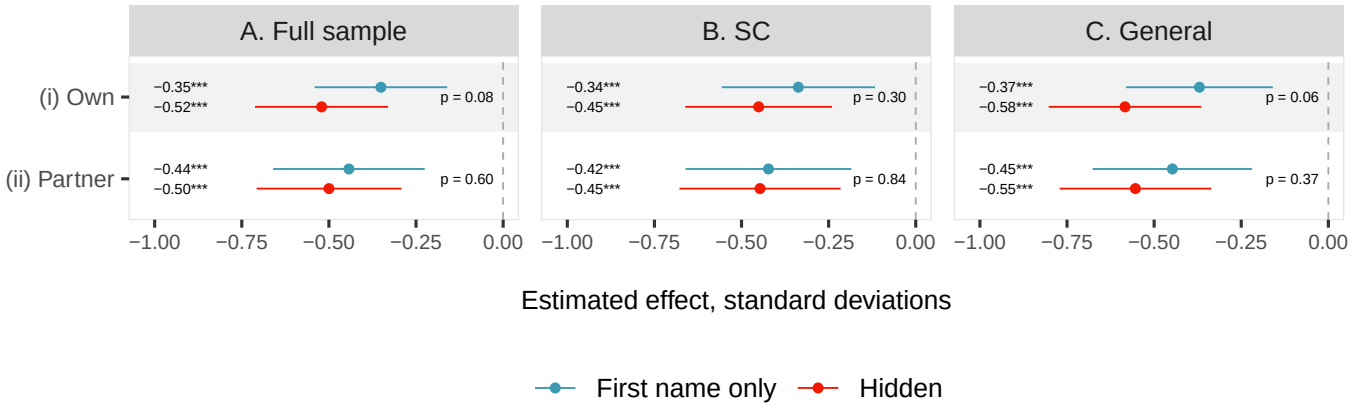
with potential implications for future economic ties (such as a higher reservation wage to work again and lower willingness to interact or collaborate).

Figure 5: Estimated treatment effects on team production.



Notes: This figure plots treatment effects on team productivity. All outcomes are counts that measure the number of correctly inputted items for each task. Counts are standardized to z-scores using the mean and standard deviation of the variable in the *Revealed* condition. The index is the simple average of the five standardized component outcomes. All regressions include date and education fixed effects, the average age of the pair, as well as mean baseline levels of productivity for the pair in both reading and typing. Standard errors are clustered at the pair level. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Figure 6: Estimated treatment effects on perceptions of own and partner productivity.



Notes: The figure reports estimated treatment effects on perceived productivity, measured separately for respondents' own performance and that of their partner. Outcomes are standardized to z-scores using the mean and standard deviation of the variable in the *Revealed* condition. All regressions control linearly for participants' age and include randomization-strata and education-bin fixed effects. Standard errors are clustered at the pair level. Horizontal bars denote 95% confidence intervals. Reported p-values correspond to tests of equality between the *First Name Only* and *Hidden* coefficients. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

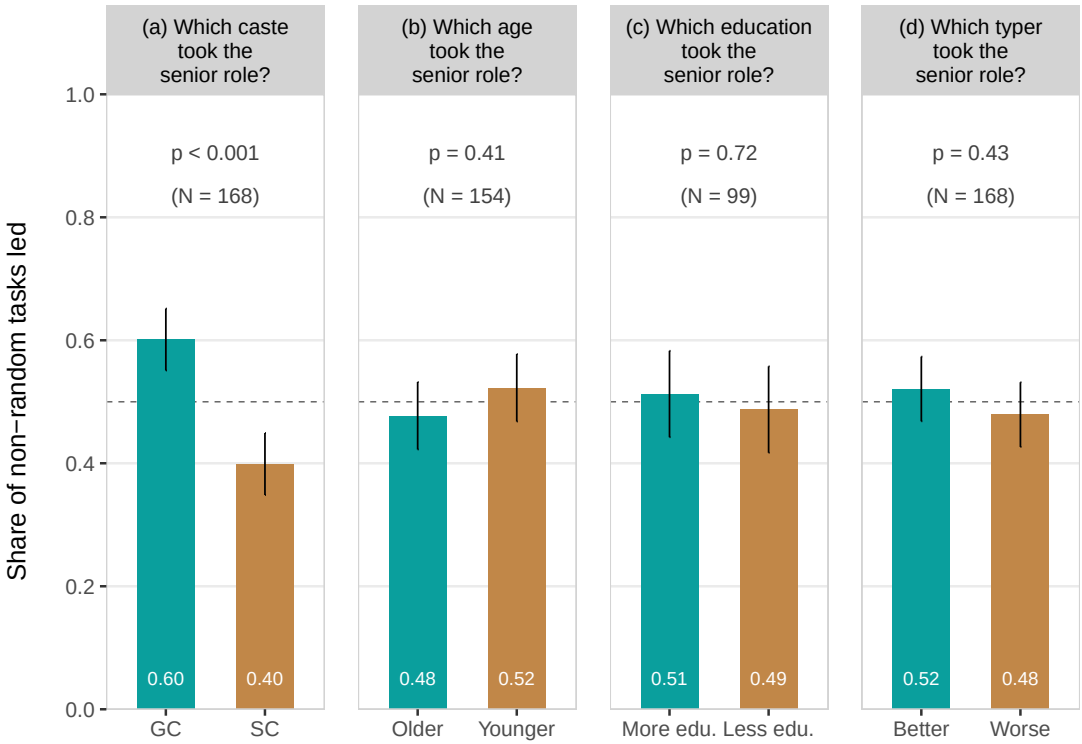
Role allocation The discrimination channel for identity-hiding makes a clear prediction: caste discrimination should be lower when caste cues are removed. We explore this idea by turning to data on ability and role assignment.

GCs and SCs perform similarly on baseline reading and tablet-typing tasks, which were designed to measure the skills required in the final workplace environment (Appendix Figure A7). The distributions for GC and SC participants substantially overlap, with GC participants having only modestly higher mean percentiles for both tasks (reading: 0.52 vs. 0.48; typing: 0.51 vs. 0.49). Despite similar ability by caste, GC participants take the senior role for 60% of tasks in *Revealed*, significantly above the 50% fair-split benchmark ($p < 0.001$, Figure 7, Panel (a)). This large caste disparity also holds when flexibly controlling for ability (Appendix Table A6). Caste identity, then, appears to be the key determinant of role assignment—we see no evidence of the senior role being taken by the older, more educated, or higher-ability partner (Figure 7, Panels (b)–(d)).

The removal of caste cues has no impact on the caste disparity (Table 5, column 1). The coefficients on both the *First Name Only* and *Hidden* conditions are small and statistically indistinguishable from zero. This suggests that caste discrimination—in which the GC partner pushes the SC partner into the junior role because of their lower-caste status—is not impacted by the removal of caste cues.[24] Similarly, the caste-concealment treatments do not give rise to meritocracy: they do not affect the probability of tasks being allocated to maximize total team productivity, which

[24]Of course, given that the allocation of roles is determined by the worker pair jointly, a related mechanism would be if the SC partner pushes the GC partner into the senior role because of their higher-caste status. Such a mechanism could reflect the SC's internalization of the caste hierarchy, or the SC pre-empting a backlash due to their expectations of caste discrimination. We think of these mechanisms as also falling within the discrimination channel.

Figure 7: Determinants of senior partner role assignment in the *Revealed* condition.



Notes: This figure reports how senior partner role assignment in the *Revealed* condition varies by four partner characteristics. Each panel shows, for one within-pair comparison, the share of the four tasks in which partners selected the senior role that was led by each partner type. The two bars in a panel are complements. The sample is restricted to GC-SC pairs. The samples in the age and education panels are restricted to pairs that differ on the relevant characteristic. The dashed line marks equal assignment. Error bars show 95% confidence intervals. *p*-values test whether the "higher-status" partner's share differs significantly from 0.5.

Table 5: Estimated treatment effects on senior partner role assignment.

	GC is senior (1)	Optimal typer is senior (2)
First Name Only	0.00	0.02
	(0.03)	(0.04)
Hidden	-0.00	0.01
	(0.03)	(0.04)
First Name Only vs. Hidden p-value	0.83	0.81
Revealed mean	0.60	0.48
N	2,028	2,028

Notes: Each column reports estimates from a linear regression. The unit of analysis is the pair-task. Tasks for which the senior role was randomly assigned (by the computer) are excluded. The outcome in column (1) is an indicator for whether the GC participant occupied the senior partner role. The outcome in column (2) is an indicator for whether the worker assigned to the senior (typing) role is the optimal typist. The optimal allocation assigns the senior role to the worker who maximizes total team productivity, defined as the assignment that yields the higher sum of within-sample percentile ranks in baseline reading and typing performance across the pair. All specifications include randomization-strata effects. Standard errors, clustered at the pair level, are reported in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

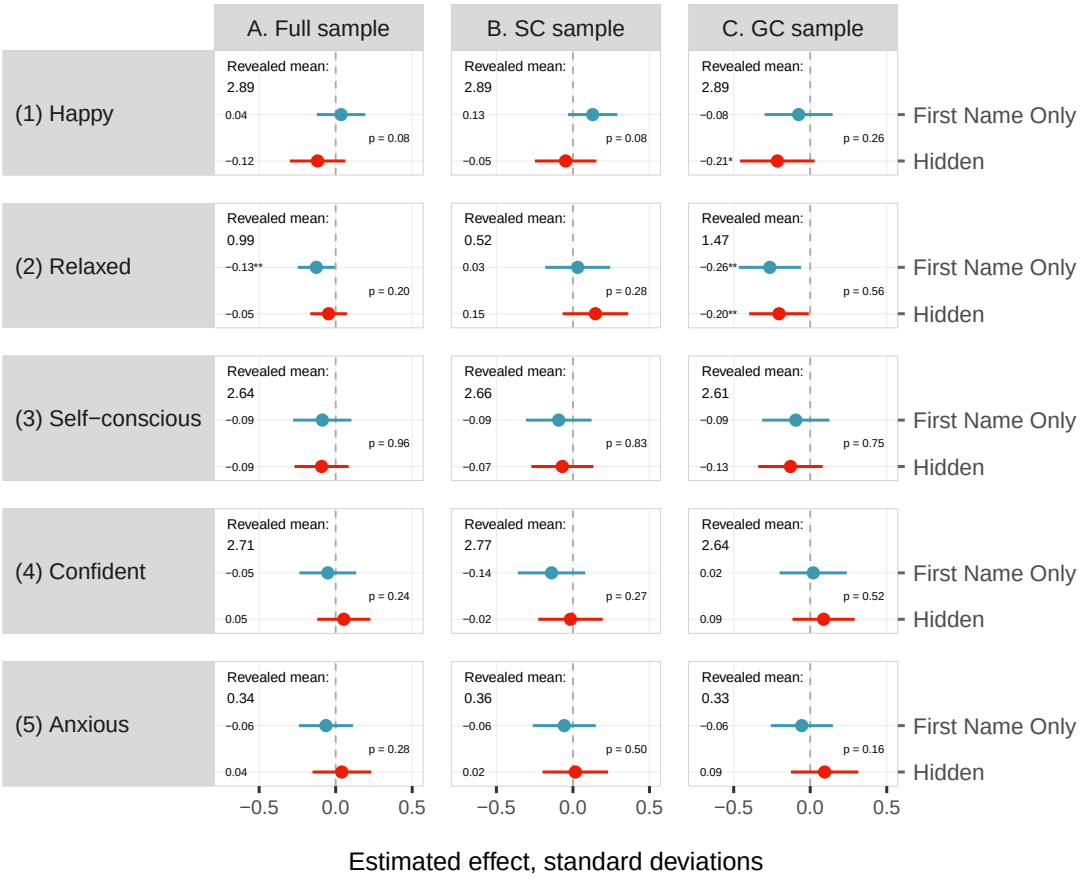
we define as the role assignment that yields the higher sum of within-sample percentile ranks in baseline reading and typing performance across the pair (column 2).

Our findings suggest that the caste disparity in task allocation is driven less by social perception (of caste identity), and more by self-perception. Indeed, GCs self-assess as more productive in the typing (i.e., senior) task despite similar measured ability, and they report being more relaxed during the interaction (see Appendix Figure A6). These differences may help explain why they are more likely to be assigned the senior role. Finally, the caste disparity in role allocation appears to be accepted by participants: pooling all treatment groups, 92% of GCs and 92% of SCs report at endline that they were happy with how the roles were assigned.

Well-being at work Beyond worker relations and productivity, one way to capture the welfare effects of identity concealment is the well-being felt by the participants at the workplace. Treatment effects on self-reported well-being tend to be small and not statistically significant (Figure 8). Overall, concealing caste identity does not substantially affect participants' subjective experience of the interaction along these affective dimensions.

We note one notable exception: there is some evidence that identity hiding has a more positive effect on happiness and feeling relaxed for SCs than for GCs, consistent with priors—under the discrimination hypothesis—that the marginalized group has more to gain (and less to lose) from the removal of caste cues. For GCs, we see negative effects: both the *First Name Only* and *Hidden* conditions reduce feelings of being relaxed at work (both $p < 0.05$), while the *Hidden* condition also reduces happiness ($p < 0.1$). Among the set of affective dimensions for which we detect effects, therefore, the costs of concealment appear to be borne more by GCs.

Figure 8: Estimated treatment effects on well-being.



Notes: Each facet panel reports estimates from an OLS regression of an outcome on indicators for assignment to the *First Name Only* and *Hidden* treatments, with *Revealed* as the omitted category. Outcomes are based on responses to the question, "At work today, to what extent did you feel...[word indicated on each row]" (0 = Not at all, 1 = A little, 2 = Some, 3 = Very). In each facet, "Revealed mean" shows the average score on the original 0–3 point scale for the *Revealed* group. Treatment effects are standardized. In Panels A, standard errors are clustered at the pair level; in Panels B and C (the SC and General subsamples), standard errors are heteroskedasticity-robust. All specifications include randomization strata fixed effects and pre-specified controls for participants' education level and age. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Mechanisms

Why does the removal of identity cues harm worker relations? We find evidence supportive of the coordination hypothesis described in the introduction. We argue that the identity uncertainty induced by our concealment treatments hinders smooth and authentic social interaction, given that participants are unsure who the person sitting across from them "really is." In contexts where misreading identity is socially consequential—because caste creates norms around hierarchy, ritual purity, and social distance—uncertainty produces guardedness and self-monitoring. We now present evidence for this interpretation.

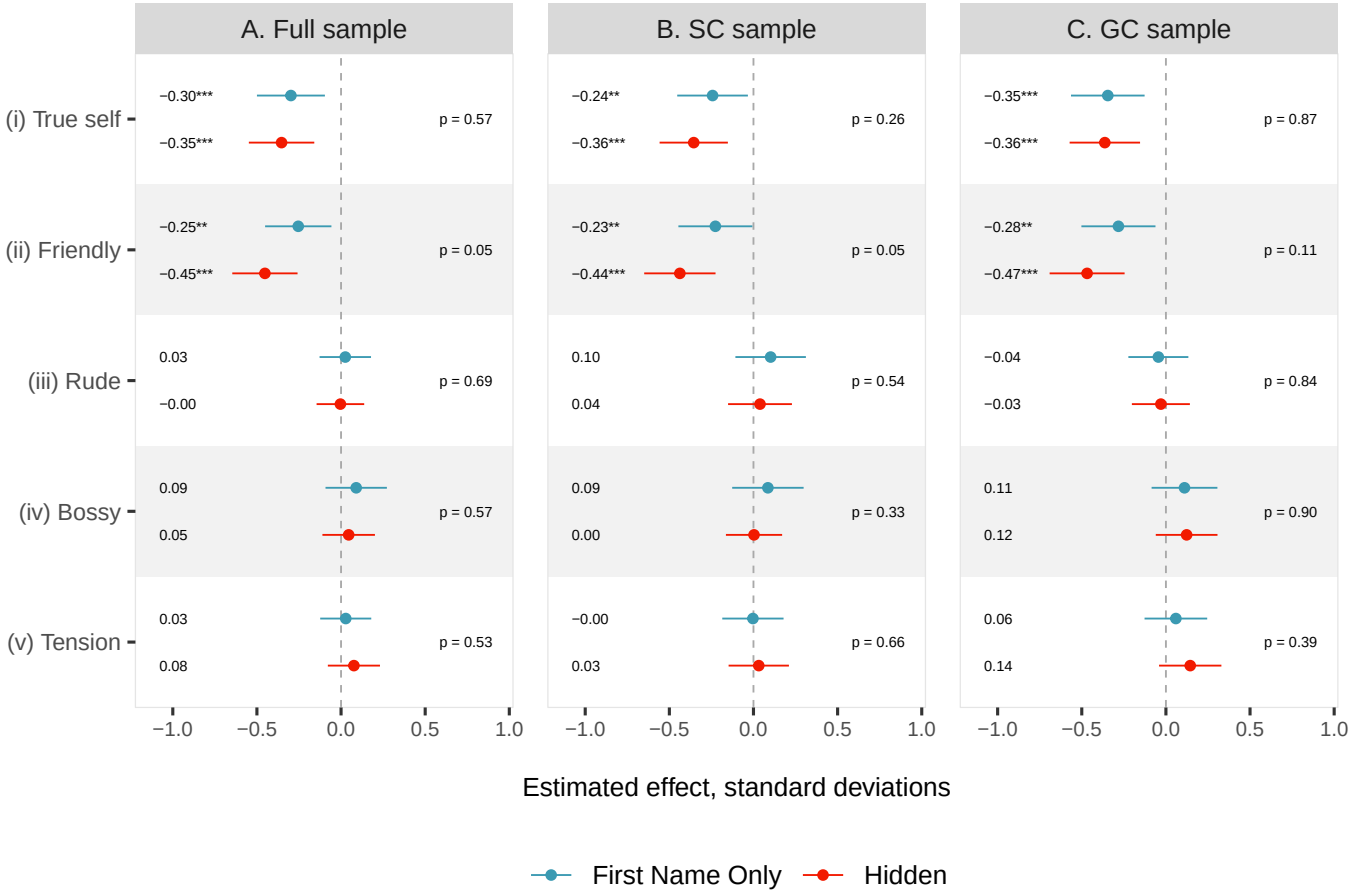
Authenticity and guardedness The loss of identity cues makes interactions feel less authentic and less friendly. In Figure 9, we analyze a set of pre-registered secondary outcomes capturing participants' subjective experiences during the interaction itself. Participants in *First Name Only* and *Hidden* report feeling substantially less able to show their "true self" to their partner ("authenticity" from now), with declines of between 0.24 and 0.36σ relative to *Revealed* (row (i), $p < 0.01$). Concomitantly, concealment also reduces perceptions that the partner was friendly by between 0.25 and 0.45σ (with *Hidden* producing a measurably larger fall, row (ii)). Across the board, the treatment effects are similar for SCs and GCs, consistent with a failure of conversational coordination that hurts all parties. In contrast, we do not find effects of concealment on overtly adversarial behaviors, such as displays of rudeness, bossiness, or tension (rows (iii)-(v), Figure 9). Rather than producing openly hostile or discriminatory interactions, identity-hiding appears to make participants more guarded, superficial, and cautious—inhibiting their willingness to interact naturally and comfortably with one another.

If caste helps structure how individuals interpret and manage social interactions, ambiguity about caste should be most disruptive for those who view caste as an important part of their personal identity. The results in Appendix Figure A8 chime with this interpretation. We estimate treatment effects of concealment, pooling *First Name Only* and *Hidden*, by the baseline importance participants attach to caste relative to other identities (being Indian, Hindu, from Uttar Pradesh, and personal relationships such as being a father, brother, or son). Participants who rank caste among their top two identities experience a negative effect on authenticity that is nearly twice as large as the effect among those who rank caste last.[25] We see a similar pattern for anxiety: because lower ranks indicate greater caste salience, the negative slope we observe (-0.087, $p < 0.1$) implies that anxiety rises more under concealment among participants for whom caste is more salient. These patterns suggest that concealment is especially disruptive for participants for whom caste is an important part of how they see themselves and navigate interactions with others.

Confidence We next probe the role of uncertainty directly, by investigating whether social relations improve when participants are more confident about their partner's caste identity. Appendix Table A10 reports treatment effects on our core social outcomes before and after controlling for participants' confidence in their caste guess and their beliefs about the partner's caste. Panel A replicates the main experimental results: relative to *Revealed*, both *First Name Only* and *Hidden* substantially depress the amount of money sent in the trust game, identity fusion, willingness to meet again, and authenticity. Panel B adds controls for caste-guess confidence, rescaled from 0 to 1, along with interactions between the respondent's caste and their guess about their partner's caste. Two findings stand out. First, confidence is strongly associated with better outcomes. Moving from no confidence to full confidence in one's caste guess is associated with sending Rs. 19 more in the trust game, reporting 1.6 more units of identity fusion

[25] We pool the first two ranks because there are few observations in the first category for caste.

Figure 9: Estimated treatment effects on subjective experiences during the interaction.



Notes: Each facet panel reports estimates from an OLS regression of an outcome on indicators for assignment to the *First Name Only* and *Hidden* treatments, with *Revealed* as the omitted category. Outcomes are self-reported measures of interaction quality: (i) ability to show one's true self (1 = not at all, 10 = completely); (ii) partner friendliness (1 = very unfriendly, 10 = very friendly); (iii) partner rudeness (1 = not at all rude, 10 = very rude); (iv) partner bossiness (1 = not at all bossy, 10 = very bossy); and (v) experienced tension (1 = no tension, 10 = a lot of tension). Treatment effects are standardized. In Panels A, standard errors are clustered at the pair level; in Panels B and C (the SC and GC subsamples), standard errors are heteroskedasticity-robust. All specifications include randomization strata fixed effects and pre-specified controls for participants' education level and age. $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

and a 50 pp higher likelihood of meeting up again, and scoring 5.9 points higher on feeling like one's true self (all $p < 0.01$). Second, the treatment effects largely disappear once we control for caste-guess confidence: all but one of the coefficients on *First Name Only* and *Hidden* become statistically indistinguishable from zero. Since confidence is measured post-treatment, this should not be read as causal mediation. But it is the regularity we would expect if concealment worsened social relations by generating uncertainty about a caste identity.[26]

We explore the role of certainty non-parametrically in Appendix Figure A9. Restricting attention to participants in the concealment arms, we examine how social outcomes vary with confidence in the partner's caste identity, using seven bins: below 50%, 50-59%, 60-69%, ..., 90-99%, and 100%, and including our pre-registered controls, caste-guess controls, and a dummy for the *Hidden* treatment, making all comparisons within-treatment. As above, we stress that this exercise is correlational, and thus more exploratory than our comparisons across treatment groups.

The confidence-relations relationship is non-linear. Outcomes improve discretely when participants report being 100% certain about their partner's caste. Participants who are fully certain show higher trust, greater identity fusion, and a higher perceived likelihood of meeting their partner again than those who are only partially certain—with especially discrete jumps at complete certainty for identity fusion and trust. This "certainty premium" hints that even a seed of doubt about group identity may be enough to keep participants guarded.[27] This would align with a model in which the social costs of misreading identity are large, making participants err on the side of caution, rather than proceeding as if their caste guess is correct.[28] Future experiments could test for this idea causally: with a *Certain* treatment group, in which participants are told that their partner is an outgroup member, along with a *Near Certain* treatment group, in which participants are told that their partner was randomly selected from a pool of 19 outgroup members and one ingroup member.

Lastly, we also find that instances of endogenous identity resolution are associated with higher trust (Appendix Table A11). Participants who report having asked their partner's caste send more in the trust game and expect their partner to send more. These associations are stronger among GC participants, in line with higher-caste participants being more uncomfortable when partner identity remains uncertain. The evidence also suggests that identity resolution can reassure SC participants: they send more to GC partners when the partner has learned their caste. Overall, these correlations point in the same direction as the experimental results: interactions are more trusting when uncertainty about caste identity is resolved.

[26] Further, using the mechanism test in Kwon and Roth (2026), we fail to reject that caste-guess confidence—binned into seven categories (0–50%, followed by 10 percentage-point increments)—fully mediates the treatment effects.

[27] We note also a parallel to the work on risk preferences of Callen et al. (2014). They find a preference for certainty in violation of expected utility theory.

[28] One useful analogy is that of being 99% confident in knowing the first name of a conversation partner. Given large social costs of misnaming an individual, the 1% doubt may well be enough to deter one from attempting to introduce the partner to others. In the case of imperfect knowledge of caste identity, 1% of doubt may be enough to avoid conversational topics and norms that would fit well with one's guess of the partner's caste.

Alternative explanations

Anonymity and information exchange A subtly different possibility is that concealment harms social relations simply by reducing the set of topics that participants discuss. Since caste is correlated with many other characteristics in this setting, instructions to hide caste may also restrict discussion of education, family background, religion, marriage, or other identity-linked topics that people use to build rapport. Appendix Figure A10 shows evidence for this in the *Hidden* condition: participants are less likely to ask directly about caste and also ask less about several other identity-linked characteristics. These participants are also less likely to report having discussed a range of topics (Figure A11).

However, reduced conversational scope cannot fully explain the negative effects of concealment. In *First Name Only*, participants are not instructed to avoid caste-revealing information, and many use this discretion. Caste is asked more often than in *Revealed*, where surnames already reveal caste (row C, Appendix Figure A10). More generally, conversations in *First Name Only* are broad and look much closer to *Revealed* than to *Hidden* in terms of topics discussed (see Figure A11). Yet worker relations remain worse in *First Name Only* than in *Revealed*. This suggests that the relevant margin is not simply whether participants exchange more information, but whether the social context within which the interaction is understood is kept intact. This pattern then helps rule out a generic anonymity or communication effect, in which prosociality increases in response to concrete information about a partner (Small and Loewenstein 2003; Andreoni and Rao 2011; Goette et al. 2012).

Experimenter demand and salience The *First Name Only* condition does not mention caste, limiting the possibility of demand effects, and yet produces negative effects that are similar to those in *Hidden*, where caste-related disclosure is explicitly restricted. Furthermore, only 3.8% of subjects correctly guessed the study's purpose, with no effects of treatment (see Appendix Figure A12 and Appendix Table A12). Of course, the larger penalties sometimes observed in *Hidden* may reflect the added effects of a narrower conversational scope and greater caste salience, on top of identity uncertainty.

Restoration of caste hierarchy We rule out the possibility that the effects are driven by familiarity with a status quo in which caste hierarchy is known. This explanation would predict that the benefits of revelation are mainly concentrated among GCs, who benefit most from the existing hierarchical order. Instead, the negative effects of concealment are large for both GCs and SCs. Revelation also does not increase the rate at which GCs take the senior partner position. This collection of findings suggests that the negative effects of concealment are not simply driven by the disruption of a familiar caste hierarchy.

Beliefs about team composition Another possibility is that the treatment effects are driven by beliefs about team composition. In *Revealed*, participants *know* that they are in GC-SC pairs. In *First Name Only* and *Hidden*, they assign some probability of being paired with a same-caste partner, and some probability of being paired with an OBC partner, closer to them in the caste hierarchy than their actual partner. But this cannot explain the negative effects of concealment. If anything, believing that one's partner's caste is less distant should *improve* social relations. Consistent with this, Appendix Table A9 shows that homogeneous-caste teams included for obfuscation exhibit higher trust and greater identity fusion than GC-SC teams when identity is revealed, while as we showed earlier, homogeneous-caste teams are more productive (Appendix Figure A5). Finally, our exploration of identity uncertainty as a mediator of treatment effects in Appendix Table A10 included controls for beliefs about team composition.

Concealment to aid productivity Incomplete revelation in *First Name Only* does not appear to reflect a strategy to maximize productivity or affective benefits from avoiding caste discussion. Although we find weak positive effects of *First Name Only* on productivity, perceived productivity nevertheless declines, suggesting that concealment is unlikely to reflect a conscious attempt to improve performance. Nor does non-disclosure appear to make participants better off emotionally: *First Name Only* participants are not happier than those in *Revealed* and report lower authenticity.

Assuming further distant social group Concealment may worsen outcomes if participants respond to an unidentified partner by assuming that they belong to an even more socially distant group than the reality (Aigner and Cain, 1977; Bronchal, 2023). For example, a GC participant may respond more positively to an SC partner once revelation rules out the possibility that the partner is Muslim. The data do not support this explanation. Muslim guesses are rare in the concealment conditions ($< 1\%$), and as we showed earlier, wrong guesses are not systematically in the most pessimistic direction: both GCs and SCs often guess that their partner is OBC.

Selective revelation

The negative effects of *First Name Only* raise a natural question: why do participants not simply reveal caste and recreate the more positive outcomes observed in *Revealed*? We address this question with a simple framework that predicts selective revelation and persistent uncertainty, both of which we observe in the data.

Conceptual framework Consider two workers who have been paired for a joint task. Each worker can either reveal caste identity or hide it, where revelation includes both explicit and implicit ways of communicating caste identity during the interaction. The experimental treatments correspond to restrictions on this choice. In *Revealed*

and *Hidden*, workers do not have a choice: caste identity is revealed in the former and hidden in the latter (provided workers follow our instructions). In *First Name Only*, workers can choose.

The reason this choice matters is that mutual revelation can be valuable because it clarifies the social context of the interaction. Knowing a partner's caste may make it easier to interpret behavior, establish common ground for conversation, and anticipate how one's own behavior will be read. For same-caste pairs, we assume mutual revelation is beneficial. For cross-caste pairs, however, the value of revelation depends on expected interaction quality. GC participants do not expect caste-based domination at work, so revelation mainly reduces uncertainty. SC participants, by contrast, may differ in their expected payoff from revelation: for some, it may facilitate smoother interaction, while for others it may expose them to hierarchy, stigma, or domination.

The potential benefits of mutual revelation, however, do not imply that disclosure will occur automatically. In this setting, revelation is also a coordination problem. If one participant reveals caste while the other does not, the interaction may become awkward or unpleasant. High-caste participants may be reluctant to signal status directly, while low-caste participants may fear exposure without reciprocal disclosure. This prevents immediate unraveling. If unilateral revelation were costless and always beneficial for high-caste participants, they would reveal, implicitly identifying non-revealing partners as low caste, and *First Name Only* would collapse into *Revealed*. Instead, because one-sided revelation is costly, participants may prefer mutual concealment to being the only person to disclose.

Together, these forces generate two possible regimes (equilibria) in *First Name Only*. In one, no one reveals caste identity: mutual concealment persists because each participant prefers not to reveal unilaterally if they expect the other to hide. In the other, GC participants reveal, while SC participants reveal only if their expected payoff from mutual revelation is sufficiently high. Revelation by SC participants therefore follows a cutoff rule: SC participants who expect disclosure to improve the interaction reveal, while those who expect it to expose them to caste domination, or to bring little benefit, conceal.

Therefore, voluntary disclosure need not undo the effects of initially obscuring caste. Some pairs can move from concealment to mutual revelation, but these pairs should be *positively* selected: they are the pairs for whom the expected gains from revelation are high and the perceived risks are low. Other pairs may remain ambiguous, either because revelation is too risky for one party or because neither participant wants to disclose first. Thus, even when identity certainty is valuable, *First Name Only* can still generate uncertainty and reduce interaction quality relative to a setting where identity is already common knowledge. Lastly, if participants can exchange cues before deciding whether to reveal, as is natural in our setting, then the awkwardness of one-sided revelation also implies that unilateral disclosure should be rare. More broadly, unresolved uncertainty should be especially costly for GC participants, for whom mutual revelation primarily resolves uncertainty. Both patterns occur in our data.

Empirical evidence on selection into revelation and the persistence of uncertainty We take these implications to the data by studying when participants choose to disclose identity-relevant information. Our framework suggests that disclosure in *First Name Only* should be selective, occurring only when it feels sufficiently safe or reciprocal. The evidence below is consistent with this logic. Revelation is indeed selective, and the option to reveal appears most valuable in interactions with greater potential to develop into positive social ties.

Disclosure is strongly predicted by socio-economic similarity within the pair (Table 6). Participants paired with someone in the same income bracket are about 18 percentage points more likely to reveal caste, and those paired with someone at the same education level are about 25 percentage points more likely to do so. Relative to a mean disclosure rate of 42 percent, these effects are large, increasing revelation by roughly 40–60 percent. By contrast, own education, income, caste status, and age have little independent predictive power. This suggests that revelation is shaped less by absolute status than by relative social distance: participants are more willing to resolve uncertainty when the partner is socially proximate, and hence when the perceived risks of disclosure are lower.

The endogenous interaction measures point in the same direction (column 4). Partner disclosure is by far the strongest predictor of own revelation, suggesting that disclosure is often reciprocal rather than an individual decision made in isolation. Partner friendliness is also positively associated with disclosure, while partner bossiness and rudeness are negatively associated with disclosure, though the latter estimates are not statistically significant. These patterns are descriptive, but they suggest that participants resolve uncertainty when the interaction appears safer, more reciprocal, and more socially promising.

We also present quantile treatment effects that provide complementary evidence to the above (Figure A13). We estimate these for two summary outcomes: an index of interaction quality and an index of downstream social outcomes.[29] The option to reveal matters most higher in the distribution, where interactions had greater potential to develop into positive social ties: outcomes in *First Name Only* approach those in *Revealed*, while *Hidden* prevents this adjustment. At the lower end, we find little evidence that concealment makes participants worse off; interaction quality does not rise, and effects on trust and cooperation are muted. The main cost of concealment thus appears to be that it limits the development of otherwise promising interactions. This pattern also helps interpret selection into revelation: participants reveal selectively in ordinary workplace interactions, a common context of GC-SC exposure. Where disclosure carries sharper risks, concealment may be more valuable and identity uncertainty more persistent.

[29]The interaction-quality index combines authenticity and reports of partner friendliness, rudeness, bossiness, and tension. The downstream social-outcomes index combines trust, willingness to meet again, identity fusion, and reservation wages for future work with the same partner.

Table 6: Who reveals their caste to their work partner?

	(1) Own chars.	(2) Pair similarity	(3) Own + similarity	(4) Own + sim. + endogenous†
Own characteristics				
General caste (=1)	0.002		−0.001	−0.013
	(0.027)		(0.026)	(0.040)
Age	0.001		−0.004	−0.003
	(0.008)		(0.008)	(0.007)
Married	0.005		0.022	−0.005
	(0.075)		(0.075)	(0.061)
Completed 12th standard	0.094*		0.031	0.028
	(0.053)		(0.051)	(0.050)
Log HH income (bracket midpoint)	−0.051		−0.031	−0.004
	(0.031)		(0.032)	(0.022)
Pair similarity				
Age gap (years)		0.000	0.002	0.001
		(0.010)	(0.010)	(0.003)
Same income bracket as partner		0.182**	0.169*	0.041
		(0.092)	(0.095)	(0.033)
Same education level as partner		0.245***	0.241***	0.065*
		(0.082)	(0.086)	(0.034)
Endogenous controls (col. 4 only)†				
Partner revealed their caste				0.690***
				(0.086)
Partner friendliness (1–10)				0.042***
				(0.016)
Partner rudeness (1–5)				−0.017
				(0.027)
Partner bossiness (1–5)				−0.005
				(0.030)
Tension with partner (1–5)				0.009
				(0.033)
Mean dep. var.	0.424	0.424	0.424	0.424
Observations	342	342	342	342
Date FE	Yes	Yes	Yes	Yes

Notes: Sample is the *First Name Only* arm throughout. Outcome is an indicator equal to 1 if the respondent revealed their caste to their work partner, measured using the partner's report to avoid same-source bias. Columns 1–3 include only pre-determined baseline characteristics. Pair similarity variables equal 1 if respondent and partner share the same value: same income bracket; same education level (both completed 12th standard or both did not). Age gap is the absolute difference in years. Own income is the log of the household income bracket midpoint. All specifications absorb date fixed effects. Standard errors (in parentheses) are clustered at the pair-level.

† Col. 4 adds controls measured at endline that are potentially endogenous: whether the partner revealed their own caste (a simultaneous decision, measured using the subject's report) and interaction quality measures (partner friendliness, rudeness, bossiness, and tension).

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Conclusion

Debates over social integration tend to center on the relative virtues of embracing group identities—the so-called "politics of recognition" (Taylor, 1992; Ambedkar, 2014)—versus suppressing them altogether, in favor of more universal, individualistic forms of belonging (Barry, 2002; Scott, 2007). This dichotomy overlooks a third, and arguably more realistic, scenario: that group identities persist, yet differ in how readily they can be discerned in the course of everyday interactions. In an experiment in North India, we varied the visibility of caste identity among pairs of workers to investigate what happens when identities are unknown or ambiguous. We find that the concealment of group membership starkly diminished trust and interpersonal bonds; it also made interactions feel less smooth and authentic, even as it left a disparity in job-role allocations intact. Our findings thus underscore identity *legibility* as a consequential and underappreciated driver of intergroup cooperation. In contexts where identities and hierarchies are braided into social routines and relationships, individuals may come to rely on them as a source of convergent expectations that help them "make sense" of unfamiliar others (Schelling, 1960; Goffman, 1959). Under these circumstances, knowledge of group identity may grease the wheels of social interaction, whereas uncertainty may throw sand in the gears.

Our findings carry at least two practical implications. First, many economic interactions rely on weak ties—especially in the informal labor markets our study sought to emulate, and which remain pervasive throughout the developing world (Granovetter, 1973). Warm interpersonal interactions set the stage for exchanging useful information, securing referrals, and accessing future work. We find that despite little effect on actual productivity, concealment lowers perceived productivity and strains partner relations. It thus seems reasonable to speculate that identity-blinding schemes in the workplace may leave a lasting imprint beyond the immediate interaction. Second, our results indicate that top-down efforts to reduce identity salience may have weak or counterproductive impacts if they do not seek to transform underlying social structures. Concealment in our experiment fails to remove the traits and status differences through which hierarchy is reproduced, even as it disrupts the social coordination that helps people navigate them.

To be sure, our findings should not be taken to imply that identity revelation is universally desirable or advisable. In myriad settings—historical and contemporary—revelation would lead to domination and harm to minority groups. Indeed, we observe traces of this in our data: disclosure is positively selected on other characteristics, consistent with the idea that identity certainty brings social risks. Evidently, there are good reasons why individuals may opt to conceal, and why institutions may choose to protect that choice. Our contribution, however, is to note that concealment can also be costly. A possible direction for future research is to quantify the extent to which individuals accurately perceive the benefits and costs of identity revelation.

A final consideration concerns time horizons. While our experiment captures the short-run effects of identity concealment, longer-term reductions in identity legibility may have different consequences. Given time, norms of social interaction might evolve, allowing individuals to move through the social world without leaning on group identity cues. Thus, the long-run effects of identity concealment may be more positive. Historical efforts to abandon caste-distinctive names in South India point in this direction: Periyar's Self-Respect movement viewed caste titles and other markers not merely as expressions of identity, but as pillars of a hierarchical social order that it sought to challenge (Venkatachalapathy, 2020; Arni, 2024).

References

Adukia, Anjali, Richard Hornbeck, Daniel Keniston, and Benjamin Lualdi, "The Social and Statistical Construction of Race," *National Bureau of Economic Research, Working Paper 33502*, 2025.

Afridi, Farzana, Amrita Dhillon, and Swati Sharma, "The Ties That Bind Us: Social Networks and Productivity in the Factory," *Journal of Economic Behavior and Organization*, 2024, *218*, 470–485.

Aigner, Dennis and Glen Cain, "Statistical Theories of Discrimination in Labor Markets," *Industrial and Labor Relations Review*, 1977, *30* (2), 175–187.

Akerlof, George and Rachel Kranton, "Economics and Identity," *Quarterly Journal of Economics*, 2000, *115* (3), 715–753.

_ and _ , *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*, Princeton: Princeton University Press, 2010.

Alesina, Alberto and Eliana La Ferrara, "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, 2005, *43* (3), 762–800.

Ambedkar, Bhimrao, *Annihilation of Caste*, London: Verso Books, 2014.

Anderson, Siwan, "Caste as an Impediment to Trade," *American Economic Journal: Applied Economics*, 2011, *3* (1), 239–263.

Andreoni, James and Justin Rao, "The Power of Asking: How Communication Affects Selfishness, Empathy, and Altruism," *Journal of Public Economics*, 2011, *95* (7-8), 513–520.

Angeli, Deivis, Ieda Matavelli, and Fernando Secco, "Expected discrimination and Job Search," *Center for Effective Global Action, Working Paper 238*, 2023.

Arni, Abhimanyu, "Taking Periyarism Seriously: The Dravidian Identity as a Universality," *Global Intellectual History*, 2024, *Advance online publication.*

Asher, Sam, Kritarth Jha, Paul Novosad, Anjali Adukia, and Brandon Tan, "Residential Segregation and Unequal Access to Local Public services in India: Evidence from 1.5m Neighborhoods," *National Bureau of Economic Research, Working Paper 34818*, 2026.

Avery, Mallory, Andreas Leibbrandt, and Joseph Vecci, "Does Artificial Intelligence Help or Hurt Gender Diversity? Evidence from Two Field Experiments on Recruitment in Tech," *CESifo Working Paper 10996*, 2024.

Banerjee, Abhijit, Esther Duflo, Maitreesh Ghatak, and Jeanne Lafortune, "Marry for What? Caste and Mate Selection in Modern India," *American Economic Journal: Microeconomics*, 2013, *5* (2), 33–72.

Barry, Brian, *Culture and Equality: An Egalitarian Critique of Multiculturalism*, Cambridge, MA: Harvard university press, 2002.

Behaghel, Luc, Bruno Crépon, and Thomas Le Barbanchon, "Unintended Effects of Anonymous Résumés," *American Economic Journal: Applied Economics*, July 2015, *7* (3), 1–27.

Berg, Joyce, John Dickhaut, and Kevin McCabe, "Trust, Reciprocity, and Social History," *Games and Economic Behavior*, 1995, *10* (1), 122–142.

Bertrand, Marianne and Esther Duflo, "Field Experiments on Discrimination," in Abhijit Banerjee and Esther Duflo, eds., *Handbook of Economic Field Experiments*, Vol. 1, Amsterdam: Elsevier, 2017, pp. 309–393.

_ **and Sendhil Mullainathan**, "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review*, 2004, *94* (4), 991–1013.

Boring, Anne, Katherine Coffman, Dylan Glover, and Marıa José González-Fuentes, "Discrimination, Rejection, and Willingness to Apply: Effects of Blind Hiring Processes," *Mimeo, Harvard University*, 2025.

Brewer, Marilynn and Norman Miller, "Beyond the Contact Hypothesis: Theoretical Perspectives on Desegregation," in Norman Miller and Marilynn Brewer, eds., *Groups in Contact: The Psychology of Desegregation*, Orlando: Academic Press, 1984, pp. 281–302.

Bronchal, Adrià, "Better the Devil You Know: The Effects of Group Identity Uncertainty on Coordination Efficiency," *Journal of Economic Behavior and Organization*, 2023, *214*, 634–656.

Burton, Frank, *The Politics of Legitimacy: Struggles in a Belfast community*, London: Routledge, 1978.

Callen, Michael, Mohammad Isaqzadeh, James Long, and Charles Sprenger, "Violence and Risk Preference: Experimental Evidence from Afghanistan," *American Economic Review*, 2014, *104* (1), 123–148.

Carpio, Lucia Del and Thomas Fujiwara, "Spillovers and Scalability in Job Ad Experiments: Evidence from Gender-Neutral Language," *Mimeo, Princeton University*, 2026.

Cassan, Guilhem, "Identity-Based Policies and Identity Manipulation: Evidence from Colonial Punjab," *American Economic Journal: Economic Policy*, 2015, *7* (4), 103–131.

_ **, Daniel Keniston, and Tatjana Kleineberg**, "A Division of Laborers: Identity and Efficiency in India," *National Bureau of Economic Research, Working Paper 28462*, 2021.

Cefala, Luisa, Supreet Kaur, Heather Schofield, and Yogita Shamdasani, "Habit Formation in Labor Supply," *Mimeo, University of California, Berkeley*, 2024.

Chakraborty, Anujit, Arkadev Ghosh, Matt Lowe, and Gareth Nellis, "Learning About Outgroups: The Impact of Broad Versus Deep Interactions," *CESifo Working Paper No. 11363*, 2024.

Chandra, Kanchan, *Why Ethnic Parties Succeed: Patronage and Ethnic Head Counts in India*, New York: Cambridge University Press, 2007.

Charness, Gary, Ramón Cobo-Reyes, Simone Meraglia, and Ángela Sánchez, "Anticipated Discrimination, Choices, and Performance: Experimental Evidence," *European Economic Review*, 2020, *127*, 103473.

Dahis, Ricardo, Ludocvica Mosillo, Emily Nix, and Nancy Qian, "Black-White Income Inequality During Jim Crow: Evidence from 'Passing' for White," *Mimeo, Monash University*, 2026.

Dovidio, John, Samuel Gaertner, and Kerry Kawakami, "Intergroup Contact: The Past, Present, and the Future," *Group Processes and Intergroup Relations*, 2003, *6* (1), 5–21.

Fix, Michael and Raymond Struyk, *Clear and Convincing Evidence: Measurement of Discrimination in America*, Washington, DC: Urban Institute Press, 1993.

Ghosh, Arkadev, "Religious Divisions and Production Technology: Experimental Evidence from India," *Journal of Political Economy*, 2025, *133* (10), 3249–3304.

_ **, Prerna Kundu, Matt Lowe, and Gareth Nellis**, "Creating Cohesive Communities: A Youth Camp Experiment in India," *Review of Economic Studies*, 2026, *93* (1), 438–475.

Goette, Lorenz, David Huffman, and Stephan Meier, "The Impact of Social Ties on Group Interactions: Evidence from Minimal Groups and Randomly Assigned Real Groups," *American Economic Journal: Microeconomics*, 2012, *4* (1), 101–115.

Goffman, Erving, *The Presentation of Self in Everyday Life*, New York: Doubleday, 1959.

Goldin, Claudia and Cecilia Rouse, "Orchestrating Impartiality: The Impact of 'Blind Auditions on Female Musicians," *American Economic Review*, 2000, *90* (4), 715–741.

Granovetter, Mark, "The Strength of Weak Ties," *American Journal of Sociology*, 1973, *78* (6), 1360–1380.

Greene, Kenneth, Erin Rossiter, Enrique Seira, and Alberto Simpser, "Interacting as Equals Reduces Partisan Polarization in Mexico," *Nature Human Behaviour*, 2025, *9* (1), 147–155.

Harris, Adam, Daniel Nielson, Lily Medina, Clara Bicalho, Michael Findley, Jeremy Weinstein, James Habyarimana, Macartan Humphreys, and Daniel Posner, "Ethnic Identification and Ethnic Deception: Experimental Evidence from Uganda, South Africa, and the United States," *Mimeo, University of California, Los Angeles*, 2024.

Hewstone, Miles and Rupert Brown, "Contact is Not Enough: An Intergroup Perspective on the 'Contact Hypothesis'," in Miles Hewstone and Rupert Brown, eds., *Contact and Conflict in Intergroup Encounters*, Oxford: Basil Blackwell, 1986, pp. 1–44.

Hjort, Jonas, "Ethnic Divisions and Production in Firms," *Quarterly Journal of Economics*, 2014, *129* (4), 1899–1946.

Hobbs, Allyson, *A Chosen Exile: A History of Racial Passing in American Life*, Cambridge, MA: Harvard University Press, 2014.

Hoff, Karla and Priyanka Pandey, "Discrimination, Social Identity, and Durable Inequalities," *American Economic Review*, 2006, *96* (2), 206–211.

Horowitz, Donald, *Ethnic Groups in Conflict*, Berkeley: University of of California Press, 2000.

Kang, Sonia, Katherine DeCelles, András Tilcsik, and Sora Jun, "Whitened Résumés: Race and Self-presentation in the Labor Market," *Administrative Science Quarterly*, 2016, *61* (3), 469–502.

Kaur, Supreet, "Nominal Wage Rigidity in Village Labor Markets," *American Economic Review*, 2019, *109* (10), 3585–3616.

Kelley, Erin Munro, Gregory Lane, Matthew Pecenco, and Edward Rubin, "Customer Discrimination in the Workplace: Evidence from Online Sales," *National Bureau of Economic Research, Working Paper 31998*, 2023.

Kline, Patrick, Evan Rose, and Christopher Walters, "Systemic Discrimination Among Large U.S. Employers," *Quarterly Journal of Economics*, 2022, *137* (4), 1963–2036.

Kwon, Soonwoo and Jonathan Roth, "Testing Mechanisms," *Review of Economic Studies*, 2026, *Advance online publication.*

Lepage, Louis-Pierre, Xiaomeng Li, and Basit Zafar, "Anticipated Gender Discrimination and Grade Disclosure," *National Bureau of Economic Research, Working Paper 30765*, 2022.

Lowe, Matt, "Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration," *American Economic Review*, 2021, *111* (6), 1807–1844.

_ , "Has Intergroup Contact Delivered?," *Annual Review of Economics*, 2025, *17.*

Miller, Norman, "Personalization and the Promise of Contact Theory," *Journal of Social Issues*, 2002, *58* (2), 387–410.

Mitchell, J. Clyde, "The Kalela Dance: Aspects of Social Relationships among Urban Africans in Northern Rhodesia," *Rhodes-Livingstone Papers*, 1956, (27).

Mitrut, Andreea, Gabriel Kreindler, Margareta Matache, Andrei Munteanu, and Cristian Pop-Eleches, "Education and Selection into Ethnic Identification: Evidence from Roma People in Romania," *National Bureau of Economic Research, Working Paper 34383*, 2025.

Moerman, Michael, "Ethnic identification in a Complex Civilization: Who are the Lue?," *American Anthropologist*, 1965, *67* (5), 1215–1230.

Mousa, Salma, "Building Social Cohesion Between Christians and Muslims Through Soccer in Post-ISIS Iraq," *Science*, 2020, *369* (6505), 866–870.

Munshi, Kaivan, "Caste and the Indian economy," *Journal of Economic Literature*, 2019, *57* (4), 781–834.

Oh, Suanna, "Does Identity Affect Labor Supply?," *American Economic Review*, 2023, *113* (8), 2055–2083.

Park, Sangyoon, "Socializing at Work: Evidence from a Field Experiment with Manufacturing Workers," *American Economic Journal: Applied Economics*, 2019, *11* (3), 424–455.

Parmar, Rajnikant, "Transacting Caste in Modern Times: Changing Social Identity Through Surnames in Urban Gujarat," *Contemporary Voice of Dalit*, 2020, *12* (2), 220–233.

Pettigrew, Thomas and Linda Tropp, "A Meta-Analytic Test of Intergroup Contact Theory," *Journal of Personality and Social Psychology*, 2006, *90* (5), 751.

Rao, Gautam, "Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools," *American Economic Review*, 2019, *109* (3), 774–809.

Robinson, Amanda, "Ethnic Visibility," *American Journal of Political Science*, 2024, *68* (4), 1234–1251.

Rose, Evan, "A Constructivist Perspective on Empirical Discrimination Research," *Journal of Economic Literature*, 2023, *61* (3), 906–923.

Ruebeck, Hannah, "Perceived Discrimination at Work," *Mimeo, MIT*, 2023.

Satyanarayanan, Kusuma and Joel Lee, *Concealing Caste: Narratives of Passing and Personhood in Dalit Literature*, Oxford: Oxford University Press, 2023.

Schelling, Thomas C., *The Strategy of Conflict*, Cambridge, MA: Harvard University Press, 1960.

Scott, Joan, *The Politics of the Veil*, Princeton, NJ: Princeton University Press, 2007.

Shukla, Soumitra, "Making the Elite: Class Discrimination at Multinationals," *Mimeo, Harvard University*, 2025.

Singh, Khushwant, *A History of the Sikhs: 1469–1839*, New Delhi: Oxford University Press, 2004.

Small, Deborah A and George Loewenstein, "Helping a Victim or Helping the Victim: Altruism and Identifiability," *Journal of Risk and Uncertainty*, 2003, *26* (1), 5–16.

Southall, Aidan, "The Illusion of Tribe," *Journal of Asian and African Studies*, 1970, *5* (1-2), 28–50.

Srinivas, M. N., *Social Change in Modern India*, Berkeley: University of California Press, 1966.

Tajfel, Henri, M. G. Billig, R. P. Bundy, and Claude Flament, "Social Categorization and Intergroup Behaviour," *European Journal of Social Psychology*, April 1971, *1* (2), 149–178.

Taylor, Charles, "The Politics of Recognition," in Amy Gutmann, ed., *Multiculturalism: Examining the Politics of Recognition*, Princeton, NJ: Princeton University Press, 1992, pp. 25–73.

Valmiki, Omprakash, *Joothan: An Untouchable's life*, New York: Columbia University Press, 2008.

Venkatachalapathy, A R, "Periyar E. V. Ramasamy," in David Ludden, ed., *Oxford Research Encyclopedia of Asian History*, Oxford University Press, 2020.

A Model

We develop a two-player Bayesian game that models the *First Name Only* treatment, where identity revelation is an endogenous choice. The model rationalizes our core empirical findings across treatments and generates additional predictions that we test in the data.

Each player $i \in \{1,2\}$ chooses an action $a_i \in \{R,H\}$, where R denotes revealing caste identity and H denotes hiding it. Action R captures both explicit and implicit ways of communicating caste identity during the interaction. The experimental treatments correspond to restrictions on this choice set: in the *Hidden* treatment players cannot choose R, while in the *Revealed* treatment they cannot choose H.

We model a player i's private type as

$$t_i = (c_i, y_i),$$

where $c_i \in \{G,L\}$ denotes caste identity (General or SC) and y_i is i's ex-ante expected payoff from a cross-caste interaction, given i's belief about facing caste-domination at work. We assume each caste occurs with probability $1/2$ and this is common knowledge.

Under the restriction $x > 0 \geq z$, the payoffs from the game are given by

$$u_i(a_i,a_j;t_i,t_j) = \begin{cases} 0 & \text{if } a_i = a_j = H \quad \text{(no revelation outcome is normalized to "0")}, \\ x & \text{if } a_i = a_j = R \text{ and } c_i = c_j \quad \text{(mutual revelation is beneficial for same-caste matches)}, \\ y_i & \text{if } a_i = a_j = R \text{ and } c_i \neq c_j \quad \text{(...but in cross-caste matches, it depends on your payoff-type)}, \\ z & \text{if } (a_i,a_j) \in \{(R,H),(H,R)\} \quad \text{(one-sided sharing is awkward, hence worse than "0")}. \end{cases}$$

One-sided sharing: We assume that one-sided identity sharing is worse than mutual hiding and this has a non-trivial testable prediction. Given participants can communicate before choosing whether to reveal or hide during the session, they should attempt to avoid reaching the one-sided sharing outcome – perhaps in favor of the pareto-dominating mutual hiding outcome – whenever possible. Thus, we should see few such outcomes.

One sided revelations being costly for both players is also significant in its capacity of preventing unraveling. Compare it to the alternative modeling assumption where unilateral revealing from the high-caste guaranteed a high payoff for them, and a low payoff for the non-revealer, as they are implicitly identified as low-caste. Such an alternative assumption would make the high-castes unilaterally reveal, and that would cause the game to unravel, as the low-caste identities would be revealed either way.

We also assume that the payoff $z \leq 0$ applies to both to the partner who revealed and who did not reveal - this is for simplicity. Assuming that the outcome is negative for both parties, but worse for the person one-sidedly sharing their identity - $z_s < z_{ns} \leq 0$ - would result in the same qualitative predictions as we currently have.

Assumptions about y_i General-caste players do not anticipate facing caste-domination, and hence $y_i = x$ deterministically. But, for lower-caste players, $y_i \sim F$ with support $[\underline{y}, x]$. Assuming $\underline{y} < 0$ guarantees that some lower-caste players perceive the mutual revelation outcome as worse than the mutual hiding outcome, in cross-caste interactions. This heterogeneity in y_i captures differences in anticipated interaction quality based on past experiences, social information, and contextual cues.

Source of y_i: Lower-caste individuals may draw on prior workplace interactions, accounts from peers or media, or signals received during the interaction itself when forming expectations about how a cross-caste interaction might unfold. While we do not observe past interactions, we have some measures of social distance between the interacting players, and we test if those are predictive of the equilibrium predictions.

For everyone else, mutual revelation is beneficial because it provides a shared social context in which individuals can interact and communicate more easily. Knowledge of a partner's social identity helps anchor expectations about norms, language, and appropriate modes of interaction, and can facilitate coordination in joint tasks. When identity is known, individuals may find it easier to interpret behavior, establish conversational common ground, and anticipate how their partner will respond. In this sense, identity revelation can reduce social uncertainty and help interactions proceed more smoothly.

A pure strategy for a General-caste player is a choice $s_G \in \{R, H\}$, since General-caste players have no payoff-relevant type-heterogeneity. By contrast, a pure strategy for a lower-caste player is a measurable map

$$s_L : [\underline{y}, x] \to \{R, H\},$$

which assigns an action to each payoff-type y. Given a lower-caste strategy s_L, let

$$p_L \equiv \Pr(s_L(y) = R) = \int \mathbf{1}\{s_L(y) = R\}\, dF(y)$$

denote the proportion of lower-caste players who reveal. Thus p_L is the endogenous disclosure rate induced by the strategy s_L.

Two opposing considerations shape disclosure dynamics in this game, as formalized in Proposition 1. First, unilateral revelation is socially costly: revealing caste when the partner does not do so yields payoff $z < 0$. Thus individuals prefer not to reveal if they expect others to conceal. As a result, a no-revelation outcome – where both sides hide caste identity – is always sustainable as an equilibrium.

Second, mutual revelation can generate benefits by clarifying the social context of the interaction. For example, if General-caste players reveal their caste, lower-caste players may also benefit from revealing, depending on their type y_i. Individuals with sufficiently low y_i anticipate unfavorable cross-caste interactions and prefer concealment, whereas those with higher y_i are willing to reveal.

These forces lead to two possible interaction regimes in pairs of worker groups. In one regime, no one reveals caste identity. In the other, General-caste players reveal, while lower-caste players reveal only if their type exceeds a threshold.

Proposition 1 (Equilibrium characterization)**.** *Assume $\underline{y} < -x + 2z$. Then the game admits exactly two pure-strategy symmetric Bayesian Nash equilibria.*

(i) No revelation. $s_G = H, \quad s_L(y) = H \ \ \forall y$.

(ii) Partial lower-caste revelation. $s_G = R$ and there exists a cutoff y^* such that

$$s_L(y) = \begin{cases} H & y < y^* \\ R & y \geq y^* \end{cases}$$

where $y^* = -p_L(x-2z)$ *and* p_L *solves*

$$p_L = 1 - F\big(-p_L(x-2z)\big).$$

The condition $\underline{y} < -x+2z$ implies that pessimistic L types block the candidate full-revelation equilibrium. All proofs are relegated to Section A.1.

A.1 Proofs

Proof. (Proposition 1)

Since caste is hidden in the First Name Only treatment, players evaluate their actions using the prior distribution over the partner's caste and disclosure behavior.

Consider first a lower-caste player of type y_i. If he reveals, then:

- with probability $\frac{1}{2}p_L$, the partner is lower-caste and reveals, yielding payoff x;
- with probability $\frac{1}{2}p_G$, the partner is General-caste and reveals, yielding payoff y_i;
- otherwise, mutual revelation does not occur, and payoff is z.

Hence

$$U_L(R \mid y_i) = \frac{1}{2}p_L x + \frac{1}{2}p_G y_i + \left(1 - \frac{1}{2}p_G - \frac{1}{2}p_L\right) z.$$

If instead he hides, then he receives payoff z whenever the partner reveals and payoff 0 otherwise:

$$U_L(H \mid y_i) = \left(\frac{1}{2}p_G + \frac{1}{2}p_L\right) z.$$

Therefore a lower-caste player reveals if and only if

$$\frac{1}{2}p_L x + \frac{1}{2}p_G y_i + \left(1 - \frac{1}{2}p_G - \frac{1}{2}p_L\right) z \geq \left(\frac{1}{2}p_G + \frac{1}{2}p_L\right) z.$$

Rearranging,

$$\frac{1}{2}p_G y_i + \frac{1}{2}p_L x + z(1 - p_G - p_L) \geq 0.$$

If $p_G > 0$, this is equivalent to

$$y_i \geq y^*(p_G, p_L) \equiv -\frac{p_L}{p_G}x - \frac{2z(1-p_G-p_L)}{p_G}.$$

The expected payoff comparison implies that, the gain from revealing is strictly increasing in y whenever $p_G > 0$. Therefore the lower-caste best response is monotone: there exists a threshold $y^* \in [\underline{y}, x]$ such that

$$s_L(y) = \begin{cases} H & \text{if } y < y^*, \\ R & \text{if } y \geq y^*. \end{cases}$$

First consider $s_G = H$. Thus $p_G = 0$, and the payoff difference for a lower-caste player is

$$U_L(R|y) - U_L(H|y) = \frac{1}{2}p_L x + (1 - p_L)z.$$

This expression does not depend on y, so all lower-caste players must take the same action. If all revealed ($p_L = 1$), then General-caste players would strictly prefer revealing since

$$U_G(R) - U_G(H) = \frac{1}{2}x > 0,$$

contradicting $p_G = 0$. Therefore the only equilibrium with $p_G = 0$ is

$$s_G = H, \qquad s_L(y) = H.$$

Next consider $s_G = R$, so $p_G = 1$. From the payoff comparison derived earlier,

$$U_G(R) - U_G(H) = \frac{x + p_L(x - 2z)}{2} > 0$$

for all $p_L \in [0,1]$, so revealing is always optimal for General-caste players.

For a lower-caste player of type y,

$$U_L(R|y) - U_L(H|y) = \frac{1}{2}\big(y + p_L(x - 2z)\big),$$

which is strictly increasing in y. Hence lower-caste best responses take cutoff form with threshold

$$y^* = -p_L(x - 2z).$$

Consistency requires

$$p_L = 1 - F(y^*) = 1 - F\big(-p_L(x - 2z)\big).$$

Define

$$\phi(p) = p - \big(1 - F(-p(x - 2z))\big).$$

Since F is continuous, ϕ is continuous on $[0,1]$. Moreover

$$\phi(0) = F(0) - 1 \leq 0, \qquad \phi(1) = F(-(x - 2z)) \geq 0.$$

Thus a solution $p_L^* \in [0,1]$ exists by the Intermediate Value Theorem.

Finally, full revelation cannot be an equilibrium because the lowest lower-caste type would reveal only if

$$\underline{y} \geq -x + 2z,$$

which is ruled out by assumption. Therefore the only equilibria are the two described above. □

B Supplementary Tables

Table A1: Table: Baseline summary statistics for the experimental sample.

	N	Mean	St. Dev.	Min.	Pctl. 25	Median	Pctl. 75	Max.
(a) Age	1198	23.30	4.34	18.00	20.00	22.00	26.00	40.00
(b) Number of siblings	1198	2.86	1.14	0.00	2.00	3.00	4.00	8.00
(c) Household size	1198	5.91	1.71	1.00	5.00	6.00	7.00	20.00
(d) Married	1198	0.27	0.45	0.00	0.00	0.00	1.00	1.00
(e) Number of children (if married)	329	1.38	0.95	0.00	1.00	1.00	2.00	9.00
(f) Household income (midpoint, INR)	1198	10659	8185	2500	2500	7500	17500	55000
(g) Completed 12th standard	1198	0.73	0.44	0.00	0.00	1.00	1.00	1.00
(h) Completed college	1198	0.20	0.40	0.00	0.00	0.00	0.00	1.00
(i) Identity ranking: Caste	1198	3.70	0.99	1.00	3.00	4.00	4.00	5.00
(j) Identity ranking: Religion	1198	3.69	1.11	1.00	3.00	4.00	5.00	5.00
(k) Identity ranking: Indian	1198	1.31	0.83	1.00	1.00	1.00	1.00	5.00
(l) Identity ranking: From U.P.	1198	3.33	1.36	1.00	2.00	3.00	5.00	5.00
(m) Identity ranking Father/Son/Brother	1198	2.97	1.16	1.00	2.00	3.00	4.00	5.00

Notes: This table presents summary statistics for variables collected at baseline. Household income (row (f)) is coded as the midpoint of the income bin, or the minimum value for the top-coded bin. Identity rankings (rows (i)–(m)) are measured on a 1 (most) to 5 (least) ranking scale. All statistics are based on the experimental sample.

Table A2: Baseline differences in means (GEN vs. SC)

	General Caste	Scheduled Caste	Diff. (GEN−SC)	*N*
(a) Age	23.07	23.59	-0.52^{*}	1,064
	(4.24)	(4.46)	(0.27)	
(b) Number of siblings	2.69	3.11	-0.42^{***}	1,064
	(1.10)	(1.16)	(0.07)	
(c) Household size	5.95	5.93	0.03	1,064
	(1.88)	(1.62)	(0.11)	
(d) Married	0.24	0.31	-0.07^{**}	1,064
	(0.43)	(0.46)	(0.03)	
(e) Number of children	1.46	1.29	0.16	296
	(0.77)	(1.08)	(0.11)	
(f) Household income (midpoint, INR)	12,307.33	9,064.85	$3{,}242.48^{***}$	1,064
	(9,325.06)	(6,437.23)	(491.27)	
(g) Completed 12th standard	0.77	0.67	0.10^{***}	1,064
	(0.42)	(0.47)	(0.03)	
(h) Completed college	0.24	0.17	0.07^{***}	1,064
	(0.43)	(0.38)	(0.02)	
(i) Identity ranking: Caste	3.88	3.51	0.37^{***}	1,064
	(0.97)	(0.96)	(0.06)	
(j) Identity ranking: Religion	3.92	3.46	0.46^{***}	1,064
	(1.04)	(1.10)	(0.07)	
(k) Identity ranking: Indian	1.19	1.45	-0.26^{***}	1,064
	(0.66)	(0.98)	(0.05)	
(l) Identity ranking: From U.P.	2.91	3.75	-0.83^{***}	1,064
	(1.27)	(1.33)	(0.08)	
(m) Identity ranking: Father/Son/Brother	3.10	2.83	0.27^{***}	1,064
	(1.00)	(1.29)	(0.07)	

Notes: This table reports baseline differences in means between General and Scheduled caste respondents. Columns (1) and (2) report group means with standard deviations in parentheses. Column (3) reports the difference in means (GEN minus SC); standard errors are shown in parentheses below. Household income is coded as the midpoint of the income bin (minimum for the top-coded bin). Identity rankings are measured on a scale from 1 (most important) to 5 (least important).
Significance levels: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A3: Sample selection

Sample	Baseline only (1)		Experimental (2)		Diff (2–1)		N	
Variable	Mean	SE	Mean	SE	Mean	SE	1	2
(a) Age	24.21	0.24	23.30	0.13	–0.919***	0.252	448	1198
(b) Number of siblings	2.99	0.06	2.86	0.03	–0.130**	0.065	447	1198
(c) Household size	5.90	0.09	5.91	0.05	0.012	0.098	432	1198
(d) Married	0.29	0.02	0.27	0.01	–0.018	0.024	479	1198
(e) Number of children	1.42	0.08	1.38	0.05	–0.041	0.094	140	329
(f) HH income (INR '000)	10.21	0.35	10.66	0.24	0.451	0.448	431	1198
(g) Completed 12th standard	0.74	0.02	0.73	0.01	–0.012	0.025	431	1198
(h) Completed college	0.25	0.02	0.20	0.01	–0.044*	0.023	431	1198
(i) Identity ranking: Caste	3.70	0.05	3.70	0.03	–0.001	0.056	421	1198
(j) Identity ranking: Religion	3.46	0.05	3.69	0.03	0.224***	0.063	421	1198
(k) Identity ranking: Indian	1.24	0.04	1.31	0.02	0.068	0.046	421	1198
(l) Identity ranking: From U.P.	3.42	0.07	3.33	0.04	–0.088	0.077	421	1198
(m) Identity ranking: Father/Son/Brother	3.18	0.06	2.97	0.03	–0.203***	0.066	421	1198

Notes: Means and standard errors (SE) by group. The first group represents the sample of individuals who participated in the baseline survey, but did not take part in the final experiment, whereas the second group represents the latter. Household income (row (f)) is coded as the midpoint of the income bin, or the minimum value for the top-coded bin and is shown in thousands of INR. Identity rankings (rows (i)–(m)) are measured on a 1 (most) to 5 (least) ranking scale. Significance levels: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A4: Randomization checks

	Control (First Name)		Revealed vs Control		*Hidden* vs Control		p: Revealed = *Hidden*
	Mean	SD	β	p-value	β	p-value	p-value
(a) Age	23.53	4.73	−0.30	0.30	−0.36	0.25	0.82
(b) Number of siblings	2.94	1.13	−0.07	0.31	−0.07	0.37	0.95
(c) Household size	5.92	1.79	0.02	0.85	0.02	0.89	0.97
(d) Married	0.30	0.46	−0.03	0.35	−0.04	0.25	0.77
(e) Number of children	1.46	1.15	−0.16	0.26	−0.15	0.35	0.98
(f) Household income (midpoint, INR)	11033	8585	−538	0.30	−562	0.31	0.96
(g) Completed 12th standard	0.72	0.45	0.01	0.74	−0.01	0.81	0.59
(h) Completed college	0.21	0.41	−0.02	0.58	−0.01	0.68	0.89
(i) Identity ranking: Caste	3.69	0.97	0.03	0.64	−0.03	0.64	0.35
(j) Identity ranking: Religion	3.70	1.11	−0.02	0.80	−0.05	0.55	0.71
(k) Identity ranking: Indian	1.34	0.87	−0.06	0.27	−0.02	0.70	0.49
(l) Identity ranking: From U.P.	3.34	1.41	0.00	0.96	0.02	0.85	0.88
(m) Identity ranking: Father/Son/Brother	2.94	1.14	0.04	0.55	0.09	0.29	0.59

Notes: Each row reports a separate regression of the covariate on treatment indicators with date fixed effects (absorbed) and standard errors clustered at the pair-level. The control group is the *First Name Only* group. Columns report control mean and standard deviation, treatment–control differences (β) for *Revealed* and *Hidden*, and the p-value for equality of treatment effects. Household income (row (f)) is coded as the midpoint of the income bin, or the minimum value for the top-coded bin. Identity rankings (rows (i)–(m)) are measured on a 1 (most) to 5 (least) ranking scale.

Table A5: Heterogeneity by caste: GEN vs SC

	(1) Identity fusion	(2) Meet up again	(3) Trust sent
General caste (GEN)			
First name only	-0.365^{***}	-0.204^{***}	-4.020^{***}
	(0.104)	(0.051)	(1.311)
Hiding	-0.544^{***}	-0.213^{***}	-4.633^{***}
	(0.107)	(0.051)	(1.353)
Mean dep. var. (Revealed)	2.777	0.741	27.349
Scheduled Caste (SC)			
First name only	-0.199^{**}	-0.101^{*}	-3.140^{**}
	(0.099)	(0.052)	(1.375)
Hiding	-0.386^{***}	-0.186^{***}	-2.928^{**}
	(0.103)	(0.051)	(1.331)
Mean dep. var. (Revealed)	2.783	0.711	26.145
p: GEN = SC, First Name Only	0.089	0.040	0.489
p: GEN = SC, Hiding	0.105	0.618	0.182
N	1,014	1,014	1,014

Notes: Omitted treatment: Revealed (GEN and SC). Standard errors are clustered at the pair-level. Each column comes from a single regression interacting treatment status with caste identity; the two coefficients shown under each caste panel heading are the interactions of that caste dummy with the *First Name Only* and *Hidden* treatment indicators respectively. All specifications control for age, and absorb randomization strata and education bin fixed effects via `reghdfe`. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A6: General-caste senior-role allocation, accounting for baseline ability

	(1)	(2)	(3)	(4)	(5)
Predicted share General senior at equal ability	0.599^{***} (0.026)	0.599^{***} (0.026)	0.599^{***} (0.025)	0.689^{***} (0.038)	0.599^{***} (0.025)
R^2	0.000	0.006	0.042	0.107	0.076
Observations	167	167	167	167	167
H_0: share = 0.50	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$
F-test: ability controls = 0	–	$p = 0.621$	$p = 0.132$	$p = 0.011$	$p = 0.121$
Ability gap controls	None	Linear	Linear	Quadratic	Quintiles
Ability level controls	None	None	Linear	Linear	None

Notes: The dependent variable is the share of tasks in which the General-caste participant takes the senior role. The table reports the predicted share of tasks in which the General-caste participant is senior when General and SC participants have equal baseline ability. Ability is measured using baseline reading and typing scores. Ability gap controls are functions of the General-minus-SC baseline ability gap; ability level controls account for the overall level of baseline ability in the pair. Standard errors are reported in parentheses. The F-test reports the joint significance of the included ability controls. $^{***}p < 0.01$.

Table A7: Logit model: Predicting caste from observable characteristics

	P(SC) (1)
Demographics	
Age	0.008*
	(0.005)
Married	0.095**
	(0.044)
Number of siblings	0.112***
	(0.015)
Household size	-0.030***
	(0.011)
Education and income	
Completed 12th grade	-0.101***
	(0.033)
College degree	-0.043
	(0.039)
Monthly household income (Rs.)	-0.000***
	(0.000)
Self-assessment and affect	
Self-assessed: better at reading	0.131***
	(0.035)
Self-conscious yesterday (1–4)	0.080***
	(0.018)
Appearance and speech cues	
Clothing quality (PC1)	-0.014
	(0.011)
Fluency/pronunciation (PC1)	0.003
	(0.011)
Shudh Hindi rating (PC1)	-0.010
	(0.014)
Observations	1,099
McFadden R^2	0.120

Notes: The outcome is an indicator equal to 1 for Scheduled Caste and 0 for General Caste. The sample excludes OBC respondents. The table reports average marginal effects from a logit model. All predictors are measured at baseline and are observable through natural workplace interaction. Clothing quality, fluency/pronunciation, and Shudh Hindi rating are each the first principal component of ratings by three independent raters. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Out-of-sample classification accuracy (50/50 Train–Test Split)

	Predicted SC	**Predicted General**	**Total**
True SC	205	98	303
True General	102	189	291
Total	307	287	594
Overall accuracy: 66.3%	(chance baseline: 50.0%)		

Notes: The logit model is estimated on a random 50% training sample (seed 12345) and evaluated on the held-out 50% test sample ($N = 594$). The predictor set is the PC1 specification from Table A7: baseline demographics, income, self-assessed ability, and self-consciousness, plus the first principal component of clothing quality, fluency/pronunciation, and Shudh Hindi ratings. A participant is predicted SC when the estimated $\Pr(\text{SC}) \geq 0.5$.

Table A9: Social relations: Homogeneous vs heterogeneous revealed teams

	(1) Trust sent	(2) Identity fusion	(3) Meet up again
Same-caste pair (=1 if HH or LL, =0 if HL)	3.157*	0.266*	0.001
	(1.913)	(0.138)	(0.072)
Mean dep. var. (HL)	26.778	2.781	0.723
Observations	425	425	425

Notes: Standard errors clustered at the pair-level in parentheses. Sample: HH-Revealed, LL-Revealed, and HL-Revealed arms. Homogeneous = 1 if same-caste pair (HH or LL), 0 if cross-caste (HL). Mean dep. var. is for the heterogeneous (HL) group. All specifications control for age and absorb randomization strata and education bin fixed effects. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table A10: Caste confidence mediates treatment effects

	(1) Trust sent	(2) Identity fusion	(3) Meet up again	(4) True self
Panel A: Treatment effects only				
First name only	-3.711^{***}	-0.286^{***}	-0.152^{***}	-0.713^{***}
	(1.221)	(0.092)	(0.046)	(0.246)
Hiding	-3.920^{***}	-0.468^{***}	-0.198^{***}	-0.846^{***}
	(1.217)	(0.097)	(0.045)	(0.237)
Mean (Revealed)	26.747	2.780	0.726	7.503
First name only vs. Hiding p-value	0.844	0.066	0.322	0.571
N	1,014	1,014	1,014	1,014
Panel B: Controlling for caste confidence and beliefs				
First name only	-1.558	-0.091	-0.068	0.005
	(1.138)	(0.089)	(0.047)	(0.207)
Hiding	-0.009	-0.126	-0.067	0.426^{*}
	(1.348)	(0.101)	(0.048)	(0.218)
Caste certainty				
Caste confidence (0–1)	19.002^{***}	1.574^{***}	0.505^{***}	5.933^{***}
	(2.620)	(0.212)	(0.093)	(0.488)
Guessed caste × own caste				
Guessed GEN × own GEN	0.128	-0.022	-0.146^{*}	0.097
	(1.569)	(0.201)	(0.083)	(0.401)
Guessed GEN × own SC	-3.526^{**}	-0.077	0.190^{***}	-0.424
	(1.566)	(0.146)	(0.064)	(0.300)
Mean (Revealed)	26.747	2.780	0.726	7.503
First name only vs. Hiding p-value	0.167	0.716	0.970	0.051
N	1,014	1,014	1,014	1,014

Notes: Omitted treatment: Revealed. All specifications control for age and absorb randomization strata and education bin fixed effects. Panel B adds caste confidence (rescaled 0–1) and guessed partner caste interacted with own caste. SEs clustered at the pair-level. Guessed SC × own GEN and Guessed SC × own SC are omitted in Panel B, because of collinearity (strata fixed effects include caste). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A11: Asking about identity increases trust

	General Caste		**Scheduled Caste**	
	(1)	(2)	(3)	(4)
	Trust: Sent	Trust: Belief	Trust: Sent	Trust: Belief
Partner asked your caste?	1.004	-1.147	4.014**	-0.078
	(2.011)	(2.057)	(2.011)	(1.857)
Asked partner their caste?	4.057**	4.463**	1.253	2.174
	(2.052)	(2.105)	(1.925)	(1.779)
Mean Dep. Var.	25.03	22.5	24.3	21.5
Observations	564	564	563	563
R-squared	0.128	0.124	0.149	0.133

Notes: Robust standard errors reported in parenthesis. Columns (1)–(2) restrict the sample to GC participants and columns (3)–(4) to SC participants. *e_did_p_ask_caste* = 1 if the partner asked about caste; *e_did_ask_p_caste* = 1 if the respondent asked the partner's caste. All specifications include date fixed effects and pre-specified controls for baseline education bin and age. $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

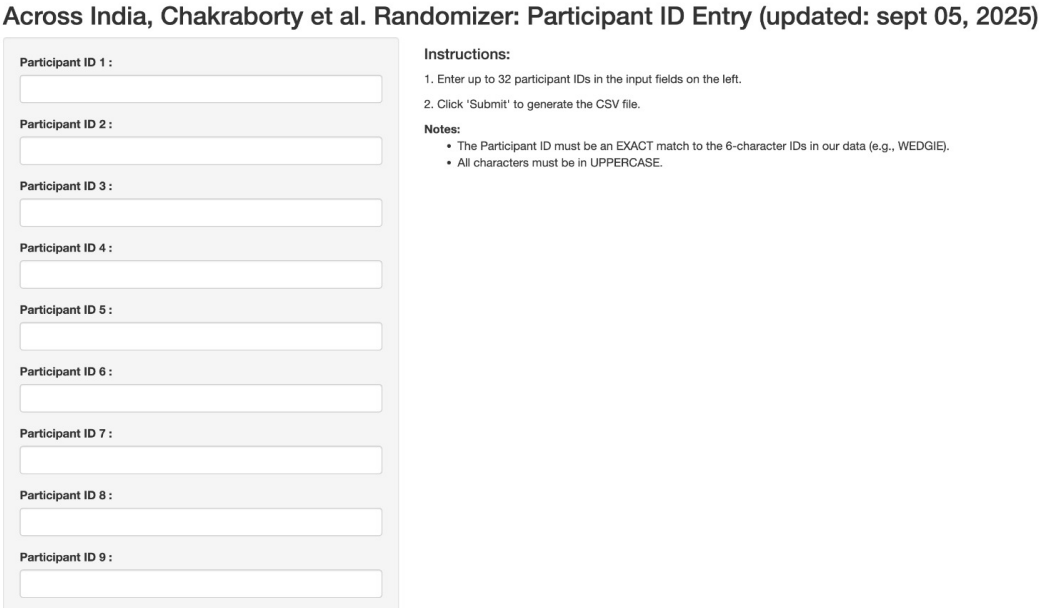
Table A12: Treatment effect on guessing study purpose

	Mentioned caste as study purpose
	(1)
First name only	0.006
	(0.014)
Hidden	0.021
	(0.016)
First Name Only vs. Hidden p-value	0.39
Mean Dep. Var. (Revealed)	0.029
Observations	1,198
R-squared	0.131

Notes: Dependent variable is a dummy coded 1 if the subject mentioned caste when asked about the purpose of the study. Standard errors clustered at the pair level are in parentheses. Regressions include randomization-strata fixed effects (date × caste category) and pre-specified controls for education bin and age. $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

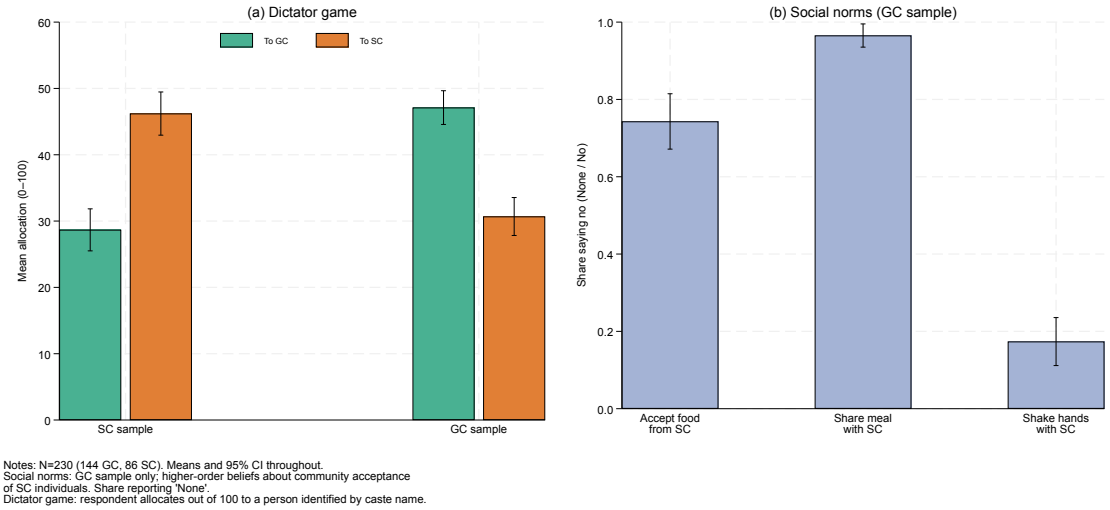
C Supplementary Figures

Figure A1: The randomizer



Notes: This figure shows a screenshot of the randomizer app we built to assign subjects to pairs and pairs to treatments each day. Subjects were given unique IDs at the end of the baseline survey, which they brought to the work site on the day they were invited. The app was linked to our baseline database, so entering a subject's ID pulled the relevant caste information and allowed the app to assign subjects appropriately. Using baseline-linked IDs ensured that neither names nor caste needed to be recorded on the day of the experiment, limiting the salience of caste during randomization.

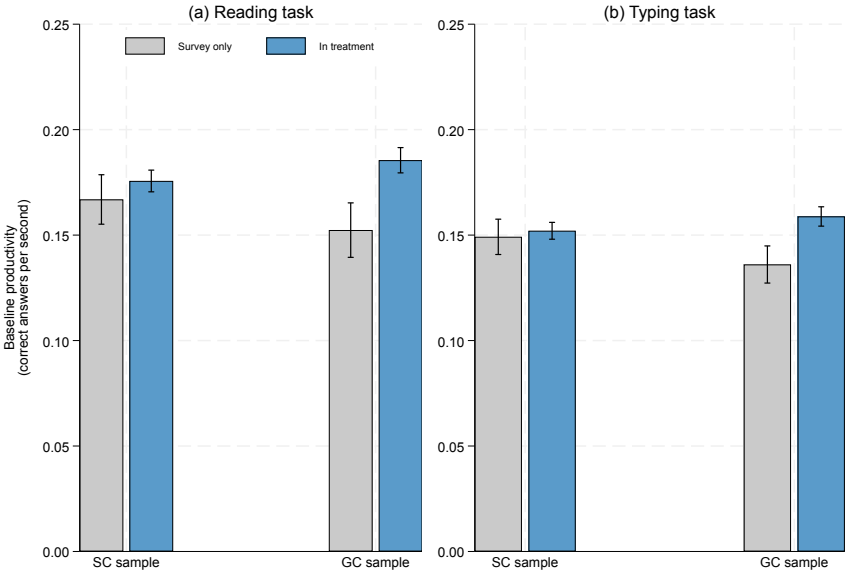
Figure A2: Caste bias in study region



Notes: N=230 (144 GC, 86 SC). Means and 95% CI throughout.
Social norms: GC sample only; higher-order beliefs about community acceptance of SC individuals. Share reporting 'None'.
Dictator game: respondent allocates out of 100 to a person identified by caste name.

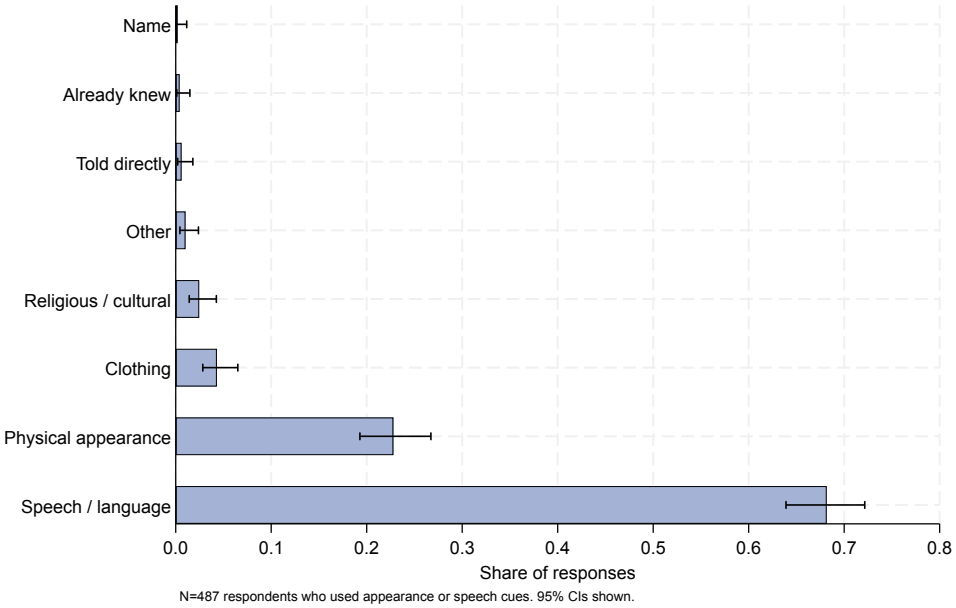
Notes: Panel (a) shows mean allocations from a dictator game in which respondents distribute 100 units to a hypothetical recipient identified by caste name (General Caste or Scheduled Caste). Bars show means by respondent caste (GC or SC), separately for giving to a GC recipient (green) and an SC recipient (orange). Panel (b) is restricted to GC respondents and displays three measures of social acceptance norms toward SC individuals: the share reporting that "none" in their village would accept food from an SC person, the share who say "none" in their family would share a meal with an SC person, and finally the share reporting that "none" in their village would shake hands with an SC person. The food and handshake items are second-order beliefs about community norms; the meal-sharing item is a self-report about their family. We asked all questions in a scoping survey conducted in the study region. Error bars are 95% confidence intervals. N = 230 (144 GC, 86 SC).

Figure A3: Sample selection on baseline productivity



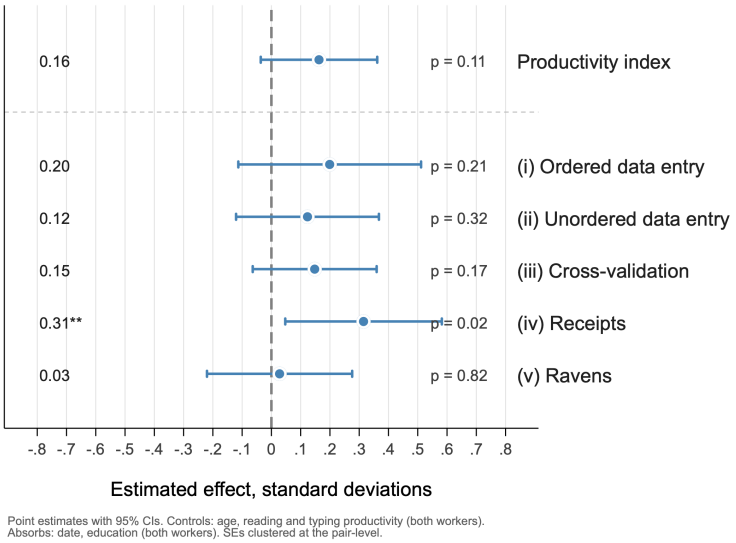
Notes: Bars show mean baseline productivity (correct answers per second) with 95% confidence intervals, separately by caste and participation status. "Survey only" denotes individuals who completed the recruitment survey but were not assigned to a treatment; "In treatment" denotes those who received a treatment assignment. Both GC and SC participants are positively selected on productivity relative to non-participants, and the pattern holds across both the reading and typing tasks.

Figure A4: Open ended responses on how subjects guessed their partner's identity



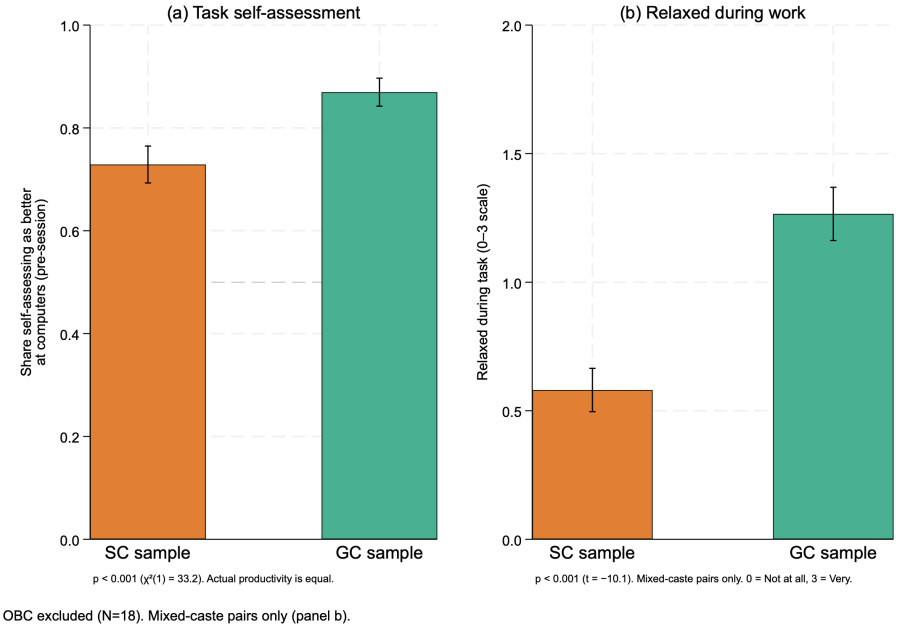
Notes: This figure summarizes qualitative responses from participants who reported that they could infer their partner's identity from appearance or speech cues. Each open-ended response is classified into one of the following categories based on the specific cue mentioned by the participant.

Figure A5: Productivity of homogeneous versus heterogeneous teams



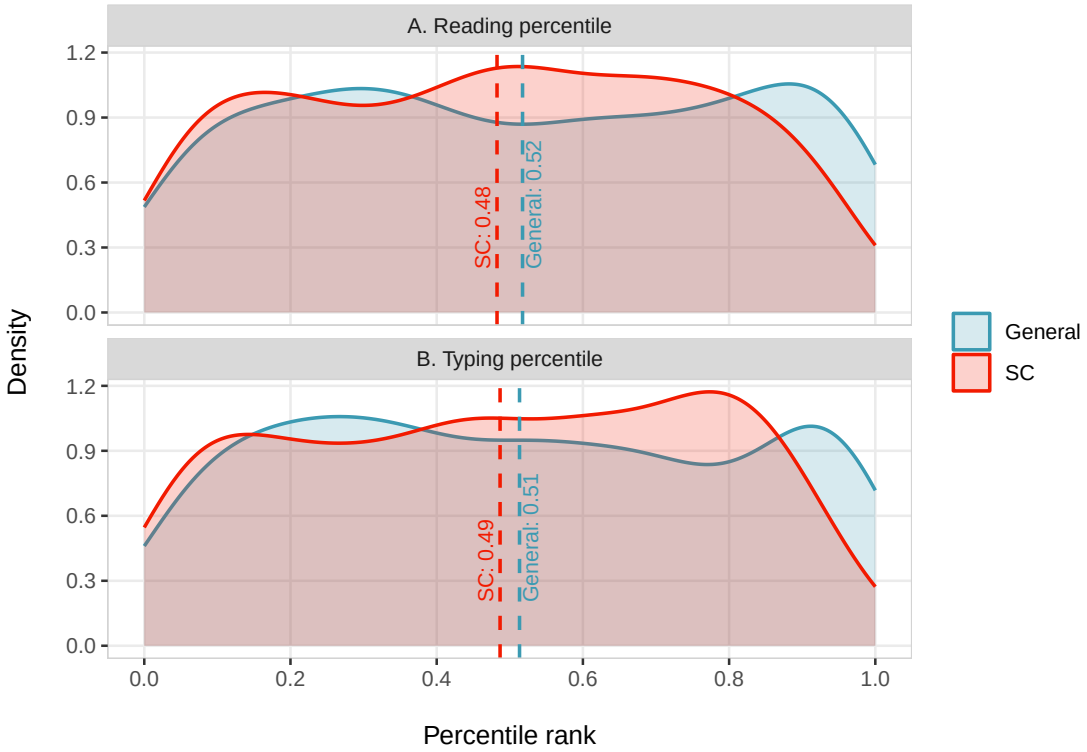
Notes: This figures compares the productivity of homogeneous versus heterogeneous teams in the *Revealed* condition. All regressions control for age and baseline productivity levels of each partner in each task (reading and typing), as well as fixed effects for education-level bin and date. Standard errors are clustered at the pair level.

Figure A6: Differences between General and Scheduled Castes in traits



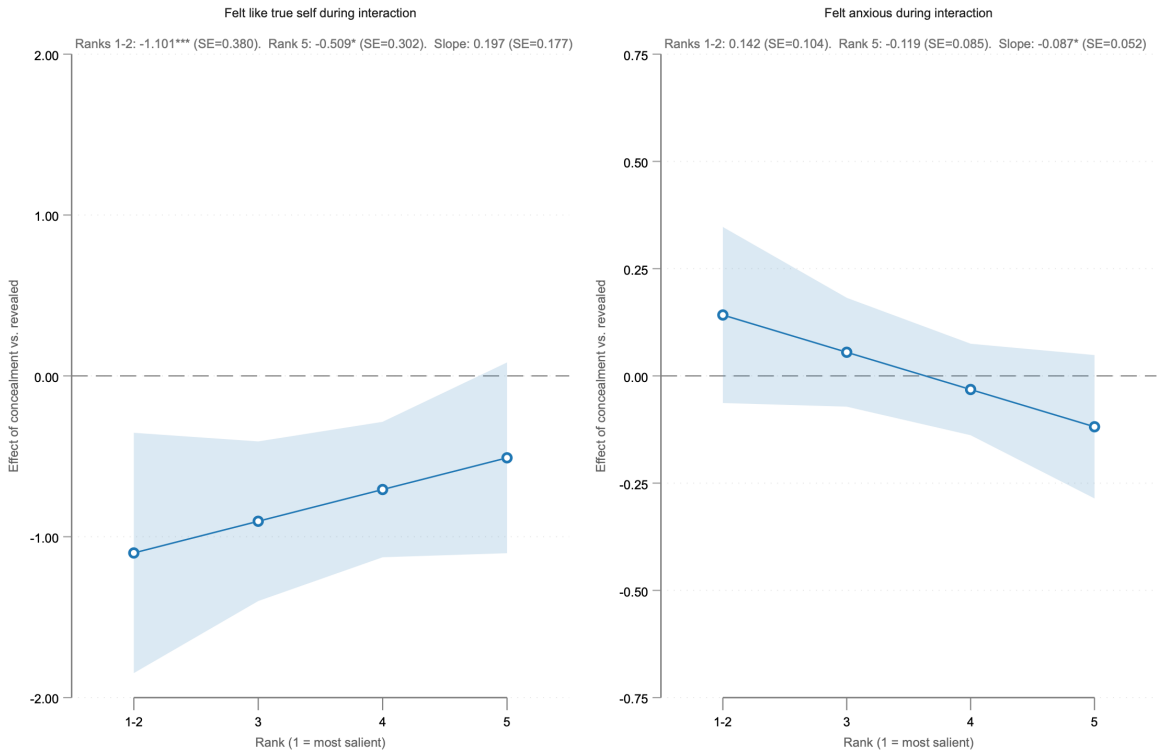
Notes: The figure compares SC and GC participants on traits measured before or during the work session. Panel (a) reports the share of participants who self-assessed as better at the computerized task than the paper task before the session. Actual ability is similar across caste groups. Panel (b) reports how relaxed participants appeared during the task, measured on a 0–3 scale, for mixed-caste pairs only. Bars show group means with 95 percent confidence intervals. Reported p-values are from tests of equality of means across caste groups. OBC participants are excluded.

Figure A7: Reading and typing ability is similar for General and Scheduled Castes



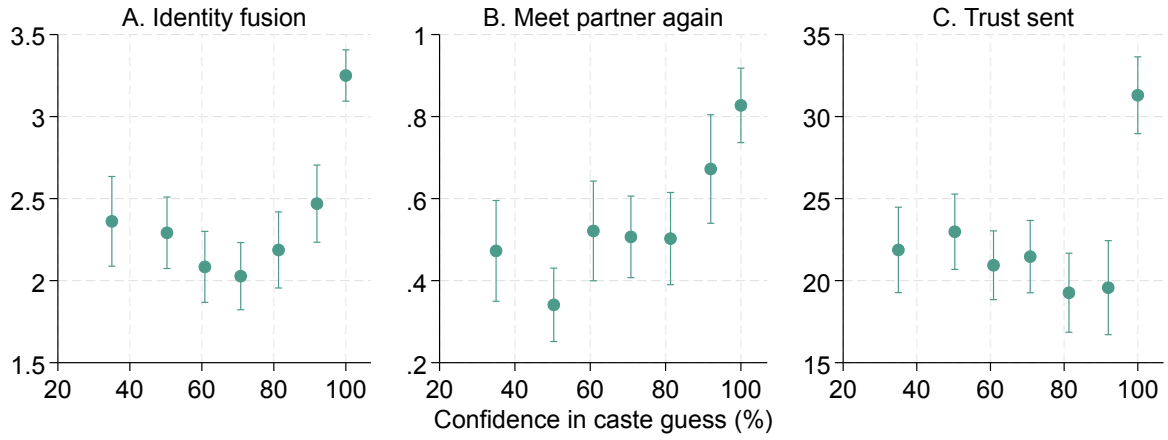
Notes: The figure shows kernel density estimates of reading and typing percentile ranks for General and SC participants using data from the baseline productivity tasks. Percentiles are constructed within task, with higher values indicating greater productivity. Vertical dashed lines indicate group means.

Figure A8: Concealment hurts more if caste identity is salient



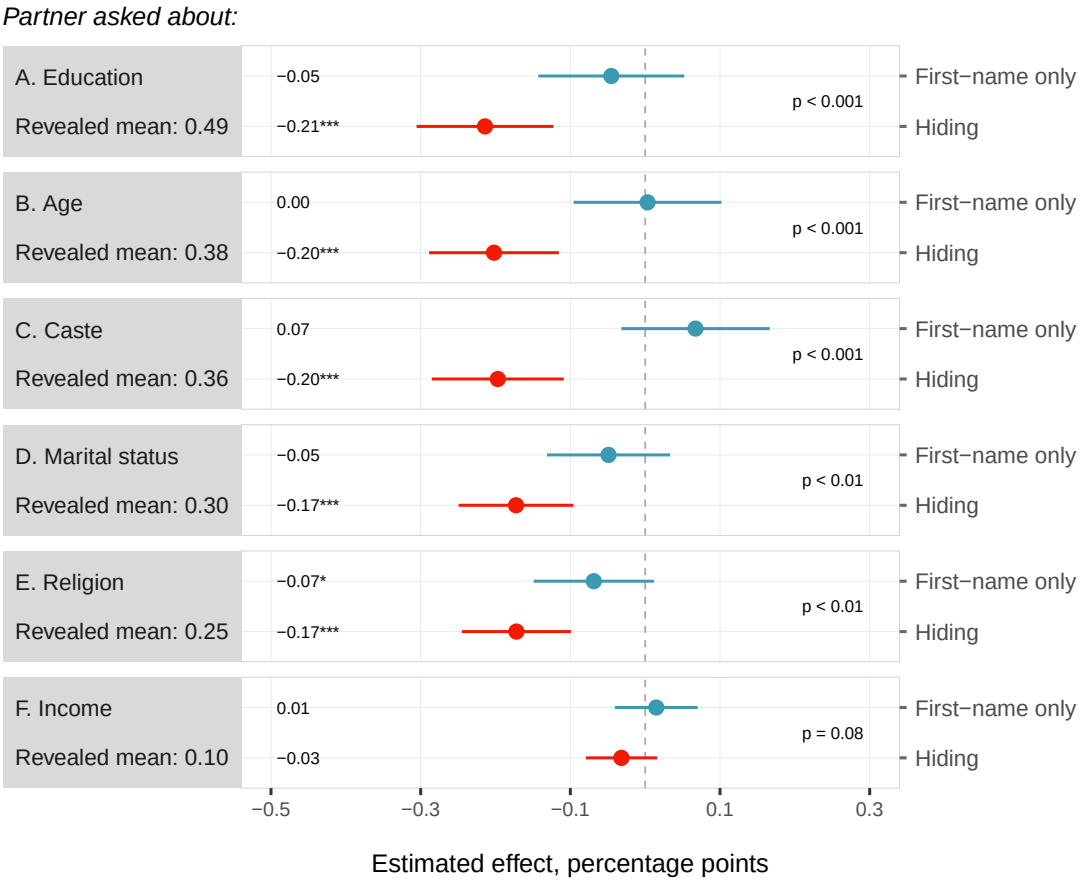
Notes: Each panel plots the marginal effect of concealment (status quo or hiding) relative to the revealed condition on the outcome indicated in the title, at each level of own caste identity salience. Identity salience is measured by participants' pre-experiment ranking of five identities (caste, religion, Indian, from U.P., and father/brother/son), where rank 1 denotes the most salient identity; ranks 1–2 are pooled due to small cell sizes. Marginal effects are derived from an OLS regression of the outcome on an indicator for concealment, the salience rank (entered continuously), and their interaction, with controls for age, education bin, partner's salience rank, and demeaned randomization strata fixed effects. Standard errors are clustered at the pair-level. Shaded bands denote 95% confidence intervals. The subtitle reports the marginal effect at ranks 1–2 and rank 5, and the estimated interaction slope (change in concealment effect per unit increase in rank). *** p<0.01, ** p<0.05, * p<0.1.

Figure A9: Is there a certainty premium in intergroup relations?



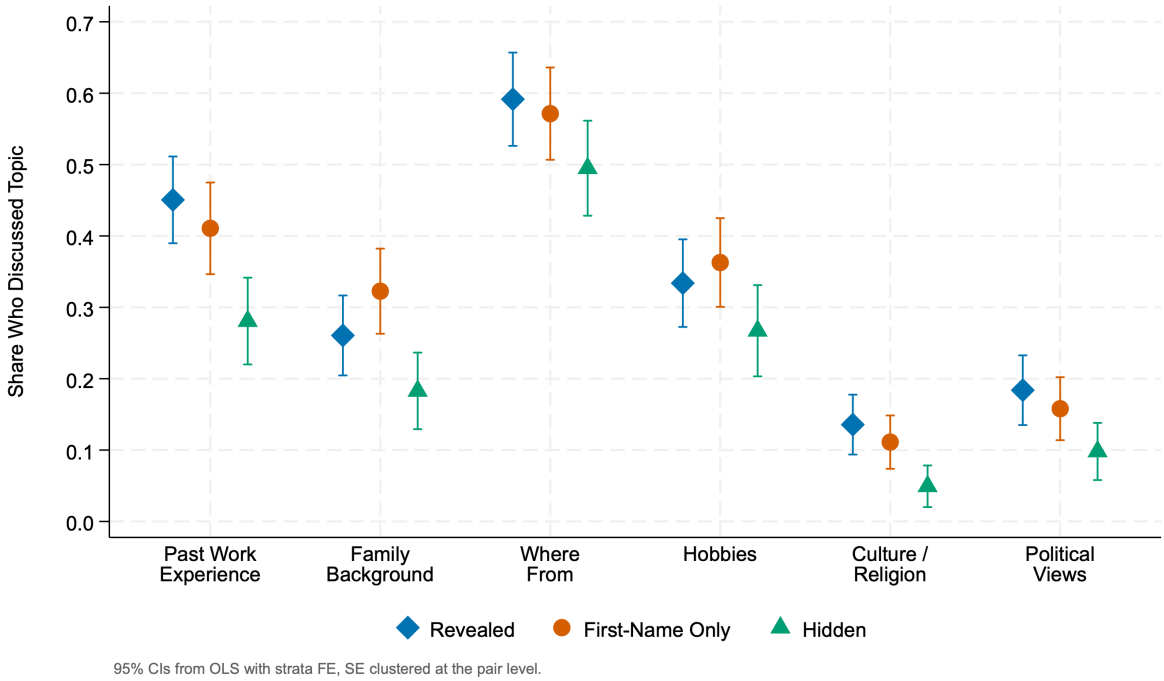
Notes: The figure plots adjusted mean predictions from OLS regressions of each outcome on confidence bins, controlling for dummy for *Hidden*, dummy for guessing partner is General-caste fully interacted with whether participant is General-caste, age, education, and randomization strata (General-caste dummy interacted with work date). The sample is restricted to the *First Name Only* and *Hidden* arms. Standard errors are clustered at the pair-level. Confidence in the caste guess is grouped into seven bins: below 50%, 50–59%, 60–69%, 70–79%, 80–89%, 90–99%, and 100%. Each point is plotted at the mean confidence within the bin; capped lines are 95% confidence intervals.

Figure A10: *Hidden* reduces questions about identity-linked topics.



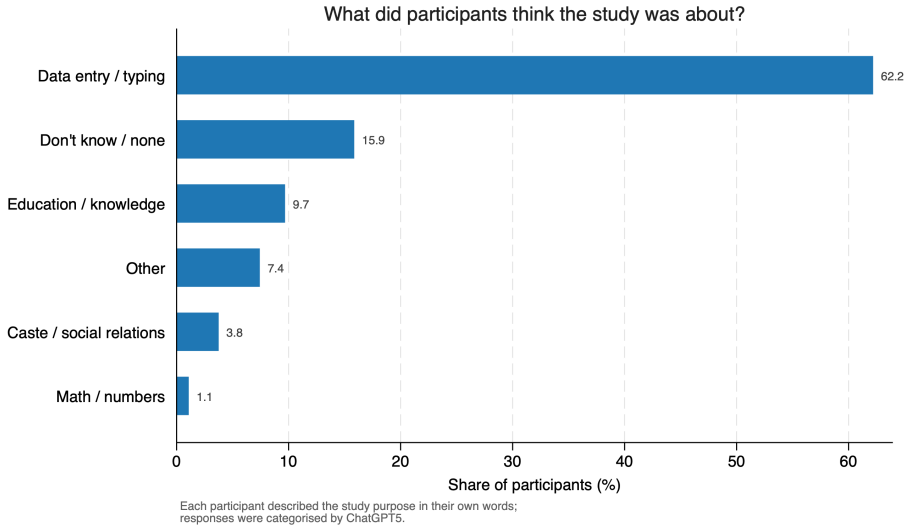
Notes: The unit of analysis is the participant. The outcomes measure whether the participant reports that their partner attempted to discover specific personal characteristics during the work session. In particular, participants were asked whether their partner tried to discover their age, education level, family income, caste (e.g., by asking their last name), marital status, or religion (0 = No, 1 = Yes). Each point reports the estimated treatment effect relative to the *Revealed* condition. The *Revealed mean* is the raw mean of the outcome for participants in the *Revealed* condition. The *First Name Only* and *Hiding* estimates are coefficients from regressions that control linearly for participants' age and include randomization-strata and education-bin fixed effects. Horizontal bars show 95% confidence intervals based on standard errors clustered at the pair level. Reported p-values test the null hypothesis that the *First Name Only* and *Hiding* treatment effects are equal. $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

Figure A11: Topics discussed with partner by treatment



Notes: Each point is the share of participants who reported discussing that topic with their partner. Estimated via OLS with no constant, including demeaned strata fixed effects; 95% confidence intervals shown. Standard errors clustered at the pair-level. All regressions control linearly for participants' age and include randomization-strata and education-bin fixed effects. Sample pools SC and GEN participants restricted to the main analysis sample.

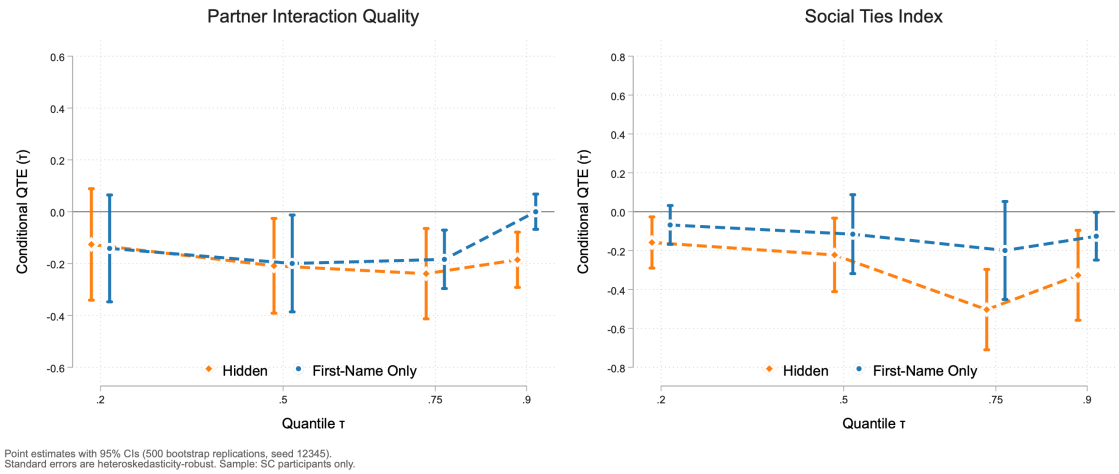
Figure A12: The purpose of the study, according to endline respondents.



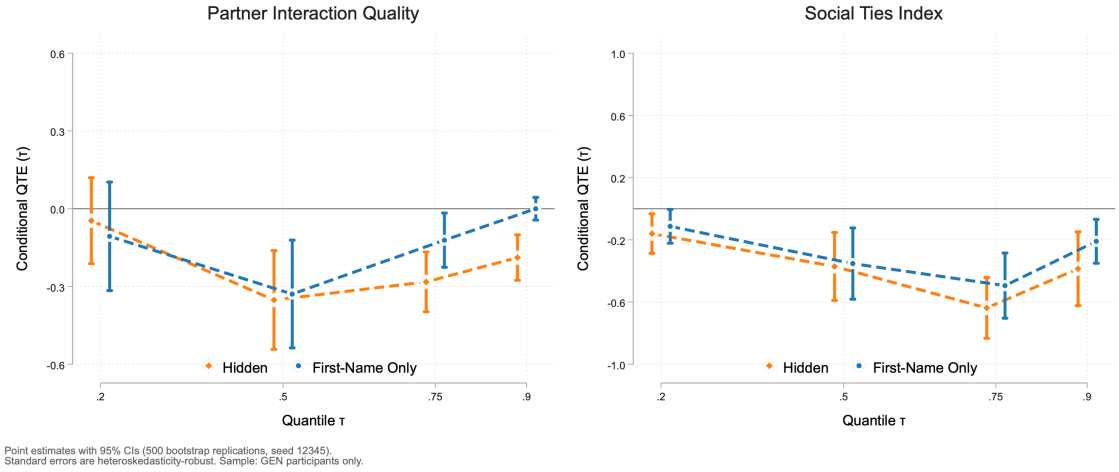
Notes: This figure summarizes endline subjects' reports of the purpose of the study. Subjects were asked at endline, "What do you think was the purpose of this research study? What was it trying to understand?" Subjects' qualitative responses were recorded by enumerators and categorized using ChatGPT5.

Figure A13: Quantile treatment effects of identity-hiding treatments on interaction quality and downstream outcomes

Panel A: Scheduled Caste (SC)



Panel B: General Caste (GEN)



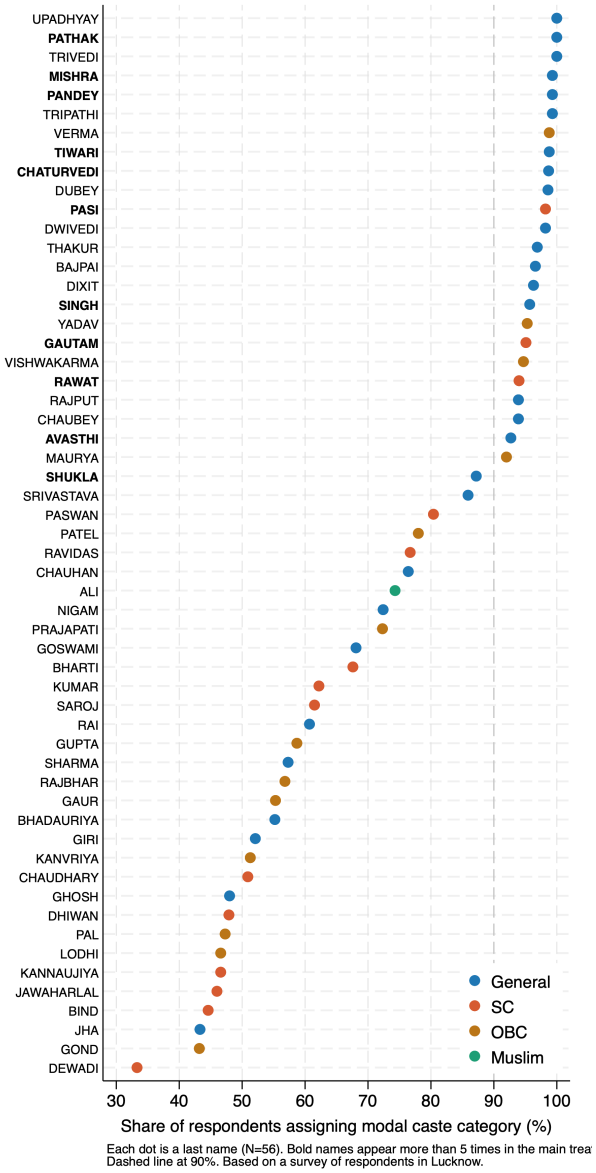
Notes: Each panel reports conditional quantile treatment effects (QTEs) of the *Hidden* and *First-name only* treatments relative to the Revealed group at different quantiles τ of the outcome distribution. Vertical lines denote 95% confidence intervals. Panel A shows results for Scheduled Caste (SC) participants; Panel B for General caste participants. The left column reports QTEs for an index summarizing the quality of interaction with the partner, constructed from responses on perceived friendliness, rudeness, tension, bossiness, and authenticity in interaction. The right column shows QTEs for an index capturing downstream trust and cooperative behavior (trust giving, trust returned, reservation wage to work with partner in the future, and identity fusion). Negative values indicate worse outcomes. Point estimates with 95% confidence intervals based on 500 bootstrap replications, clustered at the pair-level. All regressions control linearly for participants' age and include randomization-strata and education-bin fixed effects.

D Supplementary text

D.1 Selecting caste salient last names for the experiment

To validate that the last names used in our experiment are informative about caste identity, we conducted a supplementary beliefs survey with 460 residents of Lucknow (172 General caste, 194 OBC, 94 SC). Enumerators read respondents a list of 18 or 19 last names drawn from the names used in our experimental sample and asked them to guess the caste category of each name and report their confidence in that guess; across respondents, this covered 56 distinct names in total. Figure A14 plots, for each name, the share of respondents who assigned it to the modal caste category, coloured by that category and sorted from most to least informative. The results confirm that caste is highly legible from last names in this context: the majority of names are assigned to a single caste category by more than 90% of respondents, and the names that appear most frequently in our experiment — including GAUTAM, SINGH, RAWAT, MISHRA, TIWARI, SHUKLA, PANDEY, and AVASTHI (shown in bold) — all have modal shares above 87%. Thus last names displayed in the *Revealed* treatment arm provided participants with meaningful and reliable information about their partner's caste identity.

Figure A14: Caste informativeness of last names



Notes: Each dot represents one of 56 last names. The x-axis shows the share of survey respondents who assigned the name to the modal caste category, coloured by that category (blue = General, coral = SC, amber = OBC, teal = Muslim). Names are sorted from most to least informative. Bold names appear more than 5 times in the main experiment sample. Dashed vertical line at 90%. Based on a beliefs survey of 460 residents of Lucknow.

D.2 Deviations from pre-analysis plan

We followed the pre-registration in both outcome construction and empirical analysis. The deviations are minor and listed below.

- **Meal-sharing outcome.** The pre-registration included a meal-sharing measure as part of the social ties outcomes. Due to an implementation error, enumerators did not distribute the registration slips during the first two days of the study. We therefore cannot include this outcome in the analysis.
- **Well-being outcomes.** Ideally, we would have controlled for baseline well-being when estimating treatment effects on well-being measured at endline. However, we discovered an implementation error: the baseline

well-being questions were not translated into Hindi. We therefore do not use the baseline well-being measure, and omit the baseline outcome control in these regressions.

D.3 Ethical considerations

The study protocols were separately reviewed and approved by four Institutional Review Boards, at the University of California San Diego, University of California Davis, Duke University and University of British Columbia.

D.3.1 Informed consent and experimental design

We obtained informed consent from all adult participants before participation. The study was conducted in a supervised workplace setting, and participants were adequately compensated for both attendance and performance. Because the object of interest is how identity visibility shapes interaction, participants were not told in advance that caste was the focus of the study. This was necessary to limit experimenter demand effects and strategic behavior that would have altered the interactions we aimed to study. The intervention itself was deliberately light-touch, operating through small differences in how partners were introduced and, in one condition, whether they were instructed not to reveal caste-linked information.

D.3.2 Risks to participants

The main ethical concern is that revealing or concealing identity may affect participants' comfort, trust, or sense of authenticity during the interaction. We reduced these risks by conducting the study in a controlled environment and avoiding any explicit discussion of caste by the research team during the work session (with the exception of the instructions to not disclose caste in the *Hidden* condition). Participants were not asked to disclose sensitive information beyond what arose naturally in the course of interaction. The structured setting likely reduced the scope for more extreme forms of discriminatory behavior. Consistent with this, we do not find evidence for large well-being differences across the treatment arms.

D.3.3 Interpretation of results

Our findings should not be interpreted as a general normative claim that identity concealment is harmful or that disclosure is always beneficial. The experiment identifies the effects of identity legibility on a specific margin: social and workplace interactions in a formal setting. In other contexts, limiting the visibility of identity may protect individuals from more severe forms of discrimination, exclusion, or violence, which our design is not intended to measure.

More broadly, our results should not be read as placing the burden of better social relations on members of marginalized groups. If reduced identity uncertainty improves trust and collaboration, the relevant implication is not that disadvantaged individuals ought to reveal more, but that institutions should be made safer and more inclusive so that identity does not carry such acute social risk in the first place.