ANNUAL REVIEWS

Annual Review of Economics

Social Identity, Group Behavior, and Teams

Gary Charness[1] and Yan Chen[2,3]

[1]Department of Economics, University of California, Santa Barbara, California 93106, USA; email: gary.charness@ucsb.edu

[2]School of Information, University of Michigan, Ann Arbor, Michigan 48109, USA; email: yanchen@umich.edu

[3]Department of Economics, Tsinghua University, Beijing 100084, China

Annu. Rev. Econ. 2020. 12:691–713

First published as a Review in Advance on May 13, 2020

The *Annual Review of Economics* is online at economics.annualreviews.org

https://doi.org/10.1146/annurev-economics-091619-032800

Keywords

social identity, group behavior, norms, teams, experiment

Abstract

The issue of one's identity has loomed large recently and has unfortunately been used more and more as a wedge to separate subgroups. It is important to understand the ramifications of identity, both to limit the negative consequences (such as so-called identity politics) and to be able to use one's sense of identity as a positive force in the world. What are effective approaches to allow positive identities and pride about one's social identity to be reinforced for the greater good? Recent work suggests that some forms of team competition can induce greater effort, which can be applied to areas such as microlending, charitable giving, and organization of the gig economy. And yet many fascinating questions remain; for example, what is the interaction of salience, social norms, and preferences on the effects of social identity in our society?

1. INTRODUCTION

The issue of one's identity has become almost a defining feature of contemporary society, particularly in the political arena. In recent years, identity has been used more and more as a wedge to separate subgroups. It is important to understand the ramifications of identity, both to limit the negative consequences and to be able to use one's sense of identity as a positive force in the world. In this article, we discuss how social identity affects individual behavior in experimental tasks both in the laboratory and the field.

One definition of social identity refers to a person's sense of self, derived from perceived membership in social groups. When we feel that we belong to a group, we may very well derive at least a portion of our sense of identity from that group. Group identity is used to explain such phenomena as ethnic and racial conflicts (Sen 2007), discrimination, political campaigns, and the formation of human capital (Coleman 1961). Tajfel & Turner (1979) developed social-identity theory to understand the psychological basis for intergroup discrimination. According to this theory, social identity has three major components: categorization, identification, and comparison.

The first component, categorization, is the process of putting people, including ourselves, into categories (for a model of optimal categorization, see Fryer & Jackson 2008). Labeling someone as a Muslim, a female, or a soldier is a way of defining these people. Similarly, our self-image is associated with the categories to which we belong. Social-psychology experiments show that people quickly and easily put themselves and others into basic categories. The second component, identification, is the process by which we associate ourselves with certain groups. In-groups are groups with which we identify, and out-groups are ones with which we don't identify. The third component, comparison, is the process by which we compare our groups with other groups, creating a favorable bias toward the group to which we belong.

One insight from social-identity theory is that the groups to which people belong mean something to them. Once a person sees herself as part of a group, she derives self-esteem from that group membership and adopts behaviors that are consistent with the norms and stereotypes associated with that particular group identity (Shih et al. 1999). Benjamin et al. (2010) find that making ethnic, racial, or gender identity salient also causes risk and time preferences to conform to common stereotypes.

Since group identity affects individual behavior, many experiments in social psychology assess whether and to what extent people interact with in-group and out-group members differently. Most of these experiments confirm the finding by Tajfel et al. (1971) that group membership creates in-group enhancement in ways that favor the in-group at the expense of the out-group. Many of these experiments use the minimal-group paradigm. In a typical minimal-group experiment, subjects are randomly assigned to groups, which are intended to be as meaningless as possible. The subjects then assign points to anonymous members of both their own group and the other group. In these studies, subjects tend to award more points to people who are identified as in-group members. Experiments involving ratings of in-group and out-group members have found that participants tend to rate in-group members higher than out-group members.

The purpose behind the experiment by Tajfel et al. (1971) was similar to that underlying the famous Milgram experiments with simulated electric shocks—to try to understand how people could have behaved as did even ordinary Germans during the Nazi years. In this light, social identity can be seen as a negative force, and it is wise to learn how to manage it. In another light, perhaps social identity can instead be harnessed for social benefit. In fact, some recent experimental work (e.g., Ai et al. 2016, Charness & Holder 2019) has demonstrated this possibility.

Another interesting element is the issue of what activates a particular facet of one's identity or makes it salient. One's sense of social identity may depend on what is highlighted in one's

environment. Since one's identity is naturally multifaceted, there is the issue of which identity comes to the fore and the extent to which this can be influenced. Skilled politicians can bring out the worst in people by manipulating this (Sen 2007); perhaps one can instead imagine bringing out the best in people. One example of the latter is German Chancellor Angela Merkel's assertion that "Wir schaffen das!" ("We can do it!"), stated to encourage people to help millions of refugees.

In this vein, Shih et al. (1999) provide an interesting example of the effect of identity priming. They study stereotype susceptibility with a group of Asian American female undergraduates given a math test. Performance was dependent on which powerful stereotype (i.e., Asians possess excellent quantitative skills and women do not) was evoked. An open question is what determines which priming has the greatest effect when multiple identities are involved. If an Asian woman is primed about both being Asian and being female, what is the overall effect? Does the order make a difference?

Charness et al. (2014) also employ two identities in a public-goods game with endogenous network formation. One dimension is the endowment received, either high or low. Before the game, people solved anagrams in groups (on a common-fate basis) in a team-building exercise that was calibrated to yield almost certain success. It turns out that the monetary identity prevails, since high-endowment types tended to form groups with other high-endowment types rather than with their friends, with whom they participated in a successful and enjoyable task. Nevertheless, the team-building task exercise greatly increased contributions.

One issue that has largely been ignored in the literature is which aspects of identity are context dependent and which ones are not. Previous work suggests that the dimension of in-group versus out-group seems to be relatively hardwired, while priming identity seems fraught with some peril, since priming can be tricky and sensitive both to the task and to beliefs about what the priming group would do. Furthermore, much as nudges can be counterproductive if one resents being pushed, priming might well also find resistance. In effect, another element of one's identity (perhaps one's individuality or independence) could be inadvertently primed by too strident an appeal. Also, to the extent that one's group identity is not automatic, there is the potential for interventions to affect behavior. Furthermore, if a salient identity such as a stereotype is interfering with desired social policies, can one usefully introduce information to people to teach them to overcome or ignore this stereotype? Would this education be most useful when it pertains to prevailing social norms? Identity manipulation would appear to be a double-edged sword.

One might also wonder about the origins of views about how a certain identity should behave and the sensitivity of identity to context. While we cannot definitively answer such questions, this most likely reflects perceived social norms. When there are multiple salient social norms, people might choose the norm most useful for them in some sense. Charness et al. (2019) suggest that people are to some extent able to ignore social norms when they are financially inconvenient but are quite happy to go along when they coincide with their material interests. But considerably more research is needed.

Our main goal for this review is to present the main approaches, findings, and open questions in identity economics, with the hope of attracting more researchers into this new area. For an independent survey of group identity and intergroup bias with applications to labor market discrimination, we refer to Li (2020). We present some theoretical models of group membership in Section 2 and discuss methods for measuring or inducing identity in Section 3. In Section 4 we consider how being on a team affects behavior in the laboratory and in the field, and we conclude in Section 5.

2. THEORETICAL MODELS

The systematic introduction of social identity into economic analysis started with Akerlof & Kranton (2000), who presented a preference-based model of social identity. Since then, there have been two approaches in the economic models of identity: preference-based and belief-based models. We briefly introduce the major classes of models as well as their implications.

2.1. Preference-Based Models

Preference-based models fall into three categories. Earlier work by Akerlof & Kranton (2000, 2002, 2005) emphasizes the importance of prescriptions or norms associated with different social categories. Subsequent theoretical research endogenizes the choice of identities and the emergence of associated norms in various contexts (Fang & Loury 2005; Darity et al. 2006; Fryer & Jackson 2008; Shayo 2009; Carvalho 2013; Bernard et al. 2016; Akerlof 2016, 2017). Lastly, Chen & Li (2009), Basu (2010), and Chen & Chen (2011) use group-contingent social-preference models to predict how group identity might affect behavior. We introduce each of these classes of models below.

2.1.1. Social categories and norms. Akerlof & Kranton (2000) propose a neoclassical utility function, where identity is associated with different social categories, each with a prescription or norm for behavior. Deviations from the prescription cause disutility. Let C represent the set of social categories. Each person j belongs to a category, c_j. Prescriptions (or norms) P indicate the behavior appropriate for people in different social categories in different situations. A person's identity or self-image, $I_j(a_j, a_{-j}; P, \varepsilon_j)$, depends on her own action, a_j, others' actions, a_{-j}, the prescription for behavior, P, and her idiosyncratic characteristics, ε_j. Agent j's utility function is then

$$U_j = U_j(a_j, a_{-j}, I_j). \qquad 1.$$

This versatile framework has been applied to analyses of gender discrimination, the economics of poverty and social exclusion, the household division of labor (Akerlof & Kranton 2000), the economics of education (Akerlof & Kranton 2002), and the economics of organization (Akerlof & Kranton 2005).

We outline the principal-agent model formulated by Akerlof & Kranton (2005) as an application of Equation 1. In this model, a worker can take on one of two identities, $c = \{N, O\}$, where N (respectively, O) is the identity of an insider (respectively, outsider) who identifies (respectively, does not identify) with the firm, with the prescription of a high-effort activity A (respectively, low-effort activity B), with effort level $e_A > e_B$. A worker's utility function is then

$$U_j(y_j, e_j; c) = \ln y_j - e_j + I_j(c) - t_j(c)\left|e^*(c) - e_j\right|,$$

where y_j is the worker j's income, e_j is her actual effort, c is her social category, and $t_j(c)|e^*(c) - e_j|$ is the disutility from diverging from the ideal effort level for category c, $e^*(c)$.

A risk-neutral principal chooses the wage to maximize expected profit subject to the incentive-compatibility and individual-rationality constraints. The solution highlights how identity economics can be used in organization design. For example, there is less variation in insiders' wages, and the principal pays less at the top. Furthermore, under certain conditions, it is profitable for the principal to invest in motivational capital to change a worker's identity from an outsider to an insider.

These conclusions find support in a pair of experimental studies in a principal-agent framework with moral hazard, using the hidden-action trust game introduced by Charness & Dufwenberg (2006). Using natural identities (Indian villages), Dugar & Shahriar (2012) find that the efficient outcome arises more frequently when the principal and agent live in the same village than when they live in different villages. In comparison, Jiang & Li (2019) use induced artificial identities to investigate how group identity influences the principal's choice of incentive and the agent's real effort. In this setting, the principal and agent, whose group identities are randomly assigned but enhanced via collective puzzle solving, interact in a real-effort game with hidden actions under a revenue-sharing contract. They find that the principals make more generous offers to the in-group agents, who are more responsive in effort to the in-group principals' offers. These findings are consistent with those of Akerlof & Kranton (2005) on the interplay of monetary incentives and group identity in tackling the moral hazard problem.

In his presidential address to the American Economic Association, George Akerlof (2007, p. 7) discussed the exogeneity of such norms in identity models and pointed to the "incorporation of such endogeneity" as the next step. If one could induce people to internalize social norms, this would most likely enhance the effect of identity considerations on behavior.

2.1.2. Endogenous choice of group identities: social distance and status. To endogenize the choice of group identities and norms, Shayo (2009) and Bernard et al. (2016) analyze individuals' decisions to identify with social groups, focusing on the effects of social status and social distance. In Shayo (2009), if an individual i identifies with group J, her utility function is as follows:

$$U_{i,J}(a) = \pi_i(a) - \beta d_{iJ}(a) + \gamma S_J(a), \qquad 2.$$

where π_i is i's material payoffs, d_{iJ} is i's perceived distance from group J, and S_J is the status of group J. The parameters β and γ are positive weights on social distance and group status, respectively, which can vary across individuals. Each agent is characterized by a vector of attributes $q_i = (q^1, q^2, \ldots, q^H)$. A social group J is characterized by the typical attributes of its member, denoted q_J, which is the mean across group members. The perceived distance between individual i and social group J is defined as a weighted Euclidean distance function:

$$d_{iJ} = \sqrt{\sum_{h=1}^{H} w_h \left(q_i^h - q_J^h\right)^2}. \qquad 3.$$

In comparison, the status of group J is determined through social comparison with other groups along valued dimensions of comparisons. Let Π_J and $\Pi_R(J)$ be measures of group J's and its reference group's material payoffs, respectively. The status of group J can be characterized as $S_J = S[\Pi_J, \Pi_R(J), \sigma_J]$, where σ_J captures other determinants of group J's status.

Using this framework, Shayo (2009) defines the social identity equilibrium, which requires not only that players' actions be optimal given others' actions but also that each player's social identity be optimal given her social environment. He shows how social-identity considerations can explain why individuals are willing to support policies that do not maximize their individual material payoff. The driving force behind this possibility is that poor individuals can benefit from sacrificing their own material payoff, as this increases their nation's social status. In some sense, the model thus explains how adopting an identity based upon an attribute shared by others (nationality) leads to internalizing their material payoffs due to social-identity considerations.

Shayo and coauthors apply this framework to taxation and income redistribution, international trade and other domains, summarized in a survey article in this volume (Shayo 2020). Using a

similar framework, Bernard et al. (2016) present a model that directly links a group's social status to its member composition. In this model, both group status and social distance depend only on endogenous group stereotypes and on actual group members' characteristics and actions. Hence, group status does not depend on some exogenous factor, and social distance itself becomes endogenous. More specifically, in a two-type model, $\theta_i \in \{\theta_H, \theta_L\}$, a fraction $\lambda \in (0, 1)$ of agents belongs to the high type. Thus, the average type in a group K, $\lambda_K \in (0, 1)$, represents the group status or type-based stereotype. A type-based social distance is defined as $d(\theta_i, \lambda_k)$, measuring the Euclidean distance between θ_i and λ_k. Furthermore, agents can take one of two actions, $a_i \in \{0, 1\}$, where $a_i = 1$ involves a cost $c_i > 0$, which may or may not be type-specific. Thus, an action-based social distance is defined as $D(a_i, \bar{a}_K)$, the Euclidean distance between an agent's action and group k's average action. The utility function is thus defined as follows:

$$u(k, a_i|\theta_i, \lambda_k, \bar{a}_K) = U\left[\lambda_k, d\left(\theta_i, \lambda_k\right), D\left(a_i, \bar{a}_K\right)\right]. \quad 4.$$

In this model, the endogenous group stereotype (and, consequently, the endogeneity of group status and social distance to individual identification choices) gives rise to two types of equilibria. In a social free-riding equilibrium, low types free-ride on high types' contributions to group status, and group composition is representative of society as a whole. In comparison, a full-segregation equilibrium emerges if and only if the benefit to low types of social free-riding cannot compensate for the associated increase in social distance. By contrast, Shayo's (2009) model does not capture social free-riding due to the assumed structure of group status and social distance. In his model, it is actually good for the rich if the poor identify with the nation, since this yields lower taxes and a higher social status with no effect on social distance. By construction, it is not possible for the poor individuals to identify with the rich group, thereby dragging down group status and increasing social distance in Bernard et al.'s (2016) model.

To test this model in a laboratory experiment, Hett et al. (2020) measure individual identification preferences as a potential source of behavioral heterogeneity and find that both social status and social distance shape discrimination behavior. Specifically, facing a trade-off between monetary payments and belonging to different groups, subjects are willing to forego significant earnings to avoid certain groups and thereby reveal their identification preferences. These identification preferences are systematically related to behavioral heterogeneity in group-contingent social preferences (Chen & Li 2009). This illustrates the importance of identification as a choice and its relevance for explaining individual behavior.

2.1.3. Social interactions. Another path to endogenizing the identity-formation process is through models of social interactions, where the choice of identity is an equilibrium phenomenon. This is a vibrant area of theoretical research, which ties to the large literature on social networks. Here, we selectively review a few studies.

From the perspective of information processing, Fryer & Jackson (2008) build a model of how experiences are sorted into categories and characterize an optimal categorization. They further show how this categorization affects decision-making, leading to specific biases such as discrimination against minority groups.

Fang & Loury (2005) present a two-stage game, where agents choose their identity in the first stage and engage in an infinitely repeated income-risk-sharing game in the second stage. In this framework, a collective identity is adopted when, in a subgame-perfect equilibrium, agents make the same first-stage identity choices. In particular, agents interacting within relatively closed social networks may be inclined to embrace the same or similar identities, which may not be socially efficient.

Darity et al. (2006) use evolutionary game theory to model the relationship between wealth accumulation and racial identity. An interesting aspect of this model is that race is two-dimensional. One dimension is exogenous and geographically rooted (Africans versus Europeans), whereas a second dimension is endogenous and socially constructed (individualist versus racialist). People meet in pairs under various matching technologies and obtain payoffs as results of economic production. Using a replicator dynamic, the authors characterize how different matching technologies lead to the formation of identity norms that might push society toward racialism, individualism, or a mixed-identity equilibrium.

Applying this stream of theoretical approach to the economics of religion, Carvalho (2013) studies a specific religious custom, modeling veiling among Muslim women as a commitment mechanism that limits temptation to deviate from religious norms of behavior. His analysis suggests that veiling enables women to take up outside economic opportunities while preserving their reputation within the community.

Related to the endogenous choice of group identity and group norms, Akerlof (2017) proposes a theory of value formation. Agents choose values based on economic considerations as well as the desire for esteem. In a two-person simplified version of the model, each agent makes three choices: (*a*) efforts at two activities (e_{i1}, $e_{i2} \geq 0$), (*b*) whether to value achievements at these activities ($\theta_{i1}, \theta_{i2} \in \{0, 1\}$), and (*c*) whether to initiate interaction with the other agent ($x_i \in \{0, 1\}$). Interaction takes place if either player initiates it. Agent *i*'s utility is

$$U_i = -\frac{1}{2}(e_{i1} + e_{i2})^2 - kx_i + E_i\,, \qquad 5.$$

where the first term reflects the cost of exerting effort and the second term reflects the cost k of initiating social interaction, which can be positive or negative. The final term, E_i, is the utility one receives from both self-esteem and social esteem, that is, $E_i = E_i^i + G(x_1, x_2)E_i^j$, where $G = 1$ if social interaction takes place and $G = 0$ otherwise. The esteem player i grants to player j or herself depends on player j's achievement relative to others at activities valued by i.

The tension arises as agents obtain more esteem from peers if they conform in their choice of values, but they may obtain more self-esteem if they differentiate. This framework is then applied to education, peer effects, etc. Generalizing this framework to the group setting, Akerlof (2016) defines an agent's esteem for group G relative to a comparison population P in state s as

$$E_i(s, G) = N_i(s, G) - \frac{1}{|P|}\sum_{G' \in P} N_i(s, G').$$

The more general framework can be used to model the dynamics of social identity.

In another model of social interaction, Currarini et al. (2016) investigate the phenomenon of homophily in social networks and focus on the role of biases in meeting opportunities for the emergence of homophily. In their theoretical model, agents can either attempt to link only to similar types or put costly effort into searching the whole population. This results in a threshold equilibrium in which agents link to similar others if and only if their social group has a minimum size.

While most identity models based on social interaction have not been tested in the lab, a lab experiment by Currarini & Mengel (2016) considers how homophily is linked to social identity and the tendency to treat others of shared social identity more favorably (in-group bias). In-group biases are substantially decreased when participants can choose with whom to match. Self-selection cannot explain this. Instead, risk aversion plays a crucial role, in line with social-psychology theories that interpret homophily as a way to reduce subjective uncertainty (Hogg 2000).

2.1.4. Group-contingent social preferences. Another way to endogenize the group-specific norms is to generalize social-preference models by incorporating group structures. Chen & Li (2009), Basu (2010), and Chen & Chen (2011) adopt this approach.

To understand the role of social identity in determining behaviors such as reciprocity, distribution, and social-welfare-maximizing actions, Chen & Li (2009) extend the social-preference model in the text of Charness & Rabin (2002) to incorporate group identity. In this two-person version, an individual's utility function is a weighted average of her own and her match's monetary payoffs. To illustrate, let π_i and π_j be player i's and player j's monetary payoffs, respectively. Let w_i denote the weight that player i puts on j's payoff. Player i's preference is represented by

$$\begin{aligned} u_i(\pi_i, \pi_j) &= w_i\pi_j + (1 - w_i)\pi_i \\ &= (\rho r + \sigma s)\pi_j + [1 - (\rho r + \sigma s)]\pi_i, \end{aligned}$$

where $r = 1$ if $\pi_i > \pi_j$ and $r = 0$ otherwise. Similarly, $s = 1$ if $\pi_i < \pi_j$ and $s = 0$ otherwise. Therefore, the weight player i places on player j's payoff, $w_i = \rho r + \sigma s$, may depend on the comparison between i's and j's payoffs. The parameter ρ measures i's charity concern when her payoff is higher than her match's, while σ measures i's envy when her payoff is lower than her match's. Chen & Li (2009) incorporate group identity into the model by redefining the weight that player i puts on player j's payoff as $w_i^g = \rho(1 + ga)r + \sigma(1 + gb)s$, where $g = 1$ if players i and j belong to the same group and $g = 0$ otherwise. The parameters a and b capture the additional in-group effects for charity and envy, respectively. For example, when i receives a higher payoff than j, the parameter ρ measures the charity effect for an out-group match, while $\rho(1 + a)$ measures the charity effect for an in-group match. The difference a measures the additional effect of in-group identity on an individual's charity concerns. Therefore, the new utility function for player i is

$$U_i\left(\pi_i, \pi_j\right) = w_i^g\pi_j + \left(1 - w_i^g\right)\pi_i. \qquad 6.$$

Within the class of potential games, Chen & Chen (2011) use a simplified group-contingent social-preference model to incorporate social identity into players' social preferences, to demonstrate how identity can change equilibrium selection by changing the potential function (Monderer & Shapley 1996). A potential function is a global function defined on the space of pure strategy profiles such that the change in any player's payoffs from a unilateral deviation is exactly matched by the change in the potential P. Let x_i be player i's strategy in a normal-form game. Monderer & Shapley (1996) formally define potential games as games that admit a potential function P such that

$$\pi_i(x_i, x_{-i}) \geq \pi_i(x_i', x_{-i}) \Leftrightarrow P(x_i, x_{-i}) \geq P(x_i', x_{-i}), \forall i, x_i, x_i', x_{-i}. \qquad 7.$$

As noted by Monderer & Shapley (1996), the minimum-effort game (Van Huyck et al. 1990) is a potential game with the potential function

$$P(x_1, \ldots, x_n) = a \cdot min\{x_i, \ldots, x_n\} - c\sum_{i=1}^{n} x_i. \qquad 8.$$

Incorporating group identity into the payoff function of the minimum-effort game, we obtain

$$\begin{aligned} U_i(x) &= w_i^g \cdot \bar{\pi}_{-i} + \left(1 - w_i^g\right) \cdot \pi_i(x) \\ &= \min\{x_1, \ldots, x_n\} - c \cdot \left[w_i^g \cdot \bar{x}_{-i} + \left(1 - w_i^g\right) \cdot x_i\right], \end{aligned}$$

where w_i^g is player i's group-contingent other-regarding parameter, $\bar{\pi}_{-i} = \frac{\sum_{j \neq i} \pi_j(x)}{n-1}$ is the average payoff of the other players, and $\bar{x}_{-i} = \frac{\sum_{j \neq i} x_j}{n-1}$ is the average effort of the other players. Based on

estimations of w_i^g from Chen & Li (2009), we expect it to hold empirically that $w_i^I > w_i^N > w_i^O$, where $w_i^g \in \{I, N, O\}$ represents the weights on in-group, control (no group structure), and out-group members, respectively.

The transformed game defined by the utility function $U_i(x)$ is a potential function, which admits the following potential function:

$$P(x_1, \ldots, x_n) = \min\{x_1, \ldots, x_n\} - c \cdot \sum_{i=1}^{n} (1 - w_i^g) x_i,$$

which differs from Equation 7. Therefore, incorporating group identity changes the potential function, leading to a change in the potential-maximizing equilibrium from the lowest to the highest effort. To reconcile the theoretical predictions with noisy experimental data, Chen & Chen (2011) use the stochastic potentials framework and make several predictions regarding the role of group identity in normal-form games and test the first one in the lab.

1. In a minimum-effort coordination game, increases in the group-contingent social-preference parameter, w_i^g, result in a higher equilibrium effort (in the sense of first-order stochastic dominance).
2. In the class of the prisoner's dilemma and voluntary-contributions public-goods games, increases in the group-contingent social-preference parameter, w_i^g, lead to an increased likelihood of cooperation.
3. In a battle of the sexes game, increases in the group-contingent social-preference parameter, w_i^g, lead to an increased likelihood of coordination.

The experiment reported by Chen & Chen (2011) supports the prediction about the minimum-effort game. In their online appendix, they also categorize the coordination games run by other researchers, including Bornstein et al. (2002) and Weber (2006), according to potential maximization and find that they are largely consistent with the theoretical predictions. The predictions regarding other games are consistent with results reported by Eckel & Grossman (2005) and Charness et al. (2007).

2.2. Belief-Based Models

In comparison to preference-based approaches, Bénabou & Tirole (2011) model identity as beliefs in which people value and invest. In this model, people care about who they are but have imperfect self-knowledge. Thus, they infer their own values from past choices. At date zero, an individual starts with an initial endowment of some asset, forms initial beliefs about herself, and decides whether to invest (e.g., in prosocial actions with future returns). In this model, identity investments are self-signals. At date one, people update their beliefs about themselves from past choices, decide whether to reinvest, and savor or dread date-two prospects. At date two, the individual derives long-run welfare from the final stock.

This model emphasizes the management of beliefs and the cognitive mechanisms leading to identity investments, with several interesting predictions. For example, identity investments are hill-shaped with respect to prior confidence in being a high type (a moral person). High endowments trigger escalating commitment and a treadmill effect, where increases in wealth and social status induce continued investment even when the marginal return no longer justifies it. Lastly, social interactions induce both social and antisocial norms of contribution, sustained by respectively shunning free-riders or do-gooders, which explains the cross-cultural variations in punishment behavior in public-goods games (Herrmann et al. 2008).

We are not aware of experimental studies of belief-based identity models, which is clearly a promising area of future research. Beliefs about the applicable social norms and the attitude of one's group toward a choice will clearly interact with other factors such as self-interest. Once again, the context may very well matter: In some environments, these beliefs may be relatively immutable, while in others, they might be quite sensitive to the context and method by which one attempts to change them.

3. EXPERIMENTAL METHODS: LAB AND FIELD

Two main experimental methods in social identity research have been used extensively in social psychology and experimental economics. The first method induces artificial group identities in the laboratory, whereas the second method primes natural group identities.

3.1. Inducing Artificial Identities: The Minimal-Group Paradigm and Near-Minimal Groups

The first experimental method in social-identity research relies on induced group identities in the laboratory. Starting with work by Tajfel et al. (1971), a sequence of laboratory experiments using the minimal-group paradigm led to the formulation of social-identity theory (Tajfel & Turner 1979). The minimal-group paradigm is an experimental design, whereby groups are created using trivial and sometimes almost meaningless tasks. The criteria for a group to be minimal (Tajfel & Turner 1986) include the following:

1. By the group assignment rule, subjects are randomly assigned to non-overlapping groups on the basis of some trivial tasks.
2. No social interaction takes place between the subjects, where social interactions include both face-to-face and technology-mediated interactions, such as online chat.
3. Group membership is anonymous.
4. The decision task requires no link between a chooser's self-interest and her choices.

Two tasks are common in measuring in-group bias. In the first task (other-other allocation), each subject awards amounts of money to pairs of other subjects who are anonymous except for their group membership (Turner 1978, Chen & Li 2009). This task is now often called the allocation game (Lane 2016). Another frequently used task is evaluative ratings of other subjects (Mullen et al. 1992). Of the four criteria for groups to be minimal, the fourth one is the least likely to be satisfied in economics environments, where many decisions involve trade-offs between self-interest and group interest.

Summarizing 15 years of social identity research using the minimal-group paradigm, Tajfel & Turner (1986, p. 282) conclude that "the trivial, ad hoc intergroup categorization leads to in-group favoritism and discrimination against the outgroup."[1] Several factors have been found to enhance or mitigate in-group bias, for example, category salience, group status, and relevance of the comparison dimensions (Mullen et al. 1992). Furthermore, summarizing 40 years of social-psychology research on intergroup relations, Brewer (1999) concludes that in-group formation and attachment are psychologically primary while attitudes concerning out-groups are not.[2]

[1] The study by Tajfel et al. (1971) is considered to be the first experiment to use the minimal-group paradigm.
[2] Of course, this leaves the question of how robust is this minimal-group identity to other factors or circumstances in one's environment. One might suspect that this identity is not terribly strong and could be undermined by learning of evidence that conflicts with it.

More recently, however, a number of social-psychology experiments do not find in-group favoritism with minimal groups. A common feature of these studies is that they violate the fourth criterion for groups to be minimal. For example, Yamagishi & Kiyonari (2000) find that, in a modified prisoner's dilemma game with many strategies, while players cooperate more with an in-group member than with an out-group member in the simultaneous-move game, the group effect disappears in the sequential game (where all players are first movers).[3] They argue that the source of in-group favoritism in a minimal group is an expectation of generalized reciprocity with in-group members (in the simultaneous-move game). By contrast, in a sequential game where direct reciprocity is possible, group effects are eliminated.

There are two competing hypotheses in the social-psychology literature on the question of what generates group effects. The first hypothesis is pure categorization alone (Tajfel & Turner 1986), whereas the second one is the expectation of generalized reciprocity among in-group members (Yamagishi & Kiyonari 2000). The evidence from experimental economics points to the latter.

In comparison with social-psychology experiments, economics experiments on group identity almost always involve trade-offs between self-interest and group interest, thus violating the fourth criterion of the minimal-group paradigm. Experiments that satisfy the first three criteria are termed near minimal (Chen & Chen 2011). To observe intergroup differentiation in choices, experimenters usually enhance group identity by using a within-group problem-solving task, such as a puzzle-solving task (Eckel & Grossman 2005, Rong et al. 2016, Jiang & Li 2019) or an online chat to identify painters (Chen & Li 2009, Chen & Chen 2011, Currarini & Mengel 2016, Kranton et al. 2016), by using a common-fate manipulation, or by using an intergroup stag-hunt game that creates variations of both failed and successful intergroup interactions (Cason et al. 2019). The following studies provide examples of various methods for inducing and enhancing group identity.

Eckel & Grossman (2005, pp. 384–85) use induced team identity to study the effects of varying identity strength on cooperative behavior in a repeated-play public-goods game in the laboratory. They find that "just being identified with a team is, alone, insufficient to overcome self-interest." However, actions designed to enhance team identity, such as group problem-solving, contribute to higher levels of team cooperation. Their finding suggests that high degrees of team identification may limit individual shirking and free-riding in environments with a public good. McLeish & Oxoby (2007) study the effects of group identity in simple bargaining games using induced identity. They find that negative out-group opinion can reinforce in-group identity, making in-group members more cooperative with one another.

Charness et al. (2007) conduct both face-to-face and anonymous experiments, which differ by whether a player observes the identity of the other player in the game. In the face-to-face experiments, participants are randomly assigned to the row group or the column group for the duration of the session. Row (column) players go to room R (room C), written on the board. At the beginning of a period, a row player sits on one side of a long table in room R, while a column player who has arrived from room C sits across the table.[4] In this framework, the row player is a host in room R, and the column player is a guest in that same room. Two index cards for each player are placed face down on the table; each player examines these cards and passes one card face down to the experimenter. At the end of the period, the guest returns to the other room. Periods continue until each player has made a choice in each room, once as a host and once as a guest.

[3]There are no second movers in the sequential treatment, although the participants are led to believe that they exist. Every first mover is paid as if the second mover has given the same amount as the first mover.
[4]Similar events, with labels switched, simultaneously occur in the other room.

There is a 2 × 2 design. In the audience treatments, group members belonging to that room are seated in a semicircle behind the active member of their group and silently observe the players.[5] In the no-audience treatments, only the experimenter observes the players, while the inactive players wait for their turn to play in two separate, monitored holding rooms. In the feedback treatments, people in the room learn the outcome immediately after the choices are made (and this is conveyed to the respective groups in the no-audience case).[6] In the no-feedback treatments, results are revealed only at the end of the session. Importantly, payoffs have two components: the outcome of the two games one plays and the outcome of all the games played by one's group members. The first component equals the sum of the payoffs in those two games. The second component equals one-third the sum of the payoffs received by players in one's group (excluding oneself).

The anonymous experiments are known as Tajfel-style experiments: Each player knows whether the other player belongs to the same group or to the other group but does not know this person's identity. Participants choose a yellow or green slip of paper, depending on their preference and the availability of the two colors.[7] They then sit on different sides of the room, according to their color choice, and receive instructions. At the beginning of the period, one learns whether one's counterpart is someone from one's own group (in-group) or the other. At the end of the period, players learn only the outcome of the game in which they participated. This is repeated for ten periods (strangers matching). In-group or out-group membership of the other player in one's pair does not directly affect one's payoffs, and this is explained to the participants.

Both sets of experiments show strong effects. However, in contrast to the minimal-group paradigm, minimal groups alone do not affect behavior in the strategic environment. Nevertheless, salient group membership significantly increases the aggressive stance of the hosts (people who have their group members in the audience) and tends to reduce that of the guests. In the anonymous experiments, the cooperation rate was twice as high with an in-group member and a shared payoff, compared with the other three conditions in the 2 × 2 design.

Chen & Li (2009) use a near-minimal-group method to study the effects of group identity on participant social preferences. Like in classical social-psychology experiments (Tajfel et al. 1971), they assign groups by using participant painting preferences and compare this with random assignment. However, unlike in social-psychology experiments, which focus on allocation among other participants, they use a much wider class of games to systematically measure the effects of identity on various aspects of social preferences, such as distribution and reciprocity preferences. The sample of simple games is selected from Charness & Rabin (2002), which enables the authors to incorporate social identity into the social-preference model and estimate its effects on social preferences.

This experiment has five treatments and one control condition. In the treatment sessions, there are four stages. The first stage is a group assignment stage. The second stage is a collective problem-solving stage using an online chat program. The third stage is the allocation game stage, where each participant allocates tokens to two other participants. The fourth stage is a set of two-person sequential games from the work of Charness & Rabin (2002). The authors then evaluate the relative contribution of each stage to the identity-formation process by removing one stage at a time, with five treatments in total; meanwhile, subjects in the control sessions participate only in the fourth stage.

The results show that participants show a 47% increase in charity concerns and a 93% decrease in envy when they are matched with an in-group member. Likewise, participants are 19% more

[5] Therefore, row players are the audience in room R and column players are the audience in room C.
[6] The feedback consisted of the experimenter turning over the cards simultaneously for all to see.
[7] Initially, there are yellow and green slips in equal number, each sufficient for half of the group.

likely to reward an in-group match for good behavior but 13% less likely to punish an in-group match for misbehavior. Furthermore, participants are significantly more likely to choose social-welfare-maximizing actions when matched with an in-group member. All results are consistent with the hypothesis that participants are more altruistic towards an in-group match. This paper establishes that taste-based discrimination exists in a simple laboratory setting. Furthermore, the group-contingent social-preference model estimated in the paper provides one way to endogenize the exogenous norms described by Akerlof & Kranton (2000).

Is the value of a group positive or negative? Charness & Sutter (2012) point out that groups are considerably better at problem-solving in the interest of maximizing their own payoffs but are much more selfish. Hargreaves Heap & Zizzo (2009) induce artificial identities of insiders and outsiders in trust games. Interestingly, they allow participants to trade group membership. Again, trust appears to diminish in the group setting because outsiders face negative discrimination. However, the authors argue that there is a psychological benefit to group membership. While this could potentially overcome the negative welfare effects, the article concludes that the welfare effect of groups on the whole is neutral at best and potentially negative. And yet, if the environment were one involving problem-solving and cognition rather than social preferences, the effect of group membership could easily be positive.

In comparison with earlier experiments that typically estimate the sample-average intergroup preferences, Kranton et al. (2016) employ an innovative within-subject design, combining two major methods used to study group behavior: near-minimal groups and real-life groups (political party affiliations and opinions). The econometric analysis reveals considerable heterogeneity among individuals, who can be categorized into four distinct types on the basis of their series of choices. About one-third of the subjects are not what the authors term groupy, with stable allocations even when group divisions are salient. Intriguingly, these nongroupy subjects are also less likely to be affiliated with a political party, in contrast to groupy subjects, who show in-group bias in both minimal and political group treatments. The open question is whether groupiness, like risk attitude, is an individual trait that is stable across different contexts.

3.2. Natural Identities: Gender, Ethnicity, Race, and Organization Identities

Several studies using natural groups find significant effects of group identity on behavior. Bernhard et al. (2006) use a dictator-game experiment with third-party punishment in two distinct, native social groups in Papua New Guinea. They find that third parties show stronger altruism toward in-group victims and give in-group norm violators more lenient judgments. Dictators in their study are seen as upholding social norms when they transfer money to in-group members. Therefore, in-group favoritism is a strong force in altruistic norm enforcement and sharing decisions.

Goette et al. (2006) examine the effects of group membership in a prisoner's dilemma game using natural groups (platoons) in the Swiss Army. They find more cooperation when subjects interact with in-group members. In a second experiment similar to the one presented by Bernhard et al. (2006), they also find that third-party punishment is stronger when a violation affects an in-group member as opposed to an out-group member. Lastly, Tanaka et al. (2016) find that group behavior in Vietnamese village communities is affected by the status of the ethnic groups. While Khmer (poor minority) show strong in-group favoritism, Vietnamese (majority) and Chinese (rich minority) do not show in-group bias when they are matched with Khmer, but they do so when they are matched with each other. Hoff & Pandey (2006, 2014) and Afridi et al. (2015) investigate how institutionally imposed identities (e.g., caste identity in India and migrant identity in China) influence children's performance in cognitive tasks under different incentives.

Charness & Rustichini (2011) study experimentally how males and females differ in the way same-gender peers observing their action affects their social behavior. In their experiment,

participants play a prisoner's dilemma game with a partisan audience watching the choice. Two groups (separated into two rooms) participate in each session: both all-male, both all-female, or one all-male and one all-female. Each person plays the game once with an audience of the same group and once with an audience of the other group. The interaction of gender and place significantly affects behavior: Males cooperate substantially less often when observed by their peer group, while females cooperate substantially more often. One possible explanation for this pattern is that males wish to signal their formidability and females wish to signal their cooperativeness. In this way, one's sense of group identity leads to significantly different behavior across gender.

There are pros and cons for using natural group identities. For example, using induced identities might give the experimenter more control over the identity-formation process than is possible with natural group identity. On the other hand, the extent to which induced identity affects behavior depends on the strength of the social identity, which may very well be deeper and stronger with natural group identity. Goette et al. (2012) compare randomly assigned minimal groups to randomly assigned groups involving real social interactions (Swiss army platoons). The beauty of this design is that the real social groups are randomly assigned. They find that, while adding social ties leads to qualitatively similar (although stronger) in-group favoritism in cooperation, altruistic norm-enforcement patterns are qualitatively different between treatments. The latter points to beliefs in a sequential game as the potential factor, which is sensitive and context dependent.

3.3. Priming Natural Identities

Priming is an experimental technique often used in social-identity research. Priming introduces certain stimuli (primes) to activate individuals' social knowledge structures. Primes include text (e.g., questionnaire, article, or word-scrambling game), image, or audio. Research in social psychology has found that subtly making different natural social identities salient through priming can affect behavior and outcomes, such as test performance (Aronson et al. 1998), walking speed (Bargh et al. 1996), or person perception (Bargh & Pietromonaco 1982). Priming gives the experimenter the advantages of using natural social identities but with more control compared with just using natural identities.

Priming social identities can affect people's behavior and attitudes outside of their awareness and control (for a review, see Bargh & Chartrand 1999), as demonstrated in a large body of laboratory work in psychology on identity priming. Making social identities salient often induces study participants to adopt behaviors that are consistent with the identity stereotype. These effects occur even when participants are unaware of being primed. In one study, college students primed with elderly stereotypes walk more slowly as they exit the study than others (Bargh et al. 1996). Steele & Aronson (1995) find that African American students stereotyped to be poor students underperform on academic tests when asked to indicate their race prior to taking the test. These effects have also been documented in other groups such as Hispanic Americans (Aronson et al. 1998), individuals from lower socioeconomic statuses (Croizet & Claire 1998), and women in math (Spencer et al. 1999).

While activating negative stereotypes can hurt performance, activating positive stereotypes can boost performance. Shih et al. (1999) examined the performance of Asian women on a mathematics test. Women are stereotyped to have inferior quantitative skills (Benbow 1995, Hedges & Nowell 1995), while Asians are stereotyped to have superior quantitative skills (Steen 1987). As mentioned in Section 1, Shih et al. (1999) find that Asian American women perform better on a mathematics test when their ethnic identity is primed but worse when their gender identity is primed, compared with a control group with neither identity primed. In contrast, Asian Americans taking a verbal test show the reverse pattern of performance. In this case, women are stereotyped to be verbally

talented while Asians are not. Asian American women perform better on the verbal test when their gender is salient and worse when their ethnicity is made salient (Shih et al. 2006). These priming techniques have also been applied to study risk and time preferences in economics (Benjamin et al. 2010).

Chen et al. (2014) choose two ethnic groups, Caucasians and Asians, that fellow participants could differentiate by last names. For Asian participants, they focus on those with Chinese last names in order to avoid potential complex intergroup preferences among different Asian groups. They adopt the priming technique used by Shih et al. (1999) and subtly activate a social category outside of participants' awareness in the identity treatments. The stimuli are introduced through a pre-experiment questionnaire. In the ethnic-identity treatment, the questions pertain to an individual's ethnic background, family history, and cultural heritage. In the school-identity treatment, subjects are asked about which school they attend. They are then asked to reflect on their choices of schools when applying for college. Since the subjects in each experimental session study at the same university, these questions pertain to an individual's common identity of being part of her university.

In the control sessions, the questions are designed to be identity neutral, that is, related to neither the ethnic nor the school identities: Subjects are asked about their activities in leisure time. The identity-neutral questionnaire is designed to preserve direct comparability with the two identity treatments. The primes are designed to make salient the appropriate social identity and activate the constructs associated with the identity.

Subjects then play a series of sequential minimum-effort and prisoner's dilemma games. By priming a fragmenting (ethnic) identity, the researchers find that participants are significantly less likely to choose high effort in the minimum-effort games, leading to less-efficient coordination. In comparison, priming a common identity significantly increases the choice of a joint-payoff-maximizing strategy in the prisoner's dilemma game. This study is among the first in economics to empirically evaluate the effectiveness of using a common identity as a design tool to increase coordination and cooperation among an ethnically diverse group of participants.

Similar effects of priming common group identity are observed in ultimatum bargaining games (McLeish & Oxoby 2011), where subjects are most cooperative in the shared-identity-priming treatment and least cooperative in the distinctiveness-priming treatment. By contrast, in a framed field experiment in Dallas, Li et al. (2017) find that the common identity prime decreases the probability of giving to local charities in a struggling, poor neighborhood, and it increases the probability of giving only qualitatively in the low- to middle-income neighborhood. Their findings suggest that priming common group identity might have context-sensitive effects.

Cohn et al. (2017) report an experiment with bank employees to study whether the informal rules of behavior (i.e., business culture) in the banking industry promote, or at least tolerate, unethical practices to a larger degree than in other industries. Subjects are asked to toss a coin 10 times and report the number of winning tosses online. Before flipping the coins, subjects complete a survey that randomly varies a subset of the questions. Some participants receive questions that prompt them to think about their professional identity, while others are asked questions unrelated to their professional identity. Reminding bank employees of their occupational role increases the likelihood that they feel that others would break rules for personal financial gain. Those bank employees who were primed to think about their professional identity more strongly endorsed the statement that "social status is primarily determined by financial success," and greater approval of this statement was positively associated with dishonest behavior.

This evidence is consistent with the widely shared notion that the culture in financial firms is more lenient toward dishonest behavior than in other industries. The key innovation of this study is the use of identity theory to measure the behavioral effects of a (business) culture.

Identifying the impact of culture is notoriously difficult, as one cannot simply compare the behavior of people working in different firms or industries. Experimentally manipulating the saliency of professional identity within a certain group of professionals effectively eliminates issues of selection and omitted variable bias, as it allows studying otherwise identical people.

Cohn et al. (2015) describe an experiment with inmates of a maximum-security prison that examines whether criminal identity affects rule-violating behavior. The prisoners are asked to flip a coin 10 times and report the outcomes. Since these cannot be observed by anyone else, the prisoners have an incentive to misreport unsuccessful outcomes without having to fear punishment. Before flipping the coins, the prisoners receive a survey that randomly varies whether it includes questions about criminal identity or questions unrelated to that identity. Prisoners who are primed to think about themselves as criminals cheat about 60% more than the control group.

In addition, after completing the same survey as in the original experiment, prisoners are asked to complete a list of word fragments with the first words that come to their minds. Those in the criminal-identity condition are almost twice as likely to use crime-related words in the word completion task. This suggests that the criminal-identity priming increases the mental accessibility of crime-related thoughts. Finally, a placebo experiment shows that noncriminal citizens do not cheat more in response to crime-related reminders; perhaps the effect requires individuals to possess a criminal identity.

The results raise the possibility that the decision to commit a crime is not just a rational cost-benefit calculation but also depends on people's identities, that is, how they see themselves. This can have important implications for policies aimed at reducing crime and recidivism. For example, convictions are often associated with collateral punishments, such as the loss of voting rights and professional licenses. While these forms of punishment could help deter crime, they may also act as constant reminders of convicted felons' criminal identity, which in turn might increase their likelihood of reoffending.

Investigating channels of identity priming, Chang et al. (2019) demonstrate one mechanism for why certain words or ways of phrasing things can cause us to change our preferences: Framing evokes norms that then influence choice. They use a laboratory study to test the impact of describing a series of dictator games with either neutrally framed or politically charged tax-framed language. The latter makes subjects' political identities (Democratic or Republican) salient, causing those groups' social norms to become relevant. In contrast, the neutrally framed treatment does not make these identities salient, making these norms not identity dependent. The frames have a significant impact on both the norms and choices of participants. Tax-framed Democratic norms favor more equal allocations relative to neutrally framed Democratic norms, while tax-framed Republican norms favor the status quo relative to neutrally framed Republican norms. Further, neutrally framed Democratic and Republican norms are not significantly different from each other.

Priming issues also apply to racial and ethnic identity. Seemingly perversely, Peffley & Hurwitz (2007) found that when white participants were informed about racial disparities in executions, 52% strongly favored the death penalty, compared with 36% in a baseline condition. Hetey & Eberhardt (2014, p. 1949) find that "exposure to extreme racial disparities, then, can lead people to support the very policies that produce those disparities, thus perpetuating a vicious cycle." White California voters are given information about more or less extreme racial disparities in the prison population by having them view a set of photographs of incarcerated people that show either a higher (45%) or lower (25%) percentage of black inmates. When the prison population was more black, voters became significantly less likely to sign a real petition aimed at lessening the severity of California's harsh three-strikes law. Exposure to extreme racial disparities, then, can lead people to support the very policies that produce those disparities, thus perpetuating a vicious

cycle. Hetey & Eberhardt (2018, p. 183) suggest that making racial disparities salient can "trigger fear and stereotypic associations linking Blacks with crime." This unwelcome effect leads the authors to envisage three potential strategies for more effectively presenting information about racial disparities: (*a*) offering context, (*b*) challenging associations, and (*c*) highlighting institutions.

4. TEAMS

How is individual performance affected by social identity in a team setting? In recent years, there has been an increasing focus on teams (e.g., Charness & Sutter 2012). In general, parochial identity considerations tend to crowd out social preferences toward out-group members and are often detrimental for overall social welfare, even apart from any conflict that might result from identity wars. But this is not always the case, and we discuss this issue in Section 4.2.

A stronger sense of identity in a team (or group) setting may induce one to make choices that one would not make otherwise. Consider the notion that one should take one for the team or even deliver apparent electric shocks to a hapless person (Milgram 1963, 1974). Informally, an individual with a strong sense of team identity gives more weight to the overall welfare of the team and less weight to one's own individual material well-being. Being explicitly on a team will tend to focus one's identity more and presumably should lead to stronger effects.

There has been only limited study (Ai et al. 2016, Charness & Holder 2019) addressing the specific question of how one's behavior would differ across a team setting and an individual setting. Another dimension that has received even less attention is whether one has chosen to be on a team or in a group or was simply assigned to one. This seems fertile ground for new research.

4.1. Identity-Based Teams in the Laboratory

Bornstein et al. (2002) provide an early example concerning individual behavior in teams. Kugler et al. (2007) conducted a trust game where either individuals or groups of three subjects each were in the role of first or second mover. They find that groups send significantly smaller amounts (by about 20 percentage points) as first movers and return on average slightly smaller amounts. Hence, group choices are closer to the standard rationality paradigm. If first movers are groups, social welfare is significantly smaller.[8,9]

While the willingness to sacrifice one's own money to help or hurt others appears to diminish for individuals when they are in groups, in fact social welfare can be improved in coordination games where there are multiple equilibria. For example, Feri et al. (2010) consider a weakest-link coordination game with Pareto-ranked equilibria, which is a potential game. The five decision-making entities can be either individuals or groups with three members. Each player (either an individual or a three-person group) chooses an effort level between one and seven. The payoff each player receives is higher if they all choose to exert more effort, but it also gets lower—at a faster rate—the lower the minimum choice (weakest link) of all players. The social payoff increases as the minimum effort increases, but choosing a high level of effort is individually risky. They find that the three-player groups not only play more efficient high-effort equilibria more often than individuals but also are more successful in avoiding miscoordination (which in this case means picking different effort levels). These results are consistent with the predictions of Chen & Chen (2011) using a group-contingent social-preference model.

[8]Since the amount returned by the responder was simply a transfer, this does not affect the total social payoffs.
[9]Similarly, Song (2008) studied how group representatives make decisions on behalf of their group in a trust game; the representative made a decision that determined the outcome of a three-person group. Song finds that group representatives send about 20% less as first movers, much as in Kugler et al. (2007).

In the two previous examples, one's sense of group membership leads to more parochial choices, which would go against the direction of a central planner. But this can change in a competitive environment. While competition is often thought to drive out social preferences, particularly with large markets, competition between (or among) groups may actually improve individual performance when there are multiple equilibria. One nice example is offered by Bornstein et al. (2002). Two teams of seven played the minimum-effort game, where the team with a higher minimum effort receives payoffs for that period and the other team receives none; each team receives half in the event of a tie. The mean choice of effort was steady and over five with this competition; in the control treatment, the mean effort dropped to below three by the end of the 10 periods, a highly significant difference. Again, this result is consistent with the predictions that (*a*) interteam competition induces group-contingent social preferences, which changes the potential function of the game, and (*b*) potential maximization predicts convergence to a more efficient equilibrium (Chen & Chen 2011).

One wonders about how the degree of identity, w_i^g, (and whether this is chosen or imposed) would affect these results. This is another area where group-behavior experiments could serve as a basis for further identity studies. Two flexible approaches are useful for forming theoretical predictions. The first is to elicit group-specific norms by using an incentive-compatible norm elicitation method (e.g., Krupka & Weber 2013), while the second uses group-contingent social preferences (Chen & Li 2009, Chen & Chen 2011).

As discussed in Section 3.1, one approach to inducing or enhancing a sense of group identity is to use a team-building task. One purpose of team building is to improve the effectiveness of work teams within organizations. Team building has a long history in social psychology and organizational research, affecting social identity to the extent that there is a sense of shared purpose. It can consist of either sedentary joint problem-solving or activities such as ropes (or challenge) courses, where people must work together to achieve a physical goal. Buller (1986, p. 147) states, "Team-building is one of the most popular interventions in organizational development."

To our knowledge, the first research on team building in experimental economics is by Eckel & Grossman (2005) (reviewed in Section 3.1). Charness et al. (2014) use a 2×2 design to study the effect of identity when two different plausible identities are present. One factor is whether there is a team-building exercise prior to the game. In this exercise, eight people are assigned to four-person groups; in sessions without this team-building exercise, people perform this task individually. In each case, there is a monetary reward for forming a sufficient number of words from a series of letters; the threshold is set low enough so that the reward is received in every case. The second factor concerns the endowments assigned to the players. In one treatment, everyone received the same endowment (25 tokens), while in another treatment, half of the participants received twice this amount (50).

Despite the fairly minimal team-building activity, there is a striking benefit from having word-task groups, as the contribution rate is greatly increased. While endogenous group-formation per se is effective at sustaining a fairly high contribution rate, this rate is greatly enhanced (up to well over 90% of the endowment) by first having a group word-task. Just knowing that other people had all been in a team-building exercise seemed to be enough to boost contribution; when (in another treatment) only four people participated in the team-building task, contributions were definitely lower, even for the people who had been in the team-building task. Overall, team building offers promise for enhancing the provision of public goods.

4.2. Identity-Based Teams in the Field

Finally, theoretical results from social-identity models and laboratory experiments on teams lead us to predict that identity-based teams in the field might offer a useful behavioral mechanism to

increase prosocial behavior. Understanding strategies to increase prosocial behavior has important policy implications. Charities have explored various mechanisms to increase giving, such as seed money, matching gifts, and peer pressure. In comparison, an underexplored class of mechanisms uses group membership and intergroup competition to increase both participation and giving amounts. We review a few such field experiments.

To our knowledge, the earliest public-goods field experiment using team competition is one of fruit harvesting in an orange grove (Erev et al. 1993). In this setting, the authors find that team competition increases productivity, but they do not explore mechanisms that might cause such a team effect.

Ai et al. (2016) report the results of a large-scale field experiment designed to test the hypothesis that group membership can increase participation and prosocial lending for an online crowd-lending community, Kiva. The experiment uses variations on a simple email manipulation to encourage Kiva members to join a lending team, testing which types of team recommendation emails are most effective and measuring the subsequent impact on lending. Messages do increase the likelihood that a lender joins a team, and joining a team increases lending in a short window (one week) following the intervention. The impact on lending is large relative to median lender lifetime loans. The results also show that lenders are more likely to join teams based on location similarity rather than team status. Overall, the findings suggest team recommendation can be an effective behavioral mechanism to increase prosocial behavior.

To explore the mechanism through which joining a team increases giving, Chen et al. (2017) point to two mechanisms at work by using both naturally occurring data and a field experiment where experimenters post different types of messages on Kiva team forums. First, joining a team increases information sharing about specific borrowers on the team forum, which reduces team members' search costs and increases their lending. Second, joining a team increases the pressure to help improve the team's ranking on the Kiva leaderboard. Therefore, effective teams share information and coordinate their loans to reduce search costs and emphasize team competition through goal setting.

Social-identity theory and team building have many potential applications in organization design. For example, the gig economy provides workers with the benefits of autonomy and flexibility, but it does so at the expense of work identity and coworker bonds. Among the many reasons why gig workers leave their platforms, an unexplored aspect is the organization identity. Combining the predictive accuracy of machine learning with theory-based interventions, Ai et al. (2019) develop a team formation and interteam contest field experiment at a large ride-sharing platform, DiDi. The experimenters assign drivers to teams either randomly or on the basis of homophily in age, hometown location, or productivity. Having these teams compete for cash prizes, they find that (*a*) treated drivers work longer hours and earn 12% higher revenue during the contest, with a larger effect (19%) for teams composed of drivers who are more communicative and responsive, and (*b*) drivers in responsive teams continue to earn higher revenue during the two weeks after the contest ends. Together, these results show that platform designers can leverage organization identity to increase revenue and worker engagement in a gig economy.

Further research on how behavior in groups can lead to clarity concerning identity effects might focus on the importance of the strength of one's identity. Endogenous choice of group membership, competition between groups, and team-building exercises are all factors that have been shown to be important in this regard and this field still seems rather embryonic.

5. CONCLUSION

Social identity and group membership can have very strong effects on individual behavior. Here we present many experimental results and theoretical models in support of this claim. It may well

be possible to harness social identity as a positive force. Very recent work by both of us suggests that some forms of team competition can induce greater effort, which can be applied to areas such as microlending, charitable giving, and organization of the gig economy.

But, of course, there are many questions left unanswered. What determines which facet of one's identity comes to the fore in particular situations? If we can learn this, perhaps we can avoid the worst excesses of identity politics, for example. Is it possible for positive identities and pride about one's social identity to be reinforced for the greater good? Studies with identity priming indicate that one can appeal to different aspects of identity; of course, one must be careful in doing so, lest social planners become overzealous in their social engineering. In this regard, is there a socially optimal level of group identity? In times of war, high levels of patriotism might be a great idea, but at other times, this might well be a very poor idea.

More generally, how does identity transfer into behavior? It would appear that one's perception of the social norm applicable to the salient identity is a key element. Norms presumably represent some group consensus regarding underlying views or preferences. In some circumstances, multiple norms may be present, so behavior might be more sensitive to priming or context. Consider the prisoner's dilemma game labeled as either the Wall Street game or the community game or the sensitivity of choices in dictator games to the dictator's sense of entitlement. In other cases, the normative behavior applies over a broader range for the identity at issue. One might conjecture that the clearer (or more unique) the norm, the less the context may matter.

But this is really a question for future work. The study of the effects of social identity on behavior by individuals in groups or separately is still in its very early days. We welcome more research into this exciting and promising area.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank George Akerlof, Daphne Chang, Roy Chen, Alain Cohn, Luise Görges, Florian Hett, Matthew Jackson, Erin Krupka, Sherry Xin Li, and David Rojo-Arjona for helpful comments on an earlier draft and Lucy Jiang and Zhewei Song for excellent research assistance. Y.C. acknowledges financial support from the National Science Foundation via grant number SES-1838994.

LITERATURE CITED

Afridi F, Li SX, Ren Y. 2015. Social identity and inequality: the impact of China's *hukou* system. *J. Public Econ.* 123:17–29

Ai W, Chen R, Chen Y, Mei Q, Phillips W. 2016. Recommending teams promotes prosocial lending in online microfinance. *PNAS* 113(52):14944–48

Ai W, Chen Y, Mei Q, Ye J, Zhang L. 2019. *Putting teams into the gig economy: a natural field experiment at a ride-sharing platform*. Work. Pap., Univ. Mich., Ann Arbor

Akerlof GA. 2007. The missing motivation in macroeconomics. *Am. Econ. Rev.* 97(1):5–36

Akerlof GA, Kranton RE. 2000. Economics and identity. *Q. J. Econ.* 115(3):715–53

Akerlof GA, Kranton RE. 2002. Identity and schooling: some lessons for the economics of education. *J. Econ. Lit.* 40(4):1167–201

Akerlof GA, Kranton RE. 2005. Identity and the economics of organizations. *J. Econ. Perspect.* 19(1):9–32

Akerlof R. 2016. "We thinking" and its consequences. *Am. Econ. Rev* 106(5):415–19

Akerlof R. 2017. Value formation: the role of esteem. *Games Econ. Behav.* 102:1–19

Aronson J, Quinn DM, Spencer S. 1998. Stereotype threat and the academic underperformance of minorities and women. In *Prejudice: The Target's Perspective*, ed. JK Swim, C Stangor, pp. 83–103. San Diego, CA: Academic

Bargh JA, Chartrand TL. 1999. The unbearable automaticity of being. *Am. Psychol.* 54:462–79

Bargh JA, Chen M, Burrows L. 1996. Automaticity of social behavior: direct effects of trait construct and stereotype activation on action. *J. Personal. Soc. Psychol.* 71(2):230–44

Bargh JA, Pietromonaco P. 1982. Automatic information processing and social perception: the influence of trait information presented outside of conscious awareness on impression formation. *J. Personal. Soc. Psychol.* 22:293–311

Basu K. 2010. The moral basis of prosperity and oppression: altruism, other-regarding behaviour and identity. *Econ. Philos.* 26:189–216

Bénabou R, Tirole J. 2011. Identity, morals, and taboos: beliefs as assets. *Q. J. Econ.* 126(2):805–55

Benbow CP. 1995. Sex differences in mathematical reasoning ability in intellectually talented preadolescents: their nature, effects, and possible causes. *Behav. Brain Sci.* 11:169–232

Benjamin DJ, Choi JJ, Strickland AJ. 2010. Social identity and preferences. *Am. Econ. Rev.* 100(4):1913–28

Bernard M, Hett F, Mechtel M. 2016. Social identity and social free-riding. *Eur. Econ. Rev.* 90:4–17

Bernhard H, Fehr E, Fischbacher U. 2006. Group affiliation and altruistic norm enforcement. *Am. Econ. Rev.* 96(2):217–21

Bornstein G, Gneezy U, Nagel R. 2002. The effect of intergroup competition on group coordination: an experimental study. *Games Econ. Behav.* 41(1):1–25

Brewer MB. 1999. The psychology of prejudice: ingroup love and outgroup hate? *J. Soc. Issues* 55(3):429–44

Buller P. 1986. The team building-task performance relation: some conceptual and methodological refinements. *Group Organ. Manag.* 11(3):147–68

Carvalho J-P. 2013. Veiling. *Q. J. Econ.* 128(1):337–70

Cason TN, Lau S-HP, Mui V-L. 2019. Prior interaction, identity, and cooperation in the Inter-Group Prisoner's Dilemma. *J. Econ. Behav. Organ.* 166:613–29

Chang D, Chen R, Krupka E. 2019. Rhetoric matters: a social norms explanation for the anomaly of framing. *Games Econ. Behav.* 116:158–78

Charness G, Cobo-Reyes R, Jiménez N. 2014. Identities, selection, and contributions in a public-goods game. *Games Econ. Behav.* 87:322–38

Charness G, Dufwenberg M. 2006. Promises and partnership. *Econometrica* 74(6):1579–601

Charness G, Holder P. 2019. Charity in the laboratory: matching, competition, and group membership. *Manag. Sci.* 65(3):1398–407

Charness G, Naef M, Sontuoso A. 2019. Opportunistic conformism. *J. Econ. Theory* 180:100–34

Charness G, Rabin M. 2002. Understanding social preferences with simple tests. *Q. J. Econ.* 117:817–69

Charness G, Rigotti L, Rustichini A. 2007. Individual behavior and group membership. *Am. Econ. Rev.* 97:1340–52

Charness G, Rustichini A. 2011. Gender differences in cooperation with group membership. *Games Econ. Behav.* 72:77–85

Charness G, Sutter M. 2012. Groups make better self-interested decisions. *J. Econ. Perspect.* 26(3):157–76

Chen R, Chen Y. 2011. The potential of social identity for equilibrium selection. *Am. Econ. Rev.* 101(6):2562–89

Chen R, Chen Y, Liu Y, Mei Q. 2017. Does team competition increase pro-social lending? Evidence from online microfinance. *Games Econ. Behav.* 101:311–33

Chen Y, Li SX. 2009. Group identity and social preference. *Am. Econ. Rev.* 99(1):431–57

Chen Y, Li SX, Liu TX, Shih M. 2014. Which hat to wear? Impact of natural identities on coordination and cooperation. *Games Econ. Behav.* 84:58–86

Cohn A, Fehr E, Maréchal M. 2017. Do professional norms in the banking industry favor risk-taking? *Rev. Financ. Stud.* 30(11):3801–23

Cohn A, Maréchal M, Noll T. 2015. Bad boys: how criminal identity salience affects rule violation. *Rev. Econ. Stud.* 82(4):1289–308

Coleman J. 1961. *The Adolescent Society: The Social Life of the Teenager and Its Impact on Education*. New York: Free Press

Croizet J-C, Claire T. 1998. Extending the concept of stereotype threat to social class: the intellectual underperformance of students from low socioeconomic backgrounds. *Personal. Soc. Psychol. Bull.* 24:588–94
Currarini S, Matheson J, Vega-Redondo F. 2016. A simple model of homophily in social networks. *Eur. Econ. Rev.* 90:18–39
Currarini S, Mengel F. 2016. Identity, homophily, and in-group bias. *Eur. Econ. Rev.* 90:40–55
Dugar S, Shahriar Q. 2012. Group identity and the moral hazard problem: experimental evidence. *J. Econ. Manag. Strategy* 21(4):1061–81
Eckel CC, Grossman PJ. 2005. Managing diversity by creating team identity. *J. Econ. Behav. Organ.* 58(3):371–92
Erev I, Bornstein G, Galili R. 1993. Constructive intergroup competition as a solution to the free rider problem in the workplace. *J. Exp. Soc. Psychol.* 29(6):463–78
Darity WA, Mason PL, Stewart JB. 2006. The economics of identity: the origin and persistence of racial identity norms. *J. Econ. Behav. Organ.* 60(3):283–305
Fang H, Loury GC. 2005. "Dysfunctional identities" can be rational. *Am. Econ. Rev.* 95(2):104–11
Feri F, Irlenbusch B, Sutter M. 2010. Efficiency gains from team-based coordination—large-scale experimental evidence. *Am. Econ. Rev.* 100(4):1892–912
Fryer R, Jackson MO. 2008. A categorical model of cognition and biased decision making. *B.E. J. Theor. Econ.* 8(1). **https://doi.org/10.2202/1935-1704.1357**
Goette L, Huffman D, Meier S. 2006. The impact of group membership on cooperation and norm enforcement: evidence using random assignment to real social groups. *Am. Econ. Rev.* 96(2):212–16
Goette L, Huffman D, Meier S. 2012. The impact of social ties on group interactions: evidence from minimal groups and randomly assigned real groups. *Am. Econ. J. Microecon.* 4(1):101–15
Hargreaves Heap SP, Zizzo DJ. 2009. The value of groups. *Am. Econ. Rev.* 99(1):295–323
Hedges LV, Nowell A. 1995. Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science* 269:41–45
Herrmann B, Thöni C, Gächter S. 2008. Antisocial punishment across societies. *Science* 319(5868):1362–67
Hetey RC, Eberhardt JL. 2014. Racial disparities in incarceration increase acceptance of punitive policies. *Psychol. Sci.* 25(10):1949–54
Hetey RC, Eberhardt JL. 2018. The numbers don't speak for themselves: racial disparities and the persistence of inequality in the criminal justice system. *Curr. Dir. Psychol. Sci.* 27(3):183–87
Hett F, Kröll M, Mechtel M. 2020. The structure and behavioral effects of revealed social identity preferences. *Econ. J.* In press
Hoff K, Pandey P. 2006. Discrimination, social identity, and durable inequalities. *Am. Econ. Rev.* 96:206–11
Hoff K, Pandey P. 2014. Making up people: the effect of identity on preferences and performance in a modernizing society. *J. Dev. Econ.* 106:118–31
Hogg MA. 2000. Subjective uncertainty reduction through self-categorization: a motivational theory of social identity processes. *Eur. Rev. Soc. Psychol.* 11(1):223–55
Jiang J, Li SX. 2019. Group identity and partnership. *J. Econ. Behav. Organ.* 160:202–13
Kranton R, Pease M, Sanders S, Huettel S. 2016. *Groupy and not groupy behavior: deconstructing bias in social preferences*. Work. Pap., Duke Univ., Durham, NC
Krupka EL, Weber RA. 2013. Identifying social norms using coordination games: Why does dictator game sharing vary? *J. Eur. Econ. Assoc.* 11(3):495–524
Kugler T, Bornstein G, Kocher MG, Sutter M. 2007. Trust between individuals and groups: Groups are less trusting than individuals but just as trustworthy. *J. Econ. Psychol.* 28(6):646–57
Lane T. 2016. Discrimination in the laboratory: a meta-analysis of economics experiments. *Eur. Econ. Rev.* 90:375–402
Li SX. 2020. Group identity, ingroup favoritism, and discrimination. In *Handbook of Labor, Human Resources and Population Economics*, ed. K Zimmermann. Cham, Switz.: Springer. **https://doi.org/10.1007/978-3-319-57365-6_123-1**
Li SX, de Oliveira ACM, Eckel C. 2017. Common identity and the voluntary provision of public goods: an experimental investigation. *J. Econ. Behav. Organ.* 142:32–46

McLeish KN, Oxoby RJ. 2007. *Identity, cooperation, and punishment*. Discuss. Pap. 2572, Inst. Labor Econ. (IZA), Bonn, Ger.
McLeish KN, Oxoby RJ. 2011. Social interactions and the salience of social identity. *J. Econ. Psychol.* 32(1):172–78
Milgram S. 1963. Behavioral study of obedience. *J. Abnorm. Soc. Psychol.* 67(4):371–78
Milgram S. 1974. *Obedience to Authority: An Experimental View*. New York: HarperCollins
Monderer D, Shapley L. 1996. Potential games. *Games Econ. Behav.* 14:124–43
Mullen B, Brown R, Smith C. 1992. Ingroup bias as a function of salience, relevance, and status: an integration. *Eur. J. Soc. Psychol.* 22(2):103–22
Peffley M, Hurwitz J. 2007. Persuasion and resistance: race and the death penalty in America. *Am. J. Political Sci.* 51:996–1012
Rong R, Houser D, Dai AY. 2016. Money or friends: social identity and deception in networks. *Eur. Econ. Rev.* 90:55–66
Sen A. 2007. *Identity and Violence: The Illusion of Destiny*. New Delhi: Penguin Books India
Shayo M. 2009. A model of social identity with an application to political economy: nation, class, and redistribution. *Am. Political Sci. Rev.* 103(2):147–74
Shayo M. 2020. Social identity and economic policy. *Annu. Rev. Econ.* 12:355–89
Shih M, Pittinsky TL, Ambady N. 1999. Stereotype susceptibility: identity salience and shifts in quantitative performance. *Psychol. Sci.* 10(1):80–83
Shih M, Pittinsky TL, Trahan A. 2006. Domain-specific effects of stereotypes on performance. *Self Identity* 5(1):1–14
Song F. 2008. Trust and reciprocity behavior and behavioral forecasts: individuals versus group-representatives. *Games Econ. Behav.* 62(2):675–96
Spencer S, Steele CM, Quinn DM. 1999. Stereotype threat and women's math performance. *J. Exp. Soc. Psychol.* 35:4–28
Steele CM, Aronson J. 1995. Stereotype threat and the intellectual test performance of African-Americans. *J. Personal. Soc. Psychol.* 69:797–811
Steen LA. 1987. Mathematics education: a predictor of scientific competitiveness. *Science* 237:251–53
Tajfel H, Billig MG, Bundy RP, Flament CL. 1971. Social categorization and inter-group behavior. *Eur. J. Soc. Psychol.* 1:149–77
Tajfel H, Turner J. 1979. An integrative theory of intergroup conflict. In *The Social Psychology of Intergroup Relations*, ed. S Worchel, W Austin, pp. 94–109. Monterey, CA: Brooks/Cole
Tajfel H, Turner J. 1986. The social identity theory of intergroup behavior. In *The Social Psychology of Intergroup Relations*, ed. S Worchel, W Austin. Chicago: Nelson-Hall
Tanaka T, Camerer CF, Nguyen Q. 2016. Trait perceptions influence economic out-group bias: lab and field evidence from Vietnam. *Exp. Econ.* 19(3):513–34
Turner JC. 1978. Social categorization and social discrimination in the minimal group paradigm. In *Differentiation Between Social Groups: Studies in the Social Psychology of Intergroup Relations*, ed. H Tajfel, pp. 101–40. New York: Academic
Van Huyck JB, Battalio RC, Beil RO. 1990. Tacit coordination games, strategic uncertainty, and coordination failure. *Am. Econ. Rev.* 80(1):234–48
Weber RA. 2006. Managing growth to achieve efficient coordination in large groups. *Am. Econ. Rev.* 96(1):114–26
Yamagishi T, Kiyonari T. 2000. The group as the container of generalized reciprocity. *Soc. Psychol. Q.* 63(2):116–32