Ethnic Identifiability: An Experimental Approach*

James Habyarimana
Georgetown University

Macartan Humphreys
Columbia University

Daniel N. Posner
University of California, Los Angeles

Jeremy M. Weinstein
Stanford University

March 2005

Abstract

We propose a radically contingent notion of ethnic identifiability that treats identifiability as a binary relation between two individuals, conditional upon a given informational environment. We use this notion of individual identifiability to develop a language of group identifiability and group distinctness. Drawing on a sample of ninety-six undergraduate students who self declare into seven different ethnic groups, we explore the determinants of identifiability for a given ethnic demography. We find that subjects are less able to place others into ethnic categories than theories of ethnic politics assume. Subjects miscode in-group members 16 percent of the time and out-group members 33 percent of the time. They miscode out-group members as in-group members 7 percent of the time, a large number in the context of our experiment and one that indicates that errors of inclusion are more common than errors of exclusion. Subjects appear to be especially weak at incorporating new information from signals sent by individuals that may be attempting to pass. The use of signals raises the rate of correct ethnic identification to 89 percent in situations where subjects try to convince players of their true ethnic identity but lowers it to 55 percent when subjects try to pass. The net effect is a decline in identifiability, contrary to expectations from rational updating. The use of signals, we find, has a greater marginal effect when the power of signs is weak. In particular, individuals are especially poor at fooling in-group members into believing that they are out-group members and at convincing non-co-ethnics that they are in-group members.

*The authors thank Chris Crabbe for his superb programming work; Dan Young, Donna Horowitz, and Kevin Thelen for their research assistance; the Russell Sage Foundation, the Harry Frank Guggenheim Foundation, the Harvard Academy for International and Area Studies, and the International Institute at UCLA for their financial support; and the staffs of the California Social Science Experimental Laboratory (CASSEL) at UCLA and the Center for International Studies and the Law School Library at USC. Extremely helpful comments were received from participants at the 9th meeting of the Laboratory in Comparative Ethnic Processes (LiCEP), University of Wisconsin, 7-8 May 2004. Protocols for the experiment are available on request from the authors.

1 INTRODUCTION

Micro-level theories of ethnic politics commonly employ the assumption that individuals can identify the ethnic backgrounds of the people with whom they interact. This assumption is difficult to reconcile with anecdotal evidence that suggests that there is large variation in the ability of different individuals to identify others within given social schemas. It is also difficult to reconcile with research on signaling and mimicry that emphasizes the ability of individuals to use information about their identities strategically. But, although much of the study of the politics of identity emphasizes issues of placing and passing, there is no systematic research that assesses the informational and strategic conditions under which individuals distinguish in-group members from out-group members or sort out-group members into their "correct" ethnic categories.

This is likely due in part to the theoretical complexity of the notion of ethnic identifiability. On the one hand the notion of "correct identification" in any instance implies the existence of reasonable criteria for evaluating the judiciousness of decisions. But there are no such universal criteria, and the appropriateness of *ad hoc* criteria used in any given situation will likely be dependent on the strategic context of the placing. On the other hand, even if we have workable notions of correct placement, determining whether a rate of correct identification under any information condition is 'high' or 'low' depends on base rates, which themselves are sensitive to contextual factors. Beyond these theoretical issues of interpretation there are evident difficulties in observing an individual's classification of others, if indeed such classifications are made.

This paper works to fill this gap in knowledge by proposing a theoretical structure for making meaningful statements about identifiability and using experimental methods to measure identifiability for a given social structure. We propose a radically contingent notion of identifiability as a binary relation between two individuals, conditional upon a given informational environment, and use this to develop a language of group identifiability and group distinctness. Then, drawing on a sample of ninety-six undergraduate students who self declare into seven different ethnic groups, we explore the determinants of identifiability for a given ethnic demography. Specifically, we measure the ability of individuals to classify individuals whose images they are shown into the ethnic categories with which those individuals self-identify. We examine how the characteristics of the person viewing the images, the characteristics of the person whose images are viewed, and the degree of information that the former has about the latter affect the probability of a correct ethnic identification. We then examine the relationship between identification based on these physical "signs"—by which we mean aspects of the individual that are not subject to manipulation, at least in the short run—with the ability of individuals to use "signals"—by which we mean information that is more readily under the control of the subject and can be used to either communicate or hide their ethnic identity

We find that, conditional on the social demography we examine, there is substantial variation in identifiability within and across groups. Subjects are less able to distinguish in-group members from out-group members and less able to sort non-co-ethnics into their correct ethnic categories than theories of ethnic politics assume. When shown still photographs of other subjects, they miscoded in-group members as out-group members 16 percent of the time and miscoded out-group members as in-group members 7 percent of the time. The propensity of identifying out-group members as in-group members relative to identifying in-group members as out-group member is high, relative to a benchmark derived from equal likelihood of different types of errors. Subjects were even less successful at identifying the ethnic backgrounds of people from other ethnic groups. On average, subjects shown images of people from other ethnic groups miscoded the other person's ethnic background 33 percent of the time.

The use of signals however raises the rate of correct ethnic identification to 89 percent in situations where others try to convince them of his or her true ethnic identity. The identification rate drops to 55 percent when subjects are shown videos of other subjects trying to pass. Contrary to predictions of theories of rational updating, the net effect is a decline in identifiability, as respondents reject truthful attempts by players to communicate their identity and accept attempts by others to pass. The use of signals, we find, has a greater marginal effect when the power of signs is weak. In particular, individuals are especially poor at fooling in-group members into believing that they are out-group members and at convincing non-co-ethnics that they are in-group members.

2 MOTIVATION: THE PROBLEM OF ETHNIC IDENTIFICATION

From theories of in-group sanctioning (Greif 1989; Fearon and Laitin 1996) to theories that emphasize the ability of ethnic groups to police their boundaries (Barth 1969; Laitin 1995; Fearon 1999) to theories of ethnic or racial discrimination (Akerlof 1976; Becker 1971) to experimental treatments of minimal groups (Tajfel, Billig, and Bundy 1971), models of face-to-face ethnic interaction generally assume that actors can distinguish accurately between in-group members and outsiders. Even in more recent models inspired by constructivism, such as that found in Caselli and Coleman (2002), in which actors can take costly actions to change their identities, once changed, an actor's identity is assumed to be self-evident to the other players in the game. Other constructivist work that has emphasized the willingness of actors to employ different identities at different times, has been quiet about the chances that these different roles will be recognized by others.

The assumption that individuals can seamlessly identify the ethnic backgrounds of the people they encounter – or, at the very least, unproblematically distinguish in-group members from out-group members – although common, presents both theoretical and empirical problems. We consider the empirical problems first and then turn to the theoretical problems which render difficult attempts to address the empirical question.

Much anecdotal evidence suggests that people's ethnic backgrounds are oftentimes difficult to pin down (we address the question of how meaningful such statements are below). In one evocative example, Horowitz relates the following story from Sri Lanka:

> Sinhalese rioters suspected a man in a car of being a Tamil. Having stopped the car, they inquired about his peculiar accent in Sinhala, which he explained by his lengthy stay in England and his marriage to an English woman. Uncertain, but able to prevent his escape, the rioters went off to kill other Tamils, returning later to question the prospective victim further. Eventually, he was allowed to proceed on his way, even though the mob knew it risked making a mistake, which in fact it had: the man was a Tamil (2001: 130).

In this example, the attackers had difficulty coding the ethnic background of their would-be victim. Their concern was to ensure that they would not err by identifying an in-group member as an out-group member. In other contexts the concern is to ensure that out-group members are not mistakenly identified as in-group members. Recent concern with social profiling highlights the attempts (and difficulties) of correctly identifying "enemies." There have for example been an estimated two thousand backlash incidents directed at Muslims and people of Arab descent in the aftermath of the September 11 terrorist attacks in the United States. More often than not, the victims of these hate crimes turned out not to be Muslims or Arabs at all, but Sikhs, Indians, Pakistanis, Coptic Christians, and, in one case, an Iranian Jew (Human Rights Watch 2002).[1] The great lengths to which governments have historically gone to make members of particular ethnic

[1] We cannot rule out the possibility that the perpetrators of these anti-Muslim acts simply did not know that the Sikhs, Indians, and Pakistanis, and the others they attacked were not Arab Muslims, in which case the miscodings would not be

groups more readily identifiable – though, for example, the requirement that Jews wear the Star of David or that Japanese-Americans wear markers indicating their Japanese descent, or that Rwandese citizens have their official ethnic identities marked on their identity cards – further underscores the difficulty that ethnic identification often presents and the centrality that identifiability can hold for political action.

These instances suggest that the notion of perfect identification is empirically problematic. But the empirical variation in identifiability is of considerable substantive import. It opens the door to a rich set of proposition regarding the determinants of collective action. If ethnic identifiability varies across groups, then theorists of ethnic coalition building can use identifiability as a determinant of coalition choice. Theorists of in-group policing can use it to distinguish among communities with greater and lesser abilities to sanction their members, and thus greater or lesser abilities to execute business transactions, organize collectively, or prevent inter-group conflicts from degenerating into spirals of violence (Fearon and Laitin 1996). Theorists of ethnic mobilization can use it to account for variation in the ease with which political entrepreneurs may be able to organize – or organize *against* – particular communities. Theorists of ethnic violence can use it to explain the form that conflict takes.

Consider the wars in the north of Mali (1990-1995) and the south of Senegal (1982-present). The two conflicts have much in common. Both involve bids for separation by movements dominated by members of minority groups: the Tuaregs and Maures in Mali and the Diola in the Casamance region of Senegal. However, differences in the identifiability of the parties to each conflict relate to differences in how group members are mobilized and how violence is carried out. In Mali, the fact that the Tuaregs and Maures – the "whites" – are considered to be readily identifiable has meant that ethnicity could be used to pressure members of these groups, including

examples of ethnic misidentification as we treat it in this paper but simply of not being aware that there were different categories into which the would-be victims might be coded.

intellectuals living in the capital, to join the rebel movements. It has also allowed "black" sedentary groups and the Malian army to take quick reprisals against arbitrary Tuareg and Maure civilians. The result was a rapid polarization of camps and the descent of the separatist struggle into communal violence. In Senegal, by contrast, such ready association of individuals with ethnic groups has been more difficult. There, the mobilization of partisans and the targeting of reprisals has been more difficult, and the intensity of violence has been much lower (Humphreys and ag Mohamed 2005). The contrasting degree of ethnic identifiability corresponds to a sharp difference in the form of group mobilization and the scope of violence in each case – a difference that would be hard to account for if we assumed erroneously that all ethnic groups were equally identifiable.

As the foregoing discussion suggests, there is likely to be considerable empirical variation in the ability of groups to place each other; and this variation is likely to be of substantive importance. Despite this however, efforts to gauge the variation in ethnic identifiability rapidly encounter a host of deep theoretical issues, raised notably by constructivist scholars, that challenge the notion of correct or incorrect association of individuals with groups and indeed of the very notion of groups. We turn to these next.

Much recent research has emphasized the constructed nature of social categories. Some of this has also attempted to identify what it is that is specific about ethnic categories as distinct from other types of social categories. This literature raises the question of identifying in what sense it is meaningful to say that distinct ethnic groups *exist*. Both nineteenth century attempts to distinguish races "scientifically" and early twentieth century attempts to provide socioeconomic foundations to the problem of the existence of races have failed; in particular, biological arguments for fundamental differences, and arguments based on common ancestry and shared history run into logical or empirical problems (see for example Appiah 1992). These problems in providing foundations for the objective existence of races, extend also to ethnic groups.

In related work on *nations* Gellner and others have emphasized how these group categories are not pre-existing but invented. Anderson (1991), interpreting Gellner's claim as one that implies that nations do not exist at all—an implication that does *not* follow from a recognition either of the constructedness of groups of the fluidity of identity—attempts to place the existence of groups on intersubjective foundations. What is important, he argues of a nation, is that "in the minds of each [member] lives the image of their communion." This more positive vision of group existence is useful since, even if a primordialist position is logically untenable, this does not mean that people do not act *as if* they are members of a primordialist world.

The problem however is that even if the image of a communion lives in the minds of members of a group, this does not establish the existence of groups outside of the minds of individuals. Furthermore, the conditions for such a communion may be severe: in particular, images of group boundaries may not be shared. Indeed, recent political science research emphasizes that the contours and boundaries of ethnic groups and the ethnic identities of individuals are themselves a subject of contestation in a number of ways, including disagreements among in-group members about what the limits of their group is. Even if there is consensus on common membership of a collection of individuals in a community, this common recognition may not itself be recognized by out-group members, who may have rival criteria for what constitutes the group. As a result, statements by scholars about whether a person does or does not belong to a given identity group may involve adopting one individual's perspective over another's.

Note that this nominalist critique of notions of the objective existence of groups is quite distinct from some propositions about identities that have been most prominent in constructivist literatures, notably that individuals may have multiple identities and that their identification with these identities is context specific (Chandra 2001). While the salience of group identities for an individual, or the commonality of belief in a common community by a group, are certainly likely to be context dependent, the attempt by constructivist scholars to replace the notion of unique and fixed identities, with a notion of multiple and manipulable identities, does *not* resolve the problem of the objectiveness of the existence of identities. Without a clearer definition of what belonging to a group means, constructivists that claim that individuals have multiple identities, are on as weak an epistemological footing as primordialists that claim that individuals only have one.

Finally, even if the thorny problems of the degree of constructedness, multiplicity, or ontological status of groups are resolved, it proves difficult to identify criteria that indicate that the resulting group is specifically "ethnic" (Chandra 2004).

These concerns present new problems for the empirical study of ethnic identifiability. In our discussion above, which ignored these theoretical concerns, we noted that the assumption of identifiability is empirically problematic. Now we see that it is not just an empirical problem to determine whether or not ethnic groups are identifiable, since identifiability may itself be a function of which demographic structure and which notion of correct classification is employed by the researcher. And in the absence of logically coherent foundations to the existence of ethnic groups as a category, this choice appears arbitrary. This lack of foundations for the objective existence of ethnic groups does not however imply that participants do not act as if social groups exist. Rather than providing a break on empirical work, these theoretical concerns challenge empirical researchers to work with more precise notions of politically salient aspects of group membership and to study, head on, the richness that is introduced to the study of ethnic politics by the recognition of the importance of subjective beliefs rather than objective categories and of contextual contingency rather than primordial permanence.

3 AN EMPIRICALLY BASED NOTION OF IDENTIFIABILITY

To address these concerns, we propose a contingent, subjectivist notion of social classifications that treats identifiability as a binary relationship between two individuals.

We begin by defining a social demography. Given a triple $<N, I, f>$ in which N denotes a population of players, I denotes a set of categories, and f denotes a mapping from N to I (a classification rule), we say that $<N, I, f>$ is a "***social demography***" if f is a *function*: that is, if it places each $i \in N$ in one and only one element of I.

The requirement that each individual falls within only one category is not onerous. It allows for the possibility of membership in categories along multiple dimensions of identity, and it allows for membership in two or more categories along a single dimension of identity. Should, for example, an individual self-classify as both "Asian" and "Caucasian" in some set of categories, then she can self-classify in a compound category ("Asian and Caucasian") under another set of categories.[2] The restrictions on the placement rule are here clearly weak and do not imply any "objective" membership of a category or possession of any attributes. Examples of f that may form part of a social demography may for example include "i is a member of the category in I in which *the plurality of* observers in N place i, given personal criteria rules, c_j, defined over I" or "i is a member of the category in I if i's father was a member of that category."

Other than being capable of producing a partition empirically, the notion employed here does not imply any substantive consistency between the categories in I; for example that all elements in I be regional, religious or ethnic groups; indeed to incorporate the possibility that individuals have portfolios of identities, any element in I may be a combination of multiple categories on seemingly

[2] By the same token, should an individual not be classifiable under any category in some set, then she can be classified as unclassifiable in an expanded set of categories.

disparate dimensions of identity. Unlike the framework employed in Chandra (2003) the mapping need not be based on objective attributes, furthermore, it does not necessarily correspond to eligibility but rather to some empirical placement procedure. Finally, although our interest is in "ethnic demographies" as is clear from the above, the notion of identifiability we employ is not specific to ethnicity.

3.1 Identification and Identifiability

Note that in our description of social demographies, although individuals may have personal rules, c_j, for placing player *i*, given *N* and *I*, this does not imply that these rules correspond to *f*. Indeed it is this feature that allows us to generate definitions of *correct* identification which can in turn form the basis for pairwise identifiability, group "identifiability," and group distinctness, under a given social demography.

We define correct identification as follows. Given a social demography $<N, I, f>$, and a rule *c*, that maps from *N* to *I*, we say that the rule *c correctly* identifies *i* if *c* and *f* place *i* in the same category of *I*, that is, if $c(i \mid N, I) = f(i \mid N, I)$. In a world of uncertainty we treat individual placement rules as random variables. For such cases we can treat the notion of identifiability as a *probability* as follows. Given a social demography $<N, I, f>$, an information set *X*, and a placement rule *c*, that maps, probabilistically, from *N* to *I*, conditional on *X*, the ***identifiability*** of *i* under *c,* is given by the probability that *c* correctly identifies *i*, that is, by, $\Pr(c(i \mid X, N, I) = f(i \mid N, I))$. Hence, for example, if *f* is the self placement criterion, *c* is "player *j*'s best guess about *i*'s ethnicity within a given set of ethnic categories" and *X* is a name, then we can say that the identifiability of *i for j,* is the probability that *j*'s best guess about *i*'s ethnicity after learning *i*'s name, corresponds to *i*'s self-placement.

Given this definition we can then define the pairwise identifiability of *i* by *j* under demography $<N, I, f>$ and private placement rule c_j as: $\Pr(c_j(i \mid X, N, I) = f(i \mid N, I))$. We can also use these subjective microfoundations, deriving from individual decision making, to make somewhat more abstract statements about group identifiability. We define the identifiability of a group, G, for an individual j as $\mathsf{E}_{i\in G}[\Pr(c_j(i \mid X, N, I) = f(i \mid N, I))]$, and the identifiability of a given group A, for a given group, B, (with, possibly, $A=B$), as $\mathsf{E}_{j\in B}\,\mathsf{E}_{i\in A}\,[\Pr(c_j(i \mid X, N, I) = f(i \mid N, I))]$, where the membership of A and B is determined by f. As a special case, we refer below to $\mathsf{E}_{j\in N}\,\mathsf{E}_{i\in A}\,[\Pr(c_j(i \mid X, N, I) = f(i \mid N, I))]$ simply as the group's "identifiability."

3.2 Errors in Identification

With these notions of identifiability we can also generate notions of Type I and Type II errors in placement within a demography as well as notions of pairwise and global *distinctiveness*. Consider the question of whether an individual is an in-group member. We refer to the hypothesis that player i is an in-group member for a player j as j's ***in-group null*** regarding i. Given an in-group null, we say that a **Type I error** occurs if j incorrectly identifies in-group member i as an out-group member. We say that a **Type II error** occurs if player i is an out-group member who is incorrectly identified as an in-group member. The probability of *not* making a Type I error, is simply the identifiability of (in-group member) i, for j. Continuing the analogy with hypothesis testing, the probability of not making a Type 2 error is referred to as the *power* of (out-group member) i's identifiability for j.

3.3 Group Distinctness

The ***distinctness*** of groups A and B for player j is based on the expected probability that given an arbitrary individual from groups A or B, social demography $\langle N, I, f\rangle$, and information set X, that j will not misidentify i as a member of group B, or misidentify B as a member of group A.[3] More precisely we define:

[3] Although Casselli and Coleman (2002) base a theoretical model of conflict on a similar notion of distinctness, we know of no attempts to measure distinctness in this way empirically.

$$D_j(A,B \mid N, I, f X) := 1-2\mathrm{E}_{i\in(A\cup B)}\,[\,\Pr(c_j(i|X)=B|i\in A)\Pr(i\in A) + \Pr(c_j(i|X)=A|i\in B)\Pr(i\in B)]$$

Note that if members of group A are never confused with each other, then $D_j(A,B)=1$; if however, two groups are indistinguishable, for example if, $\Pr(i\in A) = \Pr(i\in B) = \Pr(c_j(i|X)=B|i\in A) = \Pr(c_j(i|X)=B|i\in A) = .5$, then $D_j(A,B)=0$.[4]

For j, a member of group A, this measure is based on the probability of not making *either* a Type I or a Type II error given an in-group null. Note that two groups may not be highly distinct for a given individual yet that person may find *individual members* of those groups highly identifiable. We illustrate the notion of distinctness and the distinction between group distinctness and individual identifiability with an example below.

Example. Consider the case where player i is drawn from group A with probability p, and from group B with probability $1-p$. Let player j be drawn from group A. Assume that A types all possess some characteristic θ, that is held by B types only with probability (or frequency) q. We let X denote the treatment wherein j observes whether or not i possesses θ. Let j's private placement rule be given by $c_j(i|X) = A$ if and only if j's posterior assessment that i is a member of group A exceeds .5. Using Bayes' rule, j's posterior is given by: $\Pr(i\in A|\theta) = p/(p + q(1-p))$ and $\Pr(i\in A|0) = 0$.

What then, is the **distinctness** of A and B for i under informational treatment X? There are two cases of interest. In the first case $\Pr(i\in A|\theta)\leq .5$. In this case, no matter what

[4] This index can, in principle take on negative values in situations where players use criteria that do worse than a random guess.

signal is received, i believes j to be member of group B. This occurs when j's prior is low, (if, $p<q/(1+q)$). In this case, the distinctness of A and B is given by: $1-2(p\times 1 + (1-p)\times 0) = 1-2p$.

In the second case, $\Pr(i\in A|\theta)>.5$, and so, upon observing θ, j infers that i is in group A. In this case with probability $(1-p)(1-q)$, j observes X=0 and correctly classifies player j as a member of B; no error is made. With probability $p+(1-p)q$ however, j observes θ and classifies i as an in-group member. In this case she makes no Type I errors but with probability $(1-p)q\,/(p+(1-p)q)$ she makes a Type II error. Since this occurs with probability $p+(1-p)q$, her net probability of making a Type II error is $(1-p)q$. Her total probability of making some error is $(1-p)q$ and so our measure of group distinctness yields $1-2(1-p)q$.

Note that in the example we considered, individual identifiability can be high even if group distinctiveness is low; the former is a function of the observed X, the latter of the distribution of X. Hence for example, letting, $p = 5/9$ and $q=9/10$, our measure of distinctiveness in this example is just $D_j(A,B)= 2/10$. Identifiability however is greater than this for some particular pairings; for example the identifiability of a B type with value $X=\theta$ is 0, but the identifiability of a B type with value X=0 is 1. The *average* identifiability of a B with a given treatment "X" depends on the frequencies of the different realizations of X among B types, and is given in this case by $0\times 9/10 + 1\times 1/10 = 1/10$.

3.4 Signs and Signals

The identifiability of individuals depends then on the information they have at their disposal, and the way they choose to act on this information. In many politically relevant contexts, the availability of this information is itself a choice and may be subject to strategic information. We distinguish between signs and signals of group membership (Gambetta 2005). Although the distinction is imperfect, for our purposes, we use **"sign"** to denote a manifestation of group membership that is beyond an individual's control, at least in the short run; and we use "**signal**" to denote an action taken by an individual in order to communicate membership of an identity group. Signaling theory suggests that the use of signals can lead to successful passing in situations in which the signal is available to multiple groups but at different rates. If a signal is available to all individuals then it may not carry any information about group membership; if it is only available to one group, then it cannot be used by another to pass as members of that group (Bacharach and Gambetta 2001). However, while signaling can lead to successful passing, its impact on identification can depend on the relative costs to individuals of making different types of errors. An example in the appendix describes the logic in more detail and illustrates how the introduction of the possibility of signaling can result—even in a context of optimal signal extraction—in a rise *or* fall in the identifiability of any given group and, correspondingly, to a rise *or* fall in the occurrence of either Type I or Type II errors. While *either* type of error can increase based on the distribution of signals and the player's priors, the introduction of new information should not result in a *net* rise in both types of error.

3.5 From Concepts to Data

The formal definition of these notions of identifiability, errors in placement, distinctness, and signs and signals, allow us to make progress in the measurement of these aspects of group structures for a given social demography. As emphasized above, and as indicated by the formal expressions we use, a group's identifiability is radically contingent—it is a function of the information available, of the social demography invoked, and of characteristics of identifiers and identified. This contingency means that the specifics of our empirical frame are important for interpreting our results. It also means however that we can seek to make more general statements about identification processes *as* *a function of the empirical frame employed*. In particular we can examine how the identifiability or distinctness of groups change as a function of the set of categories used, the criteria for membership of groups, the population under consideration, and most germane for our work in this paper, of the information made available to—and made available by—individuals.

4 Experimental Design

We now describe an experimental design that fixes a particular social demography—that is a population, *N*, a set of categories, *I*, and a function *f*—and attempts to measure the identifiability of individuals within this demography, given a collection of personal guessing rules $\{c_j\}$. The design relates identifiability to the propensity of in-group members to make Type I or Type II errors, and to engage in, or to be susceptible to, the use of signals. Note that whereas others (for example Chandra 2004) argue that identifiability may be a *defining* characteristic of an ethnic category, we treat identifiability as an empirical property of an arbitrary social demography; one that may or may not obtain and that may or may not be endogenous to ethnic politics. Indeed, we use experimental methods specifically to allow us to measure how identifiability varies with informational and strategic context.

The experimental approach we employ relates to research that has been undertaken by social psychologists that aim to study how individuals classify others into ethnic categories.

A number of early studies have emphasized the role played by prejudice, as measured through questions about the subject's awareness of and opinion about members of other ethnic groups. Allport and Kramer (1946) initiated this line of research with a study that asked a sample of university students to distinguish pictures of Jewish students from pictures of non-Jewish students. Participants were given fifteen seconds to view each photograph before being asked to identify the person as "Jewish," "Non-Jewish," or "Don't Know." Pettigrew, Allport, and Barnett (1958) added to this rudimentary design by introducing a stereoscope – a device that presents different images to the left and right eye so that participants see a single merged object. This enabled Pettigrew and his colleagues to examine how individuals classify images that combine persons from two different groups. Such work argues for example that more prejudiced individuals are more likely to classify individuals into groups against which they held prejudices; or that members of some physiologically "extreme" group are more likely to classify mixed race individuals as belonging to one or other extreme group (Allport and Kramer 1946; Secord 1959; Lent 1970). More recent research in this area has employed computer and video technologies to explore the same questions. Blasovich, Wyer, Swart, and Kibler (1997) showed photos of white, black, and 'ambiguous' individuals to participants and measured the time it took them to identify the race of the person in the photo. These authors argue that extreme groups expend more effort on "policing" group boundaries.

While much of this research does not problematize the notion of objective membership of groups ("extreme" or otherwise), the emphasis on prejudice can be interpreted in terms of our framework above as an attempt to understand the consistency between observer classification rules and the experimenter's classification rules (which are typically based on some "official" rules or on self placement criteria).

Most recently, Harris (2002), focused on the impact of observer race, gender, and experience with other races on identification choices. Using a web-based survey of university students in which participants were asked to categorize a set of photographs as white, African-American, Latino, Asian-American, American Indian, Pacific Islander, or other, he analyzed both the classifications the students made and their response times. He found that whites and Asian-Americans more quickly classified photographs and used many fewer racial groups to categorize the full set of photographs, while other minority groups were more likely to see complexity in the photographs. Harris also identified a strong relationship between a participant's experience with other races and how he or she classified the images.[5]

The experiment reported in this paper, while closest to Harris (2002), differs in our interest in understanding the effects of context on the observer decisions. We do this by experimentally manipulating the information available to participants in two ways.

First, we introduce variation in the richness of the "signs" made available to participants, by varying the information available to an observer from that which they can glean from a still photograph of the subject, to that available in video and audio imagery to the information contained in a name.[6] By exposing participants, in turn, to each of these levels of information, and asking them to guess the subject's ethnic background after each one, we are able to test the impact of informational contexts on ethnic identifiability and examine how identifiability depends on observer characteristics and group membership.

Second, we introduce variation in the use of "signals," allowing subjects to attempt to "pass." In a world of unproblematic ethnic identification, passing would not be an issue: it would be impossible. But in a world where a person's ethnic background cannot be determined with certainty without a tremendous investment in information about the person's family history, individuals will sometimes have incentives to take advantage others' uncertainty about their background to try to pass as members of groups that enjoy greater prestige, access, protection, or other benefits than their own. We test the ability of subjects to pass (and the ability of others to catch them in their attempts) by recording video clips of subjects trying to convince people that they are members of

[5] The effects of ethnic group membership have also begun to be explored by experimental economists (e.g., Fershtman and Gneezy 2001; Gil-White 2004; Ferraro and Cummings 2004). Yet in these studies the emphasis is on the effects of ethnicity on behavior in experimental games rather than on the identifiability of subjects. In fact, in all of these experiments, the identifiability of subjects as members of particular ethnic groups is simply assumed.
[6] See Isaacs (1975) for a study of the information contained in a name. In Fershtman and Gneezy (2001), names are employed as the sole marker of ethnic group affiliation.

ethnic communities that may or may not be their own. By exposing participants, to signals that may or not be concordant with signs received, we are able to test the impact of strategy on ethnic identification.

The social demography, <*N*, *I*, *f*>, that we examine has as its population, *N*, undergraduate students from the University of California, Los Angeles (UCLA) and the University of Southern California (USC) that self identify into seven ethnic groups. These ethnic groups all have large presences on both campuses and correspond to our set of categories, *I*; they are: African Americans, Arabs, Asians, Caucasians, Indians, Persian, and Latino/as. The placement rule, *f*, which we use to determine correct placements, is given by the subjects' self placement into these 7 categories. In practice of 120 students recruited into our sample, all but 3 self selected into exactly one of the 7 categories. As an alternative, we also examined a rule that placed an individual into the category that corresponded to the modal guess of other participants in the experiment.[7] The findings using the resulting demography, although not reported here, are consistent with what we find using the self-placement criterion.

Approximately 54 percent of the participants were recruited through ethnic student associations on each campus. The other 46 percent were recruited from the regular subject population of the California Social Science Experimental Lab (CASSEL) at UCLA. Although the sample can not be considered a random sample of the student population in these seven groups, participants were not identified based on evaluations of their appearance[8] and we hold it as unlikely

---

[7] In this case we measured identification success as the probability that the guess of the respondent accords with the *modal* guess of all *other* respondents. As identification rates are high on average, the two mapping rules are highly correlated, but imperfectly so. 81% of subjects are coded as being in the same group under either rule. 19% of subjects are most commonly placed by respondents in groups other than those in which they place themselves.

[8] We approached students who had signed up for other experiments at CASSEL and asked them to fill out a short screening survey that contained a question about their ethnic background. We then contacted students from our target groups and invited them to participate in the experiment.

that our recruiting methods might have biased our sample in favor of individuals who were particularly "identifiable" as members of their respective groups.[9] Of the 120 students recruited, 96 ultimately participated in the experiment.

In order to control the informational environment *X*, we collected three different images of each subject with a digital camera. Each image was designed to provide more information about the participant's ethnic background than the previous one. First we recorded a headshot. Then we recorded a brief video clip in which the participant greeted the camera and said "Hello, I am looking forward to playing the game with you." Then we recorded another brief video clip in which the participant again greeted the camera, but this time also gave his or her full name (e.g., "Hello, I am looking forward to playing the game with you. My name is John Doe."). All three images were filmed in front of an identical blue background. Participants also filled out a brief questionnaire in which we collected information about their age, gender, place of birth, parents' educational background, exposure to various media, and SAT scores.

We also randomly drew a sub-sample of our participants and invited them to record three additional videos in which they explicitly stated their ethnic backgrounds.[10] For the first two videos, we asked them to pretend that they were in a situation in which it was important that they convince the person who would view the video of their true ethnic background. By "true ethnic background" we explained that we meant the ethnic background that they used to identify themselves. We filmed two versions of this "simulation" video. In one, we asked the participant to imagine that the person who would see the video was a co-ethnic. In the other, we asked the participant to imagine that the person who would view the video was a non-co-ethnic. Finally, we asked the participants to pretend that they were in a situation in which it was important that they convince the person who would view the video that they belonged to an ethnic group different from their own. We asked them to

[9] Although we cannot rule out the possibility that the appearance of members of campus ethnic associations might be more "typical" of their groups than is that of non-members.

choose an ethnic group (other than their own) from a list of the seven groups included in the experiment, and we filmed a "dissimulation" video in which they attempted to pass as a member of that group.

The identification game was played as follows. When participants arrived at the lab they were given a card assigning them to a computer and instructed to put on a pair of headphones that we provided. Respondents were then shown a series of still images and videos of twenty-three subjects and asked to guess each subject's ethnic background. Respondents were told that they would be paid 20 cents for each correct guess. They were also told that in some images the individual playing may reveal his or her ethnic background explicitly but that they may in fact be attempting to mislead.

To ensure that respondents could only use information about the subjects that was provided during the experiment, respondents were first shown a headshot of the first subject and asked if they knew the person. If a respondent indicated that he or she knew the subject, then the subject's photograph was replaced with that of another subject.

Once the respondent indicated that he or she did not know the subject, the respondent was asked to indicate his or her best guess of the subject's ethnic background from a closed list of the seven groups participating in the experiment. Respondents were also asked to indicate their certainty about their guess on a four-point scale ranging from "a random guess" to "most certain." We also measured the respondent's certainty in a second way, by recording the response time between the moment the headshot appeared (or the video ended) and the time they entered their guess about the subject's ethnic background.

For each subject, the respondents were then shown the "greeting" and the "greeting with name" videos in turn and asked the same two questions. They were encouraged to change their answers if they saw something in the video that caused them to reassess their earlier guess.

[10] We invited 36 participants to record the additional images, of whom 32 had their images recorded.

If the respondent happened to be paired with one of the thirty-two subjects for whom we had recorded a simulation/dissimulation video, then the respondent was shown one additional video. Approximately half of the time they were shown the dissimulation video and half of the time they were shown the simulation video. If the respondent was shown the simulation video and if the subject in the video was a co-ethnic, then the respondent was shown the "co-ethnic simulation" video; if the subject was a non-co-ethnic, then the respondent was shown the "non-co-ethnic simulation" video. After seeing the video, the respondent was again asked to guess the subject's ethnic background and to indicate his or her certainty about the guess.

At the end of the session, respondents completed a questionnaire that collected information about the parts of the world they had visited and/or lived in; whether they had a roommate and, if so, of what ethnic background; and, for each of our seven ethnic groups, whether they would feel comfortable having a member of that group as a friend or a close kin by marriage. At the conclusion of the experiment, respondents were paid their winnings. In addition to their $5 show up fee, the maximum a respondent could have earned was approximately $15.60. On average, respondents earned $11.13.

5 RESULTS

5.1 EMPIRICAL VARIATION IN IDENTIFIABILITY

Table 1 reports the measured the average identifiability of members of each of the seven groups, by each of the seven groups, given the self placement criterion and information contained in headshots of subjects – roughly analogous to a situation in which the respondent views a candidate's "file." Included in brackets under each identification success rate is the total number of respondent-subject pairings of that type.

Table 1 Here

Reading down the columns in Table 1, it is clear that some ethnic groups are much more identifiable than others. Identification success rates surpassed 95 percent for one group but fell as low as 26 percent for another. This variation is confirmed in Table 2, which presents a series of probit regressions that analyze the group and co-ethnicity effects that might affect identification success rates. For ease of interpretation, we report marginal coefficient estimates (with values for the other explanatory variables set to their means). Regression disturbance terms are clustered at the subject level. Model I reports a basic model of identification success with low information (headshots only) that includes only subject group fixed effects, with the most identifiable group serving as the omitted category. The results confirm that the differences in ethnic identifiability across subject types is substantial and significantly significant, even once we account for the possibility that observations are not independent.

Table 2 Here

Beyond this variation in group level identifiability, there is considerable variation *within* groups in the identifiability of individuals; and, furthermore, there is variation *across* groups in the degree of variation in identifiability within groups. To illustrate this point, Table 3 shows the distribution of identification success rates across individual subjects *within* each ethnic group when respondents are shown headshots of subjects. In every case this observed within-group variation is significantly different from 0. There is no clear relationship however between a group's average identifiability and the variance in its identifiability: there are some highly identifiable groups that nonetheless contain some very difficult to identify members, and some poorly identifiable groups for which there is not a high degree of within-group variation.

Table 3 Here

The finding that some groups are more identifiable than others is not particularly surprising. More surprising is the finding that members of some ethnic groups are better *identifiers* than others (compare the row marginals in Table 1). The differences, varying from 65%, up to 81%, are substantively large and display statistically significant variation but account for very little of the overall variance in identifiability (see Table 2, Model II) and are considerably smaller in magnitude than the subject fixed effects.[11]

5.2 EMPIRICAL VARIATION IN GROUP DISTINCTNESS

Based on variation in the identifiability of groups, we can construct measures of the *distinctness* of groups, conditional on the social demography employed here, that takes account not simply of correct identification, but of the kinds of errors made in classification. As reported in Table 4, we find again, that there is considerable variation in distinctness across pairs of groups: some pairs are never confused for each other, while others are routinely confused.

Table 4 Here

[11] Note that in Model II there are both positive and negative significant coefficients. In such fixed effects models the significance of any individual coefficient must be evaluated relative to the omitted category; group C in this case. Although not reported here, we also find individual level determinants of the ability to identify. In particular having a friend or roommate from the subjects group increases the rate of correct identification even after controlling for co-ethnicity effects.

The binary distinctness measures reported in Table 4 clearly indicate the presence of clusters, with two groups that are highly distinct from all others (A and C), one group relative to the remaining set (E), and four groups, (B, D, F, and G) that are most commonly confused with each other.

This measure of "distance" correlates with one based on affinities (as operationalized by willingness to accept a person as a friend or into one's family by marriage). A simple correlation of our 42 measures of distinctness with the frequency of respondents willing to accept someone from the corresponding group as a friend, is -.4; the correlation with the propensity to accept the other as a member of ones family is -.46. A one standard increase in our distinctness measure is associated with a 10 percentage point fall in the propensity to accept another player as kin. We make no claims here regarding the causal relationship between these measures.

5.3 THE STRUCTURE OF ERRORS: CO-ETHNICITY AND THE PREVALENCE OF TYPE I AND TYPE II ERRORS

The design of the experiment also enabled us to assess the impact of "co-ethnicity" – a respondent being paired with a subject from the same ethnic group – on identification success rates.[12] Table 1 provides a rough sense of how identification success rates vary depending on the make-up of the pairing. For each type of subject, we highlight in bold the respondent type with the highest identification success rate. To the extent that individuals are better able to identify members of their own ethnic community than outsiders, we would expect the bold cells to be arrayed along the diagonal. With one exception we find that respondent groups are among the top two groups at identifying people from their own groups. Not surprisingly the importance of co-ethnicity varies across ethnic groups. For groups whose members are difficult to identify on average, identification success rates in co-ethnic pairings are substantially higher than in non-co-ethnic pairings. The identification success in some hard-to-identify groups is more than twice as high in co-ethnic pairings as it is for the sample as a whole, whereas the success rate in highly identifiable groups is only marginally higher for co-ethnics.

The results reported in Table 2 confirm that co-ethnicity is a powerful source of identification success when respondents are shown headshots of other subjects, even after accounting for fixed effects and the lack of independence across errors. Nonetheless, co-ethnic placements are imperfect. Players miscode in-group members 16 percent of the time. Furthermore, 7 percent of the time they miscode out-group members as in-group members. As shown in Table 5, the propensity to make these different types of errors varies across groups, with some groups more likely to err on the side of inclusion of non-co-ethics into their group and others erring on the side of excluding co-ethnics. The differences in the frequency of these errors for the population does not suggest that co-ethnics make relatively few Type I errors. In fact, the average propensity of a subject to miscode fellow group members as non-group members is 16%. The propensity to code an out-group member as an in-group member is 7%.

Table 5 Here

As with our other results in this paper, the comparability of these numbers depends on the demographic frame employed. We can get some handle on the meaning of the relative size of these numbers, however, by estimating what the difference would be in this sample were any type of misidentification equally likely. Assume, that the identifiability of individuals from each of 7 equally sized groups for a given player is p and that if the player misidentifies any other player she places that player in any other position with equal probability. In that case, the likelihood of making a Type

[12] Roughly 25 percent of the 2,203 total viewings in the experiment were co-ethnic pairings.

I error, conditional upon observing a co-ethnic, is (1-*p*), the likelihood of making a type II error, conditional upon observing a non-co-ethnic is (1-*p*)/6. Under an assumption of equal probability of errors, we expect type I errors to occur six times more than Type II errors. In fact they only occur slightly more than twice as often, suggesting a relatively disproportionate propensity to make Type II errors in our sample and a bias towards admitting non-co-ethnics to one's group.

5.4 Signs and Signals

A key innovation in our experimental design is the introduction of different levels of information. We find, across the viewings, the effect of information is significant. There is close to a 1 percentage point rise in identifiability in this sample once players speak on video. There is a further 4 percentage average rise once players speak their names (and a further 16 percentage point rise once players attempt to signal their identities, described below). The impact of simple video and aural information is statistically no different from 0. The addition of a name however results in a large increase, significant at the 99% level. The marginal effect of information on identifiability is illustrated in Figure 1.

Figure 1 Here

As with co-ethnicity, information has a more powerful impact on identification success when the subject is from a group that is not easily identifiable. One result of this is that there is greater convergence in identifiability across groups the more information there is. This effect can be perceived in Model IV of Table 2, where we replicate the complete model (from column 4) on the sample of high information viewings. While subject group effects are still strong, they are substantially reduced in size. The significance of being in a co-ethnic pairing is also substantially weakened, indicating that the advantage co-ethnics have in interpreting signs is in evidence only at low levels of information.

Next we examine the impact of information that is readily subject to strategic manipulation by our subjects. In real world situations where the stakes of misidentification are high, individuals will often collect additional information about other peoples' backgrounds before they make their decisions about how to assign these people to ethnic categories. Although still one-sided (only the subject speaks, and no opportunity is provided for the respondent to interrogate the subject), the simulation and dissimulation videos described earlier provide a better approximation of such real world interactions and permit us to investigate the ability of individuals to pass as members of groups different from their own.

A striking finding is that, on average, the use of signals leads to a *drop* in identification rates. Players in our game received rewards if and only if they guessed identities correctly. Since players should only incorporate the extra information on signals if they expected that it would lead to a rise in the chances of correct placements, we expect that the use of signals should never lead to a net fall in identification rates. All signals combined, 33%, of those that had gotten it wrong, subsequently got it right, whereas only 16% of those that had gotten it right subsequently got it wrong. This 16% represents a high number however. Because so many had gotten it right in the first place, a 16% switch by those that had been right and subsequently incorporated false information, outweighs the 33% of those that had been wrong and subsequently incorporated true information. The subsequent drop in net identifiability rates can be observed within co-ethnic pairings and within non-co-ethnic pairings, and is significantly different from 0 at the 95% level. For individuals it is composed of a rise in both Type errors and Type II errors. With such a net negative effect we would expect participants to be wary of any information given by respondents. In fact respondents do use the information to influence their decisions, but they do not do it very effectively.

Figure 1 and Table 6a reveal that rates of identification success reach their highest levels for viewings in which the subject sees a simulation video. The identification success rate across all viewings reaches 88.5 percent in the simulation sub-sample. Providing subjects with the opportunity to convince respondents of their true ethnic background has a particularly powerful impact on raising identification success rates for those groups that were most difficult to identify at lower levels of information.

But providing subjects with the opportunity to *mislead* respondents about their ethnic backgrounds substantially reduces identification success rates. As Table 6b indicates, the overall success rate drops to 55.2 percent with dissimulation, from a success rate of 77.3 percent for the dissimulation sub-sample when respondents viewed only a video greeting in which the subject provided his or her name. As with identifiability, the ability to pass is something that varies with the ethnic group of the subject.

Tables 6a and 6b Here

One possible explanation is that players had incorrect priors about the propensity of any given subject to try to mislead them. Respondents were simply told that the subjects might be giving misleading remarks. In fact truthful and misleading signals were emitted with 50% probability, each. If players had underestimated this prior they may err by accepting too many messages as true. The data however suggest that players did not fail to absorb the information that respondents could be trying to mislead them. Players showed willingness to reject new information: of those that had correctly guessed a player's identity based on video information, 70% subsequently stuck to their guns when the player tried to dissimulate; of those that had *incorrectly* guessed a player's identity based on video information and name, a full 40% stuck to their (incorrect) guesses after the subject tried to convince them of his true identity. Remarkably, exactly half (49%) of final guesses were consistent with the statements made by subjects, and hence implied a belief that subjects were telling the truth, whereas half were inconsistent with the statements by subjects and so indicate a belief that the subject was emitting misleading information. The problem it seems is not incorrect priors.

Relatedly, the problem is not that players are too gullible, but that they are gullible n the wrong ways: the respondents were *too slow to believe the subjects*: 30% of the previously correct guessers refused to believe the subjects when the subjects tried to mislead them, but 40% of the previously incorrect guessers refused to believe the subjects when the subjects attempted to pass on truthful information. Had the subjects been able to unproblematically communicate the truth, then the overall identification rate would have been 77% in the world in which players used signals compared to 75% in that in which they could not. The net failure to learn from signals stems from the weakness of signals to convey truthful information that is not already imparted by signs.

The second question of concern is: why are some players good at passing and not others? Turning to this question we find that we can account in part for the ability of individuals to pass based on the power of their signs. If signs are strong, the utility of signals for passing is limited. To confirm this we examined the relationship between the ability of a member of a group to pass using signals, given that she had failed to pass using sign information and the distinctness of the player's group, based simply on photographic information and measured as, $\min_{B \in G \backslash A} \{D_{j \in A}(A,B \mid N, I, f\, X)\}$, where G is the set of all groups. The relationship between distinctness and success is very strong, with an increase from 0 to 1 in our distinctness measure associated with a 40% percentage point drop in the ability of an individual to pass.

Other more individual level determinants of successful passing are examined in Table 7. There we employ three different definitions of successful passing. In Model I, we ask whether the respondent correctly identified the true ethnicity of the subject in spite of the fact that he or she viewed a dissimulation video. Model II asks whether the respondent, after seeing the dissimulation video, changed his or her answer from the correct to an incorrect ethnic category. Finally, Model III asks whether the respondent's guess matched the group into which the subject was trying to pass. Thus, while the first definition effectively asks "did the respondent see through the dissimulating subject's attempt to pass?" this third definition asks "did the subject fool the respondent into believing that he or she really belonged to the group in which he or she professed to be a member?"

Table 7 Here

Model I shows that some individual effects are important for successful passing. The older a subject is and the higher his or her SAT scores, the less likely the respondent will be able to correctly identify him or her when he or she is trying to pass. This plausibly stems from the fact that successful passing is a consequence of sophisticated choices. Recall that the dissimulators were asked to select a group from a list of ethnic groups and told to pretend that they were in a situation in which it was important that they convince the respondent person that they belonged to an ethnic group different from their own. Given these instructions, subjects should have chosen a group in which they thought they could reasonably expect to pass successfully. We find in fact that there is a strong and significant (negative) correlation between what groups an individual chooses to pass as, and our measure of the distinctness of the groups. Although the sample of dissimulators is small (N = 32), in the sample of groups that are chosen with some positive probability, the correlation between distinctness and choice of groups is -.62; a shift from 0 to 1 in our distinctness scale is associated with a 50 percentage point increase in the propensity to choose a group for dissimulation.

Co-ethnicity also gives a respondent a leg up, making it more difficult for the subject to deceive him or her. This is true even conditioning upon that sample of players who correctly identify a player based on sign information. We had a sample of 245 players trying to identify non-co-ethnics whose identity they had already guessed correctly based on observing a video with name. 33% of these subsequently switched their guesses (incorrectly) once the subject tried to pass. In contrast we had 109 players guessing the identity of co-ethnics whose identity they had already guessed correctly based on observing a video with name. Just over one 21% of these subsequently (and incorrectly) revised their guesses. The 12% difference of the impact of signals on identifiability between these two groups is large and significantly different at the 95% level (the associated t-statistic is 2.36). Hence, *it is especially hard to use signals to fool co-ethnics.*

We do not find however that when players try to convince other players who had previously misidentified them of their true identity, that they are any more successful when engaging with co-ethnics than with non-co-ethnics. Hence we do not find that it is especially easy, in this sample, to use signals to *correctly* convince co-ethnics.

As well as being difficult to fool a co-ethnic, a passer will be less successful at fooling members of the groups in which she is trying to pass than she will be in trying to fool members of members of third groups. We have 354 cases in which a dissimulator who had been correctly placed based on video information only, then used cues to try to pass as a member of an out-group. In the 50 cases in which she attempted to pass as a co-ethnic of the respondent, she was successful just 18% of the time. In the 304 cases in which she attempted to pass as a member of a third group, other than that of her respondent or her own, she was successful almost one third of the time. This difference is significant at the 95% level. Examination of our simulations suggests we find that the failure to be falsely accepted as an in-group member arises especially when passers use language in order to mislead. They mislead those that do not speak the language but fail to mislead those that do.

A number of these effects remain large and significant as we move to increasingly stringent definitions of success. In particular, when we ask whether subjects were able to convince respondents of their false identity (definition 3, reported in column 5), age and SAT score emerge as powerful predictors.

6 Conclusion: Implications for Theories of Ethnic Politics

Despite the theoretical complexities associated with the notions of group membership and the "correct" placement of individuals into socially constructed categories, we argue that it is possible to develop a meaningful notion of ethnic identification based on binary relations between individuals and their own placement criteria and of the distinctness of different groups, from the perspective of a given player. Using these notions, we document that, contrary to the assumption of easy identifiability, implicit or explicit in much of micro-level theoretical work on ethnic politics, there is substantial variability both within and between groups in the identifiability of their members and that the distinctness of two groups varies considerably across players and across groups. Beyond demonstrating the existence of substantial variation, we demonstrate that there is a structure to the successes and errors that occur and argue that these structures can be important to models of collective action, at least in contexts where identification based on limited information is germane.

We find that when players in this sample err regarding membership within their own self declared groups, they are more likely to make errors of inclusion than errors of exclusion. Nonetheless they are significantly better at identifying co-ethnics than they are at identifying non-co-ethnics and are better than non-co-ethnics at identifying co-ethnics. They use the information contained in speech and especially in name to improve their guesses. However they appear to be weak at incorporating new information from verbal signals sent by possible dissimulators. The net effect of signals is a decline in identifiability, contrary to expectations from rational updating. Even still, some types of passing are easier than others. Individuals find it especially hard to fool in-group members into believing that they are out-group members and to convince non-co-ethnics that they are in-group members.

Such sources of variation have substantive implications collective action within and between groups. They suggest one reason for example why collective action may be easier for some ethnic groups than for others. If identifiability is imperfect, then the ability of groups to police their members will be weakened and the advantage that ethnic groups have for collective action, particularly in anonymous environments, will disappear. To the extent that identifiability varies systematically across ethnic groups, the ability of ethnic groups to achieve collective ends should vary as well. Furthermore, the results suggest that discrimination against some groups may be easier than against others. The fact that *passing* may be easier for some individuals and groups than for others has implications for the permeability of group boundaries and the ability of groups to police them, and for the collective action benefits that come from such policing. All of these features depend on the level of information available to individuals. Information however has different implications for different types of bilateral interaction. In particular, *the costs and benefits involved in gathering information about the ethnic identity of individuals may vary across groups*, with information adding especially in the case of in-group out-group identifications and when subjects are from difficult to identify groups.

For scholars, this research has implications for both empirical and theoretical work. For empirical work, our research serves as a demonstration of a methodology for collecting rich information about one important aspect of ethnic structures. Cross-national studies of the implications of ethnic structures on various political and economic outcomes have been hampered by the weakness of data on ethnicity. Much of the data that exists relies on a reified notion of what groups are, but also fails to capture any aspects of ethnic structures beyond the number and relative sizes of groups. Our measures of identifiability, distinctness, the propensity of different types of players to make different types of errors, and the marginal impact of signs and signals on these outcomes, provide a richer class of measures of group structure. The methodology employed here also suggests a way of taking seriously the notion that categorization depends on context and moving from there to generating measures of the impact of context on categories. Finally, the measures we employ here also involve elements of ethnic structures that are endogenous to political processes and that if collected over time can serve as a measure of the evolution of ethnic demographies, chief among these is information, contained in a player's guesses, about the salience of in-group and out-group categories and the effort that individuals invest in distinguishing between them.

There are also positive implications for theoretical work. While the assumption that the ethnic backgrounds of individuals are readily apparent has facilitated theoretical analysis, it has also limited the study of important aspects of ethnic processes that should no longer be overlooked. Variation in identifiability may have real consequences, as the account of ethnic identification failure in Sri Lanka provided earlier suggests. Taking these differences seriously—and incorporating them into models of ethnic politics—is a critical next step in producing better theories that link ethnicity to cooperation and conflict.

7 REFERENCES

Akerlof, George. 1976. "The Economics of Caste and of the Rat Race and Other Woeful Tales." *Quarterly Journal of Economics* 90: 599-617.

Allport, Gordon and Bernard Kramer. 1946. "Some Roots of Prejudice." *The Journal of Psychology* 22: 9-39.

Anderson, Benedict. 1991. Imagined Communities: Reflections on the Origin and Spread of Nationalism. Revised Edition ed. London and New York: Verso.

Barth, Fredrik. 1969. *Ethnic Groups and Boundaries*. Boston: Little Brown.

Bacharach, Michael and Diego Gambetta in "Trust in Signs" in Karen Cook, editor, *Trust and Society*, New York: Russell Sage Foundation, 2001.

Blasovich, Jim, Natalie Wyer, Laura Swart, and Jeffrey Kibler. 1997. "Racism and Racial Categorization." *Journal of Personality and Social Psychology* 72: 1364-1372.

Becker, Gary S. 1971. *The Economics of Discrimination.* Chicago: University of Chicago Press.

Caselli, Francseco and Wilbur John Coleman. 2002. "On the Theory of Ethnic Conflict." Typescript. Harvard University.

Chandra, Kanchan. 2001 "Cumulative Findings in the Study of Ethnic Politics." APSA-CP, Vol. 12 (1), 2001: 7-11.

Chandra, Kanchan. 2004. *Why Ethnic Parties Succeed: Patronage and Ethnic Headcounts in India*. New York: Cambridge University Press.

Fearon, James D. 1999. "Why Ethnic Politics and 'Pork' Tend to Go Together." Typescript. Stanford University.

Fearon, James D. and David D. Laitin. 1996. "Explaining Interethnic Cooperation." *American Political Science Review* 90, 4: 715-35.

Ferraro, Paul J. and Ronald G. Cummings. 2004. "Experimental Approaches to Understanding Intercultural Conflict Over Resources." Typescript. Georgia State University.

Fershtman, Chaim and Uri Gneezy. 2001. "Discrimination in a Segmented Society: An Experimental Approach." *Quarterly Journal of Economic*s 116, 1: 351-377.

Gil-White, Francisco J. 2004. "Ultimatum Game with an Ethnicity Manipulation: Results from Bulgan Sum, Mongoloa." In Joseph Henrich et al. *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies.* New York: Oxford University Press: 260-304.

Greif, Avner. 1989. "Reputation and Coalitions in Medieval Trade." *Journal of Economic History* 49 (December): 857-882.

Harris, David. 2002. "In the Eye of the Beholder: Observed Race and Observer Characteristics." PSC Research Report No. 02-522. Population Studies Center. University of Michigan.

Horowitz, Donald L. 2001. *The Deadly Ethnic Riot.* Berkeley and Los Angeles: University of California Press.

Human Rights Watch. 2002. "'We Are Not the Enemy': Hate Crimes Against Arabs, Muslims, and Those Perceived to be Arab or Muslim After September 11." 14, 6 (November).

Humphreys, Macartan and Habaye ag Mohamed. 2005. "Senegal and Mali." Forthcoming chapter in World Bank volume on civil wars in Africa.

Isaacs, Harold. 1975. "Idols of the Tribe." In Nathan Glazer and Daniel P. Moynihan. *Ethnicity: Theory and Experience.* Cambridge, MA: Harvard University Press: 29-52.

Laitin, David D. 1995. "Marginality: A Microperspective." *Rationality and Society* 7 (January): 31-57.

Pettigrew, Thomas, Gordon Allport, and Eric Barnett. 1958. "Binocular Resolution and Perception of Race in South Africa." *British Journal of Psychology* 49: 265-278.

Secord, Paul. 1959. "Stereotyping and Favorableness in the Perception of Negro Faces." *Journal of Abnormal Social Psychology* 59: 309-315.

Tajfel, H., Billig, M. G., Bundy, R. P., and Flament, C. 1971. "Social Categorization and Intergroup Behavior." *European Journal of Social Psychology* 1: 149-178.

8 APPENDIX I

Consider a case where player *i* is drawn from group *A* with probability *p*, and from group *B* with probability 1-*p*. Assume that, her declaration, c_j, takes into account a desire to minimize the weighted sum of total Type I and Type II errors, with weights *w* and (1-*w*) respectively.

In this case, declaring *A* for all players produces no Type I errors, but Type II errors in a (1-*p*) fraction of encounters. The resulting payoff is - (1-*w*) (1-*p*). Declaring *B* for all types gives produces no Type 2 errors but produces Type 1 errors in a *p* fraction of encounters and produces a payoff of -*wp*. Declaring *A* is better than declaring *B* if and only if: $-(1-w)(1-p) > -wp$ or equivalently, defining $k \equiv (1-p-w)/wp$, if $k < 0$.

Assume now that *j* is receptive to a signal *X* about players she encounters, where $X \in (0,\theta)$. Assume furthermore that a share, α, of *A* types and a share β of *B* types are capable (with no cost) of emitting signal θ, with α > β. Assume that all types *want* to be interpreted as being of type A and hence emit θ whenever possible. Applying Bayes' rule we now have: $\Pr(i \in A|\theta) = \alpha p/(\alpha p + \beta(1-p))$ and $\Pr(i \in A|0) = (1-\alpha)p/((1-\alpha)p + (1-\beta)(1-p))$. Having observed $X=\theta$, declaring *A* dominates declaring *B* if and only if: $-(1-w)(\beta(1-p)/(\alpha p+\beta(1-p))) > -w(\alpha p/(\alpha p + \beta(1-p)))$, or equivalently, $k<(\alpha-\beta)/\beta$. And, having observed $X=0$, declaring *A* dominates if and only if: $k < (\beta-\alpha)/(1-\beta)$.

These conditions allow us to make predictions based on the information contained in *X* as a function of $(1-p-w)/wp$. Generically, there are only four cases of interest: (i) $k \in [-1, (\beta-\alpha)/(1-\beta))$, (ii) $k \in ((\beta-\alpha)/(1-\beta), 0)$, (iii) $k \in (0, (\alpha-\beta)/\beta)$ and (iv) $k \in ((\alpha-\beta)/\beta, \infty)$.

In case (i), *j* declares $i \in A$ whether or not information level *X* is provided, in this case there is not enough information in *X,* to make it worth *j*'s while taking the risk of making a Type I error. In case (iv), *j* always declares $i \in B$, there is not enough information in *X* to make it worthwhile risking a Type II error. In cases (ii) and (iii), information in *X* is sufficient to induce a change in strategies. In case (ii), in the absence of information, *j* always declares "A", but on observing *X*, switches to *B* whenever X=0. In this case, the result of observing *X* is a rise in Type 1 errors and a fall in Type II errors. In case (iii), the *X* results in a switch from always declaring *B* to declaring *A* in whenever X=θ is observed. The result is a rise in Type II errors but a fall in Type I errors. In both cases the total expected number of errors falls, from (1-*p*) to (1 - *p*)β in case (ii) and from *p* errors to (1-α)*p* in case (iii).

Tables and Figures

Table 1: Identification Success (Headshots)

Self-declared Ethnic Group of Respondent (who views subject's image)	Self-declared Ethnic Group of Subject (whose image is viewed): A	B	C	D	E	F	G	Total
A	72.41	17.65	93.48	85.19	18.18	65.85	25	65.22
	(29)	(17)	(46)	(54)	(11)	(41)	(32)	(230)
B	**89.47**	**43.75**	96.88	86.05	33.33	44.44	41.67	70.19
	(19)	**(16)**	(32)	(43)	(12)	(27)	(12)	(161)
C	80	23.68	**97.18**	82.20	55	64.10	35.14	74.12
	(50)	(38)	**(142)**	(118)	(20)	(78)	(37)	(483)
D	81.16	25	94.81	87.34	48.39	47.58	35	70.48
	(69)	(52)	(135)	(237)	(31)	(124)	(60)	(708)
E	84.62	27.78	86.96	62.69	**100**	57.14	37.50	63.48
	(13)	(18)	(23)	(27)	**(5)**	(12)	(8)	(115)
F	81.25	22.22	91.38	82.93	10	63.64	40	68.32
	(32)	(27)	(58)	(82)	(10)	(88)	(25)	(322)
G	80.56	33.33	100	**87.50**	80	**70**	**78.95**	80.98
	(36)	(12)	(34)	**(48)**	(5)	**(30)**	**(19)**	(184)
Total	80.65	26.11	95.11	84.40	44.68	57.95	38.86	70.90
	(248)	(180)	(470)	(609)	(94)	(409)	(193)	(2203)

Note: Cells in this table report the probability (as a percentage) that a typical player of the group denoted in each row will correctly identify a typical player who self identifies as a member of the column group, given the information contained in a still headshot photograph. The number of experimental observations used to calculate these probabilities are given in parentheses underneath each entry.

Table 2: Determinants of Successful Ethnic Identification

	I	II	III	IV
Subject Ethnicity				
A	-0.272		-0.284	-0.166
	[1.62]		[1.70]*	[1.22]
B	-0.725		-0.727	-0.625
	[5.27]***		[5.24]***	[4.77]***
D	-0.23		-0.234	-0.154
	[2.10]**		[2.14]**	[1.82]*
E	-0.611		-0.613	-0.45
	[4.86]***		[4.96]***	[4.57]***
F	-0.517		-0.518	-0.279
	[4.51]***		[4.52]***	[2.76]***
G	-0.655		-0.658	-0.53
	[6.58]***		[6.68]***	[6.14]***
Respondent Ethnicity				
A		-0.067	-0.057	-0.039
		[1.82]*	[1.58]	[1.15]
B		-0.016	-0.003	0.006
		[0.41]	[0.09]	[0.16]
D		-0.049	-0.04	0.021
		[1.80]*	[1.38]	[0.97]
E		-0.067	-0.059	-0.031
		[1.48]	[1.30]	[0.74]
F		-0.067	-0.046	-0.026
		[2.19]**	[1.55]	[0.85]
G		0.096	0.087	0.045
		[2.88]***	[2.63]***	[1.15]
Co-ethnicity	0.084	0.176	0.091	0.065
	[3.38]***	[5.69]***	[3.73]***	[2.74]***
Information Level	Headshot	Headshot	Headshot	Video w Name
Pseudo R2	0.21	0.03	0.22	0.15
Observations	2203	2203	2203	2203

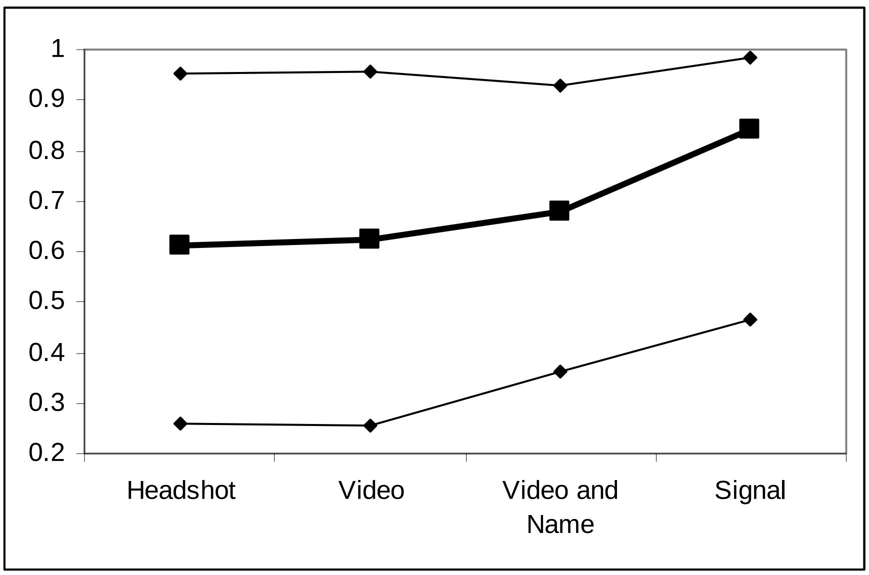
Notes: Dependent Variable is "Does the Ethnic Guess of the Respondent Match the Self-Reported Ethnic Group of the Subject?" Probit estimation, with marginal coefficient estimates (at mean values for the explanatory variables). Robust z statistics are in parentheses. Significantly different than zero at 90% (*), 95% (**) and 99% (***) confidence. Regression disturbance terms are clustered at the subject level.

Table 3: Percentage of Viewings with Correct Identification, by Group and Level of Information (Headshots)

Group	Average identifiability	*n*	Expected Standard Deviation Due to Sampling Error (95% confidence intervals in parentheses)	Observed Within Group Standard Deviation in Identifiability
A	0.86	14	0.07 (0.05, 0.10)	0.29
B	0.26	9	0.09 (0.05, 0.13)	0.30
C	0.95	25	0.05 (0.03, 0.06)	0.10
D	0.84	33	0.08 (0.06, 0.09)	0.23
E	0.44	5	0.10 (0.04, 0.17)	0.26
F	0.62	23	0.10 (0.07, 0.13)	0.31
G	0.36	11	0.10 (0.06, 0.14)	0.17

Note: The final column in this table reports the observed within-group standard deviation of the expected identifiability of each player. The second to last column reports the standard deviation that would be observed if the identifiability of all players were equal to the estimated identifiability of a typical member of the group. The 95% confidence intervals associated with this expected standard deviation demonstrate that the observed in-group variance in identifiability is significantly different from 0 for all groups.

Figure 1: The Marginal Impact of Information



Note: The thick line shows the average identifiability of the individuals in our sample conditional upon each of the information levels available from headshot to video, to video with name, to (truthful signal (under conditions where signals may be misleading). The upper line is the average identifiability under these conditions for members of the most identifiable group, the lower line shows the average identifiability for members of the least identifiable group. The first increment is not significantly different from 0, the second and third are at the 99% level.

Table 4 Distinctness

	A	C	E	B	G	D	F
A		**1.00**	**0.93**	**1.00**	**0.93**	**0.93**	0.61
C	**1.00**		**0.89**	**0.99**	**0.95**	**0.99**	**0.96**
E	**1.00**	**1.00**		0.67	**1.00**	**1.00**	**0.90**
B	**1.00**	**0.94**	0.54		0.38	0.08	0.56
G	**0.94**	**1.00**	**0.79**	0.39		**0.83**	**0.83**
D	**0.83**	**0.99**	**0.87**	0.29	0.51		0.46
F	**0.83**	**0.98**	0.51	0.46	0.44	0.61	

Note: Entries in the table show the empirical distinctness of row and column types *for a row type* based on the likelihood that row types are mistakenly classified as column types and column types are mistakenly classified as row types. A score of 1 indicates that a row player never confuses a row type for a column type. A score of 0 indicates that a row player invariable confuses one for the other.

Table 5: Propensity to Make Type I and Type II Errors

		Type I errors	Type II errors
Ethnic Group	A	**0.28**	**0.00**
	B	0.56	0.12
	C	0.03	0.01
	D	**0.13**	**0.15**
	E	0.00	0.04
	F	0.36	0.12
	G	0.21	0.07

Note: This first column in this table reports the estimated probability that a player from each of groups A-G incorrectly identifies an in-group member as an out-group member, given headshot information. This second column reports the estimated probability that a player from each of groups A-G incorrectly identifies an out-group member as an in-group member, given this information.

Table 6a: Identification Success with Simulation (Truth-Telling)

Ethnic Group of Respondent (who views subject's image)	Ethnic Group of Subject (whose image is viewed)							
	A	B	C	D	E	F	G	Total
A	50	25	100	100	100	83.33	100	88.37
	(2)	(4)	(13)	(11)	(2)	(6)	(5)	(43)
B		66.67	100	100	100	100	100	96.30
		(3)	(8)	(11)	(2)	(3)	(2)	(27)
C	80	25	97.37	78.57	100	92.86	77.78	86
	(5)	(4)	(38)	(28)	(3)	(12)	(9)	(100)
D	81.82	66.67	100	96.67	100	86.39	100	93.55
	(11)	(9)	(38)	(60)	(2)	(22)	(12)	(155)
E		40	100	100	100	100	100	83.33
		(5)	(5)	(3)	(2)	(2)	(1)	(18)
F	66.67	33.33	90.91	72.73	100	81.25	100	80.77
	(3)	(3)	(11)	(11)	(2)	(16)	(6)	(52)
G	60		100	75	100	75	100	80.65
	(5)		(6)	(12)	(1)	(4)	(3)	(31)
Total	73.08	46.43	98.32	89.71	100	86.57	94.74	88.50
	(26)	(28)	(119)	(136)	(12)	(67)	(38)	(426)

Note: Cells in this table report the probability (as a percentage) that a typical player of the group denoted in each row will correctly identify a typical subject who self identifies as a member of the column group, given the information contained in a video clip that contains the subject speaking, giving his or her (true) name, and trying to convince the row player of her true ethnicity. The number of experimental observations used to calculate these probabilities are given in parentheses underneath each entry.

Table 6b: Identification Success with Dissimulation (Trying to Pass)

Ethnic Group of Respondent (who views subject's image)	Ethnic Group of Subject (who is trying to pass as a member of a different group)							
	A	B	C	D	E	F	G	Total
A	40	0	88.89	50		0	42.86	45.45
	(5)	(3)	(9)	(14)		(6)	(7)	(44)
B	33.33	25	87.50	28.57	0	50	100	46.43
	(3)	(4)	(8)	(7)	(3)	(2)	(1)	(28)
C	40	14.29	87.18	41.67	66.67	18.18	28.57	55.21
	(5)	(7)	(39)	(24)	(3)	(11)	(7)	(96)
D	50	22.22	84.62	65	0	25	28.57	55.41
	(6)	(9)	(39)	(60)	(5)	(24)	(14)	(157)
E			100	57.14		16.67	100	54.17
			(6)	(7)		(6)	(2)	(24)
F	66.67	0	80	62.50		35.71	66.67	57.97
	(3)	(5)	(20)	(24)		(14)	(3)	(69)
G	66.67	33.33	92.31	37.50	50	50	85.71	67.50
	(3)	(3)	(13)	(8)	(2)	(4)	(7)	(40)
Total	48	15.15	86.57	55.56	21.43	25.37	48.78	55.24
	(25)	(33)	(134)	(144)	(14)	(67)	(41)	(458)

Note: Cells in this table report the probability (as a percentage) that a typical player of the group denoted in each row will correctly identify a typical subject who self identifies as a member of the column group, given the information contained in a video clip that contains the subject speaking, giving his or her (true) name, and trying to convince the row player that she is a member of some group other than her own. The number of experimental observations used to calculate these probabilities are given in parentheses underneath each entry.

Table 7: Determinants of Successful Passing

	I Dependent Variable: Does the Ethnic Guess of the Respondent Match the Self-Reported Ethnic Group of the Subject?	II Dependent Variable: Does the Respondent Change His/Her Guess from the Correct to the Incorrect Category?	III Dependent Variable: Does the Ethnic Guess of the Respondent Correspond with the Ethnic Group in which the Subject Tried to Pass?
Subject Fixed Effects			
A	-0.340	-0.197	0.138
	(1.24)	(2.64)**	(0.84)
B	-0.249	-0.013	0.084
	(1.20)	(0.11)	(0.88)
C	0.469	-0.211	-0.392
	(3.45)**	(3.55)**	(3.37)**
E	-0.261	0.148	0.087
	(2.25)*	(1.50)	(0.90)
F	-0.303	0.196	0.285
	(2.07)*	(2.10)*	(1.88)
G	-0.001	-0.071	-0.039
	(0.00)	(1.04)	(0.33)
Respondent Fixed Effects			
A Respondent	-0.040	-0.013	0.098
	(0.49)	(0.20)	(1.13)
B Respondent	0.096	-0.032	0.049
	(0.90)	(0.31)	(0.34)
C Respondent	-0.014	0.029	0.015
	(0.20)	(0.55)	(0.21)
E Respondent	0.123	-0.053	-0.009
	(1.10)	(0.57)	(0.09)
F Respondent	0.089	-0.138	-0.118
	(1.08)	(2.65)**	(1.80)
G Respondent	0.224	-0.043	-0.129
	(2.87)**	(0.67)	(1.67)
Characteristics of Subject			
Co-ethnic Pairing?	0.149	-0.066	-0.018
	(2.86)**	(1.45)	(0.33)
SAT Score	-0.221	0.014	0.168
	(2.32)*	(0.30)	(2.15)*
Subject Age	-0.098	-0.039	0.066
	(2.46)*	(1.92)	(2.32)*
Subject Female	0.125	-0.004	-0.140
	(0.95)	(0.05)	(1.33)
Father's Education	0.039	-0.027	-0.053
	(0.93)	(1.24)	(1.25)
Observations	432	432	432

Notes: Probit estimation, with marginal coefficient estimates (at mean values for the explanatory variables). Robust Z statistics are in parentheses. Significantly different than zero at 95% (*), 99% (**) confidence. Regression disturbance terms are clustered at the subject level.