Initiating trust: The conditional effects of sex and race among strangers

Catherine C. Eckel
Department of Economics
Virginia Tech
Blacksburg, VA 24061
eckelc@vt.edu

Rick K. Wilson
Department of Political Science
Rice University
Houston, TX
rkw@rice.edu

We report the results of a laboratory experiments on a one-shot two-person trust game. Two persons at different locations are paired for a single play of the game. Participants also complete two risk-assessment instruments: a survey and a decision task. The experimental treatments manipulate the information available to the subject about their counterpart. In one treatment subjects are given no information concerning their counterpart. In the second, each subject sees a photograph of his/her counterpart. We conjecture that people are strategic actors whose trusting behavior is conditional on the characteristics of a partner, and we are able to test the effect of observable characteristics on trust and reciprocity. A second experiment rates the photographs or participants in trust experiment. Those ratings are used to calibrate stereotypes about the images. Trusting behavior is significantly related to the race of the counterpart. Reciprocity is related to the race of the counterpart and the reciprocator's own sense of altruism.

Introduction.

There is widespread agreement that trust and social capital are linked in important ways. Whether the former is the basis for the latter or vice versa is a point of contention. Even so trust is viewed as important as the grease for political, economic and social transactions. Trust is assumed to be bolstered by group norms and institutions. Trust is assumed to be sustained by individual reputations. Most scholars further assume there is a reservoir of trust in society but lament its evaporation. Although much of the early work on trust focused on confidence in strangers ("generalized trust") there has been surprisingly little attention given to the interaction between strangers. As Hardin (2002) notes

> The biggest and most pervasive problem for us in trusting others is not the malign problem of dealing with cheaters, but the relatively neutral problem of often having to deal with people with whom we cannot expect to have ongoing relationships in which to ground incentives for trustworthiness. (p. 175)

If trust is important for building social capital (and this is the causal direction we prefer), then trust between strangers can be crucial for how social capital is accumulated. Trust between strangers is all about initiating a trust relationship. In this sense trust involves a truster placing himself into the hands of a trustee with the expectation that the trustee will reciprocate that trust. In a trust relationship there are mutual gains from initiating trust, but the truster is vulnerable because the trustee can choose to keep all of the gains.

Initiating trust among strangers can be important for expanding social capital. Those who are trusted can demonstrate whether they are trustworthy. Proving trustworthy may increase the likelihood of further profitable interactions. However, those who are not trusted cannot demonstrate their trustworthiness. By default they are considered untrustworthy. Moreover such individuals are omitted from relationships in which there are opportunities for mutual gain, they are left behind and they are disadvantaged.

We do not think that strangers rely on "generalized" trust in which they have an undifferentiated confidence in the trustworthiness of others. Nor do we think that strangers rely on "particularized" trust in which existing patterns of trust relationships define their expectations for trustee. Instead we suggest something akin to "conditional" trust in which guesses are made about whom to extend trust. In this sense individuals use the information available about their counterpart in order to calculate a strategy.

Unfortunately people are imperfect in their judgments about whom to trust. If people were perfectly able to read their counterparts, then trust would not be an issue. Trusters would only trust the trustworthy. But it is unlikely that there are any credible markers to note who is trustworthy and who is not. Instead, gossip, norms, reputations and institutions have stepped in to produce some guarantees for trustworthiness. Still, this does little for initiating trust among strangers. What is interesting is that strangers make decisions about whom to trust and they do so by forming judgments about the trustee. The critical question is whether those judgments, which are likely to be imperfect, are also systematically biased. If so, then flawed judgments with systematic biases can affect the ways in which social capital accumulates. Not surprisingly we focus on two obvious visual markers: gender and race or ethnicity.

We report the results of a two-phase experiment in which subjects interact with a counterpart at another location. Separate locations guarantee anonymity and preclude post-experiment interaction. In addition, our design allows us to selectively (and credibly) reveal information about the players to each other. This design is relevant for inferring the outcome of initial interactions between strangers, where agents can observe each other but are unlikely to interact in the future. We find that strangers both trust and are trustworthy. Yet the judgments made about partners are systematically biased.

Motivation.

The problem of initiating trust, especially among strangers, is important, but little attention has focused on it. Perhaps researchers share the concern by Hardin (2002) who writes:

> The devices available to us display no logic or natural order. They include taking risks with new people, reliance on reputation, the use of intermediaries in trust, quasi thick relationships that give us information and support, learning by example, escalation from minor to increasingly important matters on which we might trust someone, and, perhaps most appealingly, falling into love and friendship. (p. 133)

It is not clear whether these concepts help too much in understanding trust among strangers. Eckel and Wilson (2003) cast doubt on whether trust is based on dispositional traits concerning risk. Reputation, whether based on the past or built by anticipating the future, is unlikely to matter for people who will not interact in the future. Escalation from minor to major matters makes little sense in a one-shot setting (nor does it matter for learning by example). Strangers,

of course, fall in and out of love, but this is unlikely to direct people's willingness to trust.[1] What then, gives people confidence that they can choose trustworthy partners?

One fruitful way to consider trust is to treat it as a complete information, sequential game between two actors. In the simplest case consider the game represented in figure 1 In this game Player A has the right of a first mover. That player can choose to keep 10 units of a good (think of it as dollars) or can pass it to the second player. If kept, then Player A gets $10 and Player B gets nothing. If passed, the amount is tripled (think of this as a simple investment) and Player B is now in the position of keeping the full amount or splitting it with Player A (of course, there are many possible divisions, but here only two are considered). In a game theoretic sense, Player B need only decide between the two alternatives. Keeping everything ($30) is preferred to splitting the difference ($15). Knowing this, Player A then considers whether to keep the $10 or whether to pass it to Player B. The latter will result in Player A getting $0. Under backward induction, and assuming Player B is myopically rational, then Player A should never send anything. In a one-shot setting, where A and B know nothing about one another, it may be best that A assume B will behave myopically.

<Figure 1 About Here>

Suppose that Player A learns or observes something about Player B (that B is a nice person, that B promises to make the even split, etc.). Again, in a one-shot game, any claims made by (or inferred about) Player B are not credible. In this simplified game Player B can renege on any promise without penalty. Promises by Player B are costless and whatever A reads into B is cheap talk [see discussions by Farrell and Gibbons (1989)].

Despite the problem of cheap talk, there is substantial trust among strangers and people spend a good deal of time trying to "read" something into their counterpart. After all there is considerable gain to trusting (when reciprocated) and therefore a reason to initiate exchange. But, trusting only pays if the counterpart is trustworthy. People behave as if they are able to differentiate between trustworthy and untrustworthy counterparts. But what is it that people think they are reading? Are they reading the intentions of others, are they uncovering subtle cues that credibly signal an action or are they using a stereotype as a hypothesis about the actions of others?

---

[1] Finding oneself attracted to another and trusting them may not be too far off the mark. Eckel and Wilson (2004) offer evidence that attractiveness affects trust, possibly in a roundabout path to love via lust.

Although the concept of reading intention has attracted attention in economics [Smith (1998), Sally (2002), Camerer (2003)] we do not dwell on it here. Most of these models assume a response to an action (reading the intention behind the action) rather than how to initiate an action [see for example Rabin (1993), Levine (1998), Fehr and Schmidt (1999) or Bolton and Ockenfels (2000)]. The second approach turns on the question of credible signaling. Frank (1988) offers a simple model of a signal that can be credibly read. In particular he argues that many human emotions are impossible to mimic, are easily read and as a consequence serve to credibly commit individuals to an action. For example, if an offer is accompanied by a clear display of anger and exasperation, it is reasonable to conclude that no other offer is forthcoming. This perspective has support from psychology where Ekman (1982) argues that there are evolved, universal facial expressions that are easily read [for a cautionary note, see Fridlund (1994)]. The problem is that this approach requires clear signals that cannot be costlessly imitated [see Nowak and Sigmund (1998) or de Vos, et al. (2001)]. People often make errors in judgment or are fooled by what they see. It is unlikely that they are able to read noisy signals that might easily be copied.

We think the third approach to "reading" other people is more useful. It relies on the stereotyping literature from social psychology. A rich literature from the "minimal group" paradigm suggests that people quickly and easily differentiate between their in-group and an out-group [see Tajfel and Turner (1979, Tajfel and Turner (1978)]. Hirschfeld (1998) finds that even young children easily assign people to categories. Even so, there are many different dimensions across which people differentiate themselves (for example: academic/non-academic, male/female. Asian/non-Asian, white/red wine drinker, etc.). The salience of one's own group, the ease with which groups can be differentiated and the perceived threat of the out-group strengthens the degree to which in-group and out-group distinctions will be made [see the survey by Brewer and Brown (1998)].

As soon as people begin building categories they rely on generalizations (stereotypes) about those categories and act accordingly. Ross and Nesbitt (1991) find that people systematically overattribute the behavior of others using very simple categories. This is in line with Fiske and Neuberg (1990) whose "serial processing" model of stereotypes suggests that early experience makes it difficult to overcome biases in later interactions. Obviously stereotypes can be overcome if the target holds attributes inconsistent with the perceiver's

categories. This can be handled through a dual process model Brewer (1988), a serial processing model Fiske and Neuberg (1990) or through a connectionist model Kunda and Thagard (1996). Each model is concerned with some form of updating in which a stereotype is activated and subsequent interactions or further observation leads to strengthening or changing expectations about the stereotype.

When initiating trust with strangers there is little room for updating – one chooses to trust or not. As a consequence we think that stereotypes ought to matter a great deal for initiating trust. We focus on two simple, yet obvious, categories for sorting strangers: gender and race or ethnicity.

There is ample evidence that both categories generate stereotypes. In political science Huddy and Terkildsen (1993) point to ways that gender cues are used in voting, Alexander and Andersen (1993) look at cues tapped with respect to political leadership and Koch (2002) finds that in high visibility contests that the ideology of female candidates is overestimated (especially for the politically knowledgeable and for Republican candidates). More generally gender and how gender roles are perceived permeates political life [see the general review by Sapiro (2003)].

Likewise, stereotypes are easily triggered with race or ethnicity. When it comes to voting and the choice between candidates (particularly black and white) the results are nuanced. Terkildsen (1993) find that white voters avoid choosing black candidates. Sigelman, et al. (1995) find that race matters for candidate evaluations, but that race can be mediated by ideology. Certainly the "race card" has been used to activate stereotypes in campaigns Mendelberg (2001). In a like fashion political advertisements can be used to prime racial attitudes in elections Valentino, et al. (2002). In general see the review by Duckitt (2003).

What about in the domain of trust? Hetherington and Globetti (2002) make the case that political trust (trust in government) is important for determining support for racial policies. Arguing that Caucasians have more to lose when government implements redistributive racial policies, levels of trust should matter. They find that as trust in government increases (for Caucasians) so too does support for racial policies. Gay (2002) also focuses on political trust and asks whether having an African-American Member of Congress increases trust. Unfortunately "trust" is measured as an evaluation of legislative performance, so it only indirectly taps the concept of trust in government. What is clear is that Caucasians are less

positive in their evaluations when represented by an African-American. Whether this is a function of stereotyping, and absence of updating, or some other mechanism is left open. In a different vein, Uslaner and Conley (2003) turn toward "generalized" trust and ethnicity. Using a sample of ethnic Chinese, they find that those who are more comfortable with the broader U.S. society (presumably abandoning stereotypes of non co-ethnics) are also more likely to trust. This finding is in line with other results reported by Uslaner (2002) in which African-Americans are less likely than their Caucasian counterparts to exhibit generalized trust (see also Demaris and Yang (1994)). In general this literature points to the importance of racial categories for attitudes about trust.

There is a body of literature on trust from experimental economics. For example, Croson and Buchan (1999) find systematic sex differences between males and females as does Chaudhuri and Gangadharan (2002). By contrast Scharlemann, et al. (2001) find differences specifically due to different male/female pairings. With respect to ethnicity Fershtman and Gneezy (2001) find decreased trust (at least among males) when cued by an ethnic name. A similar finding emerges when South African high school students are cued with a photograph of their counterpart and mixed by racial groups Burns (2003). These scattered results imply that gender and race can be important categories for initiating trust (in a different context, see List (2004)).

Drawing on this disparate literature, we have several hypotheses. The first is drawn from the game theoretic literature. In a one-shot trust game the sub-game perfect equilibrium is not to trust.

H1: Trust will not be initiated with strangers.

It seems obvious that this hypothesis is false. There is a substantial literature on trust in a variety of experimental settings showing that trust is often offered and reciprocated [see for example: Kollock (1994); Berg, et al. (1995); Bolle (1998); McCabe, et al. (1998); and the overview by Camerer (2003)].

A second prediction draws on the conjecture that people look to reading a signal about others. When considering whether to trust, trusters rely on a stereotype that is activated concerning the trustee. This varies by the category information provided about the trustee. This leads to two types of hypotheses.

6

H2A: Stereotypes hold that females are more trustworthy than males.

H2B: Stereotypes hold that non-majority racial and ethnic groups are less trustworthy than Caucasians.

If H1 is false and either H2A or H2B are true, then we will observe trust initiated at different rates for different categories of individuals. This gives rise to the following related hypotheses.

H3A: Stereotypes lead to greater trust with a female counterpart.

H3B: Stereotypes lead to less trust with non-majority racial and ethnic groups.

We also take into account trustees. It seems reasonable that they should begin with similar judgments about their counterpart. But because they reciprocate only if they have been trusted, trustees can update their expectations about the counterpart. This leads to the following hypothesis concerning expectations.

H4: Exceeded expectations will be rewarded by trustees.

Trustees might expect that male or African-American counterparts will be less trusting. If trusted, trustees will update their prior expectations and respond by ignoring the stereotype.

Experimental Design.

We turn to a laboratory experimental design in order to ensure stranger interaction and to manipulate gender and race in the context of a one-shot trust game. Two distinct experiments were conducted. The first tests hypotheses concerning stereotypes about trustworthiness. The second experiment tests strangers in a one-shot trust game.

Experiment 1.

The first concern is whether individuals hold stereotypes about the trustworthiness of others. It may well be the case that gender or racial and ethnic stereotypes activate a series of expectations that then shape evaluations of trustworthiness. It is doubtful that trustworthiness is a primary part of any stereotype. However, for purposes of this analysis we want to know whether there are systematic differences across gender and race and ethnicity in attributing trustworthiness to an individual. To do so, we took photographs of subjects from the second experiment reported here. An independent sample of subjects evaluated the photographs, judging how well opposite word pair items fit the image. Those evaluations are used to test hypotheses 2A and 2B.

Design and procedure. In this experiment subjects were asked to rate photographs taken of subjects who were assigned to the *Photo* condition of the second experiment. We wanted an independent evaluation of the images so that we could confirm that a stereotype might be activated based on the same stimulus used in the second experiment. A total of 258 photos are evaluated, with 129 photos from people at Virginia Tech, 99 photos of people from Rice and 30 photos of people from North Carolina A&T. To ensure anonymity in the evaluations (such that no subject knew the person whose image was being evaluated) and to account for possible subject pool differences in judgments, subjects evaluated photos from another University.

Each subject was asked to rate between 15 and 20 photos using 15 opposite word-pair items and was paid $.25 per photo. Subjects did the evaluations on-line. The bulk of the subjects did the evaluations on their own computers at their own leisure. A subset of the subjects were brought into a laboratory and did their evaluations in a controlled setting. After completing the ratings subjects were able to pick up their payment (or if in the laboratory they were paid directly). Photos and their order were randomly assigned to each subject.

A total of 284 subjects participated: 52.5% were male and 47.5 % were female. Subjects spent an average of 84 seconds per photo (with a standard deviation of 68 seconds). While subjects were asked to rate between 15 and 20 photos this could not be enforced (only 11.3 percent failed to rate their full set). This left a total of 4560 evaluations. Because of incompletion and random assignment, the number of ratings per photo ranged from 8 to 28, with the interquartile range between 11 and 18 ratings.

Each face was evaluated using 15 opposite word pairs. A sample of the word-pairs and the screen used by subjects is shown in Figure 2, which reproduces part of an evaluation form. A complete listing of the word pairs is given in Table 1. The left/right order of the word pairs was randomly fixed prior to the experiment. The order of presentation of the word pairs was randomized across each photograph for each subject. This provides control over response-set bias in the word pairs. The rescaled ratings ranged from a low score of 1 (a negative evaluation) to a high score of 6 (a positive evaluation).

<center><Figure 2 About Here></center>
<center><Table 1 About Here></center>

Analysis. The primary concern is with the paired item, untrustworthy/trustworthy. The mean ratings are given in Table 1 for the overall evaluation as well as for categories of the evaluated photograph. Overall, the photographs are evaluated as trustworthy, with an average value of 4.09. When categories are compared there are significant differences. First, females are evaluated as more trustworthy than males, a result consistent with the Hypothesis 2A (t=14.39, df=4558, p<.001).[2] This gives credence to the idea that a stereotype can be activated in which females are viewed as more trustworthy.

In order to evaluate Hypothesis 2B each evaluator was asked to state the race/ethnicity of the photograph. In pretests we found that subjects had a difficult time categorizing some individuals (this was particularly true with Hispanic photographs). We use the evaluator's categorization in this analysis (rather than the photographed individual's self-identification). We find no support for Hypothesis 2B. Here non-majority images were expected to have lower ratings than Caucasians. Instead we find the opposite with Caucasians rated less trustworthy than non-Caucasians (t=-1.46, df=4558, p=.93 – one-tailed test based on the hypothesis). Particularly interesting is that African-Americans are rated as more trustworthy than Caucasians and the difference is striking (t=-2.73, df=3493, p=.99). This indicates that if any stereotype is activated it is that non-Caucasians are regarded as more trustworthy.

It may be that characteristics of the photographs as well as characteristics of the raters affect the evaluations. A set of regressions tested this possibility and these models are reported

---

[2] All of the analysis reported here was also repeated using a normalized score for the ratings. There is heterogeneity among raters in their evaluations – people are more or less generous in their evaluations. The ratings were normalized for each rater in order to control for that heterogeneity. The results do not change.

<center>9</center>

on Table 2. The models were estimated using random effects. This was chosen rather than fixed effects assuming that individual characteristics might interact with the various attributes of the photographs. The dependent variable is the rating of the trustworthiness of the photograph. The independent measures are dummy variables for a female photograph, an African-American, Asian or Other (non-Caucasian) photograph. Similarly there are dummy variables for the rater's own characteristics, including female, African-American, Asian and non-Caucasian "other." The baseline measures are Caucasian males. The first model (in column one) provides an estimate for the trustworthiness rating for all of the photographs. As already noted female photographs are rated more trustworthy than male photographs. African-Americans and Asians are rated slightly more trustworthy than Caucasians, but the effects are weak and not significant. Female evaluators give higher evaluations, although the effect is only a tenth of a point. Asian evaluators, on average, give lower evaluations. Little else shows up in the estimates.

<Table 2 About Here>

To account for interactions, six additional models are run, with each model estimating a specific category of the photograph. Model 2 estimates trustworthiness, but only for male images while model 3 does the same for females. Models 4 through 7 estimate coefficients for each of the racial and ethnic categories. In models 2 and 3 the only notable difference is that African-American evaluators give female photographs an extra boost in evaluations of trustworthiness. In models 4 through 7 only a few parameters affect the ratings. First, female evaluators are more likely to give higher marks for trustworthiness to males and to African-American images. Second, African-Americans evaluate African-Americans as more trustworthy than do other racial or ethnic groups. In fact this is the only instance in which there is an in-group effect. Generally, however, we find that the results from Table 1 hold up even when accounting for individual characteristics.

These results are surprising. First it is clear that there are strong stereotypes concerning males and females. Females are thought to the more trustworthy (and trusting), thereby confirming Hypothesis 2A. Females are also viewed as having a pleasant demeanor and are ascribed other positive attributes. By contrast Caucasians are viewed as less trustworthy than African-Americans or Asians. This is contrary to Hypothesis 2B and suggesting that Caucasians should be trusted at lower rates if that attribute is applied when considering whom to trust.

Experiment 2.

The second experiment tests the behavior of subjects in a simplified version of the investment or trust game, first developed by Berg, et al. (1995). All experiments were conducted over the internet. The experiment includes three main components: a trust game (under two information conditions, described below), an assessment of risk attitudes, and a survey of psychological and demographic characteristics. Subjects first complete a commonly-used survey instrument designed to measure attitudes toward risk -- the Zuckerman Sensation-Seeking Scale (SSS), form V Zuckerman (1994). Subjects earn 10 experimental laboratory dollars for completing the survey (the exchange rate was 2 lab dollars for each US dollar). In the second part of the experiment, subjects are randomly paired with another individual at a different University. Three laboratories are used, the Lab for the Study of Human Thought and Action at Virginia Tech in Blacksburg, VA, the Behavioral Research Lab at Rice University in Houston, TX, and the Computational Laboratory at North Carolina A&T. Subjects are randomly assigned to be either the first or second mover in a one-shot trust game with a specified counterpart at another site.[3] Details of the game are explained below.

After observing the characteristics of their counterparts, both first movers and second movers are asked to predict the actions of their counterparts. First movers record their predictions after their own decisions have been made, but before finding out the counterpart's move. Second movers record their predictions while awaiting the first-mover's decision. After the trust game decisions are completed, subjects turn to two risky tasks. In the first, they make a series of decisions using a computerized version of the risk instrument designed by Holt and Laury (2002). In the second, subjects face a choice between a certain outcome and a lottery over outcomes that mimic the distribution of payoffs found in previous versions of this trust experiment. In the final component of the experiment, subjects were asked to complete a three part questionnaire that collected 1) demographic information, 2) answers to survey questions designed to measure trustworthiness and altruism, and 3) debriefing information. The measures of trustworthiness and altruism were taken from Wrightsman's (1991) "Philosophies of Human Nature Scale." A total of 14 items were used, randomly ordered and presented to subjects in an on-line questionnaire. Responses to the items were arrayed on a 6-point Likert scale ranging

---

[3] The computerized experiment was run only between pairs of sites. The bulk of the experiments were run between Virginia Tech and Rice. Another series of experiments was run between Virginia Tech and North Carolina A&T.

11

from Strongly Agree to Strongly Disagree (on this scale subjects had no neutral mid-point and were forced to agree or disagree responses).

This paper focuses on behavior in the trust game and correlations with the demographic and attitudinal data; a more detailed analysis of the relationship between trust and risk is given in Eckel and Wilson (2003).

The Trust Game. In the trust game first movers choose whether to keep the $10 they earned for completing the Zuckerman risk assessment survey, or pass the $10 to their counterpart. If a subject keeps the money, then this part of the experiment is finished. Otherwise the amount is doubled and the counterpart decides among nine different allocations of the $20, ranging from ($0, $20) to ($20, $0) in $2.50 increments. In this experiment the first-movers make an all-or-nothing decision, rather than choosing to send $0-$10 in increments of $1 as with Berg, et al. (1995). The amount sent is doubled, a decision largely made for budgetary reasons.[4] The first-mover's decision is framed as a choice of whether to make a loan of $10 experimental dollars to a counterpart. The decision is framed in this way in order to provide some minimal context to the decision. The response by subjects to other abstract games made us suspect that subjects might not think of this as a "trust" game at all. We also thought that minimal context might make the game easier for subjects to reason their way through. Finally the second mover determines the allocation of the resulting $20 in fixed increments. At one extreme the second mover could take all $20 and return $0 to the first mover. Or the second mover could take $17.50 and return $2.50. The remaining allocations changed in increments of $2.50, with the final option to take $0 and give the first mover $20. These allocations include all relevant focal points (20, 0, the payoff maximizing decision; 15, 5, representing "split the surplus"; and 10, 10, representing "repay the loan").

Information treatments. Two different treatments were implemented that vary the information given to subjects about their counterparts. In the *No Information* condition subjects were given no information about their counterpart, except that they were located at another university. Pilot experiments showed that subjects were suspicious that they were actually paired with a counterpart, so a subject was randomly selected at each site, and was asked to provide a

---

[4] Several experiments have been conducted that double, rather than triple the amount sent. Notable among these is Glaeser, et al. (2000). As well, Snijders and Keren (1999) provide an interesting test of price (multiplier) sensitivity in a binary version of the trust game.

short code word. This word was then written on a large sheet of paper and photographed with the subjects at their computers in the background. Photos from both locations were uploaded to the server and appeared on the computer screens of all subjects. Eckel and Wilson (2003) detail this procedure and the extent to which it overcomes skepticism about the experiment. Other than this information, these subjects were told nothing about their counterparts.

In the second treatment, the *Photo* condition, subjects entered the laboratory and were told they would have four photographs taken. All photographs were taken with Sony Mavica digital cameras, with the individual's face centered in the picture, with the top of the head marking the top of the picture and the chin marking the bottom of the picture.[5] Lighting and background were the same at both sites. In two of the photographs the subject was instructed to give a neutral expression, while in the remaining two photographs the subject was instructed to smile. The subject was told that these photographs would be used in the experiment, and that they would get to choose which one they wanted their counterpart to see. All four photographs were uploaded to a server as soon as they were taken. When the experiment was ready to begin the subject was shown all four of the pictures and instructed to choose one that would be sent to the counterpart at the other site. Prior to playing the trust game each subject viewed the photograph selected by his/her counterpart. A sample decision screen is shown in Figure 3. In this setting the counterpart observed not only the sex of the counterpart, but also the race, ethnicity, attractiveness, demeanor, etc., of the subject. Both first and second movers viewed a photograph of their counterpart prior to making the decision.

<Figure 3 About Here>

Procedure. A total of 332 subjects, 50 percent from Virginia Tech, 41 percent from Rice University and 9 percent from North Carolina A&T, participated in 14 experimental sessions (ranging from a low of 10 to a high of 34 subjects per session).[6] Subjects were recruited from introductory classes in Principles of Economics at Virginia Tech and North Carolina A&T and from dining halls at Rice University. Volunteers were told to report at a specific time to a laboratory at their respective locations. Subjects were 56.3 percent male, and just under 90

---

[5] Images were taken at 640x480 pixels with 24 bit color resolution. This resolution was selected because it allowed a clear picture and minimized the size of the compressed jpeg files. The photos were displayed at a resolution of 300x225 pixels.

[6] The final two sessions were conducted between Virginia Tech and North Carolina A&T. The latter is an Historically Black University with a strong engineering flavor, matching Virginia Tech. Conducting these experiments at NCAT allowed us to supplement our sample of African-American students.

percent of subjects were between the age of 18 and 22. The subject pool was ethnically-diverse with 62.6 percent Caucasian, 15.1 percent African-American, 10.8 percent Asian-American, 5.4 percent Hispanic and the remaining 6.0 percent self-identified foreign nationals.

Oral instructions were read to subjects before beginning. After that, subjects went through a set of self-paced, computerized instructions. In a post-experiment questionnaire 93.7 percent strongly or very strongly agreed that the instructions were clear. Once subjects began the experiment, no talking was allowed. Subjects were asked to raise their hands if they had a question or problem, and the experimenter would answer their questions privately. When subjects completed the experiment, they were paid one at a time and in private. At that time subjects were handed a debriefing statement and asked whether they had any questions. Throughout the course of the experiment no deception was used.

Results. Table 3 reports aggregate results for trusting and reciprocity by treatment. As expected Hypothesis 1 is easily rejected. It predicts that no one will trust and yet in each experimental condition over 70 percent of first-movers made the $10 loan to their counterpart. More subjects trusted in the *Photo* condition than in the *No Information* condition, although this difference is not statistically significant (χ^2=0.36, df=1, p=0.55). Expectations were low, with subjects on average anticipating $8.47 – less than the $10.00 loaned. Those in the *Photo* treatment were more optimistic, expecting $1.00 more would be returned than those in the *No Information* condition (in a one-tailed test the difference is not statistically significant, t=1.09, df=164, p=.14).

<Table 3 About Here>

Finally, as the last column shows, these expectations were justified. On average, trusting did not pay, with second movers returning less than the initial $10 transfer. The distribution of amounts returned is shown in figure 4. Keep in mind that subjects were limited to fixed allocations of the $20 in $2.50 increments. Modally, second movers returned exactly what was loaned in both treatments: while returning $10 might seem like a fair allocation, it only means that the money was returned to the first mover and the second mover kept the full surplus. Subjects in the *No Information* condition were more likely to return amounts below $10 (38.5% as compared to 25.8%). Seven subjects in the *Photo* condition returned $15 or more, while none in the *No Information* treatment did so. On average, subjects in the *No Information* treatment

returned $7.50, while those in the *Photo* treatment returned $8.80; this difference is statistically significant at the .10 level under a Mann-Whitney test of the distributions (z = 1.70; *p* = 0.09).

<div align="center"><Figure 4 About Here></div>

From these results it appears that there is no difference between the treatments. Yet in the Photo condition subjects had much more information about their counterpart. The question is: did they use that information?

Sex, ethnicity, and trust. The data from Table 4 allow us to test Hypotheses 3A and 3B. Given what first movers could observe about their counterpart did this affect their willingness to trust? There is no support for Hypothesis 3A. Females are somewhat less likely to be trusted than men (74.1 percent to 76.1 percent), but this difference is not significant. On the other hand, there is some support for Hypothesis 3B even though non-Caucasians were imputed to be more trustworthy. As noted in Table 4 Caucasians are trusted at a much higher rate than non-Caucasians (81.5 percent to 64.6 percent). The only exception is that of self-identified Hispanics who were always trusted, although the numbers for that group are quite small. Our sample of Asian-Americans and African-Americans are large enough to give us confidence in the findings. The data show that African-Americans are trusted the least (58.3 percent) and that Asian-Americans are trusted only slightly more (66.7 percent). These findings point to discrimination in trusting.

<div align="center"><Table 4 About Here></div>

Expectations about reciprocity. It may be that there are differences in expectations triggered by the different counterparts. As noted above, first movers were optimistic. Overall, 78.9% expected that at least the $10.00 loan would be returned (with 80.6% believing this in the *Photo* condition and 73.0% believing so in the *No Information* condition). Not surprisingly, there is a strong relationship between expectations and the decision to trust. For those who did not trust, 76.7 percent expected to get back less than the $10 loan. By comparison just under 23 percent of those who trust expected to get back less than the loan. The second column of Table 4 details first mover beliefs broken out by characteristics of the second mover. These beliefs are partly in line with Hypotheses 3A and 3B. Women are expected to return more, on average. African-Americans, Asian-Americans and Foreigners are expected to return less than

<div align="center">15</div>

Caucasians. None of these differences are statistically significant, although they indicate there might be some foundation for category-based stereotypes directing strategic choice.

From these aggregate and bivariate analyses it appears that expectations, gender and race or ethnicity are accounting for some of the decisions. However, we have not yet controlled for characteristics of the first mover. For example, it may be that differences in behavior are due to dispositional traits (e.g., some people may just be more trusting). Also, we have not accounted for deeper strategic reasoning. For example the particular pairing of subjects (a male truster and a female trustee) may create different joint expectations. In order to tease out these, and similar, possibilities, we need to turn to multivariate analysis.

First-Mover Multivariate Analysis. In the models reported here we focus on the first mover's decision to trust (make the loan). Table 5 provides the variable definitions as well as descriptive statistics for the independent variables. The parameter estimates for the different models are given in Table 6. The first model includes a dummy variable for the experimental manipulation and variables tapping individual characteristics of the first mover. We find no significant difference between the treatments with the coefficient for the *Photo* treatment effectively zero. Of the individual characteristics, several are significantly related to the decision to trust. There is a weak positive relationship between the Zuckerman Sensation Seeking Scale and willingness to trust. A similar finding is detailed in Eckel and Wilson (2003) and indicates that the more risk seeking a subject the more likely the loan.

There is also a relationship between altruism and the willingness to make the loan, but not between trustworthiness and making the loan. Altruism and trustworthiness are drawn from Wrightsman's (1991) Philosophies of Human Nature scales. Each concept is based on seven attitudinal items relations to actions that could be taken with respect to strangers. The items are on a six-point scale, have been coded in the same direction and are averaged. Higher scores tap greater altruism and trustworthiness.[7] Unlike Glaeser, et al. (2000) we do not find a relationship between trust and trustworthiness. Unexpectedly, we find a strong, negative, relationship between trust and altruism. Finally, there is a strong relationship between expectations and making the loan. Not surprisingly, the more optimistic people are about what will be returned,

---

[7] The two items are positively correlated (ρ=.55). Estimated singly in each of the models trustworthiness is positive, but not significant. Altruism is negative and significant in Model 1, remains negative, but insignificant in the remaining models with or without trustworthiness.

the more likely they are to make the loan.

<center><Table 5 About Here></center>

Equally interesting in this model is what is not related to trust. While other research has found that women are less likely to trust, we do not find this is the case. The coefficient is nearly zero and is insignificant. Next, accounting for various racial and ethnic categories of the truster we find no significant effect on trust. This is estimated using dummy variables for the different groups and is based on self-identification. The omitted group is Caucasian. From this first model we conclude that there are few individual characteristics correlated with the decision to trust.

<center><Table 6 About Here></center>

Model 2 uses data from only the *Photo* condition. Here subjects are able to condition their decision based on gender and race information about the counterpart. With respect to the truster's characteristics we obtain the same findings: the Zuckerman scale and expected return have positive, significant coefficients (at least at the .10 level). The altruism scale remains negative, but is no longer significant. The gender and race or ethnicity of the first mover does not have an effect. On the other hand characteristics of the counterpart matter. First, knowing the counterpart's gender has no significant effect. Contrary to the prediction of Hypothesis 3A, the coefficient is negative, but insignificant.

Second, the race and ethnicity of the counterpart matters, although this analysis is not as straightforward. Subjects in this experiment were not asked to guess the racial or ethnic identity of their counterpart. From the analysis of data from the first experiment we discovered that many evaluators had difficulty identifying the race or ethnicity of the photographed images. We assume that subjects in this experiment also had the same difficulty. We use the modal racial or ethnic perception of the counterpart as recorded in the first experiment, with two dummy variables for African-American and Asian (Hispanics were a very small group and were often labeled as Caucasian). Caucasian is the omitted group.[8] We find that race matters if the counterpart is African-American. The coefficient for African-American is negative and significant, indicating that such a counterpart is much less likely to receive the loan. This finding

---

[8] Using the perceived racial or ethnic category seems justified because we are trying to capture the perceptions of the first movers. If the same models are run with the self-identified race or ethnicity of the second mover, we get similar results. All of the coefficients retain their signs and the African-American coefficient continues to be significant at the .05 level.

<center>17</center>

is in line with Hypothesis 3B, although it does not jibe with the findings from the first experiment showing that African-Americans trigger a stereotype attributing trustworthiness. The coefficient for Asians is also negatively signed, but the coefficient is not statistically significant. Overall these estimates point to resistance for trusting racial and ethnic minorities.[9]

Perhaps the effect we find for race is purely a function of trusters reading something else into the appearance of their counterpart. In previous studies smiles Scharlemann, et al. (2001) and attractiveness Eckel and Wilson (2004) have been shown to matter. To capture the appearance of the subject a composite measure was calculated from the evaluations of the photographs. The evaluation items were factor analyzed, yielding a single strong factor and a weak second factor (the eigenvalues were 5.71 and .95, respectively). Factor scores were generated for the first dimension and then averaged across each photograph. Negative values indicate a low evaluation and positive values indicate a high evaluation of the image. By including this measure we control for a variety of characteristics that are read into the image.

The general appearance measure is added to Model 3 and there are several notable effects. First the coefficient for the gender of the counterpart becomes more negative (and approaches significance). Second, the negative coefficient for both African-Americans and Asians increase and the latter becomes statistically significant at the .10 level. Again, this supports Hypothesis 3B. Finally, the general appearance measure is positive and significant. This indicates that first movers are responding to something non-verbal embedded in the photograph. The items that load most strongly on this factor include trusting, cooperative, honest, friendly and happy aspects of the image. This measure controls for pleasant (and unpleasant) aspects of the counterpart. While it has a strong effect it does not erase obvious features of the counterpart (e.g., race).

Models 2 and 3 from table 6 make it clear that subjects condition their decision on information about their counterpart. In the *Photo* treatment subjects focus on an obvious characteristic – the observable race of their counterpart. Subjects also make a subjective assessment of the trustworthiness of the counterpart. This seemingly subjective judgment is borne out by an independent evaluation of the image that each subject chose for the experiment.

---

[9] Hispanics constitute a very small part of these data. In the evaluation phase of the experiment we found that there were often labeled as Caucasians. Because of they are more numerous in these data we focus on African-American and Asian subjects. The results reported here hold whether Hispanic subjects are included or excluded from these data.

The more positive the evaluation the more likely the counterpart was trusted and a loan made. The question now is whether the second mover responds in the same vein.

Second Mover. In light of what is found with the first mover behavior, we turn to the second mover's decision about what to return in the event that a loan was made. There are two concerns here. First, do trustees use observable categories when deciding what to return? This follows along with Hypotheses 3A and 3B only relating to reciprocated trust on the part of second movers. Second, given that a loan is made, do trustees update their initial view of their counterpart and respond accordingly (in line with Hypothesis 4)? We investigated several alternative specifications of the dependent variable, including the dollars returned to the counterpart and we settled on a simple dummy variable for the dependent variable. We measure whether the second mover returned at least what was loaned ($10 or more). If less than the loan was returned (or nothing returned) the dependent variable is assigned a value of zero. Three different models are estimated. As with Table 6, the first model compares structural aspects of the experiment and individual characteristics. The second model focuses on observable characteristics of the truster as well as expectations by the trustee. The third model uses a two-stage estimate that accounts for initial expectations and then examines whether expectations were exceeded. Results are given in Table 7.

<Table 7 About Here>

The first model includes a dummy variable for the *Photo* condition and variables related to characteristics of the trustee including: gender, race or ethnicity and a measure of altruism. A measure of expectations is included in which the trustee was aksed how likely it was that the loan would be made, ranging along a four-point scale from Very Unlikely to Very Likely. A second measure was included to reflect exceeded expectations. It is a dummy variable equal to one if the trustee though the loan was Unlikely or Very Unlikely, but that it was made. Because the truster's decision was all or nothing and because we are estimating what was reciprocated conditional on the loan being made, we can only measure exceeded expectations. In this model we do not include any measure of risk. We have no expectation about how risk would enter the decision because there is no subsequent action by the first mover that puts the second mover at risk.

Model 1 indicates that in the *Photo* condition there is a tendency to return at least the

19

loan, although the coefficient is weak and only approaches statistically significance at the .10 level. This is consistent with other work showing that people are more generous when they view the target of their generosity [i.e., Bohnet and Frey (1999)]. Neither the gender nor the race or ethnicity matters in this model. This indicates that if trusted, females and African-American (and to a lesser extent Asians) are just as likely to return at least the loan. If given a chance to be trustworthy, these groups respond accordingly. Expectations are positively related, but insignificant. This latter implies that people update, but the effect is weak. The only individual characteristic that matters is the altruism score. Unlike the models for the truster, the coefficient for the trustee is positive. Given the stranger setting in which the trustee faces no sanctions (and has considerable social and physical distance from the first mover) this effect for altruism is not surprising.

The second model keeps characteristics of the trustee and introduces observable variables about the first mover, including an overall evaluation of the photographed image. As with the earlier estimates for the truster decision, the perceived race or ethnicity is used. In model 2 the individual characteristics of the trustee remain the same. Altruism continues to have a strong, significant effect. If expectations were exceeded the effect is more positive and approaches statistical significance, in line with Hypothesis 4. Do characteristics of the truster make any difference? Having a female counterpart increases the likelihood of returning the loan, although the effect is weak. African-American and Asian counterparts are less likely to have the loan returned to them, but only the former approaches statistical significance. The overall evaluation of the truster has only a minor positive effect, indicating that the trustee is not responding to this part o the stimulus. From this model is appears that little matters except for a trustee's personal sense of altruism.

A major problem with Model 2 is that expectations and exceeded expectations are formed at different times and through different processes. To capture this a two-stage model is estimated. In the first stage expectations are estimated as to whether the loan was likely or unlikely (a dummy variable was created for the four-point variable used to measure expectations). The regressors in the Probit equation are based on the observable characteristics of the truster. This allows a test of expectations that are formed on the basis of stereotyping. The predicted values from this equation are then used in a second Probit regression that includes trustee characteristics and whether the trustee's expectations were exceeded. Modeling the

20

process in that manner enables us to isolate expectations derived from gender, race or ethnicity and appearance from expectations that are exceeded.

Columns 3 and 4 in Table 7 present the first and second stage estimates, respectively. The first stage estimates make it clear that both gender and race are important for forming expectations. Female trusters are expected to make the loan, while African-Americans are not. This latter coefficient is quite strong and significant. An Asian counterpart does not generate any expectations, nor does the overall evaluation of the appearance of the photograph. The effects for females and African-Americans supports the thrust of Hypotheses 3A and 3B in that characteristics are attributed to strangers and that these shape expectations. In the second stage estimates these expectations are plugged into the model. They are positively related to returning the loan, as is the trustee's altruism score. Interestingly, under this model we find little support for exceeded expectations leading to the truster being rewarded. Expectations have already been formed prior to deciding whether to return the loan and African-Americans suffer.

Discussion and Conclusion.

At the outset we considered how it is that people initiate trust with strangers. Like many others we find that people are willing to trust strangers. They do so even though it is hardly the prudent game theoretic strategy. In this particular experiment, it was perhaps best not to trust, because trust, on average, did not pay. Not only was less returned to those who trusted, but even the trusters expected less to be returned than they sent.

Nonetheless, people trusted at high rates. They did so independent of information about the counterpart – trusting at roughly equivalent rates regardless of whether the trustee was anonymous or could be seen through a photographic image. Trustees, on the other hand, were slightly more likely to return more to individuals represented by photographs. This is little surprise because knowing something about one's partner often results in more pro-social behavior.

We find that there is little about people to predict whether they will trust or be trustworthy. That is, there are few predictors of an individual's disposition to trust. The exception to this is a measure of altruism. This measure taps an individual's willingness to help others. On the trustee side, where trustworthiness is unrewarded in this one-shot game with strangers, it makes sense that altruism should matter. Altruism could be considered as an

important argument for one's own utility function. On the other hand, the peculiar finding that altruism is negatively related to trust opens up a variety of possible explanations, all of which demand additional scrutiny.

If others are taken into account it is no surprise that trusters and trustees alike use categories read into the images to make decisions about whom to trust. Two categories – gender and race or ethnicity – were predicted to be the primary categories used by subjects to condition their strategy choice. Only race appears as a strong predictor, with African-Americans trusted less and rewarded less. The results for gender are quite weak. It is not surprising, given the literature on stereotyping, that race ought to be a ready category for judgment. But, what is disturbing in these data is that race does not trigger negative attributes concerning trust or trustworthiness. In fact, non-Caucasians are viewed as both more trusting and more trustworthy than Caucasians. This is hardly in line with a view that negative characteristics are attributed to a particular category (race) and that carries over into the decision to trust.

A second component to the stereotyping literature holds that if a category is used and additional information is provided about the stereotyped target, then some form of updating will occur (a similar point is true if the problem is treated as one of a signaling game). We thought that trustees, expecting no trust but finding it, would respond by updating their priors. However, we find little evidence for that position. Exceeded expectations contribute slightly to returning the value of the loan, but they are swamped by the independent effects of race in forming initial expectations.

What do we conclude from these findings? These experiments do not indicate that subjects, when considering whether to trust, are simply invoking a stereotype based on expectations of trustworthiness. The expectations are more deeply rooted. Subjects are attentive to race as a category and non-Caucasians (African-Americans in particular) are distrusted. Other visible categories are far less important.

If racial categories are routinely applied to interactions with strangers (and there is overwhelming evidence to this effect in many domains) this has an especially deleterious effect on building social capital. Initiating trust allows the trustee to demonstrate trustworthiness. This is critical for building a reputation and to building extended social exchange networks. Where trustworthiness remains undemonstrated, trusting relationships are unlikely to grow. When the basis of engaging in trust is based on a racial or ethnic marker, then it means segregating groups

of people and putting up barriers to the formation of social capital.

What is also disturbing in these findings is that further information about the counterpart (updating in terms of exceeding expectations) does not overcome initial expectations. Where those expectations are tied to racial or ethnic markers, it only enhances barriers between groups.

Our findings do not show that people are hesitant to trust strangers. In fact we find the opposite. Trust in strangers is widespread. However, that trust (and trustworthiness) is not uniformly distributed across racial and ethnic lines.

Bibliography

Alexander, Deborah, and Kristi Andersen. "Gender as a Factor in the Attribution of Leadership Traits." *Political Research Quarterly* 46 (1993): 527-45.

Berg, Joyce E., John W. Dickhaut, and Kevin. McCabe. "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10, no. 1 (1995): 122-42.

Bohnet, I., and B. S. Frey. "The Sound of Silence in Prisoner's Dilemma and Dictator Games." *Journal of Economic Behavior & Organization* 38, no. 1 (1999): 43-57.

Bolle, Friedel. "Rewarding Trust: An Experimental Study." *Theory and Decision* 45 (1998): 83-98.

Bolton, Gary E., and Axel Ockenfels. "Erc: A Theory of Equity, Reciprocity, and Competition." *American Economic Review* 90, no. 1 (2000): 166-93.

Brewer, Marilyn B. "A Dual Process Model of Impression Formation." In *Advances in Social Cognition*, edited by T. K. Srull and R. S. Wyer, 1-36. Hillsdale, NJ: Erlbaum, 1988.

Brewer, Marilyn B., and Rupert J. Brown. "Intergroup Relations." In *The Handbook of Social Psychology*, edited by Daniel T. Gilbert, Susan T. Fiske and Gardner Lindzey, 554-94. Boston, MA: McGraw-Hill, 1998.

Burns, Justine. "Insider-Outsider Distinctions in South Africa: The Impact of Race on the Behaviour of High School Students." Paper presented at the Conference on Field Experiments in Economics, Middlebury College, Middlebury VT, April 25-27 2003.

Camerer, Colin F. *Behavioral Game Theory: Experiments in Strategic Interaction*. Edited by Colin F. Camerer and Ernest Fehr, *The Roundtable Series in Behavioral Economics*. New York, NY: Russell Sage Foundation, Princeton University Press, 2003.

Chaudhuri, Ananish, and Lata Gangadharan. "Gender Differences in Trust and Reciprocity." In *Paper Presented at the American Economic Association Meetings*. Atlanta, GA, 2002.

Croson, Rachel, and N. Buchan. "Gender and Culture: International Experimental Evidence from Trust Games." *American Economic Review* 89, no. 2 (1999): 386-91.

de Vos, Henk , Rita Smaniotto, and Donald A. Elsas. "Reciprocal Altruism under Conditions of Partner Selection." *Rationality and Society* 13, no. 2 (2001): 139-89.

Demaris, Alfred, and Renxin Yang. "Race, Alienation, and Interpersonal Mistrust." *Sociological Spectrum* 14, no. 1 (1994): 327-49.

Duckitt, John. "Prejudice and Intergroup Hostility." In *Oxford Handbook of Political Psychology*, edited by David O. Sears, Leonie Huddy and Robert Jervis, 559-600. Oxford: Oxford University Press, 2003.

Eckel, Catherine C., and Rick K. Wilson. "Internet Cautions." In *Working Paper, Political Science, Rice University*. Houston, TX, 2003.

———. "Is Trust a Risky Decision?" *Journal of Economic Behavior & Organization* (forthcoming) (2003).

———. "Never Judge a Book by Its Cover: A Beauty Premium for Trust?" In *Working Paper, Political Science, Rice University*. Houston, 2004.

Ekman, Paul. *Emotion in the Human Face*. 2 ed. Cambridge: Cambridge University Press, 1982.

Farrell, J., and R. Gibbons. "Cheap Talk Can Matter in Bargaining." *Journal of Economic Theory* 48, no. 1 (1989): 221-37.

Fehr, Ernst, and Klaus M. Schmidt. "A Theory of Fairness, Competition and Cooperation." *Quarterly Journal of Economics* 114, no. 3 (1999): 817-68.

Fershtman, Chaim, and Uri A. Gneezy. "Discrimination in a Segmented Society: An Experimental Approach." *Quarterly Journal of Economics* 116, no. 1 (2001): 351-77.

Fiske, Susan T., and Steven L. Neuberg. "A Continuum of Impression Formation, from Category-Based to Individuating Processes: Influence of Inormation and Motivation on Attention and Interpretation." *Advances in Experimental Social Psychology* 23 (1990): 399-427.

Frank, Robert. *Passions within Reason: The Strategic Role of the Emotions*. New York: W. W. Norton & Co, 1988.

Fridlund, Alan J. *Human Facial Expression: An Evolutionary View*. San Diego: Academic Press, 1994.

Gay, Claudine. "Spirals of Trust? The Effect of Descriptive Representation on the Relationship between Citizens and Their Government." *American Journal of Political Science* 46, no. 4 (2002): 717-33.

Glaeser, Edward L., David Laibson, Jose A. Scheinkman, and Christine Soutter. "Measuring Trust." *Quarterly Journal Of Economics* 115, no. 3 (2000): 811-46.

Hardin, Russell. *Trust and Trustworthiness*. New York: Russell Sage Foundation, 2002.

Hetherington, Marc J., and Suzanne Globetti. "Political Trust and Racial Policy Preferences."

American Journal of Political Science 46, no. 2 (2002): 253-75.

Hirschfeld, Lawrence A. *Race in the Making: Cognition, Culture, and the Child's Consruction of Human Kinds*. Edited by Lila Gleitman, Susan Carey, Elissa Newport and Elizabeth Spelke, *Learning, Development, and Conceptual Change*. Cambridge, MA: The MIT Press, 1998.

Holt, Charles A, and Susan K Laury. "Risk Aversion and Incentive Effects." *American Economic Review* 92, no. 5 (2002): 1644-55.

Huddy, Leonie, and Nadya Terkildsen. "Gender Stereotypes and the Perception of Male and Female Candidates." *American Journal of Political Science* 37, no. 1 (1993): 119-47.

Koch, Jeffrey W. "Gender Stereotypes and Citizens' Impressions of House Candidates' Ideological Orientations." *American Journal of Political Science* 46, no. 2 (2002): 453-62.

Kollock, Peter. "The Emergence of Exchange Structures: An Experimental Study of Uncertainty, Commitment, and Trust." *American Journal of Sociology* 100, no. 2 (1994): 313-45.

Kunda, Ziva, and Paul Thagard. "Forming Impressions from Stereotypes, Traits and Behaviors: A Parallel-Constraint-Satisfaction Theory." *Psychological Review* 103, no. 2 (1996): 284-308.

Levine, David. "Modeling Altruism and Spitefulness in Experiments." *Review of Economic Dynamics* 1, no. 3 (1998): 593-622.

List, John A. "The Nature and Extent of Discrimination in the Marketplace: Evidence from the Field." *Quarterly Journal of Economics* 119, no. 1 (2004): 49-89.

McCabe, Kevin A., Stephen J. Rassenti, and Vernon L. Smith. "Reciprocity, Trust, and Payoff Privacy in Extensive Form Bargaining." *Games and Economic Behavior* 24, no. 1-2 (1998): 10-24.

Mendelberg, Tali. *The Race Card : Campaign Strategy, Implicit Messages, and the Norm of Equality*. Princeton, NJ: Princeton University Press, 2001.

Nowak, Martin A., and Karl Sigmund. "Evolution of Indirect Reciprocity by Image Scoring." *Nature* 393 (1998): 573-77.

Rabin, Matthew. "Incorporating Fairness into Game Theory and Economics." *American Economic Review* 83, no. 5 (1993): 1281-302.

Ross, Lee , and Richard E. Nesbitt. *The Person and the Situation*. New York: McGraw-Hill,

1991.

Sally, David. "Two Economic Applications of Sympathy." Review of Publisher's URL. *Journal of Law, Economics, and Organization* 18, no. 2 (2002): 455-87.

Sapiro, Virginia. "Theorizing Gender in Political Psychology Research." In *Oxford Handbook on Political Psychology*, edited by David O. Sears, Leonie Huddy and Robert Jervis, 601-34. Oxford: Oxford University Press, 2003.

Scharlemann, Jorn P. W., Catherine C. Eckel, Alex Kacelnik, and Rick K. Wilson. "The Value of a Smile: Game Theory with a Human Face." *Journal of Economic Psychology* 22, no. 5 (2001): 617-40.

Sigelman, Carol K., Lee Sigelman, Barbara J. Walkosz, and Michael Nitz. "Black Candidates, White Voters: Understanding Racial Bias in Political Perceptions." *American Journal of Political Science* 39, no. 1 (1995): 243-65.

Smith, Vernon L. "The Two Faces of Adam Smith." *Southern Economic Journal* 65, no. 1 (1998): 1-19.

Snijders, Chris, and Gideon Keren. "Determinants of Trust." In *Games and Behavior: Essays in Honor of Amnon Rapoport*, edited by David V. Budescu, Ido Erev and Rami Zwick, 355-83. Mahwah, NJ: Lawrence Erlbaum Associates, Inc., 1999.

Tajfel, Henri , and J. Turner. "An Integrative Theory of Intergroup Conflict." In *The Social Psychology of Intergroup Relations*, edited by W. G. Austin and S. Worchel, 33-47. Monterey, CA: Brooks/Cole, 1979.

Tajfel, Henri, and John Turner. "Social Categorization and Social Discrimination in the Minimal Group Paradigm." In *Differentiation between Social Groups: Studies in the Social Psychology of Intergroup Relations.* London: Academic Press, Inc., 1978.

Terkildsen, Nadya. "When White Voters Evaluate Black Candidates: The Processing Implications of Candidate Skin Color, Prejudice and Self-Monitoring." *American Journal of Political Science* 37, no. 4 (1993): 1032-53.

Uslaner, Eric M. *The Moral Foundations of Trust*. Cambridge: Cambridge University Press, 2002.

Uslaner, Eric M., and Richard S. Conley. "Civic Engagement and Particularized Trust: The Ties That Bind People to Their Ethnic Communities." *American Politics Research* 31, no. 4 (2003): 331-60.

Valentino, Nicholas A., Vincent L. Hutchings, and Ismail K. White. "Cues That Matter: How Political Ads Prime Racial Attitudes During Campaigns." *American Political Science Review* 96, no. 1 (2002): 75-90.

Wrightsman, Lawrence S. "Interpersonal Trust and Attitudes toward Human Nature." In *Measures of Personality and Social Psychological Attitudes*, edited by John P. Robinson, Phillip R. Shaver and Lawrence S. Wrightsman, 373-412. San Diego: Academic Press, 1991.

Zuckerman, Marvin. *Behavioral Expression and Biosocial Bases of Sensation Seeking*. New York: Cambridge University Press, 1994.

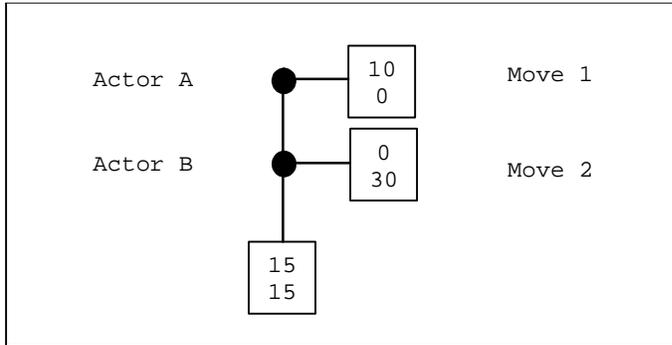Figure 1. Sample two-person, sequential move trust game.

Figure 2: A Sample Screen and Subset of the Word Pairs Used in the Evaluation Experiment

Figure 3. Sample Decision Screen for First Mover, *Photo* Condition
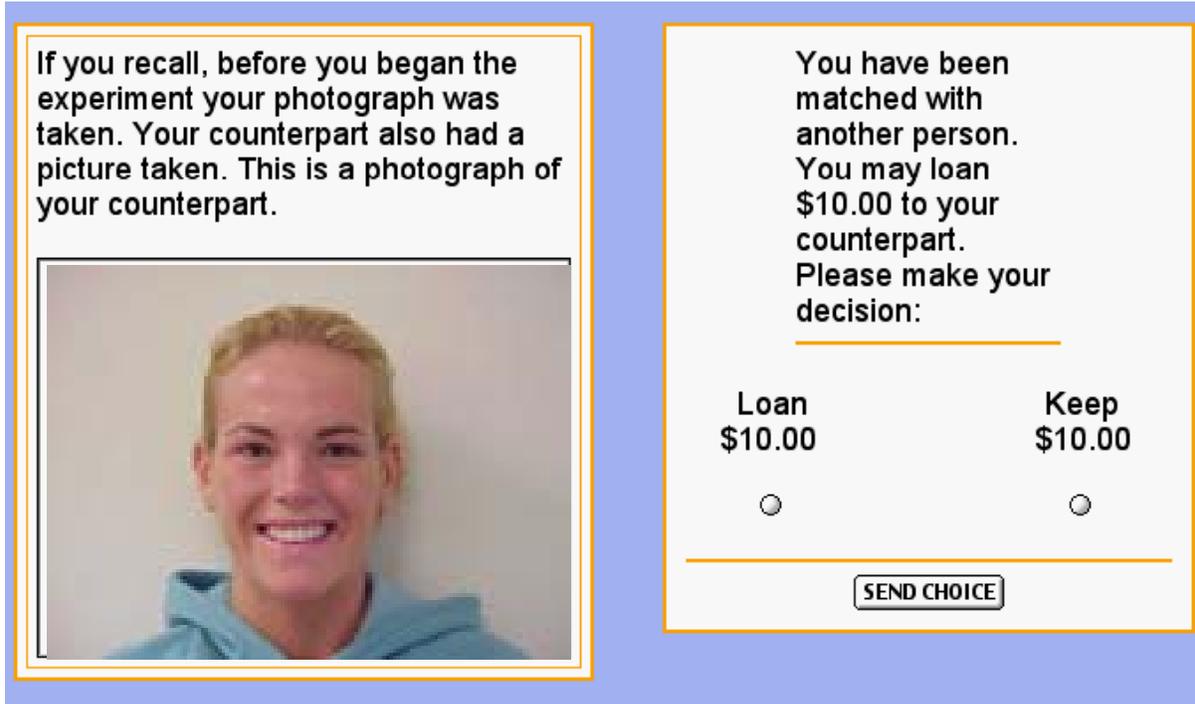

Figure 4. Bar chart of amount returned by experimental manipulation. Bars are the percentage of subjects that chose to return the amount to a first mover.
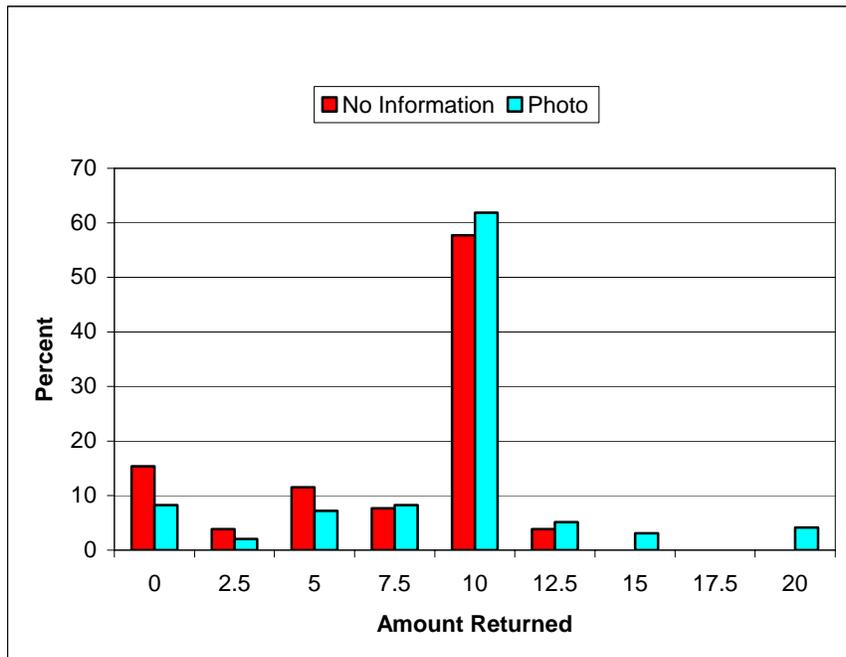
Table 1. Listing of opposite word pairs, ordered from least favorable to more favorable. Means and standard deviations (in parentheses) are given for all photos and gender and racial and ethnic categories.

	All	Male	Female	White	Black	Asian	Other
Unmotivated/ Motivated	4.06 1.21	3.87 1.24	4.31 1.11	3.97 1.19	4.17 1.25	4.33 1.25	4.05 1.17
Suspicious/ Trusting	3.71 1.28	3.46 1.28	4.06 1.20	3.72 1.27	3.89 1.30	3.60 1.32	3.53 1.24
Competitive/ Cooperative	3.37 1.40	3.18 1.35	3.64 1.41	3.42 1.37	3.30 1.47	3.13 1.43	3.43 1.35
Untrustworthy/ Trustworthy	4.01 1.13	3.80 1.14	4.28 1.05	3.98 1.11	4.10 1.16	4.06 1.14	3.91 1.13
Lazy/ Hardworking	4.03 1.23	3.86 1.26	4.26 1.14	3.89 1.20	4.09 1.26	4.48 1.20	4.16 1.19
Unintelligent/ Intelligent	4.09 1.12	4.02 1.14	4.18 1.08	3.97 1.09	4.09 1.16	4.54 1.14	4.24 1.07
Dishonest/ Honest	3.22 1.16	3.37 1.12	3.02 1.18	3.22 1.13	3.08 1.19	3.38 1.22	3.29 1.18
Disrespectful/ Respectful	4.06 1.13	3.89 1.14	4.30 1.08	4.01 1.12	4.19 1.18	4.12 1.13	4.04 1.12
Selfish/ Generous	2.85 1.15	3.04 1.17	2.60 1.08	2.92 1.13	2.78 1.21	2.67 1.12	2.80 1.13
Unattractive/ Attractive	3.36 1.25	3.20 1.21	3.58 1.26	3.34 1.24	3.55 1.23	3.16 1.27	3.33 1.24
Excitable/ Calm	3.39 1.47	3.60 1.46	3.11 1.45	3.34 1.43	3.13 1.55	3.85 1.44	3.66 1.45
Unfriendly/ Friendly	4.27 1.20	4.08 1.22	4.54 1.12	4.30 1.17	4.54 1.21	3.94 1.25	4.06 1.24
Sad/ Happy	4.20 1.18	4.04 1.17	4.42 1.16	4.20 1.15	4.44 1.23	3.98 1.19	4.02 1.15
Accepting	3.88 1.23	3.82 1.20	3.97 1.28	3.85 1.22	3.95 1.28	3.96 1.23	3.88 1.23
Irresponsible/ Dependable	3.96 1.17	3.75 1.18	4.24 1.09	3.86 1.16	4.04 1.19	4.22 1.16	4.05 1.13
Number	4560	2639	1921	2659	836	504	561

Table 2. OLS models estimating "Trustworthiness" along a 6-point scale. Higher values indicate a higher trustworthiness rating. Standard errors are in parentheses.

	Model 1 (all)	Model 2 (Males Only)	Model 3 (Females Only)	Model 4 (Caucasians Only)	Model 5 (African-Americans Only)	Model 6 (Asians Only)	Model 7 (Hispanics and Others)
Intercept	3.75*** (.044)	3.76*** (.051)	4.22*** (.055)	3.79*** (.049)	3.75*** (.073)	3.86*** (.099)	3.71*** (.086)
Image Gender 1=Female 0=Male	.470*** (.032)	--	--	.482*** (.042)	.435*** (.079)	.449*** (.101)	.570*** (.101)
Perceived African-American Photo (1=Yes 0=No)	.071 (.047)	.124* (.067)	.026 (.062)	--	--	--	--
Perceived Asian Photo (1=Yes 0=No)	.058 (.052)	.085 (.072)	.053 (.075)	--	--	--	--
Perceived Hispanic/ Other Photo (1=Yes 0=No)	-.006 (.051)	-.033 (.063)	.046 (.085)	--	--	--	--
Gender of Evaluator (1=Yes 0=No)	.108** (.051)	.106* (.061)	.111* (.065)	.101* (.059)	.201*** (.080)	.006 (.114)	.060 (.101)
African-American Evaluator (1=Yes 0=No)	.037 (.067)	-.067 (.079)	.216** (.089)	-.057 (.075)	.276** (.121)	.185 (.164)	.003 (.129)
Asian Evaluator (1=Yes 0=No)	-.183** (.081)	-.175* (.097)	-.213** (.101)	-.195** (.094)	-.225* (.124)	-.163 (.172)	-.138 (.157)
Hispanic/ Other Evaluator (1=Yes 0=No)	.059 (.094)	.150 (.114)	-.072 (.121)	-.016 (.113)	.208 (.172)	-.037 (.187)	.181 (.165)
n	4560	2639	1921	2659	836	504	561
r^2	0.05	0.01	0.02	0.05	0.06	0.04	0.02

Note: *=p<.10; **=p<.05; ***=p<.01

Table 3: Percentage of First Movers Making the Loan (Frequencies in Parentheses), Average Amount Expected to be Returned and Average Amount Returned by Second Movers (SD in parentheses)

Manipulation	% Loaning $10 (Trusting)	Amount Expected (Optimism)	Amount Returned (Reciprocity)
No Information on Counterpart	70.3% (26/37)	$7.70 ($5.78)	$7.50 ($3.93)
Photo of Counterpart	75.2% (97/129)	$8.70 ($4.66)	$9.15 ($4.05)
Total	74.1 (123/166)	$8.47 ($4.93)	$8.80 (4.07)

Table 4. First movers: percent making a trusting move (loan $10) and expectations for amount returned by characteristic of the counterpart:

	Photo	Expecations
Men	76.1 54/71	$8.45 ($5.11)
Women	74.1 43/58	$9.01 ($4.05)
Caucasian	*81.5* *66/81*	*$8.70* *($4.35)*
African-American	58.3 14/24	$8.54 ($4.79)
Asian	66.7 10/15	$8.50 ($4.80)
Hispanic	100.0 5/5	$11.00 ($5.18)
Foreign	50.0 2/4	$7.50 ($9.59)
Total Minority	*64.6* *31/48*	*$8.70* *($5.18)*

Table 5. Descriptive Statistics of Independent Variables used in Probit estimates.

Variable Definition	Mean (Std. Dev.)	Min	Max	N
Zuckerman SSS (40 item additive scale) – First Mover	20.70 (6.57)	4	36	166
Trustworthiness Scale (Wrightsman 1991, 7 item averaged scale) – First Mover	3.48 (.61)	1.71	5	166
Altruism Scale (Wrightsman 1991, 7 item averaged scale) – First Mover	3.10 (.58)	1.71	4.86	166
Altruism Scale (Wrightsman 1991, 7 item averaged scale) – Second Mover	3.17 (.55)	1.86	4.86	166
Summary Appearance measure for the First Mover	-.06 (.53)	-1.44	.98	129
Summary Appearance measure for the Second Mover	.05 (.55)	-1.55	1.41	129
Second Mover's Expectations (4-point scale from Very Unlikely to Very Likely)	2.74 (.71)	1	4	97

Table 6 : Probit regressions. Probability of a trusting move (Coefficient, (std. error))

Variable	I: No Information and Photo Treatments	II: Photo Treatment	III: Photo Treatment
Intercept	-.271 (1.029)	-.178 (1.070)	.003 (1.104)
Photo treatment	-.068 (.319)	--	--
Zuckerman SSS scale	.031 (.020)	.039* (.023)	.035 (.024)
Sex of First Mover 1=Female 0=Male	.074 (.261)	.254 (.323)	307 (.333)
African-American First Mover (1=Yes 0=No)	-.130 (.406)	-.350 (.404)	-.424 (.419)
Asian First Mover (1=Yes 0=No)	.119 (.465)	.316 (.586)	.184 (.612)
Hispanic First Mover (1=Yes 0=No)	.172 (.536)	--	--
Trustworthy Scale – First Mover	.135 (.278)	--	--
Altruism Scale – First Mover	-.523* (.286)	-.403 (.298)	-.387 (.312)
Expected Return	.204*** (.033)	.221*** (.042)	.223*** (.043)
Sex of Second Mover (1=Yes 0=No)	--	-.234 (.330)	-.494 (.367)
Perceived African-American Second Mover (1=Yes 0=No)	--	-1.019*** (.403)	-1.224*** (.429)
Perceived Asian Second Mover (1=Yes 0=No)	--	-.655 (.475)	-.856* (.493)
Summary Appearance Measure of the Counterpart	--	--	.651** (.312)
LL	-65.16	-.46.33	-44.04
Pseudo r2	0.31	.36	.39
N	166	129	129

Note: *=p<.10; **=p<.05; ***=p<.01

Table 7: Probability of a returning at least $10 once the loan is made (Probit regressions).
<<Table lists: Coefficient, (std. error)>>

Variable	I: No Information and Photo Treatments	II: Photo Treatment	Model III 1st Stage (DV Expected)	2nd Stage
Intercept	-2.423* (1.390)	-2.240 (1.664)	.500** (.208)	-2.311** (1.096)
Photo treatment	.434 (.299)	--	--	--
Sex of Second Mover 1=Female 0=Male	.166 (.265)	-.067 (.319)	--	-.064 (.310)
African-American Second Mover 1=Yes 0=No	.119 (.362)	-.097 (.420)	--	-.080 (.422)
Asian Second Mover 1=Yes 0=No	-.571 (.425)	-.770 (.501)	--	-.678 (.480)
Altruism Scale – Second Mover	.570** (.247)	.531* (.296)	--	.580** (.283)
Expect Loan	.253 (.342)	.399 (.459)	--	1.649** (.844)
Exceeded Expectations	.462 (.540)	.976 (.706)	--	.451 (.361)
Sex of First Mover 1=Female 0=Male	--	.476 (.354)	.550* (.337)	--
Perceived African-American First Mover 1=Yes 0=No	--	-.700 (.451)	-1.363*** (.404)	--
Perceived Asian First Mover 1=Yes 0=No	--	-.634 (.517)	-.105 (.497)	--
Factor1 Counterpart	--	.311 (.328)	-.048 (.297)	--
LL	-67.51	-47.09	-52.59	-49.31
Pseudo r2	.08	.15	.12	.11
N	123	97	97	97

Note: *=p<.10; **=p<.05; ***=p<.01