Exp Econ (2006) 9: 53–66
DOI 10.1007/s10683-006-4307-4

Internet cautions: Experimental games with internet partners

Catherine C. Eckel · Rick K. Wilson

Received: June 13, 2003 / Revised: May 6, 2005 / Accepted: May 12, 2005

Abstract We report the results of experiments conducted over the internet between two different laboratories. Each subject at one site is matched with a subject at another site in a trust game experiment. We investigate whether subjects believe they are really matched with another person, and suggest a methodology for ensuring that subjects' beliefs are accurate. Results show that skepticism can lead to misleading results. If subjects do not believe they are matched with a real person, they trust too much: i.e., they trust the experimenter rather than their partner.

Keywords Methodology · Experiment · Trust game · Beliefs · Internet

JEL Classification C9

1. Overview

The internet has been touted as a new and potentially valuable tool for conducting experiments. As more and more people use the internet in their daily lives, it provides a natural interface for subjects. In addition, it facilitates the guarantee of anonymity that experimentalists believe is so important in controlling the incentive structure of the experiment. However, as experimentalists move to the internet, it is important to carefully monitor subjects' beliefs. Traditionally, experimental economists have taken great care to ensure that subjects believe in the veracity of the experimental environment, instructions, and payoff structure. If subjects view the internet environment as easily manipulated, then subjects may be skeptical about

C. C. Eckel (✉)
School of Social Sciences, University of Texas at Dallas, Mail Station GR31, 2601 North Floyd Rd., Richardson, TX 75080
e-mail: eckelc@utdallas.edu

R. K. Wilson
Department of Political Science, Rice University
e-mail: rkw@rice.edu

the claims made in the instructions. In this paper we examine one element of this skepticism: whether subjects believe they are really matched with another person.

Our concern in this paper is not to compare beliefs in the lab with beliefs over the internet. As Frohlich et al. (2001) have shown, a subject's belief that his counterpart is "real" varies with the experimental environment in the lab, and a comparison of the various lab possibilities with various internet possibilities is beyond the scope of our endeavor. We compare two internet-based treatments. In the first, no information is given to subjects about their counterparts, except that they are located at another site. In the other, a procedure is introduced to reinforce subjects' perceptions that their counterparts are real. The first treatment appears to induce a high level of disbelief among the subjects, while the second substantially increases belief in a real counterpart. We show that results in a trust game differ across the two treatments, an indication that beliefs may be an important influence on behavior in internet-based experiments.

We are not the first to conduct internet experiments. Perhaps the best-known and longest-running internet experiment is the Iowa Political Stock Market (http://www.biz.uiowa.edu/iem/index.html). This site allows subjects to register and trade shares whose values depend on the outcomes of elections. (See Forsythe et al., 1992 for research results). Other, more recent efforts use the internet to conduct research that is more commonly located in traditional experimental labs. For example, Lucking-Reiley (1999) examines alternative auction mechanisms in an online experiment using "Magic" cards. However, few have directly compared internet and non-internet environments. Shavit and Sonsino (2001) conduct individual choice experiments where subjects value a series of lotteries both in a classroom and on the internet. They find that average bids and the variance of bid prices for the same set of lotteries are substantially higher on the internet. They also identify a difference in the way men and women respond to the internet environment, with men showing a greater difference between the two environments than women. Anderhub et al. (2001) also investigate individual choice tasks, using exactly the same software over the internet and in the lab. They find similar results on average, but with larger variance for subjects on the internet. Guth et al. (2003) examine differences between using a mail-back response and an internet response to the Ultimatum game. They find significant differences in the behavior of subjects, with those using the internet being more greedy and more opportunistic than those relying on the mail. Those studies offer some caution about jumping on the internet bandwagon.

The internet also is an increasingly popular source for classroom experiments. Charles Holt has compiled many such experiments which can be run from his Veconlab website (http://veconlab.econ.virginia.edu/admin.htm). Others also provide access to internet-based classroom experiments, including Aplia (http://www.aplia.com), a company founded by Paul Romer, and EconPort at the Economic Science Laboratory at the University of Arizona (http://www.econport.org:8080/econport).

The experiment reported here matches subjects in experimental laboratories at two different universities. Because subjects are seated in labs, we are able to use standard experimental control procedures. The internet environment is used to match each subject with a counterpart at the other site. This procedure has the twin advantage of providing strict control over the conduct of the experiment and ensuring anonymity for our subjects.

2. Internet experiments: Promises and problems

Use of the internet has both great promise and potential problems for experimentalists. Its primary advantage is access to a larger, more diverse subject pool. The main potential shortcoming is the loss of control over the experimental environment. Several of the more obvious promises and problems are detailed below:

Promise 1. Using the internet, it is possible to run experiments with a large number of participants. Plott (2000) argues that the number of participants that can be squeezed into a lab limits most laboratory tests of market theories. He and his colleagues have conducted large scale internet-based experiments with as many as 70 subjects at a time participating from multiple locations, confirming competitive equilibrium theory for larger markets (Asparouhova et al., 2003; Bossearts and Plott, 2004; Bossearts, Plott and Zame, 2002, 2003).

Promise 2. By using subjects at different sites simultaneously, there is the possibility of comparing decisions across different countries or cultures, including their interactions. While cross-national comparisons are now almost commonplace (Roth et al., 1991; Hayashi et al., 1999; Buchan et al., 2002), we are unaware of any studies where subjects interact in real time across national borders.

Promise 3. Most importantly, the internet provides a greater guarantee of anonymity, especially when subjects are paired across different sites. A concern in many experiments is that subjects may act as if they are engaged in repeated games, even when brought into the lab for a single decision. When subjects are recruited to a single site, they may consider the possibility of running into their counterparts during some point in their college career. Subjects at completely different sites, who are unlikely to ever come into contact with one another, are much less likely to bring those same expectations with them.

While these promises hold a great deal of potential benefit for experimentalists, they are offset by a number of potential problems.

Problem 1. Coordination of subjects can be more difficult. For many experiments, subjects must participate simultaneously in real time. While this is not important for some types of experiments (large auctions, etc.), it is crucial for negotiation and bargaining games. Asparouhova et al. (2003) solve this problem by having a prearranged start time for an experiment, and subjects log on at that time. This practice is also followed in the classroom experiments conducted by Aplia. Likewise, it is possible to allow subjects a short window of time during which decisions can be submitted, with feedback provided at a later point.

Problem 2. It may be difficult to convince subjects that the payoffs are "real" and that they will be paid. Equally problematic are the logistics of paying subjects. In the lab, the experimenter can show subjects a stack of cash and credibly promise immediate payment. Internet experiments conducted outside of a lab require subjects to believe that they will be issued a check. Alternatively, subjects may be required to register for an account into which earnings are paid; however this practice may compromise anonymity.

Problem 3. It is difficult to monitor the identity of subjects. The anonymity granted by the internet also means that people can construct artificial identities. In addition, a subject may discuss his/her decisions with other people in the room, so that the decision maker is no longer an individual, but rather a group. Both problems are a concern for test of individual differences in behavior by sex, race, age, etc.

Problem 4. Subjects may make decisions too quickly, or may be distracted. If experiments are being conducted in an individual's own environment, then there may be considerable confusion in the background—a dog barking or child crying who needs attention. A decision maker may be using drugs or alcohol, or may rush through the experiment in his haste to meet some other obligation. This loss of control over the experimental environment is likely to increase the noise in experimental data; in some cases it may also introduce a serious bias.

Problem 5. Subjects may exhibit wariness and lack of trust toward others on the internet. Several studies of on-line transactions point to the absence of trust and the impediment that this poses for electronic commerce (Resnick and Zeckhauser, 2001). This can only compound the problem of studying bargaining behavior in controlled laboratory settings.

Problem 6. Finally a subject may not believe there is another subject involved in the experiment. If subjects are paired with another person, and yet they do not believe this to be true, then it is not clear what the subjects are responding to. They might believe they are playing against the experimenter or against the computer. Absent knowing the type of beliefs there is a loss of experimental control.

It turns out that many of these problems can be handled through standard experimental design in the laboratory. That is, if subjects are brought into the lab, then it is possible for an experimenter to observe the subjects, ensure that they have common experiences (a quiet, if not sterile environment), and ensure that they are seeing the same material. However, this does not take advantage of the promise of the internet where subjects can be run across many different environments.

Running experiments where many labs at different locations are linked via the internet solves several of these problems. Experimenters at all locations can use a common protocol and can maintain control over the subjects and the experimental environment. However, running in many linked locations does not overcome the problem that subjects may not believe they are matched with a real counterpart. The question we pose is whether these beliefs are problematic in internet-based experiments where subjects are at different sites, and if so, whether they can be addressed by a relatively simple protocol change.

3. Beliefs

Subjects may behave very differently when they believe they are matched with another person rather than a computer programmed by the experimenter. Blount (1995) showed that subjects responded quite differently to ultimatum game offers depending on whether they were generated by a human counterpart or by a computer. Subjects were more likely to reject low offers from a human, in effect punishing them for their unfair behavior. An innovative double anonymous procedure introduced by Hoffman et al. (1994, 1996) paired each subject with a counterpart in another room. The authors varied their procedure to alter the degree of social distance (or anonymity), and showed that the greater the anonymity, the more selfishly subjects behaved. Frohlich et al. (2001) argue that one reason for Hoffman et al.'s results could be that with sufficient social (and physical) distance, subjects no longer believe their partner exists. They in turn find that subjects who doubt the existence of a counterpart are more likely to behave selfishly. Beliefs are correlated with the treatments, and so confound the effect of greater anonymity.

To compound the problem, some subjects may have more sensitive beliefs than others. Shavit et al. (2001) find that women and men react differently to the internet environment. While on average both groups are less risk averse on the internet than in classroom experiments, this difference is substantially more pronounced for men than for women. Indeed, men are often risk-seeking on average in their valuations of lotteries on the internet, but always risk-averse in the classroom. For women, the results differ less across environments, and the pattern of differences is mixed. This might be due to differences in their beliefs about what the experimenters tell them.[1]

From these and other experiments we can conclude that beliefs are important, and that it is sometimes difficult to convince subjects, even when they are in the same lab, that their counterparts are real. A similar situation may occur with experiments over the internet, though as far as we know this possibility has not been studied. In this paper we report results of experiments that where information about beliefs is collected, and a procedure is developed and tested to enhance subjects' beliefs in a real counterpart at another location.

4. Design

4.1. Information treatments

The experiment we report here arose from an anomalous result in a set of pilot experiments that we conducted to assess the effectiveness of a procedure to match subjects at two different universities (See Eckel and Wilson, 2003, 2004). The experiment was designed to test subjects' behavior in a variation on the standard trust game (Berg et al., 1995). We conducted two sessions with 10 and 13 pairs of subjects, respectively. Our data showed an unusually high level of trust—all but one of the 23 senders trusted—and analysis of our debriefing questions showed that a large proportion of the subjects did not believe they were matched with a real person. We took this to mean that their trust was in the experimenter, not the counterpart.

To test the hypothesis that the result might be due to differences in beliefs induced by the internet environment, we designed a new treatment. We developed a procedure that uses *group photos* to convince subjects that they are matched with real counterparts. In this procedure, the two sites exchange group photos and a pair of code words. Of course, in this treatment, as well as the *no-information* treatment, subjects were in fact paired with subjects at a remote site: No deception was used in the experiment. Three sessions were conducted with 11, 11, and 15 pairs of subjects respectively.

The data we report are for these two treatments, which vary the information given to the subjects about the sites. In our first two sessions, the *no-information* treatment, we revealed no information about the other site. Subjects were told they were paired with another subject at a different University. At the beginning of the experiment they were told to check one of two boxes on the computer indicating their location–Rice students were told to check an "R" and Virginia Tech students were told to check a "V." The names of the Universities were not provided for fear of inducing an uncontrolled for in-group/out-group effect. Following the experiment some students were curious as to the identity of the matched University, with a variety of different guesses. However, in debriefings it was also clear that subjects in this condition frequently doubted that there was another University.

In the *group photo* treatment, subjects entered the laboratory in exactly the same manner. Once subjects were seated and both sites were ready to begin, one subject was chosen at random and asked to provide a "code word" of up to five letters. This code word was transmitted to the other site, where it was printed on a poster board and photographed with the lab and participants at the other site visible in the photograph. (Faces of the persons at the other site were not shown in the photograph.) Both photographs were uploaded to the server.

[1] This pattern of results is similar to that found in Holt and Laury (2002) for high and low-stakes gamble choices. In their data, men are risk-seeking in the low-stakes data, and risk averse in the high-stakes data. If internet subjects are less likely to believe that the payoffs in the experiment are real, then they might appear more risk-seeking in the internet experiments because they perceive that the stakes are low.

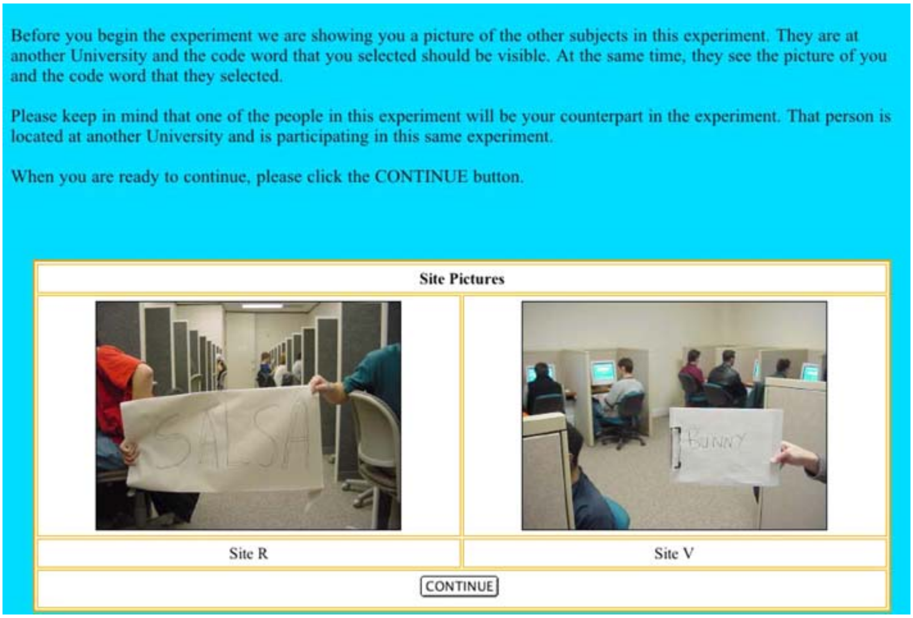

Fig. 1 Group photo screen with code words

Subjects at each site then saw, on their computer screen, the photo of the other site with their own code word clearly visible as well as the photograph of themselves holding up the other site's codeword. (See Figure 1 for an example of this screen.) This treatment was instituted to allow subjects to verify the existence of the other site and the fact that there were subjects at the other site. No other information was provided about the counterpart.[2]

4.2. Experimental components

The experiment included two main components: a trust game and an assessment of risk attitudes analyzed in Eckel and Wilson (2004). Subjects first completed a survey designed to measure attitudes toward risk, Zuckerman Sensation-Seeking Scale (SSS), form V (Zuckerman, 1994). Subjects earned 10 experimental laboratory dollars for completing the survey (the exchange rate was 2 lab dollars for each US dollar). For the second part of the experiment, subjects were randomly paired with another individual at a different site.[3] In addition subjects were randomly assigned to be either the first or second mover in a one-shot trust game. In this trust game, a variation on Berg, Dickhaut and McCabe (1995) (BDM), first movers could keep the $10 they had earned, or pass the entire $10 to their counterpart. If passed, the amount was doubled and the counterpart then decided among 9 different allocations of the

[2] Eckel and Wilson (2003, 2004) report treatments in which subjects were able to see a photograph of their counterpart. In the data reported here subjects knew nothing about their counterpart. They received no information about the sex or race of their counterpart nor did they view a photo.

[3] One site was the Lab for the Study of Human Thought and Action (LSHTA) at Virginia Tech in Blacksburg, Virginia; the other was the Behavioral Research Lab at Rice University in Houston, TX. Both are off-campus facilities, dedicated to behavioral research and not used for classroom instruction. See Eckel and Wilson (2003) for a more extensive discussion of the full design and analysis for the two-site experiments.

$20, ranging from ($0, $20) to ($20, $0) in $2.50 increments. While this game has many Nash equilibria, the single subgame perfect Nash equilibrium is for the first mover to decide to not make the loan, with the second mover receiving nothing. Previous results from variations on this game indicate that a large fraction of subjects trust, and trust is just reciprocated on average (Berg et al., 1995; Glaeser et al., 2000; Croson and Buchan, 1999; Scharleman et al., 2002)

This game differs from BDM in three ways. First, our first movers made an all-or-nothing decision, while the BDM subjects chose to send $0–$10 in increments of $1; similarly the second-mover's decision was restricted to fixed increments of $2.50 while theirs was again $1. Second, in our experiment the amount sent was doubled, while in theirs the amount sent was tripled. Finally, our decision was framed as a "loan", while theirs used neutral language. Instead of a decision to keep or send any portion of a $10 allocation, our first-mover's decision is framed as a choice of whether to make a loan of 10 experimental dollars to a counterpart. If the loan is made, the $10 is "invested" and doubled; the second mover then determines the allocation of the resulting $20 in fixed increments.[4]

Both first movers and second movers were asked to predict the actions of their counterparts. First movers recorded their prediction after their own decision was made, but before finding out the counterpart's move. Even those not making the loan were asked what the counterpart would have returned had the $10 been sent. Similarly, before being told whether the loan was offered, the second mover was asked to predict what the first mover would do. Once both subjects finished their tasks, the outcome was revealed. Subjects were then asked to type a brief answer to the question, "We are very interested in what you thought about the decision problem that you just completed. In the space below please tell us what kind of situation this problem reminds you of."

After the trust game decisions were completed, subjects finished the risk-assessment element of the experiment by making a series of risky choices. The risky decision task replicates the paired-lottery choice instrument designed by Holt and Laury (2002), but in a computerized environment.[5] Finally, subjects respond to a questionnaire that collects demographic data and measures their attitudes toward trust and altruism.

5. Results

5.1. First-movers

In the no-information treatment, 96% of subjects (22 of 23) chose to trust by making the loan, while in the photo treatment, 70% (26 of 37) made the loan. Using a binomial probability

[4] Our reasons for making these changes are elaborated in our other cited papers. The binary first-mover choice was designed in part to eliminate the complicating effect found in other studies of a positive relationship between the amount sent and the percent returned. (See Bellemare and Kröger, 2004; Ashraf, Bohnet and Piankov, 2004; Eckel and Wilson, 2005; Andreoni and Petrie, 2004). Snijders and Keren (1999) use a binary version of the trust game that is similar to the all-or-nothing decision used here by first movers. McCabe et al. (2001), and the references therein report results of discrete trust games. The decision to double rather than triple was primarily for budgetary reasons, and is not unprecedented. Glaeser et al. (2000) double the amount sent in the trust game and obtain results similar to those reported by BDM. Other recent work on the trust game appears in a special issue of the *Journal of Economic Behavior and Organization* (December 2004, 55(4)).

[5] An analysis of the risk instrument and the relationship between risk and trust can be found in Eckel and Wilson (2004), but is omitted from the present discussion.

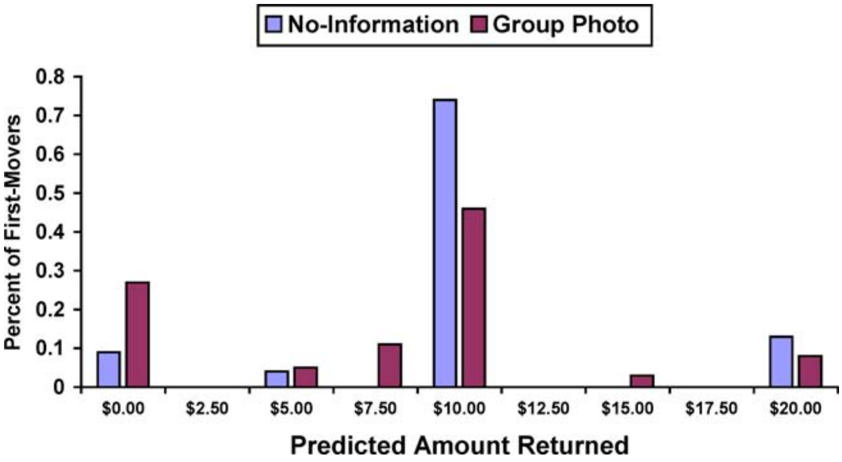

Fig. 2 Predictions by first-movers of the amount the second-mover will return (out of $20)

test the difference is statistically significant ($p = .0026$).[6] Clearly the group photo procedure affected subjects' decisions. Figure 2 shows a histogram of the predictions by the first movers of the amount their counterpart will send back. Subjects in the no-information treatment expected $10.22 to be returned on average, while subjects in the photo treatment expected $7.70 on average. The difference in mean forecasts for these two treatments approaches statistical significance ($t = 1.73$, $df = 58$, $p = .09$). Although both distributions have a strong mode at $10.00 (returning the amount loaned), the mode is weaker in the group photo distribution, which also has greater weight on $0.00 returned. A nonparametric test shows that the distributions resulting from the two treatments are statistically different (Wilcoxon $z = 2.00$, $p = .045$). Overall the difference in distributions is consistent with the observed pattern in trust behavior. Those in the no information condition are more likely to make the loan, justified by their greater expectation of receiving some return on the loan.

A histogram of the amount returned by the second movers is given in Figure 3. Subjects in the group photo treatment returned slightly less on average ($7.50) than those in the no-information treatment ($8.07), though the difference is not statistically significant ($t = .48$, $df = 58$, $p = .63$). The Wilcoxon test shows a similar result ($z = .53$, $p = .59$).

We also find no difference across the treatments concerning second mover's expectations about whether the loan will be made. In the no-information treatment, 34.8 percent of second-movers expected to receive the loan; in the group photo treatment 32.4 percent expected the loan (under a binomial probability test, $p = .851$). In the aggregate expectations affect the amount returned. If subjects do not expect their counterpart to make the loan and yet that loan is made, then an additional $2.21 on average is returned ($t = 1.72$, $p = .09$, $df = 46$).[7] However, this result is not robust across statistical test. A Wilcoxon test shows no significant difference in the distribution of amounts returned ($z = 1.44$, $p > .15$). Second movers may be somewhat more likely to reward unexpectedly generous behavior; this is explored further below.

[6] Note that similar results are obtained using the normal approximation to the binomial ($t = 3.034$, $p = .0037$), as suggested by the reviewer.

[7] Here the data is pooled across the two treatments.

Table 1 Characteristics of subjects by experimental treatment

Variable	No group photo	Group photo
Percentage of males	58.7%	56.8%
Percentage white	67.4%	67.6%
Percentage at modal family income ($50–80,000)	37.0%	31.1%
Average number of siblings	1.78	1.88
Average grade point average	3.06	3.13

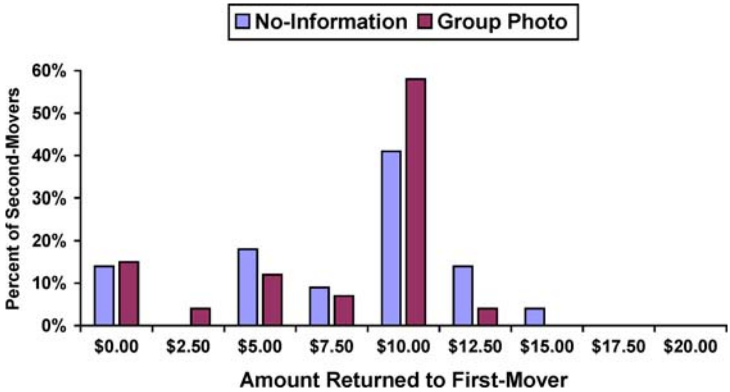

Fig. 3 Amounts returned by the 2nd mover by treatment

We next analyze the possible causes of the lower level of trust in the group photo treatment. We can rule out demographic differences as an important determinant, as the two groups of participants were drawn from the same subject pool and are virtually identical. Table 1 presents summary data on the subjects in the two treatments. There are no differences in terms of the percentage of white subjects or males across the treatments. There is a minor difference between the groups in terms of the modal family income, but this difference is not statistically different (and applying a simple χ^2 test across the 5 income categories recorded in the experiment, we cannot reject the null hypothesis that the two treatments are the same). Even the average number of siblings in the family and the average grade point average of subjects are not different across the experimental treatments. Because the *group photo* treatment was run after the *no information* treatment it is possible there could have been subject pool differences, but we do not find them.

We next examine the beliefs of the subjects about their counterparts, and the relationship between those beliefs and choices. As a manipulation check, at the end of the experiment and following all decisions, subjects were asked to respond to several statements including: "My counterpart in the experiment was a real person." The response categories were on a six-point scale ranging from Strongly Disagree to Strongly Agree in which there was no neutral category. Figure 4 presents the distribution of responses to the statement broken out by experimental treatment.[8] Subjects in the group photo treatment were significantly more likely to believe that they were facing a real counterpart. In that treatment 82.4 percent agreed that

[8] Figure 4 pools both the first mover and second mover data. Separate analysis indicates there are no differences between the two groups within treatments.

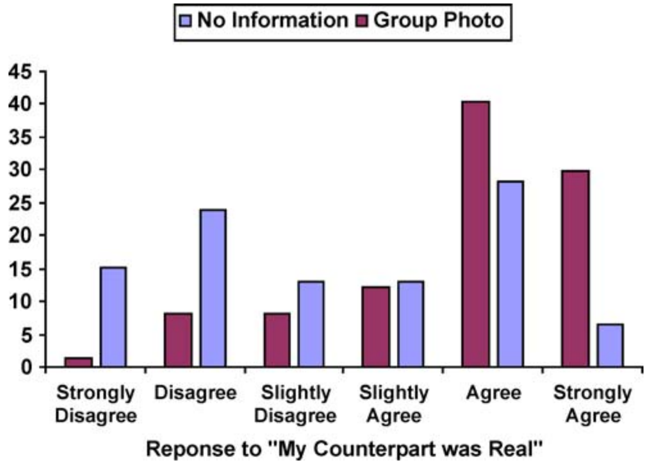

Fig. 4 Belief that counterpart is real by all subjects

they were matched with a real counterpart compared with 48 percent in the no-information treatment. Those in the no information condition were more skeptical that they were paired with another person, with over 15 percent strongly disagreeing with the statement.

The treatment has a significant effect both on trust, on first-mover's expectations about the amount returned by their counterparts, and on whether subjects believed in a real counterpart. How might beliefs have affected decision making in this game? First, beliefs may have affected expectations of first movers about the amount they were likely to receive. During informal debriefings following the experiment, several subjects in the No Information condition commented that they did not believe they were paired with another subject and that they "trusted" the experimenter not to "screw" them in the experiment. For first movers, the correlation between beliefs about whether the counterpart was real and expectations was negative (− .26) in the No Photo condition, while the same correlation is positive in the Photo condition (.16). This means that subjects who thought their counterparts were real in the first condition expected less to be returned, while the reverse was true in the photo condition. It is possible that this is due to an effect of the procedure on second-order beliefs: i.e., whether the first mover believes that the second mover believes that the first mover is real. If I as a first mover believe my counterpart is real, but am uncertain about her beliefs, then I might expect less to be returned, for example. Thus the effect of beliefs on expectations could be complex. Subjects might believe they are matched with the experimenter or another student (or a computer algorithm?); and subjects might expect more or less trustworthiness from the experimenter as compared with members of their own cohort. Since we do not have a within-subjects design, and we did not ask subjects about the trustworthiness of the experimenter, we are unable to say how these beliefs might have affected any particular subject.

We next ask whether differences in beliefs had a direct impact on behavior in the games apart from their impact on expectations. To test this, we estimate a multivariate model predicting whether a first-mover made the loan. Table 2 reports two models, with controls for the treatment, beliefs of the first mover about whether the counterpart was a real subject, and, in Model 2, the expected amount returned. (We also estimated models with demographic variables and interactions, but these were never statistically significant in any model, individually or in combination.) *Treatment* is a dummy variable equal to 1 for the Group Photo

Table 2 Probit analysis of the decision to make the loan. Dependent variable = 1 if the subject made the loan (standard errors in parentheses, probabilities in italics)

	Model 1	Model 2
Intercept	1.30	0.65
	(.70)	(0.88)
	p = .06	*p* = .46
Treatment	−1.29	−1.16
1 = Group photo	(.53)	(0.68)
0 = no information	*p* = .02	*p* = .09
Beliefs that	0.11	− 0.06
counterpart is real	(.15)	(0.18)
	p = .46	*p* = .74
Expected amount	–	0.19
returned		(0.05)
	–	*p* < .01
LL	− 26.37	− 25.25
Pseudo r^2	.122	.41
N	60	60

treatment. *Beliefs* are based on a six-point Likert scale, from strongly disagree (1) to strongly agree (6), for the statement "My counterpart is a real person." In the second model we add the *Expected Amount Returned*.

Model 1 estimates the likelihood of making a loan contingent on the treatment and beliefs about whether the counterpart is real. The Group Photo treatment is negatively related to the decision to make the loan. (Recall that 22 of 23 first movers made the loan in the No Information condition.) The coefficient for beliefs is positive but insignificant. In Model 2 we introduce the expected amount returned, which is positively and significantly related to the decision to make the loan. This variable is known to vary by treatment, but its introduction does not eliminate the treatment effect nor strengthen the effect of beliefs. Thus while there is some indication that beliefs affect the outcome of the game, the exact nature of the effect is not clear.

To reiterate a point made above, two factors seem to be at work. In the no-photo treatment, subjects may be less likely to think they are matched with a real person, but whether their trust is greater in the experimenter or in a fellow experimental subject is unknown. Beliefs may affect subjects differently depending on what they believe about the experimenter, a factor we did not measure. Thus some subjects may expect that the experimenter will not betray them, and others may hold different expectations, leading to an ambiguous or insignificant direct effect of beliefs.

We next turn to an analysis of the second-mover data and ask whether the treatment and/or beliefs affect how much of the loan was returned. Table 3 contains estimates of the amount returned; the data include only second-movers who were trusted. These estimates show a negative but insignificant treatment effect. Belief in a real counterpart carries a positive but insignificant sign: neither the treatment nor beliefs significantly affect the amount returned by the second mover. Second movers appear to be unaffected by the environment of the experiment, and respond only to whether they receive a loan, and whether that loan was expected.

The negative coefficient on expectations means that receiving a loan when one was unexpected (a surprise loan) results in higher returns. Pleasant surprises are rewarded. While this result was not evident in the original investment game study (Berg et al., 1995), the

Table 3 OLS estimate of the amount returned by the 2nd mover

	Model (both treatments)
Intercept	8.276 (1.629) $p = .00$
Treatment 1 = Group photo 0 = No information	− 1.106 (1.450) $p = .45$
Loan was expected (1 = 2nd mover expected to receive the loan)	− 2.347 (1.320) $p = .08$
Beliefs that counterpart is real (1 = 2nd mover believed in a real 1st mover)	0.117 (.450) $p = .70$
r^2	.09
N	48

positive relationship between the return rate and the amount sent has been documented in other studies. Bellemare and Kröger (2004) find this pattern in experiments conducted with a representative sample of Dutch adults, as do Ashraf, Bohnet and Piankov (2004) for college students in three countries–Russia, South Africa and the US. See also Eckel and Wilson (2005), and Andreoni and Petrie (2004) for similar results.

6. Conclusion

Our results suggest that caution should be used when translating experiments to an internet environment. When matched with another person over the internet, subjects may not believe they are playing with a real person. Assuring subjects that their counterparts are real is important, and may affect the experimental results and their interpretation.

It is not surprising that subjects would behave differently if they believe they are playing with a computer (Blount,1995). It should not be surprising that subjects may play differently if they believe that the experimenter controls their payoffs. The point to note is that when subjects do not believe the information in the instructions, the experimenter loses control of the experimental environment in ways that can affect the result of the experiment. A disbelieving subject is forming expectations, not about her counterpart, but rather about the experimenter. Such false expectations are outside the intended game, and can lead to incorrect inferences from the experimental data. The procedure we introduce avoids that possibility.

We show that our procedure alleviates the concerns that subjects have about the existence of their counterparts. It is not enough to tell them that they are paired with a real counterpart (as we continually reminded subjects in our instructions). Instead subjects need to be given a credible means of verifying the existence of their counterparts.

The behavior of second-movers does not appear to be affected by their beliefs, but rather respond only to the fact of the loan, and whether that loan was unexpected. It is also possible that the subjects think of their response as a measure of "trustworthiness," and that they are willing to incur similar costs to show their trustworthiness not only to another subject, but also to the experimenters if they do not believe their counterpart is real.

We are cautiously optimistic about the ability of experimentalists to utilize the internet to extend our subject pools and explore subject interactions while maintaining anonymity. The ability to assure anonymity for subjects, while at the same time displaying their personal photograph, is exceedingly useful in many bargaining games. The capacity to increase the scale of experiments by using multiple experimental laboratories is also promising. As with any new technology, different challenges will be posed and methodologies will have to adjust. The technique we develop leads subjects to believe that their counterparts are real. Even if this technique is not used, we find it is important to make certain that subjects believe they have a real counterpart.

Acknowledgments We thank Selhan Garip, Martin Johnson and Nancy Martorano for research assistance. This research was supported by an NSF Infrastructure Award, SES 0094800. Additional support for Eckel was provided under NSF-SES 98-19671, and for Wilson under NSF- SES 98-19943. Eckel was also supported by the John D. and Catherine T. MacArthur Foundation, Network on Preferences and Norms. Experiments were conducted in the Lab for the Study of Human Thought and Action at Virginia Tech (http://lshta.vt.edu) , and the Behavioral Research Lab at Rice University.

References

Andreoni, J., & Ragan, P. (2004). Trusting appearances and reciprocating looks: Experimental evidence on gender and race preferences. Unpublished manuscript, Georgia State University.

Anderhub V., Rudolf, M., & Carsten, S. (2001). Design and evaluation of an economic experiment via the internet. *Journal of Economic Behavior and Organization*, *46*(2), 227–247.

Ashraf, N., Iris, B., & Piankov N. (2004). Is trust a bad investment? Kennedy School, Harvard University, Working Paper.

Asparouhova, E., Peter, B., & Charles, R. P. (2003) Excess demand and equilibration in multi-security financial markets: The empirical evidence *Journal of Financial Markets*, *6*, 1–21.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity and social history. *Games and Economic Behavior*, *10*, 122–142.

Bellemare, C., & Kröger, S. (2004). On representative social capital. Tilburg University, CentER Working Paper 2004–57.

Blount, S. (1995). When social outcomes aren't fair: The effects of causal attributions on preferences. *Organizational Behavior and Human Decision Processes*, *63*(2), 131–144.

Bossaerts, P., & Plott, C. (2004). Basic principles of asset pricing theory: Evidence from large-scale experimental financial markets. *Review of Finance*, *8*(2), 135–169.

Bossaerts, P., & Plott, C., Zame, W. R. (2003). Prices and portfolio choices in financial markets: Theory and experimental evidence, CalTech Working Paper.

Bossaerts, P., Plott, C., & Zame, W. R. (2002). Prices and portfolio choices in financial markets: Econometric evidence, CalTech working paper.

Buchan, N. R., Croson, Rachel T. A., & Dawes, R. M. (2002). Swift neighbors and persistent strangers: A cross-cultural investigation of trust and reciprocity in social exchange. *American Journal of Sociology*, *108*(1), 168–206.

Burnham, T., & McCabe, K. et al. (2000). Friend-or-foe intentionality priming in an extensive form trust game. *Journal of Economic Behavior & Organization*, *43*(1), 57–73.

Croson, R., & Nancy, R. B. (1999). Gender and culture: International experimental evidence from trust games. *American Economic Review*, *89*(2), 386–391.

Eckel, C. C., & Wilson, R. K. (2003). Conditional trust: Sex, race, and facial expressions in a trust game. paper prepared for the conference on trust and institutions, Harvard University, April 24–26.

Eckel, C. C., & Wilson, R. K. (2004). Is trust a risky decision? . *Journal of Economic Behavior and Organization*, *55*(4), 447–466.

Eckel, C. C., & Wilson, R. K. (2005). Detecting trustworthiness: Does beauty confound intuition? unpublished manuscript, department of economics, Virginia Tech, March.

Forsythe, R. et al. (1992). Anatomy of an experimental political stock market. *American Economic Review*, *82*(5), 1142–1161.

Frohlich N., Oppenheimer, J., & Bernard Moore, J. (2001). Some doubts about measuring self-interest using dictator experiments: The costs of anonymity. *Journal of Economic Behavior and Organization*, *46*(3), 271–90.

Glaeser, E. L., & Laibson, D. et al. (2000). Measuring trust. *Quarterly Journal Of Economics*, *115*(3), 811–846.

Hayashi, N., Ostrom, E., Walker, J., & Yamagishi, T. (1999). Reciprocity, trust, and the sense of control - A cross-societal study. *Rationality and Society*, *11*(1), pp 27–46.

Hoffman, E., McCabe, K., Shachat, K., & Smith, V. (1994). Preference, property rights and anonymity in bargaining games. *Games and Economic Behavior*, *7*(3), 346–80.

Hoffman, E., McCabe, K., & Smith, V. (1996). 'Social distance and other-regarding behavior in dictator games.' *American Economic Review*, *86*(3), 653–60.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review.* 92(5), 1644–55.

Lucking-Reiley, D. (1999). Using field experiments to test equivalence between auction formats: Magic on the internet. *American Economic Review*, *89*(5), 1063–1080.

McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2000). A functional imaging study of cooperation in two-person reciprocal exchange. In *Proceedings of the National Academy of Science*, *97*, 3777–3781.

Plott, C. R. (2000). Markets and information gathering tools, *Southern Economic Journal*, *67*(1), 1–15.

Resnick, P., & Zeckhauser, R. (2001). Trust among strangers in internet transactions: Empirical analysis of ebay's reputation system. Draft prepared for NBER workshop on empirical studies of electronic commerce.

Roth Alvin, E., & Ockenfels, A. (2002). Last-minute bidding and the rules for ending second-price auctions: Evidence from eBay and amazon auctions on the internet. *American Economic Review*, *92*(4), 1093–1103.

Roth Alvin, E., Prasnikar, V., Mashiro, O.-F., & Shumel, Z. (1991). Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburg, and Tokyo: An experimental study. *American Economic Review*, *81*(5), pp. 1068–95.

Scharlemann, J. P. W., Eckel, C. C., Kacelnik, A., & Wilson, R. K. (2001). The value of a smile: Game theory with a human face. *Journal of Economic Psychology*, *22*(5), 617–640.

Shavit, T., Sonsino, D., & Benzion, U. (2001). A comparative study of lotteries-evaluation in class and on the web. *Journal of Economic Psychology*, *22*(4), 483–491.

Snijders, C., & Keren, G. (1999). Determinants of trust. In D. V. Budescu, I. Erev and R. Zwick (Eds.), *Games and behavior: Essays in honor of Amnon Rapoport,* Mahwah, NJ pp. 355–383.

Werner, G., Carsten, S., & Matthias S. (forthcoming) Fairness in the mail and opportunism in the internet - A newspaper experiment on ultimatum bargaining, *German Economic Review*, *4*(2), 243–265.

Zuckerman, M. (1994). *Behavioral expressions and biosocial bases of sensation seeking.* New York, Cambridge University Press.