Appendix to On the Transportability of Laboratory Results

Felix Bader, Bastian Baumeister, Roger Berger, and Marc Keuschnigg*

*Corresponding author: marc.keuschnigg@liu.se

Sociological Methods & Research

A1 Socio-demographics across Samples and Control Variables

Table A1 summarizes the different socio-demographic backgrounds of our participants as well as relevant control variables capturing properties of our design potentially affecting elicited behavior. We include these characteristics in our calculations of conditional means (see Appendix A2 for the full models). We indicate binary measures with (0,1); all other measures are continuous. We test, further, for significant differences in sample composition, focusing on the seven pairwise comparisons that are crucial to our main analysis.

Gender. Women are broadly overrepresented in both student pools in Leipzig and Munich. The nationwide sample, in which gender is one of the criteria for stratified sampling, has a balanced sex ratio. Women, however, make up only a third of the MTurk sample.

Age. On average, student participants are 22 years old, the broader population 45, and crowdworkers 33. Because age often correlates non-linearly with elicited behavior in the lab (see, for an overview, Fréchette 2016), we categorize age into intervals 18–22, 23–29, 30–45, and ≥46 for inclusion in our regressions (see Appendix A2).

Education. Participants in our student pools differ with respect to study fields. Among Leipzig students, most pursue a major in humanities (e.g., cultural studies, literature, pedagogy). In Munich, most major in other programs such as social sciences, law, or STEM fields. We particularly differentiate between students of economics and of the humanities. Prior studies have shown that the former behave more selfishly in incentivized games (Etzioni 2015). We use all other study fields as our reference category. In the nationwide sample, 60% have at least a high-school diploma and 9% are currently enrolled in a study program. Among crowdworkers, 90% have at least secondary education and 12% are currently students.

Employment. For student subjects, we code 1 if they have a side job (50% in Leipzig and 63% in Munich). In the non-student samples, we code 1 for the fully-employed. This is true for 58% in the nationwide sample and for 69% of crowdworkers.

Income. We measure monthly disposable income using the PPP$-adjusted household income divided by the square root of household size. On average, students are significantly more affluent in Munich (PPP$1,282) than in Leipzig (PPP$870). Both groups have substantially less income at their disposal than the average German (PPP$2,727). MTurk participants take a middle position with an average income of PPP$1,562. In our regressions, we include a log transformation of this variable.

Parenthood. A number of observational studies find that parents differ in prosocial behavior compared to individuals without children (see Wiepking and Bekkers 2012 for an overview). Parenthood is rare among students (1–2%). In the nationwide sample, 57% have at least one child. This rate is 43% for crowdworkers.

Experience. Prior experience with laboratory games varies considerably across studies. Both student subjects and members of the broader population on average had only taken 1–3 social experiments prior to participation. For MTurk workers, this number amounts to 66 on average.

Understanding. For each game, participants had to answer two control questions following their decisions (see Appendix A6 for examples). Monitoring the understanding of instructions we allowed for repeated submissions, each time informing participants whether their entries were correct or incorrect. Our binary indicator of misunderstanding takes the value 1 for each subject providing more than three incorrect entries in either the Dictator Game (DG) or the Trust Game (TG). 85% of those coded 1 struggled with questions on TG. The fraction of participants scoring 1 is lowest among students (15%), but relatively high in both the broader population (43%) and the MTurk sample (34%).

Surroundings during online participation. We collected self-reported data on the physical and social surrounding during online participation. Most online subjects participated from their home, although this rate is significantly smaller for the two student samples (62%) than for either the broader population (85%) or the MTurk sample (87%). Only a few reported that others were observing them while making their decisions: Rates are 9% for both students and crowdworkers, and 6% for the broader population.

Decision sequence. To account for sequence effects, we include in our regressions a binary variable indicating whether the current decision was not the focal subject's first. Note that we randomized participants to specific sequences of games and first- and second-mover roles.

The tabulation finally demonstrates effective randomization of student participants to modes of data collection (see columns "a vs c" and "b vs d"). With the exception of age (p =.026 for the Munich pool), differences in socio-demographics between lab and online participants are non-significant.

Table A1. Socio-demographics across samples.

Variable	Measure	Study 1 Lab (a) Leipzig	Study 1 Lab (b) Munich	a vs b *p*	Study 2 Online (c) Leipzig	Study 2 Online (d) Munich	c vs d *p*	a vs c *p*	b vs d *p*
Gender	Female (0,1)	0.70	0.66	.279	0.68	0.70	.690	.746	.353
Age	Years since birth	22.34	22.21	.553	22.31	21.53	.037	.920	.026
Students' major	Economics (0,1)	0.12	0.15	.209	0.10	0.12	.568	.524	.419
	Humanities (0,1)	0.47	0.37	.006	0.53	0.43	.132	.250	.227
	Other (0,1)	0.41	0.48	.061	0.37	0.44	.244	.433	.513
Employment	Side job (0,1)	0.51	0.63	.001	0.48	0.64	.009	.498	.789
Income	Monthly disposable income	879.85	1290.88	<.001	841.25	1253.13	<.001	.177	.268
Parenthood	≥1 Child (0,1)	0.02	0.01	.432	0.01	0.01	.967	.218	.606
Experience	Number of prior lab experiments	0.52	2.61	<.001	0.34	1.64	.002	.286	.062
Control questions	>3 Mistakes in DG or TG (0,1)	0.13	0.17	.082	0.12	0.16	.459	.905	.663
Online participation	At home (0,1)				0.67	0.57	.120		
	Observed by others (0,1)				0.06	0.13	.056		
Decision sequence	DG not first decision (0,1)	0.68	0.64	.278	0.66	0.68	.714	.632	.464
	TG trustor not first decision (0,1)	0.83	0.83	.775	0.86	0.86	.996	.477	.366
N		362	351		122	115			

Variable	Measure	Study 3 Nationwide (e)	c+d vs e *p*	Study 4 MTurk (f)	c+d vs f *p*	e vs f *p*
Gender	Female (0,1)	0.49	<.001	0.35	<.001	<.001
Age	Years since birth	44.61	<.001	32.98	<.001	<.001
Education	≥High-school diploma (0,1)	0.60		0.90		<.001
	Currently student (0,1)	0.09		0.12		.035
Employment	Fully employed (0,1)	0.58	.534	0.69	<.001	<.001
Income	Monthly disposable income	2727.03	<.001	1561.55	<.001	<.001
Parenthood	≥1 Child (0,1)	0.57	<.001	0.43	<.001	<.001
Experience	Number of prior lab experiments	1.18	.410	65.52	.005	.005
Control questions	>3 Mistakes in DG or TG (0,1)	0.43	<.001	0.34	<.001	<.001
Online participation	At home (0,1)	0.85	<.001	0.87	<.001	.321
	Observed by others (0,1)	0.06	.128	0.09	.970	.034
Decision sequence	DG not first decision (0,1)	0.66	.955	0.50	<.001	<.001
	TG trustor not first decision (0,1)	0.67	<.001	0.50	<.001	<.001
N		1,223		491		

Note: We report means and *p*-values from *t*-tests for between-sample differences.

A2 Correlates of Elicited Behavior

The conditional means reported in the main text are predicted values ($\hat{Y}$) for each Study obtained from pooled OLS regressions. Table A2 displays these models. The dependent variable is first-mover behavior in either the DG or TG. We allowed individual transfers in each game to be multiples of 10% of the endowment (including 0%). Hence, $Y \in [0, 100]$. Because our dependent variables are both left and right censored we also ran tobit regressions, which yield comparable results. We repeat the predicted values mentioned in the text at the top of the table. The lower part shows how elicited behavior correlates with socio-demographic characteristics and how properties of our design affect subjects' choices. The reported coefficients rest on non-experimental variation and remain descriptive, even when controlling for all known confounders.

First, there are apparent differences between games regarding the potential influence of individual characteristics: Dictator behavior appears amenable to socio-demographic influence. TG, on the other hand, remains less affected by individual characteristics. Differences in model determination (R^2) reflect this fact.

Second, we find noteworthy results for single variables. Women, although behaving similarly to men in DG, are significantly less trusting in TG (p<.001). Compared to subjects aged 18–22, older participants share considerably less in both DG and TG. For the latter game, this difference is only significant for those aged 46 and older (p=.011). With respect to students' study field, our results are consistent with prior evidence (Etzioni 2015), indicating that economics majors behave more selfishly in incentivized games: Compared to majors in other areas, they give almost 11.0 percentage points less of their endowment in DG (p<.001) and transfer 8.3 percentage points less in TG (p=.010). Students of the humanities, in contrast, give 2.7 percentage points more in DG (p=.018) than their counterparts from other disciplines. Affluent participants, in turn, are less giving: A 10% higher disposable income associates with 0.2 percentage points lower dictator allocations (p=.013). Again, there is no such association in TG. Experimental experience, finally, correlates negatively with dictator allocations (p<.001). This finding is in line with prior results suggesting that frequent participation in social experiments provides the opportunity to calibrate one's responses and to arrive at behavior more consistent with the standard economic theory (Levitt and List 2007; Rand et al. 2014). In our data, however, this association is extremely weak (with each prior participation average allocation decreases by 0.005 percentage points) and absent for prosocial behavior as measured in TG.

Third, we show how specific properties of our design correlate with participants' choices. Subjects having difficulties in understanding the instructions exhibit considerably more prosocial behavior in DG (p<.001) but not in TG (p=.676). We address the crucial issue of comprehension in Appendix A3. Here, we stress the importance of controlling for differences in understanding to attain conditional means unaffected by some participants' faulty responses. Online subjects being observed by others during participation feature higher rates of prosocial behavior than those making their decisions in isolation (p=.024 in DG, p=.046 in TG). Finally, our analyses control for the sequence of games and roles (note that this variable is a randomized treatment). Not making the focal decision at first has a significant

Table A2. Correlates of elicited behavior

Predicted means		DG		TG	
		$\hat{Y}$		$\hat{Y}$	
Study 1	Leipzig (lab)	42.296		52.379	
	Munich (lab)	36.377		48.607	
Study 2	Leipzig (online)	42.938		48.261	
	Munich (online)	35.814		48.554	
Study 3	Nationwide	47.691		58.808	
Study 4	MTurk	33.721		42.371	
		β	t	β	t
Gender	Female	0.583	.77	−5.674***	5.01
Age	18–22 (reference)				
	23–29	−4.445***	4.13	−1.639	.99
	30–45	−4.136*	2.41	−4.387	1.80
	≥46	−3.982*	2.25	−6.585*	2.54
Students' major	Other (reference)				
	Economics	−10.890***	5.23	−8.325*	2.57
	Humanities	2.695*	2.37	−1.790	.95
Education	≥High-school diploma	0.263	0.27	−0.398	.28
	Currently student	−2.721	1.31	1.482	.56
Income	log(Monthly disposable income)	−2.018*	2.50	−0.714	.76
Employment	Fully employed	−0.540	.69	−1.966	1.70
Parenthood	≥1 Child	1.811	1.68	0.329	1.22
Experience	Number of prior lab experiments	−0.005***	6.34	−0.003	1.25
Control questions	>3 Mistakes in DG or TG	4.350***	5.30	−0.473	.42
Online participation	At home	−1.429	1.34	−0.500	.31
	Observed by others	4.248*	2.26	4.413*	2.00
Decision sequence	Current decision not first	−1.575*	2.12	−4.504***	3.88
N		2,664		2,664	
R^2		0.139		0.059	

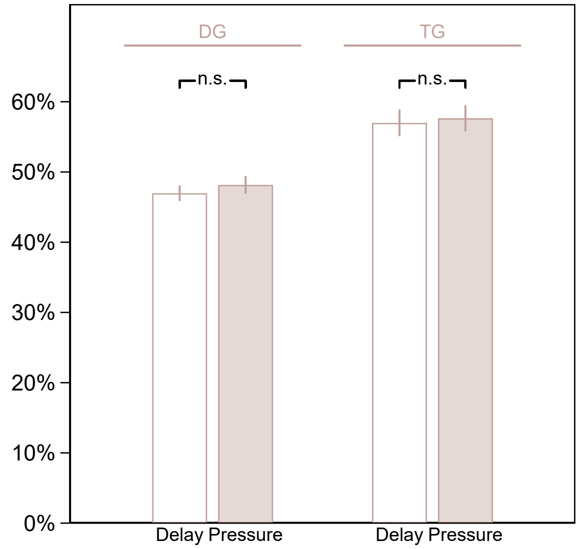
Note: OLS regressions on first-mover behavior in the Dictator Game (DG) and the Trust Game (TG), respectively. We report unstandardized coefficients and t-values calculated from robust standard errors. *** $p<.001$, ** $p<.01$, * $p<.05$.

negative effect on prosocial behavior in both DG (p=.034) and TG ($p<.001$). This effect is most pronounced for placing trust, where having had previous decisions reduces sharing by 4.5 percentage points on average.

A3 Comprehension

In our studies, we place participants into decision situations in which making mistakes and showing prosocial behavior are, on average, confounded (Andreoni 1995; Recalde, Riedl, and Vesterlund 2018): Ambiguous individual responses in the middle of the choice set indicate sharing rates of 50% of the endowment. The same holds for the average over many random responses on a 0–100% continuum. Hence, "mistakes and non-standard preferences move

Figure A1: Treatment effects of time pressure



Note: Blank bars show unconditional means of first-mover transfers under time delay in the Dictator Game (DG) and the Trust Game (TG), respectively. Shaded bars represent unconditional means under time pressure. We include 95% confidence intervals and two pairwise comparisons (t-tests). n.s.=nonsignificant.

in the same direction (away from predictions for perfectly optimizing and selfish agents)" (Fréchette 2016:472).

To test whether difficulties in understanding drive prosocial choices in our nationwide sample, we randomized participants in Study 3 into a time-limit treatment. Under *time pressure*, we asked 50% of our participants, after they had read their instructions, to take their decisions within 10 seconds in DG and within 15 seconds in TG, adjusting time pressure for the complexity of the decision situation. 17.6% of exposed subjects did not comply with our time limit; to avoid selection bias (cf., Tinghög et al. 2013), we allowed entries outside the time frame and include non-compliers in our analysis. The other half faced *time delay*, under which they could only take each decision after a waiting time of 20 seconds.

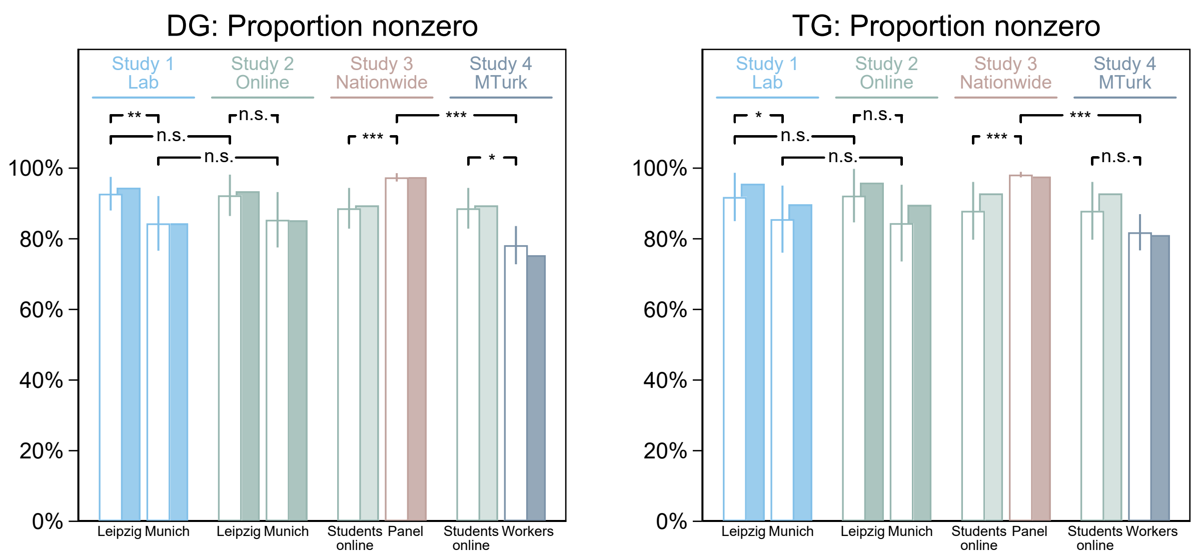
If differences in comprehension explain prosocial behavior in our broader population sample, behavior should map students' behavior more closely under the time-delay condition, permitting active deliberation and reconsideration of instructions. Under time pressure, one can plausibly assume that participants will deliberate less and make more mistakes. The latter should evoke random responses and thus move sample averages toward the middle of the choice set.

We find no statistically significant effect of the time-pressure treatment in either DG or TG (Figure A1), suggesting that differences in comprehension do not explain higher rates of prosocial behavior among the broader population.

A4 Replication: Proportions of Nonzero Transfers

Differences between our samples' mean transfers in DG and TG partly stem from the proportions of nonzero transfers. In contrast to the standard economic theory, which predicts zero transfers in both decision situations (e.g., Engel 2011, Johnson and Mislin 2011), 90.5% of our subjects shared positive amounts in DG and 92.8% invested in TG. Confirming our results, we find behavioral differences similar to our main findings when, alternatively, focusing on proportions of nonzero transfers (Figure A2). We report p-values of two-sided z-tests.

Figure A2: Proportions of Nonzero Transfers



Note: Shaded bars show unconditional proportions of nonzero first-mover transfers in the Dictator Game (DG) and the Trust Game (TG), respectively. Blank bars represent conditional proportions obtained from logistic regressions keeping underlying socio-demographics constant. We include 95% confidence intervals and seven pairwise comparisons (z-tests). *** p<.001, ** p<.01, * p<.05, n.s.=nonsignificant.

Among the Leipzig students who participated in lab sessions, 94.5% shared positive amounts in DG; this rate is 84.3% in Munich. This difference between student pools remains when controlling for socio-demographics (z=3.23, p=.001). In TG, the difference is smaller (95.6% in Leipzig and 89.7% in Munich transferred positive amounts), but—in contrast to the results on mean transfers—the difference remains significant under conditioning on socio-demographics (z=2.10, p=.036). Corroborating our finding for mean transfers, pool differences in the rates of nonzero transfers are larger in DG.

Again, we find no mode effect. Within locations, rates of nonzero transfers in online sessions are similar to and not significantly different from those in the laboratory (DG in Leipzig z=0.14, p=.891 and in Munich z=0.24, p=.810; TG in Leipzig z=0.09, p=.925 and in Munich z=0.24, p=.814).

The broader population sample, again, shows the highest rates of prosocial behavior in both decision situations—even when holding socio-demographics constant: 97.4% transfer a positive amount in DG (compared to students online 88.6%; z=3.48, p<.001) and 98.2% in TG (z=3.68, p<.001). Like before, our implementation at MTurk returns the lowest rates of prosocial behavior in both decision situations: Adjusted for differences in socio-demographics, only 78.2% transfer a positive amount in DG and 81.8% in TG. In DG, these rates are significantly lower than those obtained in both the broader population (z=9.48, p<.001) and the pooled online student sample (z=2.18, p=.029). In TG, only the difference to the broader population sample is significant (z=9.45, p<.001).

A5 Replication: Trustees and Ultimatum Game

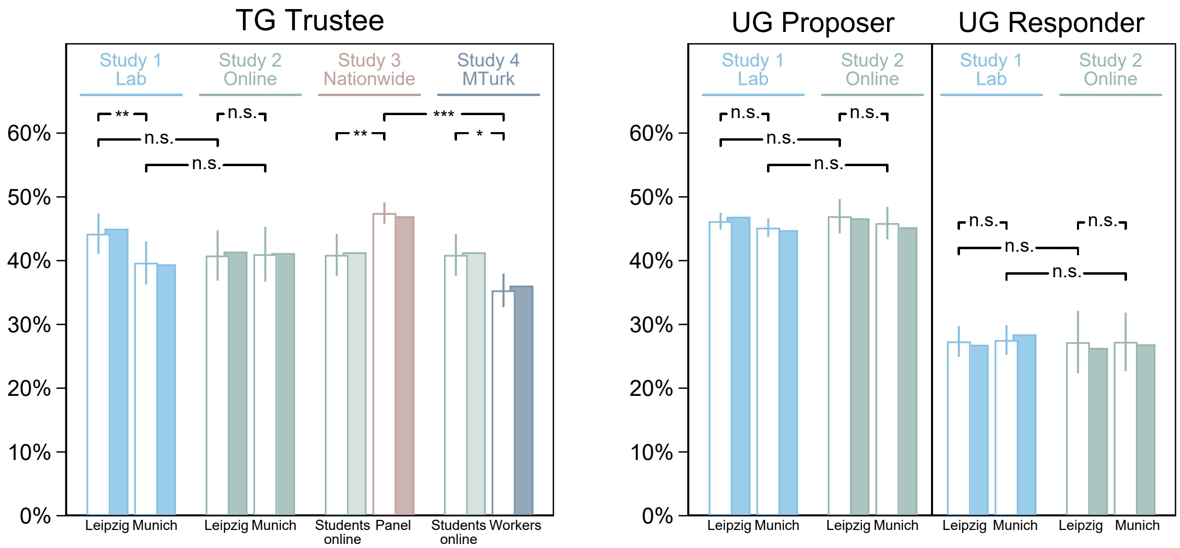
We can further substantiate our findings using second-mover behavior in TG. After receiving a monetary transfer (partly from a trustor and partly from the experimenter), a trustee can decide how much of the pie (0–100%) she passes back to the trustor. To elicit back-transfers in the TG, we employed the strategy method (Rauhut and Winter 2010), asking trustees to respond to all possible transfers. We refrained from using the standard measure of trustee behavior (second player's response to actual transfer) in order not to censor the response variable.

In Studies 1 and 2, we further ran the Ultimatum Game (UG) as a third decision situation. We dropped UG in our non-student Studies 3 and 4 to sustain online participants' attentiveness and to cap processing times to ranges typical for the respective panel or platform. In UG, a proposer receives a stake and can decide how much of the pie (0–100%) she offers to a responder. The responder can then accept (and both receive their share) or decline the offer (and both receive nothing). In contrast to DG, UG introduces fear of sanction into first movers' decision situation. Proposers need to take into account responders' option of rejecting offers that fall short of a "fair" split. Responders' acceptance or rejection of offers, in turn, measures willingness to pay for the punishment of proposers' unfair behavior. Again, we measured response behavior in UG using the strategy method: Here, we asked second movers to accept or reject each feasible offer, allowing us to determine responders' minimal acceptable offer (MAO).

Figure A3 summarizes our findings for second movers in TG and for both roles in UG. Pooled across Studies 1–4, trustees return on average 43.2% of their received transfers. In his compendium on experimental games, Camerer (2003:86-88) reports trustees as typically returning about 100% of trustors' unmultiplied investment. This, depending on the multiplication factor, amounts to 33–50% of responders' actual received transfer. Pooled across Studies 1 and 2, mean offer in UG is 45.9% of the individual endowment. Responders' MAO on average amounts to 27.4%. A meta-analysis of 75 UG-results reports mean offers of 40.4% (Oosterbeek, Sloof, and van de Kuilen 2004), with responders typically rejecting offers below 20% about half of the time (Camerer 2003:49).

Most important for our study is that the transportability of quantitative results fails for another laboratory measure of prosocial behavior: For trustees' back-transfers, four out of seven pairwise comparisons reveal significant differences. On the other hand, for the decision situations faced by both UG proposers and responders, not a single pairwise comparison is significant.

Figure A3: Quantitative results from additional decision situations.



Note: Shaded bars show unconditional means of second-mover transfers in the Trust Game (TG) and first-mover and second-mover behavior in the Ultimatum Game (UG). Blank bars represent conditional means keeping underlying socio-demographics constant. We include 95% confidence intervals and seven pairwise comparisons (t-tests). *** p<.001, ** p<.01, * p<.05, n.s.=nonsignificant.

In our two student pools (Studies 1 and 2), we find significantly different rates of trustees' back-transfers (t=2.60, p=.009). In contrast, *pool generalizability* holds for the UG: Both proposers' offers (t=1.13, p=.258) and responders' MAOs (t=0.11, p=.911) are not significantly different between locations. The cross-location gap we were able to replicate for second movers in TG is absent for UG irrespective of whether we control for socio-demographic differences across students pools.

With respect to *sample generalizability*, trustees reproduce our earlier finding on the limited transportability of quantitative results: Back-transfers in the nationwide sample differ significantly from the student-online benchmark (t=2.98, p=.003). Similarly, comparing the student-online benchmark to participants at MTurk, we found *context generalizability* violated (t=2.33, p=.020). These results further call into question the transportability of quantitative results in social science lab research.

Corroborating our favorable result on *mode generalizability*, we find no measurement effects of lab vs. online data collection (Study 2). At both locations, data collected from trustees (t=1.65, p=.100 in Leipzig; t=0.59, p=.557 in Munich), proposers (t=0.48, p=.630 in Leipzig; t=0.45, p=.653 in Munich), and responders (t=0.04, p=.969 in Leipzig; t=0.10, p=.918 in Munich) do not differ, on average, between lab and online sessions.

A6 Instructions

Dictator Game Instructions

Task 1

Please read these instructions carefully.

For this decision task, you have been randomly matched to interact with another Mechanical Turk worker.
None of you will know the other participant's worker ID.

We will provide $2 to each pair of players.

Player A decides how to divide this money between him- or herself and Player B.
Player A must allocate between $0 and the total $2 to Player B.

Player B takes home as a bonus whatever Player A allocates to him or her.
Player A gets whatever he or she does not allocate to Player B.

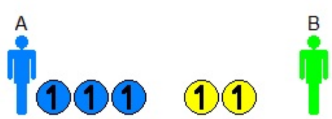
You will see an example on the next page.

Next

Dictator Game Animation

Task 1

Here is an example:



Once Player A has made a decision, the interaction is over.
Then, neither player will be able to affect the other's earnings.

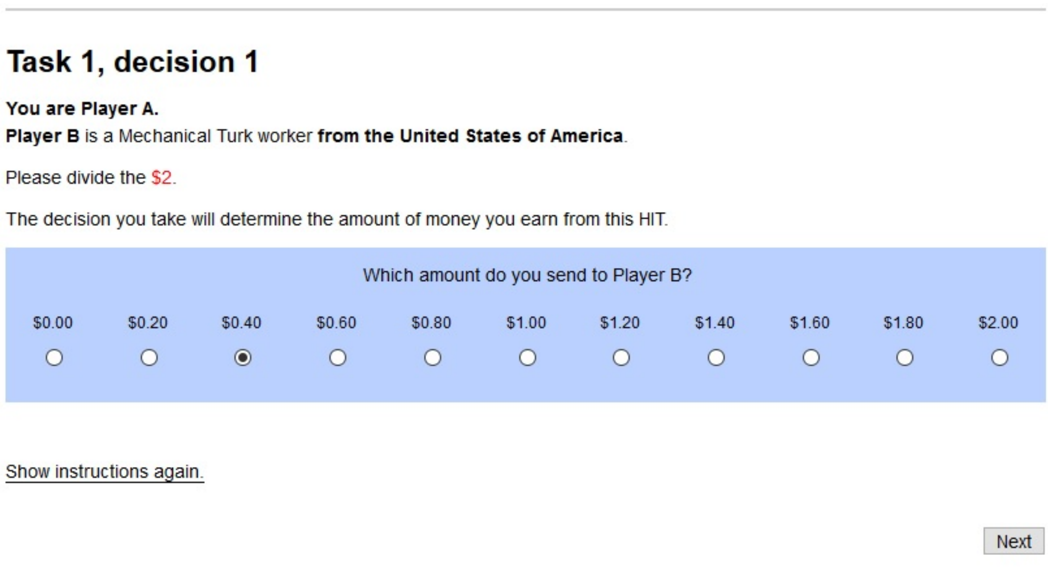
You will be Player A. You will take three decisions in this task.

Each time, you will be matched to another Mechanical Turk worker.
You will not know the other participants' worker IDs. The other participants will not know your worker ID.

Back

Next

Dictator Game Decision



Task 1, decision 1

You are Player A.
Player B is a Mechanical Turk worker **from the United States of America**.

Please divide the $2.

The decision you take will determine the amount of money you earn from this HIT.

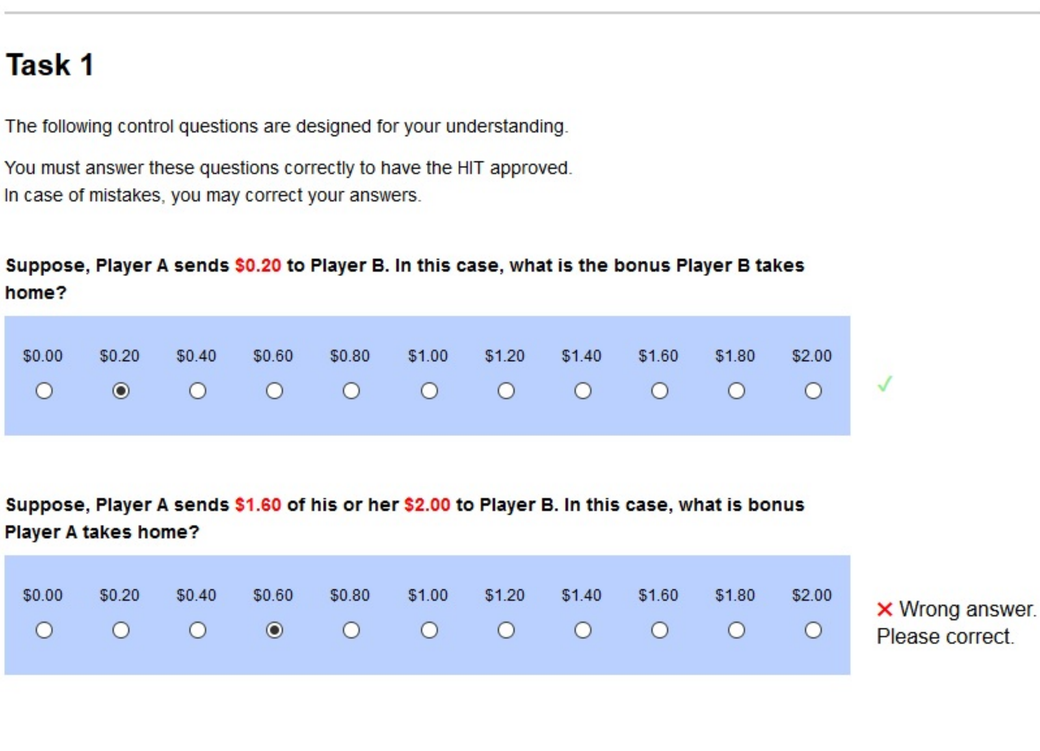
Which amount do you send to Player B?

$0.00	$0.20	$0.40	$0.60	$0.80	$1.00	$1.20	$1.40	$1.60	$1.80	$2.00
○	○	◉	○	○	○	○	○	○	○	○

Show instructions again.

Next

Dictator Game Control Questions



Task 1

The following control questions are designed for your understanding.

You must answer these questions correctly to have the HIT approved.
In case of mistakes, you may correct your answers.

Suppose, Player A sends $0.20 to Player B. In this case, what is the bonus Player B takes home?

$0.00	$0.20	$0.40	$0.60	$0.80	$1.00	$1.20	$1.40	$1.60	$1.80	$2.00
○	◉	○	○	○	○	○	○	○	○	○

✓

Suppose, Player A sends $1.60 of his or her $2.00 to Player B. In this case, what is bonus Player A takes home?

$0.00	$0.20	$0.40	$0.60	$0.80	$1.00	$1.20	$1.40	$1.60	$1.80	$2.00
○	○	○	◉	○	○	○	○	○	○	○

✕ Wrong answer. Please correct.

Trust Game Instructions

Task 2

Please read these instructions carefully.

For this decision task, you have been randomly matched to interact with another Mechanical Turk worker, different from the persons you interacted with before.
None of you will know the other participant's worker ID.

We will provide $1 to each player.

Player A decides how to divide his or her amount between him- or herself and Player B.
Player A must allocate between $0 and the total $1 to Player B.
This website then doubles the allocated amount and sends it to Player B.

Remember, Player B has (just like Player A) already received $1 from the experimenter.
In addition, Player B receives the doubled amount allocated to him or her by Player A.
Then, Player B decides how much of the doubled amount (between 0% and 100%) he or she sends back to Player A.

Player A takes home as a bonus $1 minus his or her allocation to Player B plus the amount Player B sends back.
Player B gets $1 plus the doubled amount allocated by Player A minus the amount he or she sends back to Player A.

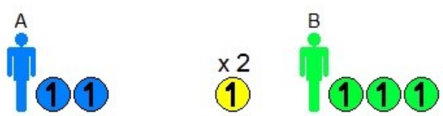
You will see an example on the next page.

Next

Trust Game Animation

Task 2

Here is an example:



Once both Players have made their decisions, the interaction is over.
Then, neither player will be able to affect the other’s earnings.

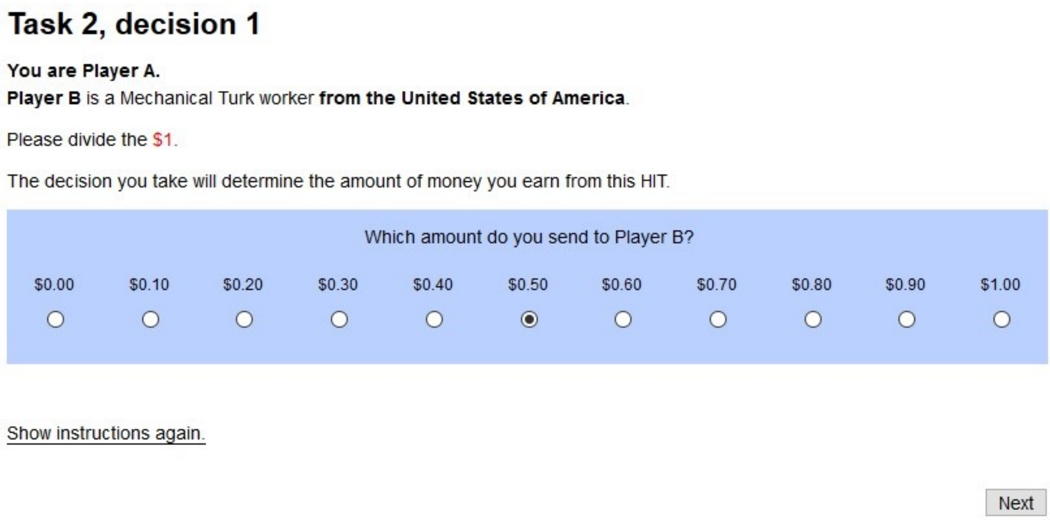
In this task, you will participate in two roles: First you will take three decisions as Player A.
Each time, you will be matched to another Mechanical Turk worker.

Then, paired with a new Mechanical Turk worker, as Player B.
You will not know the other participants' worker IDs. The other participants will not know your worker ID.

Back

Next

Trustor Decision



Task 2, decision 1

You are Player A.
Player B is a Mechanical Turk worker **from the United States of America**.

Please divide the $1.

The decision you take will determine the amount of money you earn from this HIT.

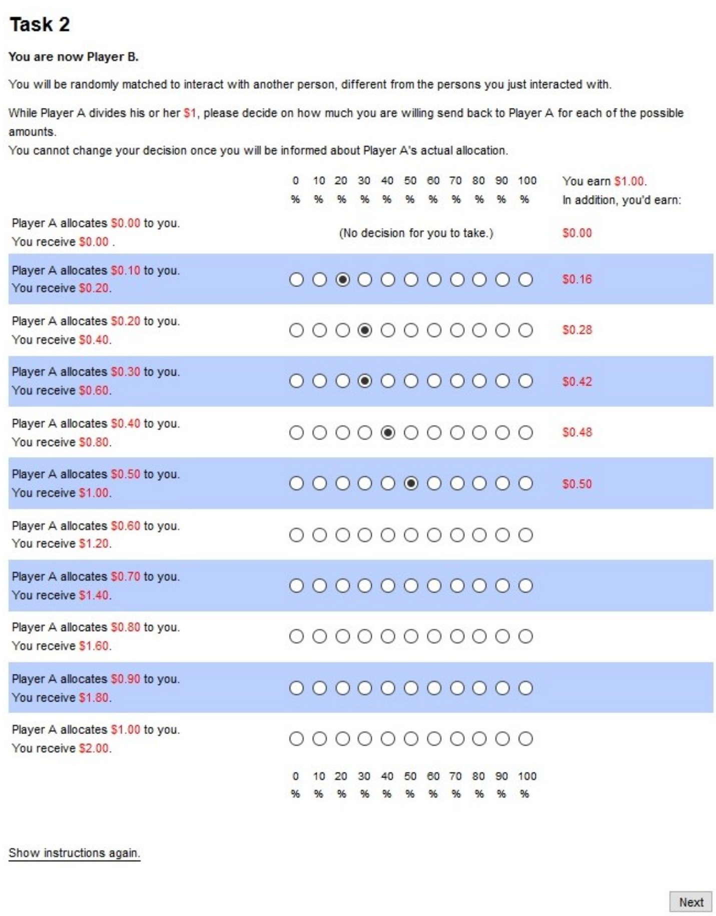
Which amount do you send to Player B?

$0.00	$0.10	$0.20	$0.30	$0.40	$0.50	$0.60	$0.70	$0.80	$0.90	$1.00

Show instructions again.

Next

Trustee Decision



Task 2

You are now Player B.

You will be randomly matched to interact with another person, different from the persons you just interacted with.

While Player A divides his or her $1, please decide on how much you are willing send back to Player A for each of the possible amounts.
You cannot change your decision once you will be informed about Player A's actual allocation.

	0 %	10 %	20 %	30 %	40 %	50 %	60 %	70 %	80 %	90 %	100 %	You earn $1.00. In addition, you'd earn:
Player A allocates $0.00 to you. You receive $0.00 .	(No decision for you to take.)											$0.00
Player A allocates $0.10 to you. You receive $0.20.												$0.16
Player A allocates $0.20 to you. You receive $0.40.												$0.28
Player A allocates $0.30 to you. You receive $0.60.												$0.42
Player A allocates $0.40 to you. You receive $0.80.												$0.48
Player A allocates $0.50 to you. You receive $1.00.												$0.50
Player A allocates $0.60 to you. You receive $1.20.												
Player A allocates $0.70 to you. You receive $1.40.												
Player A allocates $0.80 to you. You receive $1.60.												
Player A allocates $0.90 to you. You receive $1.80.												
Player A allocates $1.00 to you. You receive $2.00.												
	0 %	10 %	20 %	30 %	40 %	50 %	60 %	70 %	80 %	90 %	100 %	

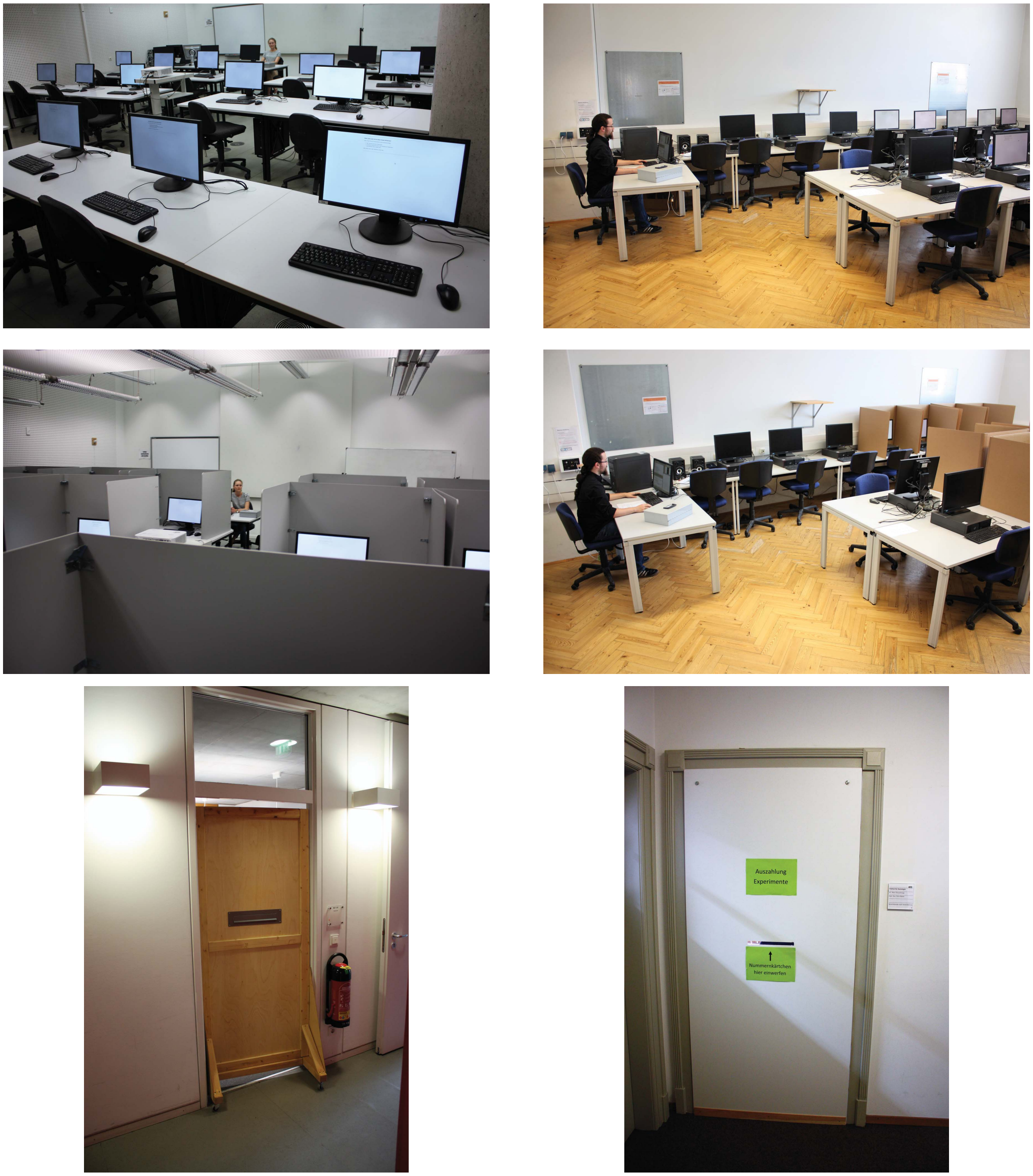
Show instructions again.

Next

A7 Anonymization procedures in the two laboratories

Leipzig

Munich



Note: The top row shows the low-anonymity condition (no blinds, payoff at the experimenter's desk). The middle row shows the blinds used to create anonymity toward other participants (standard anonymity). The bottom row shows our payment setup to create anonymity toward the experimenter (high anonymity).

References

Andreoni, James 1995. "Cooperation in Public-goods Experiments: Kindness or Confusion?" *American Economic Review* 85(4):891-904.

Camerer, Colin F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction.* New York: Sage.

Engel, Christoph 2011. "Dictator Games: A Meta Study." *Experimental Economics* 14(4):583-610.

Etzioni, Amitai 2015. "The Moral Effects of Economic Teaching." *Sociological Forum* 30(1):228-33.

Fréchette, Guillaume R. 2016. "Experimental Economics Across Subject Populations." Pp. 435-80 in *The Handbook of Experimental Economics, Vol. 2.*, edited by J. Kagel and A. Roth. Princeton, NJ: Princeton University Press.

Johnson, Noel D. and Alexandra A. Mislin 2011. "Trust Games: A Meta-analysis." *Journal of Economic Psychology* 32(5):865-89.

Levitt, Steven D. and John A. List 2007. "What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?" *Journal of Economic Perspectives* 21(2):153-74.

Oosterbeek, Hessel, Randolph Sloof, and Gijs van de Kuilen 2004. "Cultural Differences in Ultimatum Game Experiments: Evidence from a Meta-analysis." *Experimental Economics* 7(2):171-88.

Rand, David G., Alexander Peysakhovich, Gordon T. Kraft-Todd, George E. Newman, Owen Wurzbacher, Martin A. Nowak, and Joshua D. Greene 2014. "Social Heuristics Shape Intuitive Cooperation." *Nature Communications* 5:3677.

Rauhut, Heiko and Fabian Winter 2010. "A Sociological Perspective on Measuring Social Norms by Means of Strategy Method Experiments." *Social Science Research* 39(6):1181-94.

Recalde, Maria P., Arno Riedl, and Lise Vesterlund 2018. "Error Prone Inference from Response Time: The Case of Intuitive Generosity in Public-good Games." *Journal of Public Economics* 160:132-47.

Tinghög, Gustav, David Andersson, Caroline Bonn, Harald Böttiger, Camilla Josephson, Gustaf Lundgren, Daniel Västfjäll, Michael Kirchler, and Magnus Johannesson 2013. "Intuition and Cooperation Reconsidered." *Nature* 498:E1-E2.

Wiepking, Pamala and René Bekkers 2012. "Who Gives? A Literature Review of Predictors of Charitable Giving. Part Two: Gender, Family Composition and Income." *Voluntary Sector Review* 3(2):217-45.