Social norms as anchor points for trust*

Benjamin Oyebode† and Nicky Nicholls‡

March 31, 2020

Abstract

We investigate whether a social norm communication can be used to nudge decision makers towards increased trust. An experiment was conducted using a modified Berg, Dickhaut and McCabe (1995) trust game where subjects played two rounds of the game. Prior to the second round, a social norm of sending behaviour was communicated to participants. The impact of the social norm communication on trusting behaviour varied with initial transfers: for those who initially transferred amounts below this point, transfers (indicative of trust) increased. However, this impact was almost entirely offset by decreased transfers among those who had initially transferred amounts above this point, suggesting that the social norm communication acted as an anchor point towards which subsequent transfers were drawn. The result in our study was that mean trusting behaviour did not change significantly with the social norm communication. The net impact of an anchor point would be expected to vary with the distribution of behaviour prior to the social norm communication, risking a negative net impact in some cases. Our findings therefore suggest a need for caution in using broadly targeted social norms communications.

Key words: trust game, anchoring, experiment, social norms
JEL classification: A13, C24, C91, D91

*This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

†University of Pretoria

‡Corresponding Author: Department of Economics, Tukkiewerf Building, University of Pretoria, Lynnwood Road, Pretoria, 0002, RSA, email: nicky.nicholls@up.ac.za

1 Introduction

Trust has long been established as having a central role in the functioning of society, from macro settings, including societies' economic performance and ensuring that democratic governments remain viable (Fukuyama, 1995; Putnam et al., 1994) and promoting social order for democratic institutions to function (Barber, 1983); to micro settings where trust provides a social lubricant to increase cooperation between individuals (Arrow, 1974; Luhmann, 1979). Given the importance of trust at all levels of a well-functioning society, investigating ways in which levels of trust might be increased becomes an important area of study.

The concept of "nudges" entered the popular discourse with Thaler and Sunstein's best-selling book by the same title (Thaler and Sunstein, 2008). Sunstein (2014) defines nudges as "liberty preserving approaches that steer people in particular directions, but that also allow them to go their own way" (p.583). These usually take the form of small, inexpensive changes (such as changes to default rules or providing information) resulting in cost effective ways of bringing about behaviour change, often with more impact than far more expensive incentive-based interventions.

Social norms communications, where decision makers are provided with information about others' behaviours, are noted by Sunstein (2014) as an effective form of nudging. Such communications have been used with success in a variety of contexts. However, very few studies consider social norms in the context of trust decisions.

In this study we explore the impact of a social norms communication on trust. We conducted a trust game experiment (using the Berg et al. (1995) methodology), where our participants played two rounds of the game. Senders in trust games are endowed with an amount of money, from which they can choose whether to send a (zero or positive) amount to a receiver player. The sent amount is tripled by the experimenter, and the receiver can then choose to return any of the tripled amount to the sender. Positive sent amounts are seen as indicative of trust. In our experiment, the first round of the trust game had no communication on expected sent amounts, while the second round introduced a social norms communication, where respondents were told that "most people send at least 75% of their endowment".[1] Our hypothesis, based on the successful use of social norms to modify behaviour in other contexts, was that trust (measured in sent amounts in the game) would increase following this communication of positive trusting norms.

Our main finding was that, although respondents with lower initial trust levels did respond to this social norms communication by increasing the amount sent in the trust game, this increase was almost entirely offset, as respondents with higher initial trust levels decreased the amount sent following the social norms communication. Average sent amounts therefore did not increase significantly. This finding implies an important caution for policies attempting to influence behaviour in this way. If social norms communications cannot be targeted exclusively towards those with initial behaviour at the lower end of the distribution,

[1] There are a number of studies where high transfers (average above 75% of the endowment) are seen in trust games, such that the deliberately vague statement we included as our social norm would in fact be true of specific experiments (see, for example, the seminal work by Glaeser et al. (2000) with average transfer of 83%; as well as papers by Swope et al. (2008) with average transfer of 89% and Lazzarini et al. (2005) with transfers of 86% on average, *inter alia*). We selected a social norm greater than the typical average transfer seen in trust games, as we wanted to see whether transfers would increase with a higher social norm.

they risk also reducing positive behaviour in those on the higher end of the distribution. In our study the net result was a marginal and non-significant positive change. However, it is easy to imagine a situation where the negative impact on the higher end of the distribution might more than offset the positive impact on the lower end, producing a negative net impact on behaviour.

The remainder of the paper proceeds as follows: Section 2 discusses existing literature on social norms communications, particularly with reference to trust; Section 3 describes our experiment set-up and data; results are reported in Section 4; and Section 5 concludes.

2 Literature

Social norms communications have been used with success in a wide range of contexts. Examples include reducing alcohol consumption in US colleges (Wechsler et al., 2003); reducing antibiotic prescriptions from General Practitioners (Hallsworth et al., 2016); increasing willingness to purchase healthy food (Aldrovandi et al., 2015); reducing sexual violence against women (Fabiano et al., 2003); increasing female labour force participation in Saudi Arabia (Bursztyn et al., 2018) and reducing energy consumption (Ayres et al., 2013). Related to social norms, Duflo and Saez (2002) found evidence of peer effects in retirement savings outcomes; Lerner and Malmendier (2013) noted peer effects in entrepreneurial risk taking and Ahern et al. (2014) found peer effects on attitudes to risk.

Prentice and Miller (1993) argued that people make decisions about their behaviours at least partly based on their beliefs (including their beliefs about others' decisions), which might not be accurate. Social norms communications address this possible error in decision making by providing information on others' behaviour or beliefs. A few reasons have been put forward in the literature for the social norm impact on behaviour: this might be based on a belief that peer behaviour reflects private information held by peers that is relevant to payoffs (Banerjee, 1992; Ellison and Fudenberg, 1993). Alternatively, it might reflect a taste for social conformity (Festinger, 1954) or a belief that conformity is necessary to avoid loss of social status and utility (Akerlof, 1980).

Another possible explanation for conformity to social norms is that such information provides decision makers with an anchor point about normative behaviour, from which the decision maker can adjust upwards or downwards. Tversky and Kahneman (1974) introduced the idea that people make use of heuristics, or cognitive shortcuts, in making decisions, with the consequence that such decisions might suffer from systematic biases. Their research points to the tendency of decision makers, when given a number on which to "anchor" a decision, to start their decision making process at the anchor point, and then adjust away from that point. They note that adjustment is usually not sufficient to reach an accurate answer, where an objectively correct answer exists.

We propose that social norm communications function as anchor points for behaviour. After being presented with a social norm anchor point, decision makers might start their next behavioural decision at the anchor point, and adjust away (likely insufficiently, per Tversky and Kahneman, 1974) based on their preferences. The result would be that behaviour after the introduction of the anchor point would be closer to the anchor point than behaviour based on preferences alone.

Very few studies have investigated the impact of social norms on trust decisions. Ahern et al. (2014) looked at the impact of interactions with a randomly assigned peer group on trust, where trust was measured using a standard survey question based on the World Values Survey. Although they found no change in reports of trust on this survey question, it should be noted that survey questions are widely acknowledged to be less predictive of behaviour than incentivized behavioural decisions. Wei et al. (2016) and Wei et al. (2019) used a modified version of the Berg et al. (1995) trust game as an incentive compatable experiment to examine the impact of peer trust decisions on individuals' decisions on trust, noting higher trust where other players were believed to have shown trust.

The Berg et al. (1995) trust game (also referred to as an investment game) proceeds as follows: there are 2 players: a sender and a receiver. The sender is endowed with an amount of money, from which they can choose to transfer a portion (from 0 to the full amount) to the receiver. The transferred amount is tripled by the experimenter. The receiver can then choose to return some of the tripled amount received to the sender. Because of the multiplier, payoffs to the 2 players combined (social outcomes) are maximised when the sender sends the full endowment. However, the pure self-interest sub-game perfect equilibrium derived through backward induction is for the receiver to keep the full tripled sent amount, returning nothing to the sender; and for the sender, anticipating this, to send nothing. The decision of the sender to send a strictly positive proportion of the endowment (resulting in a social improvement for the pair of players) is therefore typically viewed as an indication of trust; while the decision by the receiver to return some of the tripled sent amount is viewed as an indication of trustworthiness/reciprocity. Without trust and reciprocity, the sender is best off keeping the full endowment. With trust and reciprocity, both players can be made better off. This characteristic of the game therefore fits well with the literature on the benefits of trust and reciprocity for societies.

Berg et al. (1995) noted an average sent amount of just over half (52%) of the endowment, and an average return amount of just under half of the initial endowment (31% of the tripled sent amount), suggesting evidence of trust and reciprocity. Although the sender was marginally worse off for having shared the endowment, the receiver was far better off, resulting in a clear social gain from trust and reciprocity. A meta-analysis of trust game studies conducted by Johnson and Mislin (2011) considered 162 replications of the trust/investment game, and noted that the original Berg et al. (1995) finding of positive transfers and returns is common across a wide range of variations of the game: on average, 50% of the endowment was sent and 37% of the multiplied received amount was returned.

Wei et al. (2016) modified this trust game to a series of 70 binary decisions where decision makers could decide whether to trust (by sending the full endowment to the receiver) or not to trust (by keeping the full endowment) each receiver. Receivers would have the option of keeping the full (trebled) sent amount or returning half of it. For each decision, participants were given some information about other players' decisions regarding the receiver. Specifically, they were given feedback about the number of other players (out of four) who had chosen to trust each receiver. This feedback was given as either "3 to 4 other players", "2 other players" or "0 to 1 other players". The receiver and the other players about whom information was given were hypothetical, although this was only revealed to decision makers at the end of the experiment. These authors reported significant differences in likelihood to trust based on other players' decisions.

A few other experiments have included some kind of anchor point in trust games. Gneezy et al. (2000) set up an experiment to establish whether positive transfers in trust games were based on investment motives or distributional motives. They tested this by having varying upper bounds on the amount the receiver could return (a zero upper bound would then be similar to a dictator game, where there would be no return on the investment in the sent amount). These upper bounds impacted sent amounts, with more money being sent where more repayment was feasible. Schotter and Sopher (2006) also used trust games, considering the impact of intergenerational advice on sent and returned amounts in the trust game. Players from one round of this game gave advice to players in the next round. Incentive compatibility was made possible by players being paid based on the outcomes of the games in which they participated as a decision maker; as well as a (discounted) payment based on the outcomes of the games in which they gave advice. The authors noted that less was sent when advice was given, but that decision makers sent more (on average) than the advised amount. Schotter and Sopher (2006) concluded that the advice acted as an anchor point from which decision makers adjust upwards, but (in line with Tversky and Kahneman, 1974) not enough to compensate for the very low anchor point.

Our paper contributes to this literature by testing a social norms communication as an anchor point to see how trust (and reciprocity) decisions in a trust game adjust in response to this communication.

3 Methodology

3.1 Experiment

A modified Berg et al. (1995) trust game was used for this research. The strategy method (Selten, 1967) was used, such that players played both roles in the game. Participants played two rounds of the game, allowing for within-subject analysis of the results. The first round was played with no communication other than the instructions for the game. Preceding the second round, the social norms communication was announced: participants were simply told by the experimenter that "most people send at least 75% of their endowment".[2] The communication took the form of a comment made by the experimenter to all participants in the room while the pages to complete for the second round of the game were being distributed to participants. Since no feedback was given between rounds, learning effects were minimised.

Senders were endowed with 100 South African Rand (ZAR).[3] In each round, participants were asked to play the role of sender and receiver, such that they were simultaneously participating in two games. In the game where they were the sender, they simply had to

[2]Prior to the main game, a pilot game had been tested with a smaller group of participants to get a better idea about initial transfer amounts, from which we could select an anchor point to use for our social norms communication. For the purposes of our experiment, we were interested in whether a positive social norm would increase trust (measured as sent amount in the game). To that end, we wanted to select a point above the existing levels of trust (sent amounts). The pilot game had an average transfer amount of 60%, leading to our selecting 75% as the anchor point for the main experiment.

[3]In July 2018, when the experiment was run, ZAR100 was equivalent to approximately 7.50 US dollars or 6.50 Euros.

decide on a transfer amount (in multiples of ZAR10, including the possibilities of sending nothing or the full endowment) from their ZAR100 endowment. Since actual decisions were not revealed during the game, in the role of receiver, participants were asked to state the amount they would return (any whole number including zero could be chosen) for each of the 10 possible positive tripled transfers that they could have received from the sender with whom they were paired.

In line with the ethics requirements of the University of Pretoria, participants were guaranteed anonymity in the game. Players were therefore given two random numbers to use in the game. Their own number (their sender number for one game and receiver number for another) was noted for them on the top of their answer sheet, where the Round 1 answer sheets were randomly handed out at the start of the experiment. A receiver number for the person who would receive and respond to their transfer decision was given to them in a closed envelope to ensure that they could not identify this person during the experiment. As well as preserving anonymity, this mechanism made it clear to participants that they had been paired with a real receiver (the number in their envelope) for their sender decision; and with a real sender (their number would appear in the receiver envelope given to another player), making the game credible.

Findings in experimental economics have suggested that hypothetical questions do not always result in elicitation of accurate or behaviourally predictive answers from experiment participants (see, for example, Harrison et al., 2005; Harrison, 2006; Neill et al., 1994). To make our experiment incentive compatible, participants were told that five games would be selected at random to be paid in real money according to the sender and receiver's choices in that game. Sender numbers were randomly selected, and the selected senders were asked to open their envelopes to reveal the number of the receiver in that game. In this way 10[4] of the 91 subjects were paid based on their decisions. To ensure that the random selection of paid games was credible and transparent, numbers were selected in front of the participants using a bingo cage containing the decision maker numbers of all participants. This method, whereby a subset of participants are paid for their decisions, has been widely used in experimental research (for example Kahneman et al., 1991). Participants were encouraged at the start of the experiment to make all decisions as if they were to be paid in real money, as this would mean they would be satisfied with their choices should their game be selected for payment.

3.2 Participants

Participants were undergraduate economics students at the University of Pretoria. Students interested in participating in a decision making experiment were invited to stay after class to participate in the study. Our sample comprises the 91 students who elected to participate.[5]

[4]It would be theoretically possible for the random draw to result in fewer than 10 unique subjects being paid, since random draws of five sender numbers could see a subject being paid first based on their own sender number and second by being the receiver of another selected sender player's transfer. In our game, the random draw resulted in 10 unique subjects being paid.

[5]The pilot study to establish an appropriate level for the social norms communication was conducted with a smaller group (n=50) of honours economics students.

3.3 Data

Our trust game data consist of sending decisions and receiver return decisions for Rounds 1 (no social norms communication) and 2 (following social norms communication) of the game. Recall that for return decisions, respondents reported the amount they would choose to return for each of 10 possible positive transfers (tripled by the experimenter).[6] We summarise this data by reporting the mean percentage returned in each round.[7] Table 1 presents these average amounts for each round. T-tests show that neither the average sent amount (p=0.32) nor the average percent returned (p=0.15) was statistically significantly different in Round 2 versus Round 1.

Table 1 – Mean percent sent and average percent returned by Round

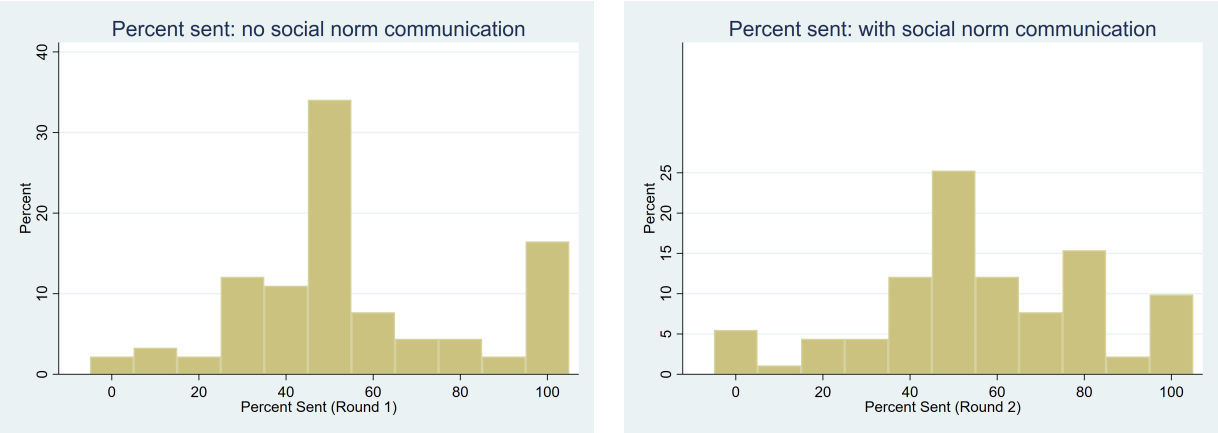
	Round 1	Round 2	t-statistic
Percent sent (out of R100)	55.49	56.59	0.476
Average percent returned	40.38	41.69	1.051
N	91	91	

Figures 1 and 2 present the distributions of this data for percent sent and average percent returned. Comparing Figure 1(a) to (b) shows some decrease in probability mass at 50% alongside increases in transfers of 60%, 70% and 80%, suggesting that the social norms communication might have acted as a positive anchor where transfers were initially below the R75 social norm. However, a concurrent decrease in transfers of 100% from Figure 1(a) to (b) is also seen, implying that the social norm communication also reduced transfers among those whose initial sent amounts were above this point.

[6] 76% of return decisions show (weakly) monotonic increases in return amount as sent amount increases.

[7] The mean percentage returned does not vary a lot by the sent amount: the smallest mean percentage returned for a transfer amount was 39.4% (for R10 sent amount) and the maximum mean percentage returned for a transfer amount was 42.6% (for R100 sent amount). Although the minimum and maximum percentages returned align with the minimum and maximum positive sent amounts, changes in the mean percentage returned do not increase monotonically. For example the second highest mean percentage returned was 42.3% (for R60 sent amount). A reviewer suggested taking an "expected value" approach to calculating the return percentage by weighting return percentages according to the distribution of senders' decisions. All results are robust to this alternative specification.

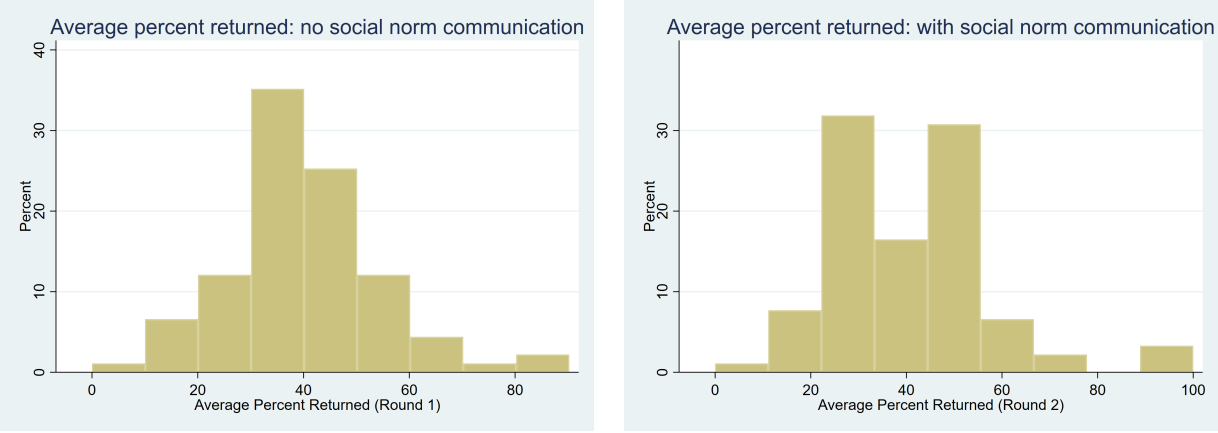
Figure 1 – Distributions of percentage sent pre- and post- social norms communication



(a) Pre social norms **(b)** Post social norms

Figure 2 shows some changes to the distribution of average percentages returned, with a shift from having the largest probability mass concentrated around 40% in Figure 2(a) to having more probability mass around 30% and 50% returned in Figure 2(b).

Figure 2 – Distributions of average percent returned pre- and post- social norms communication



(a) Pre social norms **(b)** Post social norms

Based on the data in Figure 1, we explore the possibility of a social norm communication acting as an anchor point by impacting trust decisions in different directions, depending on whether initial behaviour was above or below the anchor point.

3.4 Estimation strategy

Given some deviations from normality seen in Figures 1 and 2, we start by using non-parametric tests (Wilcoxon rank-sum tests) to confirm the t-test findings of no significant

changes in average sending (trusting) and returning (reciprocity) behaviour, as presented in Table 1.

We next consider the anchor point hypothesis based on Figure 1. We hypothesise that for participants with initial transfers below the social norms anchor point, transfers (trusting behaviour) would increase following the social norms communication; while for those with initial transfers above the social norms anchor point, transfers would decrease following the social norms communication. We first use Wilcoxon rank-sum tests to compare Round 1 and 2 decisions for these two groups of decision makers separately.

Finally, as a robustness test for the findings from the Wilcoxon tests, we estimate a very simple model. Our model shows the extent to which the post social norm communication sent amount for an individual can be predicted from the initial sent amount. To test the bidirectional anchor point hypothesis, we also include a dummy variable (HighTransfer) for whether the initial transfer (prior to the social norm communication) was above the social norm anchor point.[8] That is, we estimate the following regression:

$$PostNormTransfer_i = \beta_0 + \beta_1 InitialSent_i + \beta_2 HighTransfer_i \quad (1)$$

We do not expect a significant impact of our sent amount social norms message on the average percent returned, particularly because of our use of the strategy method of response elicitation. As our respondents reported a return amount for all 10 possible transfers, an expectation of higher transfers would be unlikely to impact this.

To confirm the extent to which return amounts after the social norm communication are impacted by return amounts prior to this communication, and by whether the initial transfer was above the social norm anchor point, we estimate the following regression, where the dummy variable is the same as in Equation 1:

$$PostNormReturn_i = \beta_0 + \beta_1 InitialReturn_i + \beta_2 HighTransfer_i \quad (2)$$

Since both sent and return percentages are limited to amounts between 0 and 100, we use Tobit regressions (for censored data) for these estimations. These are widely used in literature on trust games and other games using similarly censored data.

4 Results

Although skewness and kurtosis tests in Stata do not suggest dramatic differences from normal distributions in our data ($p>0.16$), the histograms shown in Figures 1 and 2 suggest non-normal distributions. We therefore start by comparing Round 1 and Round 2 sent amounts and average percentages returned using non-parametric tests (Wilcoxon rank-sum tests). Table 2 confirms the parametric results from the t-tests reported in Table 1.

The shifts in probability mass noted in Figures 1 and 2 suggested a possible unintended consequence of the social norms communication: while lower transfers appear to have increased with the introduction of this social norms communication, higher transfers seem to have concurrently decreased. We hypothesised that the social norm communication acted as

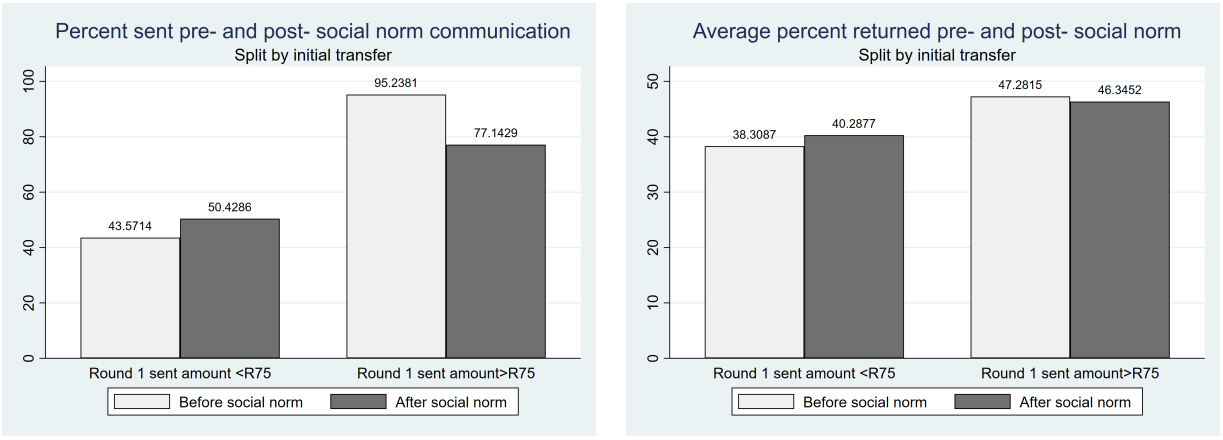
[8] Recall that the transfer cannot exactly equal the R75 social norm point, since transfers were reported in multiples of R10.

Table 2 – Mean percent sent and average percent returned by Round: Wilcoxon rank-sum test

	Round 1	Round 2	z-statistic
Percent sent (out of R100)	55.49	56.59	0.768
Average percent returned	40.38	41.69	0.407
N	91	91	

an anchor on sent amounts (trust behaviour), increasing transfers for those with initial sent amounts below the norm; but decreasing transfers for those with initial sent amounts above the norm. We investigate this possibility by considering separately decisions among participants with Round 1 sent amounts below the anchor point of R75 and those with Round 1 sent amounts above this anchor point.

Figure 3 – Mean percent sent and returned pre- and post- social norm communication: low versus high initial transfers



(a) Pre social norms

(b) Post social norms

Figure 3 shows the results of this analysis for sent and received amounts. In Figure 3(a) we note that for initial transfers below the social norm anchor point (n=70), transfers increased significantly following this anchor point (z=2.094, p=0.036). We also note that for initial transfers above the social norms anchor point (n=21), transfers decreased significantly after the social norm communication (z=-3.188, p=0.0014). Worth noting in Figure 3(a) is the relative magnitude of these changes: for those initially below the anchor point, transfers increased by R6.86 on average. For those initially above the anchor point, the decrease in transfers was far larger: R19.10 on average. In our sample there were far fewer decision makers in the second group, resulting in an overall slight (and insignificant) increase of R1.10 on average. However, had our sample been composed of more generous/trusting initial senders, we might have seen a negative net effect of our social norms communication.

In Figure 3(b) we compare decision makers' average percentage returned before and after the social norms communication. As with the sent amounts, we separate decision makers according to their initial sending behaviour, since this was directly targeted in the social norms communication. Although decision makers whose initial sent amount was below

the anchor point responded to the social norm communication with directional increases in average percentages returned to the anchor point, this increase was not significant (z=0.616, p=0.538). For those whose initial sending behaviour was higher than the anchor point, we see a very slight directional decrease in returned amounts following the social norm communication. Again, this change was not statistically significant (z=-0.252, p=0.801). The magnitudes of the changes in percentage returned are far smaller than those for sent amounts: those with lower initial sent amounts increased their average returns by less than 2% on average; while those with higher initial sent amounts decreased their average returns by less than 1% on average.

Finally, as a check for the robustness of the finding of different impacts of the social norms anchor point, depending on whether initial transfers were above or below this point, Table 3 presents the results of our Tobit regressions (with standard errors clustered at the individual level). In line with the Wilcoxon results, we note a negative and significant coefficient for high initial transfers (sent amounts before the social norms communication that were above the social norm anchor point). For return percentages the coefficient is negative but not significant. Unsurprisingly, since individual subject characteristics are common to both decisions, decisions in the second round of the game are strongly associated with decisions in the first round. We therefore see significant positive coefficients on initial sent amounts and initial return percentages.

Table 3 – Tobit Regressions: sending and returns after social norm communication

	Post Norm Sent	Post Norm Return
High Transfer	-20.032**	-2.007
	(8.814)	(3.680)
Initial Sent	0.979***	
	(0.126)	
Initial Return		0.866***
		(0.065)
Constant	7.515	7.132**
	(5.658)	(2.982)
F	34.75***	91.22***
N	91	91
Standard errors in parentheses		
* p<.10	** p<.05	*** p<.01

5 Discussion

We investigated the use of a social norms communication to increase trusting behaviour (measured in sent amounts in a Berg et al. (1995) trust game). Although we saw no significant aggregate impact of this communication on either trusting behaviour or reciprocity, we did note an interesting bidirectional impact of the social norm treatment. This communication of a benchmark for the normative sent amount appears to have acted as an anchor point for decision making. For those who initially sent amounts below this norm, we did see a significant aggregate increase in sent amounts, implying an increase in trusting behaviour on learning that most people show higher trust. However, for those who had initially shown levels of trust (sending behaviour) above the norm communicated, the anchor point reduced transfers (effectively reducing the level of trust shown by these people).

Tversky and Kahneman (1974) proposed that when people are shown an anchor point, they start at this anchor point and adjust (usually insuffiently for an accurate solution) from that point. This helps to explain our findings: the social norm communication appears to draw decision makers to evaluate their previous decisions using the social norm as a starting point, and adjusting away from the norm based on their appetite for trust. Insufficient adjustment away from an anchor point explains how Round 2 (post-anchor) trust decisions for those with low initial trust do not adjust all the way back down to Round 1 levels, resulting in increased trust for these people. Similarly, Round 2 trust for those with high initial trust does not adjust all the way back up to Round 1 levels, resulting in decreased trust for this group.

Because our sample included fewer decision makers with initial transfers above the anchor point, the net impact on trust was a small positive, but insignificant change. However, we noted that the aggregate downward adjustment for an initially high trust decision maker was greater than the aggregate upward adjustment for an initially low trust decision maker. This suggests an important caution for using social norms communications: had our sample included more initially high trust decision makers, we might have seen a negative net impact on trust from this communication. Policy makers looking to use social norms communications should be aware of this risk. Where social norms communications can be targeted exclusively to those with behaviour weaker than the normative level desired, these might be more effective in inducing desired behavioural change.

References

Ahern, K. R., Duchin, R., and Shumway, T. (2014). Peer effects in risk aversion and trust. *The Review of Financial Studies*, 27(11):3213–3240.

Akerlof, G. A. (1980). A theory of social custom, of which unemployment may be one consequence. *The Quarterly Journal of Economics*, 94(4):749–775.

Aldrovandi, S., Brown, G. D., and Wood, A. M. (2015). Social norms and rank-based nudging: Changing willingness to pay for healthy food. *Journal of Experimental Psychology: Applied*, 21(3):242–254.

Arrow, K. J. (1974). *The limits of organization*. New York: WW Norton & Company.

Ayres, I., Raseman, S., and Shih, A. (2013). Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. *The Journal of Law, Economics, and Organization*, 29(5):992–1022.

Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817.

Barber, B. (1983). *The logic and limits of trust*, volume 96. New Jersey: Rutgers University Press.

Berg, J., Dickhaut, J., and McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1):122–142.

Bursztyn, L., González, A. L., and Yanagizawa-Drott, D. (2018). *Misperceived social norms: Female labor force participation in Saudi Arabia*. Technical report, National Bureau of Economic Research.

Duflo, E. and Saez, E. (2002). Participation and investment decisions in a retirement plan: The influence of colleagues' choices. *Journal of Public Economics*, 85(1):121–148.

Ellison, G. and Fudenberg, D. (1993). Rules of thumb for social learning. *Journal of Political Economy*, 101(4):612–643.

Fabiano, P. M., Perkins, H. W., Berkowitz, A., Linkenbach, J., and Stark, C. (2003). Engaging men as social justice allies in ending violence against women: Evidence for a social norms approach. *Journal of American College Health*, 52(3):105–112.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7(2):117–140.

Fukuyama, F. (1995). *Trust: The social virtues and the creation of prosperity*, volume 99. New York: Free press New York.

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., and Soutter, C. L. (2000). Measuring trust. *The Quarterly Journal of Economics*, 115(3):811–846.

Gneezy, U., Güth, W., and Verboven, F. (2000). Presents or investments? An experimental analysis. *Journal of Economic Psychology*, 21(5):481–493.

Hallsworth, M., Chadborn, T., Sallis, A., Sanders, M., Berry, D., Greaves, F., Clements, L., and Davies, S. C. (2016). Provision of social norm feedback to high prescribers of antibiotics in general practice: A pragmatic national randomised controlled trial. *The Lancet*, 387(10029):1743–1752.

Harrison, G. W. (2006). Making choice studies incentive compatible. In Kanninen, B. J., editor, *Valuing environmental amenities using stated choice studies*, pages 67–110. New York: Springer.

Harrison, G. W., Johnson, E., McInnes, M. M., and Rutström, E. E. (2005). Risk aversion and incentive effects: Comment. *American Economic Review*, 95(3):897–901.

Johnson, N. D. and Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, 32(5):865–889.

Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic perspectives*, 5(1):193–206.

Lazzarini, S. G., Madalozzo, R., Artes, R., and Siqueira, J. d. O. (2005). Measuring trust: An experiment in Brazil. *Brazilian Journal of Applied Economics*, 9(2):153–169.

Lerner, J. and Malmendier, U. (2013). With a little help from my (random) friends: Success and failure in post-business school entrepreneurship. *The Review of Financial Studies*, 26(10):2411–2452.

Luhmann, N. (1979). *Trust and Power*. New York: Wiley.

Neill, H. R., Cummings, R. G., Ganderton, P. T., Harrison, G. W., and McGuckin, T. (1994). Hypothetical surveys and real economic commitments. *Land economics*, 70(2):145–154.

Prentice, D. A. and Miller, D. T. (1993). Pluralistic ignorance and alcohol use on campus: Some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology*, 64(2):243–256.

Putnam, R. D., Leonardi, R., and Nanetti, R. Y. (1994). *Making democracy work: Civic traditions in modern Italy*. New Jersey: Princeton University Press.

Schotter, A. and Sopher, B. (2006). Trust and trustworthiness in games: An experimental study of intergenerational advice. *Experimental Economics*, 9(2):123–145.

Selten, R. (1967). Die strategiemethode zur erforschung des eingeschränkt rationalen verhaltens im rahmen eines oligopolexperiments. In *Sauermann, H. (Ed.), Beitrage zur experimentellen Wirtshaftsforschung*, pages 136–168. Tubingen: Paul Siebeck.

Sunstein, C. R. (2014). Nudging: a very short guide. *Journal of Consumer Policy*, 37(4):583–588.

Swope, K. J., Cadigan, J., Schmitt, P. M., and Shupp, R. (2008). Personality preferences in laboratory economics experiments. *The Journal of Socio-Economics*, 37(3):998–1009.

Thaler, R. H. and Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Connecticut: Yale University Press.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157):1124–1131.

Wechsler, H., Nelson, T. E., Lee, J. E., Seibring, M., Lewis, C., and Keeling, R. P. (2003). Perception and reality: A national evaluation of social norms marketing interventions to reduce college students' heavy alcohol use. *Journal of Studies on Alcohol*, 64(4):484–494.

Wei, Z., Zhao, Z., and Zheng, Y. (2016). Moderating effects of social value orientation on the effect of social influence in prosocial decisions. *Frontiers in Psychology*, 7(952):1–9.

Wei, Z., Zhao, Z., and Zheng, Y. (2019). Following the majority: Social influence in trusting behavior. *Frontiers in Neuroscience*, 13:89–96.