Tasmanian School of Business and Economics
University of Tasmania

Discussion Paper Series N 2021-08

Do We Prefer Praise from Acquaintances or Strangers? An Experiment on Esteem-Seeking in One-Shot versus Repeated Interactions.

Paul Blacklow
University of Tasmania, Australia

Hugh Sibly
University of Tasmania, Australia

Amy Beth Corman
University of Melbourne, Australia

ISBN 978-1-922708-12-0

Do We Prefer Praise from Acquaintances or Strangers? An Experiment on Esteem-Seeking in One-Shot versus Repeated Interactions.

Paul Blacklow
Tasmanian School of Business and Economics
University of Tasmania
Hobart 7000 TAS
Australia
Paul.Blacklow@utas.edu.au

Hugh Sibly
Tasmanian School of Business and Economics
University of Tasmania
Hobart 7000 TAS
Australia
Hugh.Sibly@gmail.com

Amy Beth Corman
Faculty of Business and Economics
University of Melbourne
Parkville 3010 VIC
Australia
abcorman@unimelb.edu.au

10 December 2021

Abstract

We present an experiment designed to identify whether repeated interactions between people, relative to one-shot interactions, influences the extent they undertake costly actions to improve their social image. We expected these actions to be reduced in repeated interaction but, in contrast, we find they were increased. Gender differences are critical to our findings, with females more likely to spend some money to improve their social image than males irrespective of treatment, but those males who spend, spend significantly more when interactions are repeated. Repeating interactions, and gender, also influence the formation of feedback participants provide to one another.

JEL Codes: C92

Keywords: Social image, esteem seeking, Experimental Economics

Do we prefer praise from acquaintances or strangers? An experiment on esteem-seeking in one-shot versus repeated interactions

1. Introduction

People care about their social image, and how others react to it. The importance of social image has achieved prominence in recent years, both in the popular (e.g. Oaks 2019) and academic literature (e.g. Bursztyn and Jensen, 2017; Friedrichsen and Engelmann, 2018; Meyer and Tripodi, 2021). The increased awareness of the role of social image in shaping societal outcomes is driven in part by the rise of social media and its benefits (e.g Nisar, Prabhakar and Strakova, 2019) and harms (e.g Lukianoff and Haidt, 2018). Social image is of potential interest to economists as it can influence people's resource allocation: for example, the time spent on social media or (more traditionally) their expenditure on fashionable clothing. Identifying the underlying factors that drive people to devote their resources to influencing their social image is potentially an important economic issue.

It would be natural to expect that the frequency a person engages with others may influence the extent to which they are willing to spend to improve their image. The experimental literature demonstrates that people act more cooperatively in repeated games than in one-shot games (e.g. the prisoner's dilemma games studied by Andreoni and Miller, 1993; Cooper et al, 1996). This suggests that people may care differently about the image they project, and the response to that image, in an interaction with a stranger, compared to that of an acquaintance (e.g. the barista at a coffee shop visited once while on holiday versus the barista in their local coffee shop).

As a consequence, there will be a difference in the willingness to pay for improvements to one's social image if one is interacting with a stranger or an acquaintance. In the field, it might then be expected that there is a systematic difference in an individual's expenditure on their social image depending on the fraction of that person's interactions that are one-off or repeated. For example, how much time a person spends on social media – and the nature of views expressed – may depend on whether they interact with other individuals (knowingly) repeatedly or not (anonymously).

We refer to efforts to improve one's social image as esteem seeking. In this paper we report a laboratory experiment aimed at addressing this question, particularly if repeated interactions change the extent of esteem seeking, and how the resultant change in social image is evaluated. Esteem seeking might either be more competitive or cooperative when people interact repeatedly. The repetition of interactions increases the benefits of cooperation, and thus costly esteem seeking would be suppressed relative to a one-shot interaction. As noted above, this is what is experimentally observed in prisoners' dilemma games as well as oligopoly games (Horstmann et. al., 2018). Alternatively, repetition might result in cascading competition (a 'rat-race') resulting in higher esteem seeking relative to the one-shot interaction.

We measure the difference between esteem seeking in a one-shot and repeated interaction by pairing participants in each round of a 5-round experiment. The paring of participants (our treatments) is either the same across all 5 rounds or different in each round (being randomly assigned with no repetitions). In both treatments, each round begins with a real effort task. Participants also decide how many 'credits' they wish to purchase. Their matched partner is then shown their score which is defined by their performance on the real effort task plus the number of credits they purchase. A participant's score is the experimental analogue of their real-world social image. Differences in esteem seeking across treatments are captured by differences in performance and credit purchases across treatments.

To provide additional saliency to the observed score, matched partners provide feedback on the participant's score using a five-point Likert scale (from very poor to very good). Feedback is our experimental analogue of others' responses to real-world esteem-seeking activities: for example, the response of other people to a social media post or a new clothes purchase.

Our prior expectation (discussed in detail in section 3) is that people will behave more cooperatively in the repeated interactions than in the one-shot interaction. However, we find the opposite: there is greater esteem seeking when people are repeatedly matched than when they are randomly matched. We also find significant gender effects. Women are more likely to buy credits, however men are more likely to behave competitively, particularly in repeated interactions.

The use of credits to manipulate social image has been used previously in the literature (Charness et al 2014, Blacklow et al 2021). As noted in the results section, where they are comparable, our results are consistent with previous findings. Like Charness et al (2014) and Blacklow et al (2021), our experiment involves esteem seeking based around an activity (the real effort task) that has only private benefits. The field studies of Burstyn, et al (2018) and Kirchler and Palan (2018) similarly investigate activities with exclusively private benefits (the demand for high-status credit cards and compliments on sales service respectively). However much of the literature, including the theoretical literature (Brennan and Pettit, 2004; Benabou and Tirole, 2006; Ellingsen and Johannesson, 2008) and the experimental literature (Lacetera and Macis 2010; Friedrichsen and Engelmann, 2018), deal with status or esteem seeking using activities that have a public good dimension.

2. The Experiment

In this experiment there are 5 rounds. Participants are paired within each round. There are two between-subject treatments. In the random matching treatment (RaMT) each participant is matched with a different, randomly-selected, partner in each of the 5 rounds. In the repeated matching treatment (ReMT) each participant is matched with the same, randomly selected partner in each of the 5 rounds. Each round consists of the following activities, which are given in the order of their occurrence:

1. Participants undertake the slider task, which lasts four minutes. The number of sliders a participant places correctly in each round determines their performance in that round.
2. A participant then determines their score, which is equal to their performance on the slider task plus any credits they purchase. The participant's payment for the round is equal to their performance minus the credits bought. The payment for a correctly placed slider, and the cost of a credit, is $A0.10.
3. The score is then observed by the participant's partner. The partner then provides feedback to the participant on their score, via a 5-point Likert scale.

It is common knowledge that participants can buy credits to increase their score.

Prior to the commencement of the experiment, participants read an onscreen set of instructions that explained in detail the structure of the experiment. They were also provided with a printed version of these instructions, which are reproduced in Appendix 2 for the RaMT. Participants then undertook a quiz to ensure they understood these instructions. All participants had a practice round with the slider task prior to the commencement of the experiment.

After all five rounds were completed participants undertook a questionnaire. They were asked a series of demographic questions (age, gender, citizenship, ethnicity, enrolment degree). They were also asked a series of questions on their psychological makeup: a 15-question survey measuring the "big 5" personality traits (extraversion, agreeableness, conscientiousness, neuroticism, openness) and a 10-question survey measuring locus of control.[1] These two questionnaires were based on questions used in the 2004 survey of the German Socio-Economic Panel.

The 264 (134 in the RaMT and 130 in the ReMT) participants were students from the University of Melbourne recruited randomly using ORSEE (Greiner, 2004) from a pool of more than 6000 volunteers from a range of academic backgrounds. All participants received a turn-up payment of $A10. The experiment was conducted in the Experimental Economics Laboratory at the University of Melbourne between May-July 2016 and December 2019 and programmed using z-Tree (Fischbacher, 2007).

3. Conjectures

The RaMT, being a series of one-shot interactions, plays the role of our benchmark treatment. We are interested in determining how repeated interactions modify participants' behaviour relative to this benchmark.

First, consider behaviour in a one-shot interaction between two individuals. Suppose these individuals evaluate social image/esteem, at least in part, with reference to that of the other.

[1] A description and discussion of the big 5 personality traits is given by McCrae and John (1992). Lefcourt (1991) provides a discussion and review of locus of control measures.

In this event, esteem seeking has a negative externality. If one person raises effort or buys credits to improve their social image, this reduces the value of others' social image/esteem. As a result, esteem seeking between a pair of individuals could be modelled as a prisoner's dilemma – with the resultant equilibrium level of esteem seeking inefficiently high. In particular, it would be expected that participants may buy credits or increase performance beyond its optimal level, i.e., the level where the marginal cost of effort is equal to the marginal utility of income.

Repeated interactions between the two individuals may modify their behaviour in two ways. The first possibility is that the two participants may behave cooperatively. That is, they both tacitly 'agree' to limit their esteem-seeking behaviour to reduce the cost associated with excessive esteem seeking. In this event, we might expect low or zero credits to be purchased, and performance on the real effort task to be limited to the optimal level of performance. Cooperative participants in repeated interactions may also tacitly agree to give better feedback to each other than they would in a one-shot interaction.

A second possibility is that the two participants may behave competitively. In this scenario, a participant would attempt to raise their relative score in a round by raising their effort and credit purchases. This may induce their partner to do the same in the next round and so on. As a result, participants would be drawn into a `rat race'. In this event, it would be expected that higher credit purchases and higher performance would occur than would occur in the one-shot interaction. Competitive participants may also reduce their feedback to each other in order to reduce the perception of the performance of their rival in their `rat race'.

The experimental literature suggests a prediction on which of these two hypotheses will be supported by our experimental findings. We noted above the analogy between esteem seeking between two participants and the prisoner's dilemma. An extensive literature on the finitely repeated prisoners' dilemma indicates that a significant fraction of people behave cooperatively, particularly in the early rounds of the game. Similarly, Grimalda et al (2016) find that social image concerns can increase cooperation. Additionally, tacit cooperation is observed in a wide variety of findings in oligopoly experiments (see Horstmann et al, 2019, for a recent meta-analysis). Thus we adopt the following conjecture:

> Conjecture 1: When esteem-seeking, participants in the ReMT behave more cooperatively than in the RaMT: i) the proportion who buy credits and their average credit purchase is lower in the ReMT than the RaMT, and ii) Average performance is lower in the ReMT than the RaMT.

If conjectures 1 holds then the average score will be lower in the ReMT than the RaMT.

The literature also suggests that females are less competitive than males (Niederle and Vesterlund, 2007). In our experiment, we might expect that females are less motivated to achieve a high performance and purchase credits to create a high score for appearance's sake. Further, Jones and Linardi (2014), find that women are more motivated than men to choose a social signal which makes them appear average. If this were the case in our

experiment, then females would be less motivated than males to stand out by achieving a high score. With these previous findings in mind, we conjecture:

> Conjecture 2: i) The proportion who buy credits and their credit purchase is higher for males than females. ii) Performance is higher for males than females.

If people care about the feedback they receive, in each round both matched participants could benefit by 'tacitly colluding' to provide good feedback to one another. It would be expected that if participants behaved cooperatively by limiting esteem-seeking in repeated interactions, then they would also cooperate by offering better feedback in the ReMT than the RaMT. Correspondingly, if we saw more extensive esteem seeking in the ReMT than the RaMT, we would expect to see worse feedback in the ReMT than the RaMT. We thus adopt the following conjecture:

> Conjecture 3. Participants strategically provide feedback in a way that that matches their behaviour.

If conjecture 1 holds, then participant feedback will be more positive in the ReMT than the RaMT. Furthermore, if conjecture 2 holds, feedback from males would be more competitive than from females, thus it is less positive for males than females.

4. Results

Of the 134 participants in the RaMT, 57% were female, and of the 130 participants in the ReMT, 54% were female. One participant in RaMT did not identify themselves as male or female. This participant is not included in the results below as (i) one representative of a group is insufficient to provide any statistical insight into the group's behaviour and (ii) to maintain the confidentiality of that participant.

First consider the purchase of credits, which is a key indicator that a participant is esteem seeking. Table 1 provides the number of participants who purchase credits in at least one round and at least two rounds. The proportion who buy credits in at least one round is greater in the RaMT than in the ReMT ($p=0.061$)[2], and this is the case for both males and females ($p=0.126$ and $p=0.155$). The proportion who buy credits in only one round is significantly greater for in the RaMT than in the ReMT ($p=0.003$) across all participants and for females ($p=0.008$). The proportion of males who buy in only one round is also larger in RaMT than in the ReMT ($p=0.074$).

Further, using Pearson's chi-squared test for independence, we reject the hypothesis that for all participants the number of times credits are purchased is independent of matching

[2] All tests of all tests of "lesser/greater" or "lower/higher" are one-tailed t-tests. All tests of "difference" are two tailed t-tests. All t-tests are conducted assuming equal variances and using standard errors clustered by participant.

(p=0.024). This suggests that some participants in the repeated matching treatment are worried about starting a 'rat race', and thus do not purchase even once.

A greater proportion of females than males buy credits in both treatments, and when the treatments are combined (p=0.148 for the RaMT, p=0.093 for the ReMT, and p=0.044 for the combined treatments).

Note that the proportion of all participants who buy credits at least twice is not significantly different across treatments, with 22% of all participants buying credits at least twice. This figure is consistent with the proportion of subjects who paid to modify the public reporting of their performance in the studies of Charness et al (2014) and Blacklow et al (2021).

Table 1: Credit Buying Behaviour by Treatment and Gender

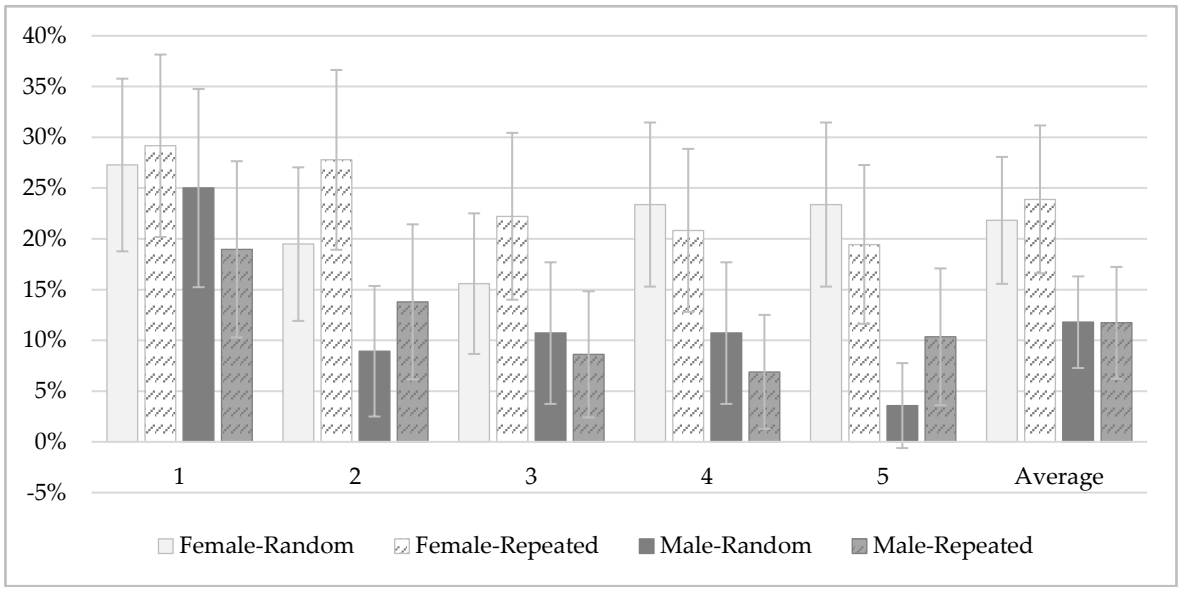
Treat-ment	Gender	# partici-pants	# bought credits at least once	% bought credits at least once	% bought credits only once	# bought credits at least twice	% bought credits at least twice
Random Matching	Females	77	33	43%	18%	19	25%
	Males	56	19	34%	18%	9	16%
	Total	**133**	**52**	**39%**	**18%**	**28**	**21%**
Repeated Matching	Females	72	25	35%	6%	21	29%
	Males	58	14	24%	8%	9	16%
	Total	**130**	**39**	**30%**	**7%**	**30**	**23%**
Total	**All**	**263**	**91**	**35%**	**13%**	**58**	**22%**

Now consider the proportion of participants who buy credits in each round. Figure 1 shows these proportions disaggregated by treatment and gender. As it is most insightful to show data across rounds when it is disaggregated by treatment and gender, we will do so for the remaining figures in the paper. Figures showing the data by round individually disaggregated by gender and disaggregated by treatment are presented in Appendix 1.

In the data presented in Figure 1, there are no significant treatment effects observed for the average purchase of credits within the rounds or on average. However, there is some evidence that males buy less in the RaMT than ReMT in round 5 (p=0.077). This finding is not consistent with conjecture 1i).

There are strong gender effects in the data in Figure 1. A greater proportion of females buy credits, both overall (p=0.001) and in each treatment (RaMT: p=0.017 and ReMT: p=0.015). Taken with the data in table 1, this finding does not support conjecture 2 i.).

Figure 1: Proportion who Purchased Credits by Treatment and Gender



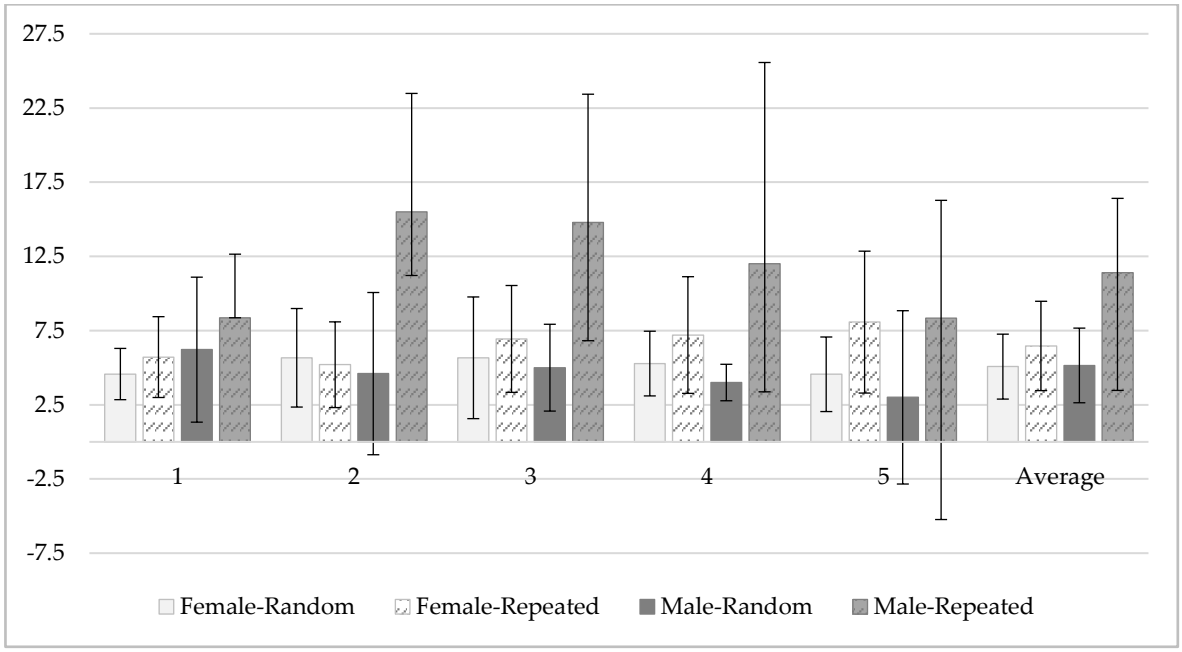
Note: Bars show 90% confidence intervals for individual rounds. While non-overlapping confidence intervals suggest a statistically significant difference, tests for differences discussed in the text include covariances and use standard errors clustered by participant, which results in the two tests not being comparable.

In the ReMT there is evidence of a decline in the proportion of participants buying credits over rounds for both genders. The proportion who buy credits is higher in round 1 compared to round 5 ($p= 0.017$ for females and $p=0.066$ for males) and higher in round 2 compared to round 4 ($p=0.048$ for females and $p=0.022$ for males). In the RaMT, however, the genders behave differently from each other. While significantly fewer males buy credits in round 5 compared to round 1 as is in the ReMT ($p<0.001$), the proportion of females in the RaMT who buy credits in round 5 is not significantly different to that in round 1 ($p=0.235$). Indeed, the proportion of females purchasing credits was numerically (though not significantly) greater in rounds 4 and 5 than either round 2 and 3. This finding is consistent with the hypothesis that females perceived an ongoing reputational advantage for buying credit in the RaMT.

Figure 2 shows the average number of credits bought by those who bought credits. The most prominent feature of this data is the purchase of credits by males in the ReMT, which is significantly greater than that of males in the RaMT in rounds 2 ($p=0.023$) and 3 ($p=0.022$) and for the average of the 5 rounds ($p=0.031$). The purchase of credits by males in the ReMT is significantly greater than that of females in the ReMT and RaMT for rounds 2 and 3 ($p=0.020$ and 0.010) and on average over the five rounds ($p=0.030$). This finding is consistent with conjecture 2i), but not conjecture 1i).

There are no significant changes in the average credit purchases across rounds in either treatment or by gender, except for females in ReMT who bought significantly more credits in round 4 than round 2 ($p=0.030$).

Figure 2: Number of Credits Purchased by Treatment and Gender

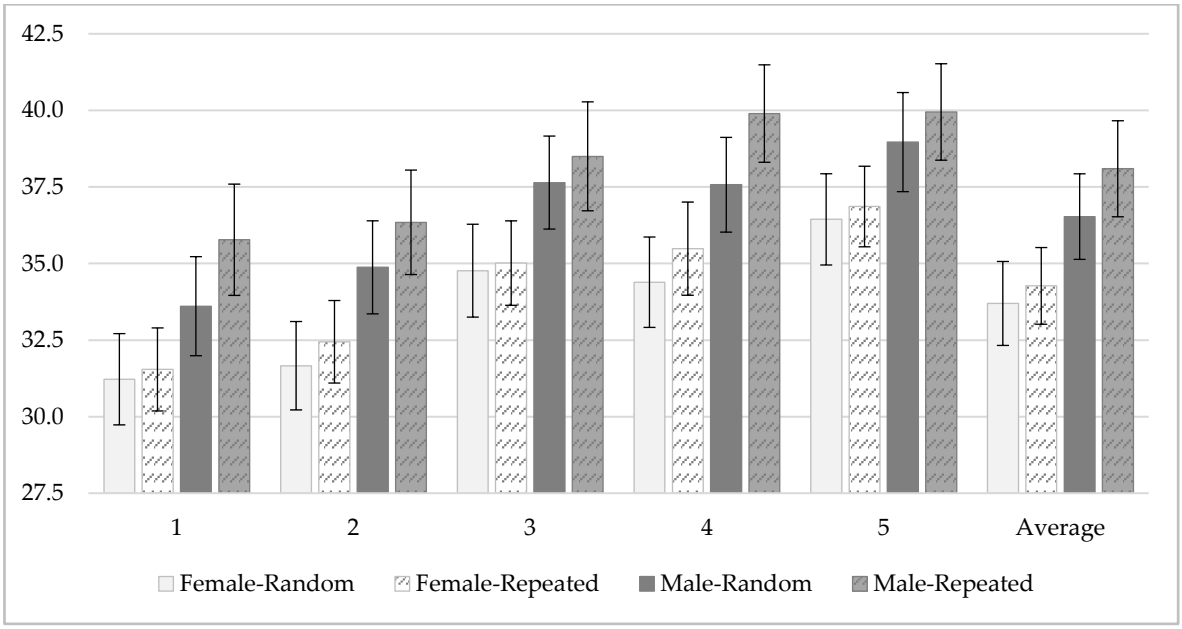


Note: See caption on figure 1.

Figure 3 shows average participant performance across rounds. Average performance is higher in the ReMT than in the RaMT, though this difference is only significant for male performance in round 4 ($p=0.041$) and round 1 ($p=0.069$). This finding does not support conjecture 1ii)

Males have significantly higher performance in each treatment and each round than do females. This finding supports conjecture 2ii). The performance of both males and females in both treatments monotonically increases across rounds. Performance for both males and females in both treatments is significantly greater in round 4 compared to round 2 ($p<0.001$ in all cases) and in round 5 compared to round 1 ($p<0.001$ in all cases).

Figure 3: Performance by Treatment and Gender



Note: See caption on figure 1

The number of credits purchased by a participant may be influenced by that participant's performance in addition to gender and treatment. For example, performance appears higher in the ReMT (though not always significantly so), and this might cause participants to buy less credits in that treatment. To control for these simultaneous effects, we conduct a regression using a random-effects (RE) Heckman model, in which the number of credits purchased by those who buy credits is determined by a performance measure (described below), gender, treatment, and all cross-terms of these variables. We use a Heckman model to be able to make conclusions about all participants since participants that purchase credits in similar settings have been shown to be different to non-buyers (see Blacklow et. al, 2021). We employ a RE model to allow for the panel structure of the data and additional heterogeneity in the error term for participants.

The performance measure used in the regression, for each round r, is the sum of scaled performance in all rounds other than r, where an individual's scaled performance in round r is their performance divided by the average performance of all other participants in that round. Scaling in this way controls for performance growth across rounds, while avoiding endogeneity between the purchase of credits and performance in a given round. This scaled performance measure proxies an individual's willingness to supply effort.

The marginal effects of this regression analysis for credit purchases are given in Table 2. The regression estimates for both the selection and outcome equation that yield these marginal effects are provided in Table A1 in the appendix. First observe that the analysis finds that participants who buy credits in the ReMT buy significantly more credits than in the RaMT ($p=0.026$). This finding is consistent with those from the summary statistics in Figure 2. There is also some evidence that females buy more credits in the ReMT than the RaMT

(p=0.078), as was found by the analysis of figure 2. However, there is no evidence of any other gender effects in the regression taken on the overall sample.

Table 2: Purchase of Credits: Marginal Effects

`			(1) Full Sample	(2) RaMT	(3) ReMT
Variable		*Sample*	*Marg. Eff. (std.errors)*	*Marg. Eff. (std.errors)*	*Marg. Eff. (std.errors)*
Performance measure	For	All Obs	0.332	-1.782*	2.359***
			(0.514)	(0.961)	(0.610)
	For	Female	0.963	-3.312***	1.566**
			(0.951)	(1.082)	(0.692)
	for	Male	-0.492	0.321	3.343***
			(0.968)	(1.196)	(1.102)
	for	ReMT	-1.080		
			(0.768)		
	for	RaMT	1.778***		
			(0.676)		
ReMT	for	All Obs	2.157**		
			(0.965)		
	for	Female	2.409*		
			(1.366)		
	for	Male	1.827		
			(1.391)		
Female	for	All Obs	-0.014	0.180	-1.131
			(1.848)	(1.998)	(1.034)
	for	ReMT	-0.302		
			(1.832)		
	for	RaMT	0.280		
			(2.336)		
log-likelihood			-1172.39	-585.93	-561.92
sample size			1315	665	650

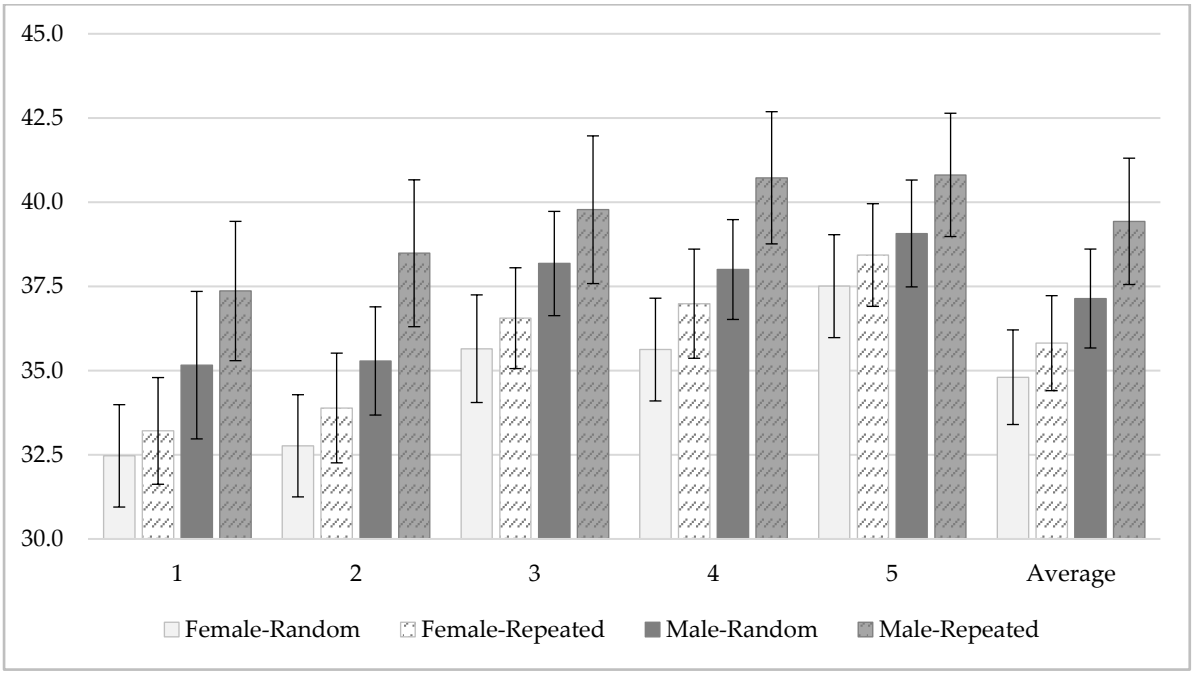
Notes: The *** denote statistically different from zero at 1%, ** at 5% and * at 10% level of significance. Standard errors were computed using the delta method from a random-effects Heckman model. For more information see Appendix 1 and for full estimation results see Table A1 in Appendix 1.

However, the estimation on each treatment separately (as shown in columns 2 and 3 in Table 2), indicates that gender interacts differently with performance across treatments. In the ReMT sample, higher performance leads to significantly higher purchases of credits by both men and women. In contrast, in the RaMT, a higher performance leads to a significantly lower credit purchase by women. Further, there is no evidence that performance affects the

credit purchases of men in the RaMT. This finding thus reinforces the earlier findings that participants are more motivated to buy credits in the ReMT than the RaMT, and that gender influences these choices.

The findings on how participants' scores are related to treatment and gender are of interest because a participant's score represents their social image within the context of the experiment. Figure 4 shows the average score across rounds given treatment and gender. For males, score is higher in the ReMT than in the RaMT, notably in round 2 ($p=0.025$), round 4 ($p=0.033$), and the round average ($p=0.057$). For both genders, score is higher in all rounds in the ReMT than in the RaMT, and noticeably so in round 2 ($p=0.029$), round 4 ($p=0.027$), round 5 ($p=0.099$) and the average ($p=0.051$).

Figure 4: Score by Treatment and Gender



Note: See caption on figure 1.

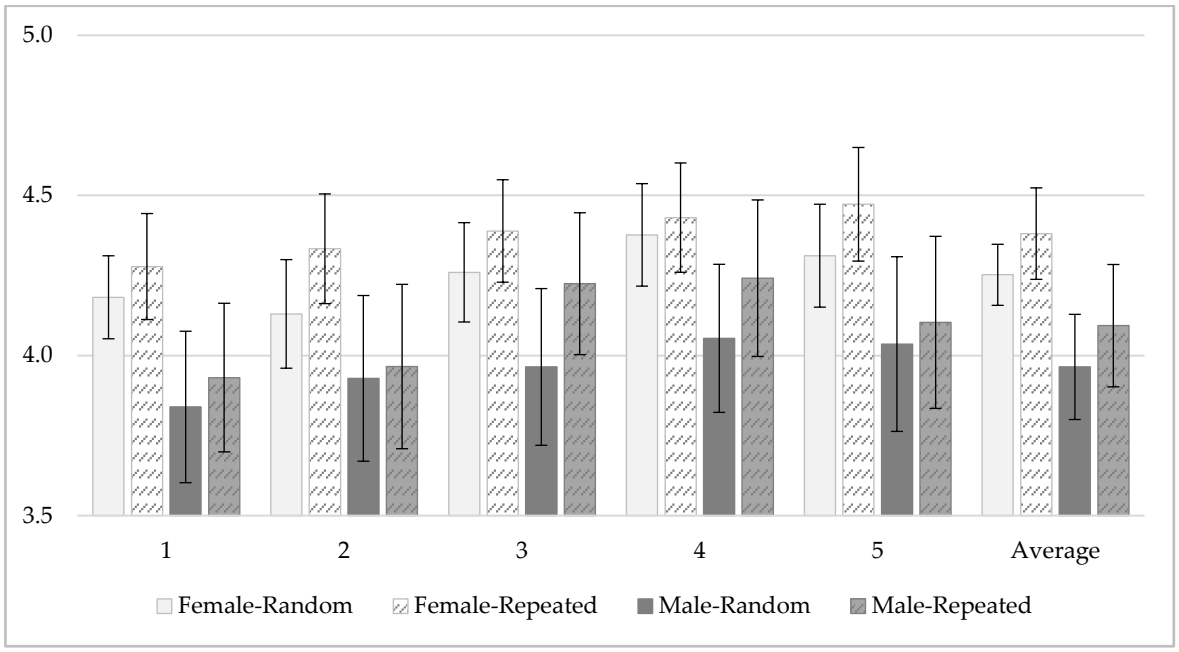
Males have a significantly higher score than females in each treatment. Although more females buy credits, this is offset by the higher performance of males on the real effort task and the greater number of credits purchased by those males who do buy them. This finding is somewhat consistent with those of Jones and Linardi (2014): female purchase of credits acts to compensate for their relatively low performance, thus making their score appear more average.

The score for both males and females in both treatments is significantly greater in round 4 compared to round 2 ($p<0.001$ in all cases) and in round 5 compared to round 1 ($p<0.001$ in all cases). The increase in score across rounds is primarily due to the increase in performance across rounds.

Figure 5 summarises the average feedback given by participants in each round by treatment and gender. In figure 5, feedback of 5 is very good and 1 is very bad. Note that, for each gender, in each round, the feedback in the ReMT is more positive than that in the RaMT, though in all but two rounds this difference is not statistically significant. There is some evidence that feedback is better in the ReMT than RaMT for females in round 2 (p=0.080) and for males in round 3 (p=0.094).

Females give more positive feedback than males in all rounds in both treatments. In particular, the feedback from females is significantly greater than males in the ReMT in rounds 1, 2 and 5 (p=0.022, p=0.024 and p=0.028) and overall (p=0.024). Feedback from females is also significantly greater in the RaMT in round 1 (p=0.017), round 3 (p=0.045) round 4 (p=0.028) and overall (p=0.007).

Figure 5: Feedback by Treatment and Gender



Note: See caption on figure 1.

The summary statistics presented in Figure 5 cannot be used to isolate treatment or gender effects on feedback, as feedback may be influenced by a participant's own score and that of their partners, and these may, in turn, be influenced to some extent by treatment and gender. To identify treatment and gender effects we estimate a random effects (RE) ordered probit, which includes as determinants, treatment, partner's score, the difference between the partner and own scores, whether the participant buys credits, gender and round. The difference between partner and own scores are included in the estimation as (i) it has been argued by psychologists that social comparisons are an integral part of human cognition (e.g. Festinger, 1954; Olson, Herman, and Zannan, 1986; Wood, 1996; Arigo et al, 2020) and (ii) there is experimental evidence that social comparisons influence behaviour (e.g. Gächter and Christian Thöni, 2010; McDonald et. al. 2013; Bursztyn et al., 2014; Cohn et. al., 2014). In

this case, a participant's assessment of the partner score may be influenced by how it compares to their own score. The experiment reported in Blacklow et al (2021) also measures participant feedback on performance, and analyses how that feedback is influenced by partner score, and gender. However, their experimental design did not allow for the study of the impact on the feedback of repeated interactions and score differences.

Table 3 provides the key results from estimating a RE ordered probit for feedback, with the full results being reported in Table A2 in Appendix 1. Column 1 in Table 3 shows results from the regression when applied to the entire data, including both treatments. In line with data in Figure 5, the regression yields a positive coefficient on ReMT, though this is not significant. Consequently, the analysis does not find evidence that feedback is higher when people have repeated, relative to random, interactions. Such a finding suggests that ongoing relationships have neither a strong competitive nor collusive impact on feedback. This, in turn, might arise because people are lying-averse (Mazar, Amir, and Ariely, 2008; Fischbacher and Föllmi-Heusi, 2013) and usually attempt to give feedback they believe is accurate.

Intuitively it might be expected that a participant's feedback is positively related to their partner's score. Though it might also be the case that a participant discounts the information provided by their partner's score, as there is a possibility that it is 'inflated' through the purchase of credits. The results in Table 3 indicate that is not the case, with all regressions showing that in both treatments feedback is statistically significantly related to score. This finding is consistent with the findings of Blacklow et al (2021).

The amount by which the partner's scores exceeds the own score is significant and positively related to feedback in regression (1) in Table 3 over the whole sample. However, it is only significant when the regression is taken on the data from the ReMT and not when taken on the data from the RaMT, suggesting that the relative magnitude of a participant's score with that of their partner is not a determinant of their feedback in the RaMT. In this instance, social comparisons play a greater role in the response to repeated interactions compared to random interactions. As discussed in section 3, social comparisons have the potential to lead to competitive behaviour. We observed earlier in this section that competitive behaviour was observed in the ReMT and not in the RaMT. Thus, our results suggest that competitive behaviour is a characteristic of credit purchase, performance and feedback in the ReMT, but is not a characteristic of any of these measures in the RaMT.

Table 3: Ordered Probit for Feedback

	Full Sample (1)	RaMT (2)	ReMT (3)
	Coef. / (std.errors)	*Coef. / (std.errors)*	*Coef. / (std.errors)*
ReMT	0.272		
	(0.176)		
Buy	0.016	-0.029	0.097
	(0.139)	(0.158)	(0.245)
Partner's Score	0.084***	0.092***	0.082***
	(0.014)	(0.016)	(0.028)
Partner's less Own Score	0.019**	0.006	0.035***
	(0.008)	(0.009)	(0.013)
Female	0.388**	0.416**	0.354
	(0.171)	(0.191)	(0.314)
Round 2	0.066	0.040	0.099
	(0.106)	(0.157)	(0.134)
Round 3	0.118	-0.031	0.299
	(0.115)	(0.145)	(0.199)
Round 4	0.284**	0.211	0.363
	(0.117)	(0.132)	(0.235)
Round 5	0.120	0.019	0.222
	(0.133)	(0.160)	(0.249)
log likelihood	-1244.22	-671.06	-560.08
sample size	1315	665	650

Notes: Figures in parentheses ( ) are standard errors clustered by participant. The *** denote statistically different from zero at 1%, ** at 5% and * at 10% level of significance. The full estimation results are in Table A2 in Appendix 1.

As suggested by the data in Figure 5, females provide statistically significant (at the 5% level) higher feedback than males when the regression is taken over the entire data set. This is also true when the regression is conducted over the data from the RaMT. These findings are also consistent with those of Blacklow et al (2021). However, there is no evidence that females provide more positive feedback than males in the ReMT. Both male and female participants in the ReMT face ongoing consequences (apart from the final round) of poorly formulated feedback, and this might encourage both genders to provide similar levels of feedback.

Overall, conjecture 3 is not confirmed as stated. Conjecture 3 proposes that, because participants behave more competitively in the ReMT than the RaMT, they would offer more negative feedback in the former. However, treatment has not been shown to significantly affect feedback. Nonetheless, the partner less own score positively influences feedback in the ReMT but not in the RaMT. This points to competitiveness in the ReMT that is not present in the RaMT. If females exhibit greater cooperation in esteem seeking than males, then conjecture 3 would predict they would offer more positive feedback. We do find that females offer higher feedback, but only in the RaMT. Females thus are more competitive when providing feedback when interactions are repeated, than when they are random.

5. Discussion

This paper reports on an experiment designed to identify how repeated interactions between people, relative to one-shot interaction, change their esteem seeking behaviour. Our broad conclusion is that esteem-seeking behaviour is increased when interactions are repeated. We interpret this finding as arising because, in the ReMT, participants are in a type of 'rat-race', where each participant's esteem seeking reinforces their partner's incentive to esteems seek. While the average number of credits purchased is higher in the ReMT, we also noted a lower propensity to buy credits in ReMT. We interpret this as indicating that at least some participants recognise they may be drawn into a rat race in the ReMT, and have less inclination to buy credits to avoid that.

The finding that some participants behave more competitively in repeated interactions is at odds with our expectations, which are based on the previous experimental literature. Notably, more cooperation is observed in the finitely repeated prisoners' dilemma compared to the one-shot implementations of these games. This suggests that people are framing the acquisition of money and the acquisition of esteem in quite different ways, with the former promoting cooperative behaviour while the latter promoting competitive ('rat-race' type) behaviour.

Our findings on esteem seeking are consistent with observed online behaviour. Conversations on social media (which are repeated interactions) are notorious for spiralling out of control. Some people, though, avoid social media, or avoid commenting on others' posts or responding to comments on their posts. Others avoid social media altogether to avoid the conflict associated with escalating exchanges.

Our results also have parallels in the physical world. People dress more formally, and with more of an eye for appearance when going to important social occasions or work, than when they are (for example) going shopping on Saturday mornings. The former examples are ones usually involving repeated interactions, while the latter usually involves one-shot interactions.

We proposed that competitive behaviour in esteem seeking would be paralleled by correspondingly negative feedback: participants would attempt to make themselves look better by denigrating their partners. Indeed, Charness et; al. (2014) found that subjects were willing to expend resources in order to damage (the equivalent of) their partners' score. However, we found that repeated interactions do not have a significant independent impact on feedback when compared to one-shot interactions. We attribute this finding to lying-aversion: in each round participants feel obligated to give an honest assessment of their partner's score in that round.

If our finding that feedback is independent of treatment were externally valid, it would be expected that, for example, two rivals would give the same assessments of each other's work as would an independent observer, irrespective of their rivalry. Those with bitter experience of such assessment processes may doubt the external validity of this conclusion. Given the importance and prevalence of assessment processes, it is therefore a finding which would

usefully be followed up by field experiments. It would similarly be useful to find external validation of our finding that a person's relative score also influences their assessment of other's performance. If it were, there would be systematic differences in assessment made by high and low performers: high performers would 'mark down' low performers and low performers would 'mark up' high performers. In this way, the method of allocating assessors to individuals would influence the outcome, and possibly the objectivity of the assessment process.

We observe important gender differences in our results. Males buy significantly more credits in the ReMT than in the RaMT in some rounds, whereas there is not a significant difference in purchases between treatments for females. Thus the higher credit purchases in the ReMT is primarily due to males. Additionally, performance, and thus score was higher for males, in the ReMT than in the RaMT.

These gender effects in credit purchase are identified directly from the summary statistics, which do not control for the potential impact of performance on credit purchases. We find that once we control for performance, these findings change. In contrast to the conclusion from summary statistics, there is some evidence that females purchase more in the ReMT than the RaMT, whereas we do not find evidence that males do. The increase in the purchase of credits by females in the ReMT may be motivated by a desire to compensate for their relative low performance on the real effort task in that more competitive environment.

Nonetheless, the lower score by females in both treatments suggests that, consistent with expectation (which were based on the literature), females act more cooperatively than males. However, we did observe that a greater proportion of females bought credits than males. This also may have been an attempt by those females to reduce the gap in scores with the males.

Gender effects are also observed in the feedback. In line with their more cooperative behaviour, females provided more positive feedback than males. However, when the data is separated by treatment, this finding only occurred significantly in the RaMT. This finding may arise because males act more competitively with regard to feedback in the RaMT, as that behaviour does not have ramifications in future rounds. This suggests that males are influenced more by strategic considerations than females when providing feedback.

This study was conducted under conditions of anonymity. It would be expected that removing that anonymity would increase the extent to which individuals esteem-seek. In future work, it would be of interest to determine how the extent of esteem seeking changes once the condition of anonymity has been removed. Would esteem-seeking be ubiquitous or remain the behaviour of a minority?

In a similar vein, the repeated interactions within the experiment were designed to mimic social interactions between strangers and acquaintances, and not those with close personal relationships. Clearly many, if not most, repeated interactions occurring between people are with friends and family. One would imagine that esteem-seeking, if it exists between those with close personal relationships, manifests in different and more complex ways than that

between strangers and acquaintances. A study of such esteem seeking would no doubt require a different methodology than the laboratory experiment discussed in this paper, though would be necessary to gain a complete understanding of the economics of esteem-seeking.

References:

Andreoni, J., and Miller, J.H. (1993) Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence. *Economic Journal*, 103(418), 570-585.

Arigo Danielle, Mogle Jacqueline A., Brown Megan M., Pasko Kristen, Travers Laura, Sweeder Logan, Smyth Joshua M. (2020) "Methods to Assess Social Comparison Processes Within Persons in Daily Life: A Scoping Review", Frontiers in Psychology, 10, DOI=10.3389/fpsyg.2019.02909

Benabou, R. and Tirole, J. (2006) "Incentives and Prosocial Behavior", *American Economic Review*, December, v. 96, iss. 5, pp. 1652-78

Brennann, G. and Pettit, P. (2004) The economy of esteem: An essay on civil and political society, Oxford and New York: Oxford University Press

Bursztyn, L., Florian Ederer, Bruno Ferman, and Noam Yuchtman, (2014) "Understanding Mechanisms Underlying Peer Effects: Evidence from a Field Experiment on Financial Decisions," *Econometrica*, 82 (4), 1273{1301.

Bursztyn, L., Fermany, B., Fiorinz, S., Kanzx, M., Rao, G. (2018), "Status Goods: Experimental Evidence from Platinum Credit Cards" *Quarterly Journal of Economics*, Volume 133, Issue 3, 1 August 2018, Pages 1561–1595

Leonardo Bursztyn and Robert Jensen (2017), "Social Image and Economic Behavior in the Field: Identifying, Understanding, and Shaping Social Pressure" *Annual Review of Economics*, 9:1, 131-153.

Charness, G, Masclet, D. Villeval, M.C. (2014) "The Dark Side of Competition for Status" *Management Science*, 60, January, pp. 38-55.

Alain Cohn, Ernst Fehr, Benedikt Herrmann and Frédéric Schneider (2014) "Social Comparison and Effort Provision: Evidence from a Field Experiment", *Journal of the European Economic Association*, Volume 12, Issue 4, 1 August, Pages 877–898

Cooper, R., DeJong, D.V., Forsyth, R. & T.W. Ross, T.W. (1996) Cooperation without Reputation: Experimental Evidence from Prisoner's Dilemma Games. *Games and Economic Behavior*, 12(2), 187-218.

Dohmen, T., Falk, A., Fliessbach, K. Sunde, U., Weber, B. (2011) "Relative versus absolute income, joy of winning, and gender: Brain imaging evidence", *Journal of Public Economics*, 95 279–285

Ellingsen, T. and Johannesson, J., (2008) "Pride and Prejudice: The Human Side of Incentive Theory", *American Economic Review*, June 2008, vol. 98, iss. 3, pp. 990-1008

Festinger, L., (1954) "A Theory of Social Comparison Processes", *Human Relations*, 7, 117-140.

Fischbacher, U., (2007) "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments", *Experimental Economics*, 10(2), 171-178.

Fischbacher, Urs, and Franziska Föllmi-Heusi. 2013 "Lies in disguise: An experimental study on cheating" *Journal of the European Economic Association,* vol. 11 (3), pp. 525–47.

Friedrichsen, J. and Engelmann, D. (2018) "Who cares about social image?", *European Economic Review*, Volume 110, Pages 61-77,

Simon Gächter and Christian Thöni (2010) "Social comparison and performance: Experimental evidence on the fair wage–effort hypothesis", *Journal of Economic Behavior & Organization, vol* 76, pp. 531–543.

Gill, D. and Prowse, V. (2012) 'A Structural Analysis of Disappointment Aversion, in a Real Effort Competition.' *American Economic Review*, 102(1): 469--503

Grimalda, G., Pondorfer, A. & Tracer, D. Social image concerns promote cooperation more than altruistic punishment. *Nat Commun* 7, 12288 (2016). https://doi.org/10.1038/ncomms12288

Greiner, B., 2004. The Online Recruitment System ORSEE 2.0 - A Guide for the Organization of Experiments in Economics. University of Cologne, Working Paper Series in Economics 10.

Niklas Horstmann, Jan Krämer and Daniel Schnurr, (2018) "Number Effects and Tacit Collusion in Experimental Oligopolies", *The Journal of Industrial Economics*, Vol. 66, Issue 3 pp. 650-700

John, O., and S. Srivastava (1999): "Big Five Trait Taxonomy. History, Measurement, and Theoretical Perspectives," in *Handbook of Personality*. Theory and Research, ed. by L. Pervin, and O. John, pp. 102–139. Guilford, New York.

Jones, Daniel and Sera Linardi, 2014, "Wallflowers: Experimental evidence of an aversion to standing out", *Management Science*, vol. 60, pp. 1757-1771.

Kirchler, M. and Palan, S. (2018), "Immaterial and monetary gifts in economic transactions: evidence from the field" *Experimental Economics*, 21: 205-230

Lacetera, Nicola and Macis, Mario (2010) "Social image concerns and prosocial behavior: Field evidence from a nonlinear incentive scheme", *Journal of Economic Behavior & Organization*, Volume 76, Issue 2, November pp. 225-237.

Greg Lukianoff; Jonathan Haidt (2018). *The Coddling of the American Mind: How Good Intentions and Bad Ideas Are Setting Up a Generation for Failure.* Penguin Publishing Group. ISBN 978-0-7352-2489-6.

Mazar, Nina, On Amir, and Dan Ariely, 2008, "The dishonesty of honest people: A theory of self-concept maintenance." *Journal of Marketing Research*, vol. 45(6), pp. 633–44.

McDonald, I., Nikiforakis, N., Olekalns, N., and Sibly, H., (2013) "Social Comparisons and Reference Group Formation: Some Experimental Evidence", *Games and Economic Behavior*, Volume 79, May, Pages 75–89

Myer, C. J. and Tripodi, E. (2021) "Image Concerns in pledges to give blood: Evidence from a field experiment", *Journal of Economic Psychology*, Volume 87, December, 102434

Niederle, Muriel and Lise Vesterlund, 2007, "Do women shy away from competition? Do men compete too much?", *Quarterly Journal of Economics*, vol. 122(3), pp. 1067-1101.

Nisar, TM, G Prabhakar, L Strakova (2019) "Social media information benefits, knowledge management and smart organizations", *Journal of Business Research*, Volume 94, January, pp. 264-272

Oaks, Kelly (2019) "The complicated truth about social media and body image" BBC Future, 12 March, available at: https://www.bbc.com/future/article/20190311-how-social-media-affects-body-image. Accessed 6 November 2021

Olson, J.M., Herman, C.P., and Zanna, M.P., (1986) *Relative Deprivation and Social Comparison: The Ontario Symposium*, Volume 4, Lawrence Erlbaum Associates.

Rotter, J. (1966): "Generalized expectancies for internal versus external control of reinforcement," *Psychological Monographs: General & Applied*, 80 (1), 1–28.

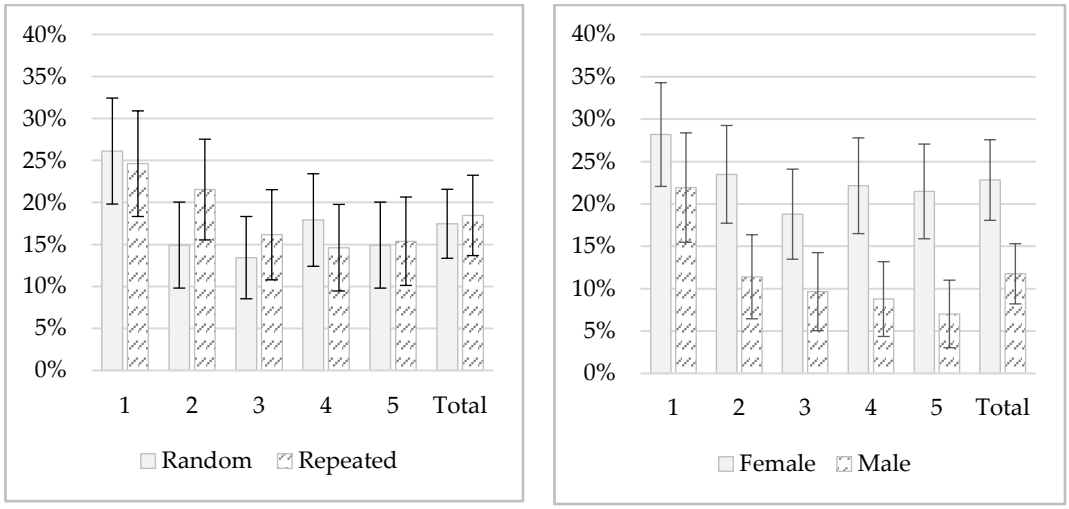
Wood, J. V. (1996), "What is Social Comparison and How Should We Study it?", *Personality and Social Psychology Bulletin*, Vol 22, Issue 5, pp. 520-537 https://doi.org/10.1177/0146167296225009

Appendix 1: Results with treatment combined and gender combined.

The left panel of Figure A1 shows the proportion of participants buying credits in each round in the RaMT and ReMT. Averaged across rounds, there is no evidence that the proportion is greater in RaMT than in the ReMT. In fact, there is some evidence the proportion buying credits in the RaMT is higher than in the ReMT in round 2 ($p=0.083$). There is a significant decline in the proportion buying credits between rounds 1 and 5 ($p=0.003$ for the RaMT and $p=0.005$ for the ReMT), but there is no change in the proportion buying credits in either treatment between rounds 2 and both rounds 4 and 5.

The right panel of Figure A1 shows the proportion of participants buying credits in each round disaggregated by gender. The gender effects discussed in the text with respect to Figure 1 are also apparent in this data, with females buying significantly more credits than males on average. There is a significant decline in the number of credits bought by males between rounds 1 and 5 ($p<0.001$) but the decline by females is not significant ($p=0.029$).

Figure A1: Proportion Buying Credits by Treatment (left) and Gender (right)

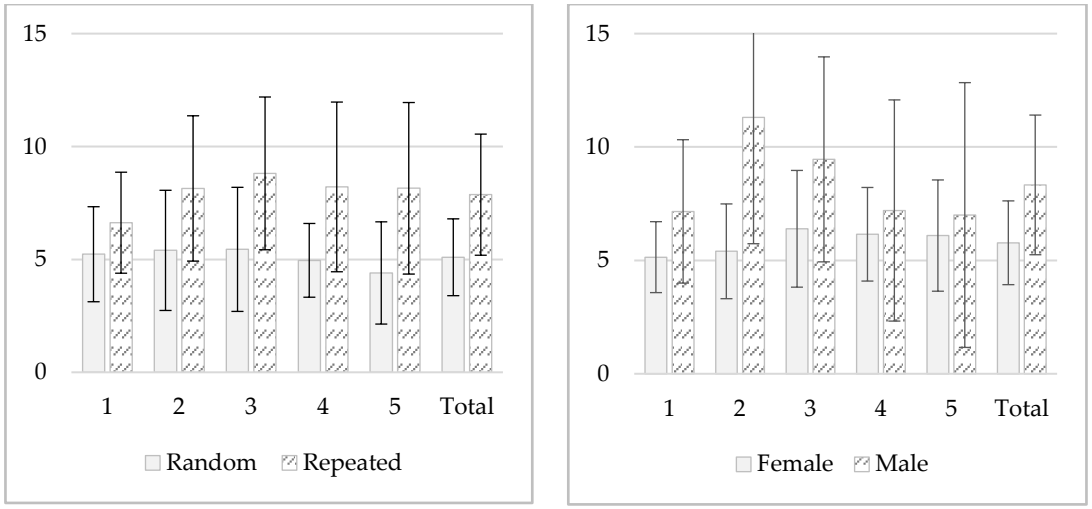


Note: See caption on figure 1.

The left panel of Figure A2 shows the average number of credits purchased by those who purchase them in each round in the RaMT and ReMT. The number of credits purchased is always greater in the ReMT than the RaMT, notably in rounds 3 ($p=0.092$), 4 ($p=0.086$), round 5 ($p=0.073$) and on average ($p=0.074$).

The right panel of Figure A2b shows the number of credits purchased in each round disaggregated by gender. The average number of credits purchased by males is greater than that by females. This difference is significant in round 2 ($p-=0.040$), but not in other rounds.

Figure A2: Number of Credits Purchased by Treatment (left) and Gender (right)

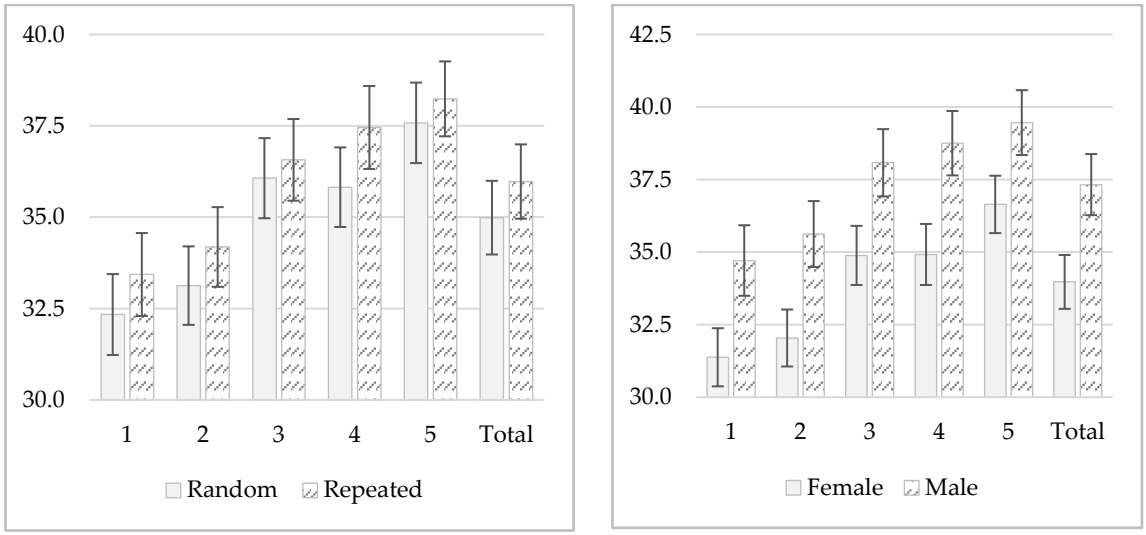


Note: See caption on figure 1

The left panel of Figure A3 on the following page, shows the performance in each round in the RaMT and ReMT. The performance is always greater in the ReMT than the RaMT but is only significantly so in round 4 ($p=0.043$).

The right panel of Figure A3 shows the performance in each round disaggregated by gender. The average performance by males is significantly higher than females ($p<0.001$ in all rounds).

Figure A3: Performance by Treatment (left) and Gender (right)



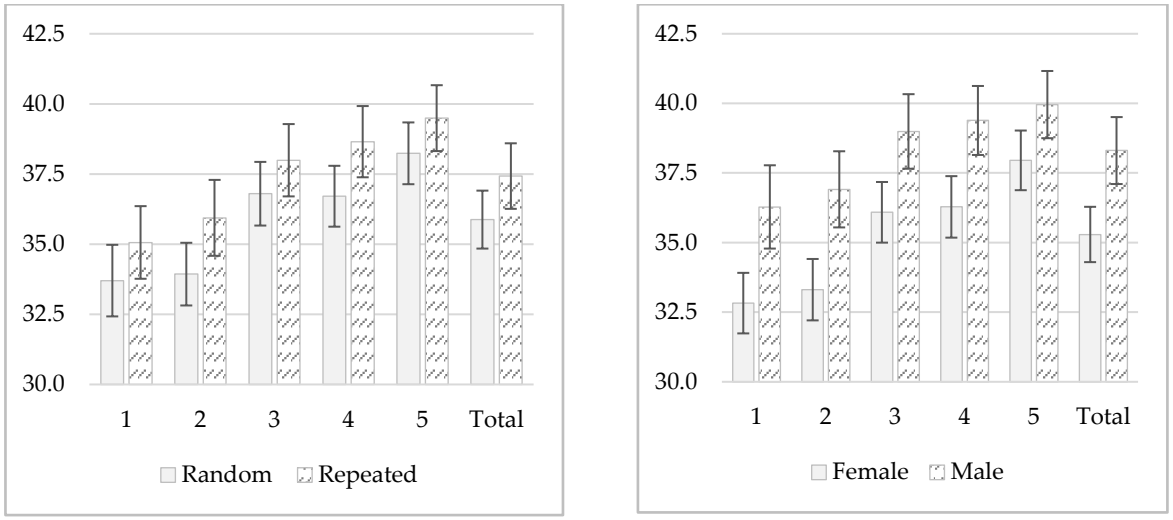
Note: See caption on figure 1

Figure A4 shows the average score in each round in the RaMT and ReMT in the left panel. The average score is always greater in the ReMT than the RaMT, and significantly so in

round 2 (p=0.029) and round 4 (p=0.027) with additional evidence of that difference in round 5 (p=0.099) and in the average across rounds (p=0.051)

The right panel of Figure A4 shows the average score in each round disaggregated by gender. The average score by males is significantly more than females in all rounds and for the average across rounds (p<0.001). Note there is a significant increase in score across rounds, reflecting the increase in effort across rounds.

Figure A4: Score by Treatment (left) and by Gender (right)

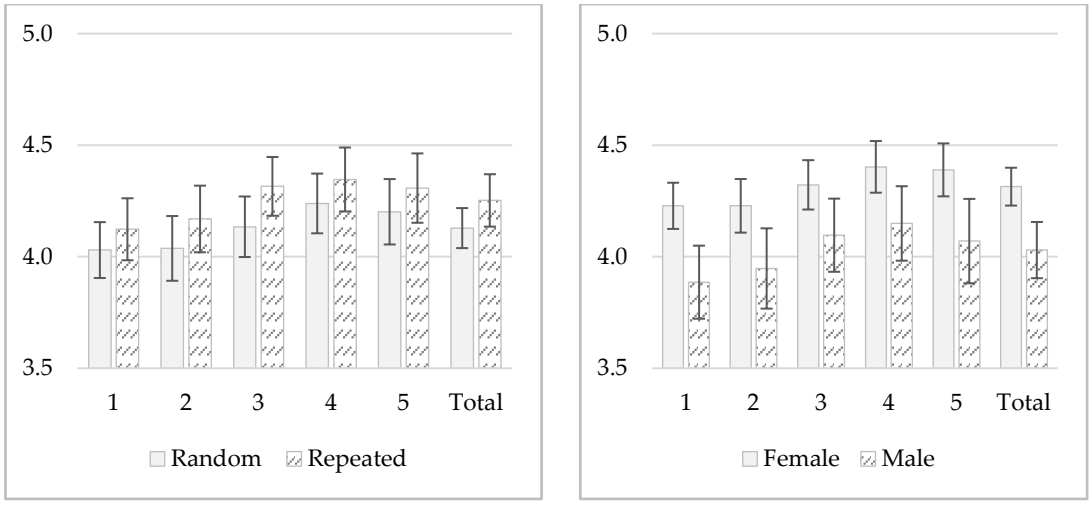


Note: See caption on figure 1

The left panel of Figure A5 shows the average feedback in each round in the RaMT and ReMT. The average feedback is always greater in the ReMT than the RaMT, notably in round 3 (p=0.057) and on average for the rounds (p=0.084).

The right panel of Figure A5 shows the average feedback in each round disaggregated by gender. The average feedback provided by females is significantly higher (more positive) than that provided by males in each round at the 5% level of significance and higher on average for the rounds (p=0.001).

Figure A5: Feedback by Treatment (left) and by Gender (right)



Note: See caption on figure 1

The data collected from the experiment for the 263 participants over 5 rounds was arranged in a panel format. We estimate a Heckman model with random effects (RE) by participant, in both the selection or "Buy" equation, and the outcome or "Credits" purchased equation. We include fixed effects for rounds with dummy variables for rounds 2 to 5 (round 1 being the base). Since individual participants did not experience both treatments and we were interested in gender differences, individual fixed-effects estimation was not suitable. Our main variables of interest: ReMT, female and our performance measure, are included in both equations, as well as interactions: ReMT x female, ReMT x Performance and Female x Performance. The base case for the estimations is a male participant, with RaMT treatment in round 1. With limited variables available we include age, the locus of control and the big 5 personality variables as additional variables in the selection equation. The non-adaptive Gauss-Hermite quadrature with three integration points was used to obtain the estimates for the full sample (1) RaMT treatment sample (2) and ReMT treatment sample (3). The standard errors are clustered by participant.

The estimates from the "Credits" and "Buy" equations are presented in Table A1. Note that for the sample as a whole and the ReMT treatment, the correlation between the error term of Buy and Credits equation at a participant level is significant at the 1% level, justifying the use of the RE Heckman model. The participant variance of the RE terms for Credits and Buy are significant at the 5% and 1% level providing additional evidence in support of the RE model.

To investigate the effects of our treatment, measures of performance, and gender we estimate a RE ordered probit model for the ordinal feedback that each participant provides to their partner. We include fixed effects for rounds with dummy variables for rounds 2 to 5 (round 1 being the base) to control for changes in feedback over rounds and random effects at a participant level to allow for the panel structure of the data. The estimated coefficients and cut points of the internet and other estimation details are provided in Table A2 on the following page. The participant variance of the RE terms for feedback is significant at the 1% level across all three samples, providing support for the RE model.

Table A1: Full Estimation Results for RE Heckman for Credits Purchased and Buy Decision

	Full Sample (1)		RaMT (2)		ReMT (3)	
	Credits	**Buy**	**Credits**	**Buy**	**Credits**	**Buy**
ReMT	-9.548**	-0.782				
	(4.001)	(1.147)				
Female	-6.092	-1.865	14.420***	-1.948**	6.053	-2.138
	(5.860)	(1.854)	(4.288)	(0.954)	(5.303)	(1.399)
ReMT x Female	0.582	0.416				
	(1.981)	(0.956)				
Performance	-1.905	-0.609	0.321	-0.676***	3.343***	-0.248
	(1.314)	(0.432)	(1.196)	(0.225)	(1.102)	(0.193)
ReMT x Performance	2.858***	0.090				
	(1.021)	(0.384)				
Female x Performance	1.455	0.531	-3.633***	0.680**	-1.777	0.454
	(1.633)	(0.484)	(1.207)	(0.283)	(1.325)	(0.321)
Round 2	0.495	-0.572***	0.197	-0.866***	2.041	-0.208
	(1.080)	(0.163)	(1.495)	(0.258)	(1.292)	(0.281)
Round 3	0.959	-0.881***	0.426	-0.990***	2.023*	-0.754**
	(1.142)	(0.182)	(2.047)	(0.280)	(1.205)	(0.307)
Round 4	0.199	-0.708***	-0.348	-0.605**	0.418	-0.971***
	(1.124)	(0.186)	(1.640)	(0.261)	(1.199)	(0.309)
Round 5	0.835	-0.820***	-0.512	-0.874***	1.647	-0.811**
	(1.410)	(0.217)	(1.963)	(0.294)	(1.673)	(0.346)
Constant	11.708**	0.197	5.104	-1.161	5.981	-3.064***
	(4.721)	(1.711)	(5.551)	(1.114)	(4.764)	(1.008)
Age		-0.013		0.045**		-0.011
		(0.023)		(0.021)		(0.035)
Locus of Control		-0.069*		-0.022*		-0.110***
		(0.039)		(0.011)		(0.027)
Extraversion		0.136*		0.143***		0.214***
		(0.077)		(0.038)		(0.048)
Agreeableness		0.052		0.022		-0.049
		(0.046)		(0.043)		(0.041)
Conscientiousness		0.010		-0.297***		0.101
		(0.044)		(0.057)		(0.093)
Neuroticism		-0.008		0.032		-0.030
		(0.055)		(0.026)		(0.044)
Openness		-0.035		0.092***		-0.007
		(0.053)		(0.024)		(0.093)
var(e.Credits)	25.699***	1.000 [na]	23.031**	1.000 [na]	24.304***	1.000 [na]
	(7.266)	(0.000)	(10.401)	(0.000)	(8.642)	(0.000)
corr(e.Buy,e.Credits)	-0.455		-0.486		-0.712***	
	(0.300)		(0.567)		(0.180)	
var(u)	68.220**	2.265***	13.868	2.926***	62.458***	4.198***
	(31.794)	(0.382)	(13.995)	(0.743)	(8.583)	(1.353)
corr(u.Buy,u.Credits)	0.428***		0.107		-0.497***	
	(0.160)		(0.246)		(0.065)	
log likelihood	-1172.39		-585.93		-561.92	
sample size	1315		665		650	

Notes: Figures in parentheses ( ) are standard errors clustered by participant. The *** denote statistically different from zero at 1%, ** at 5% and * at 10% level of significance.

Table A2: Full Estimation Results of RE Ordered Probit for Feedback

	Full Sample (1)	RaMT (2)	ReMT (3)
	Coef. / (std.errors)	*Coef. / (std.errors)*	*Coef. / (std.errors)*
ReMT	0.272		
	(0.176)		
Buy	0.016	-0.029	0.097
	(0.139)	(0.158)	(0.245)
Partner's Score	0.084***	0.092***	0.082***
	(0.014)	(0.016)	(0.028)
Partner's less Own Score	0.019**	0.006	0.035***
	(0.008)	(0.009)	(0.013)
Female	0.388**	0.416**	0.354
	(0.171)	(0.191)	(0.314)
Round 2	0.066	0.040	0.099
	(0.106)	(0.157)	(0.134)
Round 3	0.118	-0.031	0.299
	(0.115)	(0.145)	(0.199)
Round 4	0.284**	0.211	0.363
	(0.117)	(0.132)	(0.235)
Round 5	0.120	0.019	0.222
	(0.133)	(0.160)	(0.249)
cut1	0.225	0.490	-0.249
	(0.419)	(0.446)	(0.882)
cut2	0.810*	1.205**	0.193
	(0.435)	(0.485)	(0.887)
cut3	1.969***	2.290***	1.464
	(0.478)	(0.543)	(0.920)
cut4	3.642***	3.868***	3.276***
	(0.533)	(0.625)	(0.966)
σ_u^2	1.574***	0.849***	2.802***
	(0.350)	(0.263)	(0.861)
log likelihood	-1244.22	-671.06	-560.08
sample size	1315	665	650

Notes: Figures in parentheses ( ) are standard errors clustered by participant. The *** denote statistically different from zero at 1%, ** at 5% and * at 10% level of significance.

Appendix 2: Printed Instructions from the RaMT

Summary of Instructions

These printed instructions are identical to the instructions on the computer screen and are provided for your reference.

About this experiment:

All participants receive the same information about the way this experiment is conducted.
As with all economics experiments, we will not use deception of any kind. Everything stated in these instructions is true.
You will receive payment for participating in this experiment. The amount of money you actually leave with will depend on your actions in the experiment. You will be paid privately in cash at the end of the session.

Payment:

Throughout the experiment we will refer to payments in Exp$ or Experimental Dollars.
At the end of the experiment these will be converted into Australian Dollars at the rate of **10** Exp$ = AU$1.00.
The experiment has five rounds, and you will be paid your earnings from each round. You will also be paid a show up fee of AU$10.00 which is paid to all participants.

Experimental Rules:

If you do not comply with these rules you will not be paid your earnings from this experiment. You will not be allowed to participate in future sessions.
1. Talking is not permitted during the experiment. You must not share any information with others during the experiment.
2. Only the experiment windows are permitted to be open during the experiment. You are not permitted to operate other software such as email or internet during the experiment. This also applies to your own devices including mobile phones.
3. If you have a question, please raise your hand and an experimenter will come answer your question. Experimenters can answer questions about procedures but cannot provide you with advice about decisions. You must make decisions by yourself.

Slider Task:

In this experiment you must undertake a series of "slider tasks" which works as follows:
Each slider is initially positioned at 0 and can be moved as far as 100.
Each slider has a blue number to the left showing the target value.
Each slider has a black number to its right showing its current position.
You can use the mouse in any way you like to move each slider. You can readjust the position of each slider as many times as you wish.
You need to make the current position match the target value for each slider. Everyone has the same targets.

Trial Slider Task:

On the screen are 3 sliders. The purpose of this trial is to ensure that you understand how to place the slider at the target. This trial will not end until you have successfully placed all three sliders.
You will not be paid for the trial.

Structure of the experiment:

This experiment has a practise round and then 5 rounds for which you will be paid. In each round participants undertake a slider task. Performance on the slider task will play a role in how much you are paid.
In each round of the slider task you will be presented with a screen with 48 sliders. You will have 4 minutes (240 seconds) to complete as many sliders as you can.

You will be randomly matched with one other participant in every round. You will be matched with a different person each round. You will never learn who the person is that you are matched with.

Your payment and your score:

In each round your matched participant will see your score for that round, and give you feedback on your score.

Your score is determined in the following way in each round. You will be credited with 1 Exp$ for each slider you correctly place in the available time. Before your score is shown to your matched participant, you will have the opportunity to spend the Exp$ you earned on the slider task in that round to buy credits to increase the score that your matched participant sees. Each credit will cost you 1 Exp$. You can, if you choose, spend up to the total amount of Exp$ that you have earned in that stage.

Your matched participant will see your score in each round which is calculated by:
Score = the number of sliders correctly placed + any credits bought

Your payment in each round is calculated by:
Payment = the number of sliders correctly placed - any credits bought

Your matched player will only ever see your score.
The number of sliders you place and the number of credits you buy are never revealed to other participants.

Feedback:

In each round, the pair of matched participants see each other's score in that round. That means that you will see your matched participant's score and they will see yours. You will provide feedback to your matched participant and they will provide feedback to you.

Each participant will then provide feedback to their matched participant on their score by answering the following question:

You must now select the feedback you would like you send your matched participant about their score from the following options: Very Good Good

Neither good nor bad
Bad
Very Bad

The participant will then see this feedback. For example if the matched participant chose 'Neither good nor bad ' then the participant would see the feedback as follows:

Your matched participant thinks your score in this round is: Neither good nor bad

ID:

You will be randomly assigned an ID for the entire experiment from the list below.

GREEN
YELLOW
WHITE
RED
PURPLE
CYAN
BROWN
ORANGE
MAGENTA
BLACK
GREY
BLUE
RECTANGLE
TRIANGLE
HEXAGON
CIRCLE
PENTAGON
TRAPEZOID
DIAMOND
SQUARE
MOON
OCTAGON
STAR
OVAL

You will never learn which participants are assigned to which IDs.
You will be shown your randomly assigned ID after you complete the quiz.