Does being observed enhance self-control?

Florian Diekert and Kjell Arne Brekke: January 11, 2019

Man is a social animal, and it is not surprising that behavior in a group differs from behavior in solitude. In a review of the literature, Charness and Sutter (2012) claim that groups are more selfish, cognitively more sophisticated and have better self-control. The aim of the current study is to look into the last claim.

What does it take to be a group? Here, we study the weakest possible group effect, two persons who never meet and are anonymous to each other, but where one of them observes some of the actions of the other and where they have a joint interest as their payment depends on the performance of both. The question is whether the fact that the person is being observed, although anonymously, will affect their ability to self-control. Alternatively, the presence of a partner with shared interest, may affect self-control.

There is a large literature on the effects of peers on productivity. Falk and Ichino (2006) find that a pair of workers have higher and more equal productivity if they sit in the same room and can observe each other, than if they are in different rooms and work in solitude. In a meta-study of 35 published paper, Herbst and Mas (2015) find that both laboratory and field studies show a significantly higher productivity for workers working alongside others rather than working in solitude. Note that while higher productivity in groups is an indication of better self-control, the studies are not measuring self-control as such but productivity.

A different strand of literature is social facilitation theory due to Zajonc (1965). The presence of an audience leads to better performance in some cases – simple repetitive tasks – and worse in other – complex tasks. While there is a rich empirical literature on this theory, we have not found any studies on the effect of an audience on self-control.

While the literature cited above hints at a positive effect of an audience on self-control, none of the studies directly targets the question. The idea is not new, though. For example, a common recommendation for people who want to quit smoking is to tell a friend, presumably because an audience will enhance self-control. But based on this rather anecdotal evidence, it is a daring claim that the minimal audience that we propose in our experiment should have profound effects on self-control. The reason why we still put forward this hypothesis is the results from a previous similar study (Diekert and Brekke, 2019).

In that study subjects where guessing the most popular answers to some questions under time scarcity. Subjects could borrow time from future question at a 100% interest rate, and most subjects indeed did do so, even when this obviously was not optimal. This is consistent with a previous study by Shah et al (2012). The new feature in our design (Diekert and Brekke, 2019) was that subject could take time

from a partner playing after them, and where the players' earning was proportional to the total number of correct answers from them both. Subject did take time from their partner; this is also not optimal. We then compared two conditions, one where the partner was a real person and one where the partner was a computer who would score the same number of correct answers as a real person left with the same amount of time. While players did not take significantly more from their computer partner than from a real partner, this difference was largest when in the start of the game and declined towards the end. In the beginning subjects are borrowing from themselves. Only when they have exhausted their own time they start taking from the partner, and only then does the message appear on the screen pointing out that they are taking from their partner. Thus, subjects without real partner borrow more from themselves than they take from their partner.

We think it is reasonable to consider the amount of borrowing in the first four periods as an issue of self-control. While borrowing is quite obviously suboptimal, due to the high interest rate, it is very easy to become absorbed in the game with the short time budget. Thus, it takes some effort to get out of the tunnel and press the forward button. Players are more able to do this when a real player follows after them. This real second player will get a time budget that depends on the first players; hence, the second player can infer something about the first player's action. Because the real second player will also earn the same as the first player, pro-social preferences may also enhance self-control.

Design

We follow the main features of the design in our previous study (Diekert and Brekke, 2019) which again is based on Shah (2012).[1] Participants are recruited through Amazon Mechanical Turk to play a guessing game known from the TV-show Family Feud. The participants will guess the five most popular answers to questions such as "Name five things you take on a picnic". Participants are matched in pairs and paid for each correct answer they, or their partner, give.

Initially, participants have 15 seconds per questions, but they can borrow time from future questions. This has a 100% interest rate, thus each second used takes two seconds from the time budget for future questions. The first player in the pair can further, when their time is exhausted, take time from their partner. The same 100% interest rate applies; each second they spend after their budget is exhausted will reduce the partner's budget with two seconds.

We have four treatments in a 2 x 2 design. The treatments differ along the dimensions **observability** and **group composition**. In all treatments, the first player can take time from the second player with a 100% interest rate. For **group composition,** the two alternatives are either *Pair (P)*, where there are two players in a sequence, or *Alone (A),* where the second player is a computer that will score like a

---

[1] Details on experimental design of our previous study can be found in the corresponding study plan: https://www.socialscienceregistry.org/trials/2582/history/23269

real player given the same amount of remaining time. For **observability**, the two alternatives are either *Observable (O)* or *Non-observable (NO).* In the observable treatments, the first player will be told that the second player will be informed about the first players borrowing behavior. The second player will be told about total borrowing in round 1-4 as well as total borrowing in all rounds, and how much time the player has taken from the second player. As borrowing in rounds 1-4 is a key outcome variable we will not highlight this unequally across treatment, thus instruction to first player will only state second players will learn about their borrowing behavior. In the *Alone* x *Observable (AO)* treatment, this second player is not in the same pair, but a player who plays the same game after the first player has completed.

The 2x2 design is thus as follows:

\ Group type Observability	Pair	Alone
Observable	**PO** A real partner will be informed about first players borrowing behavior	**AO** Second player virtual, but some real player will be informed about first players borrowing behavior
Non - observable	**PNO** A real partner who only learns his own time budget with no information about partners behavior	**ANO** Virtual second player and no-one learn about first players borrowing behavior.

Definition – key variables

The main hypothesis in this project is about the amount of self-control for the first player. Consider first the hypothetical case where the player cannot take time from the second player. The results both the results from Shah et al (2012) as well as our own initial results show that the optimal time use is to use 15 second on each question. Any borrowing from the players own time budget is inefficient. The temptation to borrow is largest at the beginning, then there is much time left. Thus, a possible measure of self-control is the borrowing in the first rounds. In the instructions, we focus in particular on the first four rounds, so total borrowing in the first four rounds would be a measure. The next question is how to deal with the possibility to take from the second player.

With the possibility to take from a second player, there are two possible motives for borrowing in the first rounds. (1) The player borrows with the intention to take time from the partner. (2) The player borrows from a lack of self-control. For case (1), consider a player who decides to borrow half the time budget of the partner, but otherwise have perfect self-control and is able to spend the exact same time on all rounds. This would imply spending about 16 seconds per round, rather than 15 seconds. This amount of borrowing in the first four rounds is very small relative to the amount of borrowing observed in the corresponding treatment of our previous study (there, participants borrowed 20.6 seconds, difference is strongly significant, *p<0.001*). Note also that the first player will first borrow from his own future time budget. The message that he is borrowing from his partner will only appear

when his own time is entirely exhausted. We thus think that the total amount of borrowing in the first four rounds is a good measure of self-control, and that variations in the intention to take from the partner will represent negligible noise, even if it should correlate with the treatment.

Main outcome variable: $B_4 \sim$ The remaining time budget after 4 rounds (=initial time budget minus what has been used in the first four rounds (at the task or for paying interest), the lower the more time a participant borrows). Measure of self-control; a higher value of B_4 is better self-control.)

The following variables are included as control-variables in potential regressions on the main effect.

Total time budgeting: $B_{12} \sim$ The remaining time budget after 12 rounds (=initial time budget minus what has been used in the twelve rounds of the game (at the task or for paying interest), the lower the more time a participant borrows, a negative value for the first player shows how much he/she borrowed from the second player, B_{12} cannot be negative for second players).

Score; individual and pair: $S_i \sim$ The score of player $i=1,2$; number between 0 (when no answer to no question has been found) and 60 (when all answers to all questions have been found). $S \sim$ Score for the pair.

Gender: G

Age: A

Know Family Feud: K

Pair: $P \sim$ Dummy for solitude treatments ($P = 0$ for alone, $P = 1$ for pairs.)

Observed: $O \sim$ Dummy for treatments where later players observe borrowing. ($O = 1$ when observed, $O = 0$ otherwise.)

Hypotheses

We have two main hypotheses and one less central hypothesis.

Hypothesis A. $E(B_4|O = 1) < E(B_4|O = 0)$
Hypothesis B. $E(B_4|P = 1) < E(B_4|P = 0)$
Hypothesis C. $E\big(B_4 \,\big|\, P = 1 \,\&\, O = 1\big) < E(B_4|P = 0 \,\&\, O = 0)$

The first hypothesis is that being observed improves self-control, and the second is that knowing that you share payment with a real person also improves self-control. Note that, unless there is a strong negative interaction effect, the last hypothesis follows from the two previous. Our results from a

previous study strongly indicate C, but neither our results nor the literature gives any basis for a decomposition. It is even possible that both A and B are wrong but hypothesis C holds, hence we include this as well. As better self-control also improves the score that can be achieved, we expect a similar relationship as in Hypothesis A-C with score, S, replacing borrowing B_4.

Additional hypothesis – peer effects

Second players will observe the productivity of the first player. Peer effects would suggest that a participant will be more productive when following a productive first player compared with following an unproductive one. A problem here is that first players will differ in how much they take from their partner. The one who takes the most have spent more time and thus should be able to score more. Still, as argued above, the high interest rate implies that player who take the maximum from their partner only get about a second extra per round.

Hypothesis D. In a regression on second players performance: $S_2 = aS_1 + bT + cS_1 \cdot O$, the coefficient, c, for being an observer, has a significant positive coefficient.

Tests and regressions

We will test all three hypotheses in pairwise t-tests, while we test hypothesis D using a linear regression as stated.

We pointed out above that perfect self-control combined with an intention to take time from the partner will have negligible effect on the time allocated to each round. But many players may not understand this. With an intention to take time from the partner they may thus borrow more than what would be consistent with equal allocation of time across rounds. Thus, an intention to take from a partner could explain borrowing in the initial rounds. We thus intend to do a regression on B_4 including control variables *A, K, T, G,* as well as the treatment variables *P, O,* and the interaction *PO* as explanatories.

Under hypothesis A we then predict a significant negative effect of O while hypothesis B suggest a significant effect of P. Note that hypothesis C is a combination of A and B as well as a possible interaction effect. We do not have a specific hypothesis about the interaction effect.

Power calculations

In our initial experiment, the current treatments were not included as they are now. But, two of the treatment from that study are somewhat close and can be used as an estimate of expected effect size. Using time left after period 4, we find an effect size *d*=0.3. Below we give power for a two means t-test with 200 observation in each group.

The effect that is most similar to the effect in the previous experiment is hypothesis C. Using "JS" and "VSP" in Diekert and Brekke (2019) as a basis, and assuming the same effect size, we get a power of 85%. Note that hypothesis C uses only data from two of the four treatments, thus only half the sample. Assuming further that one of the two hypotheses A or B account for 2/3 of the effect, we get a power of 80% for the stronger of these effects. The smaller effect reduces the power, but on the other hand, we can use the entire sample of 800 observations.

While the effect in Diekert and Brekke (2019) may have been spuriously high, the design now highlights the difference much more: We point out that it is smart to borrow as little as possible in the first four rounds and we highlight the observability much more than before. Thus, we expect that this will enhance the effect and we think that the experiment has reasonable power.

We present no separate power calculations for hypothesis D as this is not the main target of the experiment.

Exploratory work

The experiment is much richer than the tests we plan to conduct. Thus, there is ample scope for further exploratory studies. Do subjects behave differently when they start borrowing from their partner – will they start skipping questions? If so, does such behavior differ across treatments? How does borrowing behavior and self-control differ with participant's characteristics such as age, gender, or education?

References

Charness, G., & Sutter, M. (2012). Groups make better self-interested decisions. Journal of Economic Perspectives, 26(3), 157-76.

Diekert and Brekke (2019): Groups discipline resource use under scarcity. Unpublished

Falk, A., & Ichino, A. (2006). Clean evidence on peer effects. Journal of labor economics, 24(1), 39-57.

Herbst, D., & Mas, A. (2015). Peer effects on worker output in the laboratory generalize to the field. Science, 350(6260), 545-549.

Shah, A. K., Mullainathan, S., and Shafir, E. (2012). Some consequences of having too little. *Science*, 338(6107):682-685.

Zajonc, R. B. (1965). Social facilitation. Science, 149(3681), 269-274.