

Social norms or low-cost heuristics: An experimental investigation of imitative behavior

(Discussion Paper No. 2014-2)

Reply to Referee #1's comment

March 28, 2014

Response Letter

We would like to thank the Reviewer for reading carefully our paper, submitting her/his valuable comments and also for grasping the novelty of our paper. Below we report the original comments and observations (in italics) and explain, through a list of point-by-point replies, how the revised version of our paper would deal with the concerns raised by the Reviewer.

Response to Reviewer #1

1. Given the title of the paper, the authors must have been disappointed with their findings. As they say in their abstract "our results do not provide strong evidence for imitative behavior". I am rather surprised that they did not change their title as a consequence. Be that as it may, they carried out an interesting and unusual experiment. Subjects were asked to perform a cognitively demanding task - that of deciding which of 16 figures was closest to a given, target, figure. Their design was 2 by 2 - one dimension being the cognitive difficulty of the task (either 'high' or 'low'), and the other being information given about a particular suggested figure (either 'one of the most likely' or 'one of the most likely and chosen by a majority in a previous run of the experiment'); these latter are referred to as 'default' and 'majority'.

We agree with the Referee on the need of changing the title in light of our results. Thus, we have modified the title into "Social norms or low-cost heuristics: An experimental investigation of choice behavior". By replacing "imitative" with "choice", we believe that the new version of the title is less misleading about the results of our paper.

2. One problem is that, once subjects click on this default or majority card, they are stuck with it and can no longer go on exploring the other 16 cards. Obviously a lot of exploration was carried out by the majority of the subjects,

and therefore I am not surprised by the numbers in Table 5: out of 154 subjects only 8 chose the default/majority card. Although these numbers increase as we go through 'low-default', 'high-default', 'low-majority' and 'high-majority', they are far too small to do any serious statistical testing, despite the authors' valiant attempts to do so in the first paragraph on page 17. This implies that one hypothesis advanced in the paper cannot be reliably investigated. I cannot help feeling that the way that this was set up was not too clever: subjects were told, in addition to the information that the majority card "represents one of the 8 best cards appearing covered on the screen" (which was also the case in the default treatment), they are told that the default card "has been chosen in a previous experimental session by the majority of participants". Although I do not have the Instructions for that treatment to hand, if this is so, then it seems a very weak incentive for subjects. If it was meant to induce social imitation, it is difficult to imagine subjects feeling any kind of social connection with subjects "in a previous experimental session". I cannot help feeling that this why they did not observe much social imitation.

We would like to clarify the way in which participants can interact with the default or majority card. Experimental subjects cannot inspect this card. The rationale for this is that we want them to rely only on the information about this card that we provide them in the instructions. This allows us to test subjects' beliefs about the goodness of the card, without the confounding factor of subjects being influenced by what they see on the card.

We agree with the Reviewer that we do not observe statistically significant imitative behavior, even though in previous experimental literature the information about what a majority did in a previous experimental session created some sort of peer effects. Nevertheless, we believe that our results regarding the beliefs about the quality of the default or majority card are quite interesting, since participants think that the majority in the high-cost treatment performed the best choice, but the card representing the majority card is not chosen more frequently. The results about the beliefs are statistically significant and led us to conclude that the imitative heuristic was weaker than the timing heuristic (saliency of more recent memories), that our results clearly show in Section 3.3.

3. So this leaves us with a comparison of the low and high treatments. Here one the results are in line with what one might expect: payoff is lower in the high cognitive difficulty treatments. Yet at the same time (perhaps this is part of the same hypothesis) time spent is lower. There are minor differences between the numbers of uncovered cards and in the timing of decisions (though I would regard a significance level of 9.6% - foot of page 18 - to be pushing it a bit) but they are hardly dramatic.

Since the experimental task implemented in the paper is novel, we could not take for granted the difficulty of the task in the two versions (more or less squares in the figures). Hence, we believe that this result is not trivial, since it validates our novel experimental task, which can be further exploited in other future experiments in which one wants to model a more difficult versus an easier condition.

4. *I wonder if the subjects were given pen and paper to record their thoughts about the figures as they worked through them? I somehow feel not. Then they have to remember what they found and what they thought.*

We report in the paper, in Figure 2, the page participants faced on their desks, together with a pen. As we write in Section 2.1, “At the beginning of each session, participants are endowed with a printed page consisting of the contours of the 16 cards divided into rows of four, as they appear on each computer screen. The idea is to facilitate subjects in remembering the fitness of the cards, by allowing them to take notes related to them. An example of the page is provided in Figure 2.”

We can further and more clearly highlight this point in the modified version of the paper. As the Reviewer states, we thought it would be difficult for them to remember the goodness of each card. At the end of each session, we removed the papers from the desks and we stored them. Some papers contain drawings made by participants, but the majority of the papers stored which are not blank report indeed different kinds of notes related to the fitness of the card: some subjects wrote percentages on the cards in the papers, others pluses and minuses, etc. In the revised version of the paper we will hence include the percentage of participants, in each treatment, that used the papers provided to them in order to better remember the fitness of the cards inspected. This would help to shed more light on how participants came up with their decisions, as the Reviewer suggests in the following point.

5. *Given the nature of the experiment, I would have thought that one key interest would have been in the process by which subjects came to their decisions. Unfortunately it is difficult to see how the data recorded (presumably the figures clicked on and their timing) can tell us anything about that. Indeed the whole design of the experiment seems to inevitably imply that one can infer nothing about what they were thinking about as they searched for the best figure. True one can infer something rather superficial about the effect of changing the cognitive difficulty, but that is all.*

The experimental design allows us to keep track of the temporal sequence of the cards inspected by subjects, as well as the time spent inspecting each card. We can hence record the number of times each participant inspects each card, how many times they changed their mind, and also the pattern in which they inspect the cards on the screen (for instance, from top left to bottom right). We believe that this feature of the experiment is relevant since it allows us to report an important heuristic we found in our data: no matter what information subjects received, and the beliefs they had about the default or majority card, experimental subjects chose more frequently the most recently inspected cards. The paper contains a selection of the most important patterns that emerge from the data. Nevertheless, the revised version of the paper will contain a more detailed analysis of participants’ choices.

6. *I think that subjects only played the game once. This is a pity as practice and repetition may have changed behaviour.*

The aim of the experiment was to test choice behavior when the given al-

ternative is not repeated over time. Indeed, we were interested in studying a decisional process which is always new, and thus does not have room for learning dynamics. Nevertheless, it might be interesting to study how participants behave when learning dynamics are taken into account. We will introduce this point in the concluding section of the paper.

7. I am unsure about the belief elicitation task and about its validity. I quote from the Instructions: "After you have made your choice ... the computer will ask you to specify the amount (number?) of squares for which you think the DEFAULT card (that you cannot visualize) differs with respect to the TARGET card. In case your answer is correct, you will receive an additional 1 euro to your final earning." It seems from the reported statistics that 45 of the 154 subjects got the answer right. I find this amazing, unless I have misunderstood what was going on. Perhaps something has been lost in translation? Subjects cannot see the default card and have absolutely no idea how many squares are in the default card or where they are. There is not even any information about the possible number of squares, as different figures have, according to the Instructions, "a different number of squares". How can 45 out of 154 get it spot on? I would have thought that the chances were close to zero. If I had been a subject I would have thought that, and would have guessed at random. But even if I had some information, the incentives are weak - one euro for a spot-on guess and zero otherwise. I need to know more.

At the end of the day, it is difficult to decide what to recommend. I like the task posed to the subjects and the software. But I think that the way that social imitation was meant to be induced is remarkably weak - and undermines the whole purpose of the experiment. Changing the cognitive difficulty is also a clever idea, but the results are rather superficial - in that they tell us that subjects find more difficult the task that is cognitively difficult, but not why nor what they do about it. I would recommend publishing it as an interesting experiment, but without insightful results.

We thank the Reviewer for pointing out that the part of the risk elicitation task is not clear. We will include a better explanation of this task in the paper. In the low-cost treatments, each of the 16 cards differs from the target card for 5 to 20 squares that changed position. So, among the 16 cards, there is one card with 5 squares changed, another one with 6 squares changed, and so on. Each card corresponds to a different number of squares that changed position with respect to the target card. At the end of the experiment, once the task has been completed, we ask subjects the following question "By what amount of squares do you think that the default (or majority) card differs from the target?"

In the low-cost treatments, numbers from 5 to 16 were displayed and subjects had to click on one of them, in order to end the experiment. In the high-cost treatments, since the 16 cards could differ from the target card for 20 to 80 squares (20, 24, 28, 32, and so on, until 80), the numbers were displayed accordingly and subjects were asked to choose one of them.

Hence, subjects were not asked to write the amounts on their own, without any prior information of the possible numbers, but were asked to select one among the numbers displayed. In both scenarios, participants were not told that the default/ majority card was chosen at random. In the default treatments, they

were told that it was one among the 8 best cards on the screen, not that it was chosen at random among these 8 cards. Similarly, in the majority treatments, they were told that the majority card was one of the 8 best cards, and also, that it was chosen by the majority of subjects in a previous session of the experiment. Thus, it is possible that participants made up their own beliefs. Still, 45 out of 154 participants guessed the fitness of the default/ majority card correctly. This was quite surprising also for us. The way in which we explain this, and we will stress this point more deeply in the new version of the paper, is that the information we provided to participants allowed them to form correct beliefs about the quality of the default or majority card. On top of this, even if they thought that the default or majority card was the best one, they did not choose it in most of the cases. This allows us to state that, on average, they have higher expectations about their inspection, and also, that the prevailing heuristic that subjects follow is the temporal one.