Review of: "Trust is the best policy. Game theoretical analysis of bias in elicitation procedures in linguistics"

Zacharias Maniadis[1]

1 University of Southampton

Referee Report for QEIOS

'Trust is the best policy. Game theoretical analysis of bias in elicitation procedures in linguistics.'

This article employs game theory to assess whether elicitation processes of intuitive judgements in linguistics can meaningfully employ non-naïve participants. In particular, when 'consultants' familiar with linguistics are asked to make an assessment about a statement, it is likely that their knowledge will bias their assessment. The authors use formal mathematical theories to show that this should not be a major concern in most cases. I first need to note that the domain of linguistics is far beyond my range of knowledge. Accordingly, I will only comment on the application of game theory, and from the perspective of social science. I have several concerns, discussed below.

- First of all, in page 3 the authors seem to underappreciate the importance of inconsistency in elicitations of human judgement. A very large methodological literature in psychology - on the measurement of constructs - examines issues such as interrater reliability, consistency of judgements across time, susceptibility of judgements on arbitrary anchors, etc. These are important considerations for empirically examining the reliability and meaningfulness of elicitations, and perhaps this literature could inform the discussion in the beginning of this article.

- In page 4, where the basic components of a game are presented, the language could be improved to avoid confusion and make the discussion consistent with existing conventions. 'Optimal strategy' should be replaced with 'best response' which explicitly indicates that the strategy is optimal conditional on a particular choice of the opponent. Similarly, a Nash Equilibrium is a set of strategies, one for each player, that are best responses to each other. In addition, the phrase '*her best strategy is to assume that there is a 50% probability for Driver 2 to turn left ..*'. is problematic, since strategies do not involve choosing beliefs, and even if they did, assuming 50-50 play is not desirable in itself. Alternative wording could be used to explain mixed Nash equilibria. Similarly, in page 6, the meaning of the sentence '*Since games with imperfect information have … to pursue the optimal strategy*' is unclear.

- The analysis of Table 3 in page 6 (positively biased consultants) seems particularly problematic, since this situation should not be modelled as a game. The linguist has fixed expectations about what they expect to find, and this is not a decision variable (a strategy). Since the linguist is not choosing a strategy, the game-theoretic analysis that follows is not very meaningful. Put in evolutionary terms, the distribution of expectations in the 'population of linguists' (for *P* or *notP*) is fixed, and it is not an equilibrium outcome. The same point holds for the analysis of Table 5 in Page 9.

- Another issue concerns the distributional assumptions made in the article. For instance, assuming that the consultant

has no information about the researcher's expectations (so 50-50 beliefs are warranted) is crucial in driving the results of the first set of games. A large literature on "researchers' bias" and "experimenter demand effects" indicates that this is not always achieved in social science research. Similarly, the simple assumption of page 10 that there are equal shares of each type of consultants (at least according to linguists' beliefs) could be weakened and the analysis could become more realistic and interesting. The (non-justified) assumption of equal probability also underlines the analysis of Table 9 in page 14. Interestingly, in page 16, the authors drop the approach of agnosticism in terms of relative probabilities on the population, claiming that pure liars are extremely unlikely in the population. Such a more refined approach in terms of discussing the plausibility of different distributions and performing robustness checks would be useful in the other parts of the analysis as well. It is worth noting that some of the issues discussed in this paragraph are discussed in part 5, but in a rather shallow way.

- One also needs to note that game theory and decision theory are branches of mathematics, and they are useful when they can provide insights into complex problems that are not prone to obvious solutions and conclusions. Accordingly, analyses such as the ones of Table 7 can be omitted, since they add few new insights, but tend to make the exposition heavier.

- Finally, the article needs a careful proofread, as there are grammatical errors and exposition problems: for instance, the awkward editing of the last equation in page 11 (the one for A(v)).