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

Daniel Stone[1]

1 Bowdoin College

Potential competing interests: No potential competing interests to declare.

While I appreciate the motivation here - I like the idea of applying game theory to a wide variety of topics including disciplines beyond the social sciences - there are several basic errors (in addition to the numerous typos noticed by other reviewers) and I don't think we learn much from this article in its current state. Let's start with the paragraph after Table 1. The author writes “there are two optimal strategies for each player.” This is incorrect, or at least misleading (Left is a best reply to Left, and Right is a best reply to Right, but it isn't correct to refer to either or both strategies as simply “optimal strategies”). More problematically, the next sentence is: “The driving game has perfect information, since each player knows the other player’s available and optimal strategies.” This is in fact a game of imperfect information since the players do not know the other's action before acting themselves. (That is, in the extensive form representation of this game, not all information sets are singletons.) The next sentence is also a major problem: "The players’ optimal strategy is the so-called **Nash equilibrium**." A Nash equilibrium is a strategy profile, one strategy for each player, such that each best responds to the other(s). It's incorrect to refer to a single “optimal strategy” as a Nash equilibrium, and would also be incorrect to refer to “optimal strategies.” (A Nash equilibrium can be inefficient!)

I'll now jump to Section 3. The idea of analyzing a game with a possibly biased “consultant” is interesting but the game does not seem to be fully specified or analyzed in a standard way. If there are multiple types of consultants (or linguists), then it is standard to assume a probability distribution over these types that is common knowledge. This distribution is then used to characterize the linguist's beliefs about the consultant's actions in equilibrium (we would not just say “If the consultant is genuinely in the dark about the linguist’s expectations” - the point of game theory is to figure out what the consultant will think about the linguist's expectations, given the assumed structure of the game. This should also allow the author to characterize whether the linguist knows “whether the judgment is stabilized or not” (the answer will depend on the parameters of the game and perhaps the equilibrium solution concept, and the particular equilibrium if there are multiple equilibria).

A more minor comment is that the author appears to use unnecessarily obscure words (why not just write “truthful” instead of "vertiginal"? Wouldn't “analogically” be better as "analogously"?) and to report extremely simple equations that are normally omitted like -0.5 + (1-0.5) = 0, and again there are substantial typos (e.g. “where linguist’s strategies are in the first column and consultant’s strategies in the first row” should I believe be “where the linguist's strategies are the rows and consultant's strategies are the columns”). But my main suggestion is the author represent the game(s) with biased consultants in the extensive form and then solve for equilibrium/ia throughout so the analysis is more standard and

interpretation becomes much more clear.