Review of: "An Empirical Study of Goal Intentions and Monetary Compensation for Reviewers in Information Science"

Michael Falta[1]

1 University of Canterbury

Potential competing interests: No potential competing interests to declare.

The article uses a questionnaire to empirically test a small set of extrinsic and intrinsic incentives that may play a role when academics decide to accept a manuscript review in the field of information science.

Dear authors and readers:

My overall conclusion, after an about 8h analysis, is that the article makes no contribution to the understanding of the stated research question because the instrument is flawed to a degree that renders the data unusable; as a consequence, no value judgment about the drawn inferences can be made. I will substantiate below with the intention to demonstrate to the authors what research elements to pay more attention to in their next project (Points 1. to 4.). In Point 5, I summarize other aspects that would improve the paper.

1. Q4: “I work hard in the peer review process (0: completely disagree; 10: completely agree)”. Leading question. The subjective perception of what it means to ‘work hard’ is strongly culturally and physiologically influenced. It is also invalid to assume that two ‘0’ responses mean that those individuals' efforts in the review process are minimal (and to the same degree).

→ How to improve: If the researchers wanted to capture, say, ‘effort put into reviewing a manuscript', a better measuring question is: “How many hours do you spend, on average, to review a 15-page article?” Further specifications on the contents of the article may be needed.

→ Questions 5 to 8 suffer from the same problem of self-assessed perception based on leading words. E.g., Q5 “The quality of my reviewer report is very high.”

2. Q1: “When receiving an invitation to review a high-difficulty manuscript, my goal intentions for the evaluation of the manuscript are: (Check all that apply) …[a couple of options]". Several problems here. Firstly, there is no explanation of what a ‘high-difficulty manuscript' is. What is perceived to be difficult by one person may not be perceived anywhere close to that difficulty level by another person. Perceptions about the difficulty of a task are strongly correlated with the experience of a person, for example. Secondly, is it fair to say that not all manuscripts submitted to the 4 journals are highly complex (a better word for ‘difficulty’)? What about those? Third, the academic review process is assumed to be a so-called ‘blind peer review process’ for good reason: when people know one another, they often mutually scratch their

backs, which does not help to make objective decisions at the editor's desk. Why then would one of the response options be ‘networking opportunities’? Fourth, the response options are incomplete and miss a major ‘goal intention’: duty. It is generally accepted that professors at universities have a responsibility to volunteer and do take their role in society seriously. At many universities, this expectation is explicitly stated in the job role descriptions.

3. Q2: “In a competitive deadline scenario, I prefer to: … write review … write grant [application] … both”. The hidden ambiguity in this question is as follows. The words ‘competitive deadline scenario’ imply that there is only enough time for one, the grant application or review. However, given that the third response allows that both documents may be written, it negates my earlier interpretation, i.e., there in fact is enough time to do both. Well, if so, then one may think about which one will be done first, if that is of any relevance. To finish the circular reasoning: It also means that because both reports can be written, it is not a ‘competitive’ scenario any longer; the question would then not ask about preference ranking of two options but rather about working diligence (although there is enough time to do both, those who answer with doing only one of the two reports cannot be bothered to do both).

4. Q3. I am not quoting this question here; please refer to the article. This question has the following major flaws. Firstly, it contains a comparative ‘I am more likely to': it is generally agreed on not to use comparatives in questionnaires when no explicit comparison is stated. Secondly, the term ‘evaluation goals’ is explained by the same set of ‘goal intention’ aspects. That is confusing, as goal intentions are in the reviewer decision process (Fig. 1) listed at Date 0, but the evaluation goals are nowhere to be found. Third, the answer options are $0, $10, $50, $100, $150, $200. Here it should be stated which dollars, given the international demographic of the respondents. Furthermore, some of these amounts are both immaterial and immaterially different from other options.

→ If the intent is to measure the amount of money offered/required that would sway a potential reviewer to accept a review from not accepting it, simply ask “Assume that you are about to click to deny accepting a review invitation for …. reason. What amount of money would change your mind?” And to cover those respondents who are not incentivised by money, simply ask “Did money play any role in the last x reviews that you have accepted (and a 2nd question for the ‘not accepted’ case): yes or no?”. (Note: The original passage contains a mix of bullet points and arrows, which may not be preserved in the corrected version depending on the surrounding content and formatting.)

5.1. Replace the tables with the bar graphs from the appendix. Much better visual assessment of data. Less duplication, i.e., no need for the appendix.

5.2. Improve level of analysis. In scientific reporting, it is not a useful strategy to calculate statistics, put them into a table, and then merely refer in the text to a table where the reader can self-assess what the numbers mean. E.g., “… Table 2 shows the correlation between …[a and b]." The next sentence then is about Table 3. What you should do is to comment on your findings. Explain them, compare them against expectations, comment on how what is found influences the rest of the analysis. To the authors' credit, 3 paragraphs later, they do conclude as follows: “…given Q1.1 and Q1.2 are positively correlated … if a reviewer chooses ”Quality Control", with a high probability, they will also choose “Helping the profession”". However, the correlation between these variables is 0.12! I don't suggest describing a correlation of 0.12 as

‘high probability’.

5.3. “From Question 5 to Question 8 of the survey, we also found that the majority of those surveyed strongly agreed that the quality of their review reports was very high”. A challenge to the researchers: How could you measure a tad more objectively the quality of a reviewer's work?

5.4. Explain all variables in expressions. E.g., the y with the bars below and above the why. Don't let the reader do the work to find out that it is the range over which the survey responses span.

5.5. “Table 6 also shows a p-value of 0.033 … with a Pearson correlation value equal to -0.169. This means that there is a negative correlation between them.” This is similar to Point 5.2. Give your analysis and importance of this result.

5.6. It is acceptable practice that widely known methods and statistics need not be explained (e.g., p-values). However, lesser known concepts need at least a reference but really warrant explicit discussion. E.g., ‘power analysis' around Table 8, and Pillai's Trace (Table 10).

Out.