Antecedents of Crowdsourcing in Science: Scale Development and Initial Testing

Regina Lenart-Gansiniec
Jagiellonian University in Krakow, Krakow, Poland
regina.lenart-gansiniec@uj.edu.pl

Abstract: Crowdsourcing in science has become a growing field of research interest. Although the importance of crowdsourcing in science along with its barriers and requirements have been well recognized in the literature, many questions remain open. One of them refers to factors that determine the intention of academics to use crowdsourcing in science. Besides, previous research is held back by the lack of a rigorous measuring instrument. The aim of the article is to identify the antecedents of crowdsourcing in science, as well as the construction and initial testing of the measuring scale. Identifying the antecedents of crowdsourcing in science followed a combined approach, i.e., systematic literature review and nine focus group discussions were used. The process of developing the scale and its initial testing includes a pilot study conducted among 66 researchers. The final scale consists of 38 items covering the five antecedents of crowdsourcing in science.

Keywords: crowdsourcing in science, antecedents, measurement scale

1. Introduction

Crowdsourcing in science is seen as an alternative to research projects (Lukyanenko et al. 2019), a strategy for organizing the work of researchers (Lukyanenko et al. 2019), a research tool (Law et al. 2017), a new way of modern research activity that enables creation of knowledge. Previous research on the antecedents of crowdsourcing in science was conducted from the perspective of biology, medicine, physics, geography, anthropology, scientific information, and engineering (Lukyanenko et al. 2019), and was limited only to scientific crowdfunding (Tucker et al., 2018) or public involvement in scientific activities (Poliakoff & Webb 2007). At the same time, the literature points out that the specificity of crowdsourcing in science depends on the type of discipline (Beck et al. 2020). For this reason, the antecedents of crowdsourcing in science in the context of management sciences may differ from others. For example, Beck et al. (2020 p. 19) believe that "studies of university-industry collaboration consistently reveal that fields including the applied sciences and parts of the social sciences, such as economics and management studies, are more prone to collaborations with the private sector, patenting, and spinout formation". In turn, the motivation to use crowdsourcing in science differs among researchers and depends on the discipline and field (Beck et al., 2020). Therefore, the research problem was formulated in the form of two research questions:
RQ1. What factors influence academic teachers to reach for crowdsourcing in science?
RQ2. How can antecedents of crowdsourcing in science be measured?

In answer to the first question (RQ1), a combined approach was adopted, i.e., deductive and inductive. This approach results from the fact that "considering that item generation may be the most important part of the scale development process, future measures should be developed using the appropriate definition of the conceptual domain based on the combination of both deductive and inductive approaches" (Morgado et al. 2017, p. 10). First, as a part of the deductive approach, a list of antecedents of crowdsourcing in science was generated based on a systematic literature review. Then, in order to detail, develop and supplement them, nine focus group interviews (induction) were conducted. In this way, we were able to provide a theoretically and empirically grounded set of antecedents in crowdsourcing in science.

Then, to answer the second research question (RQ2), it was concluded that "in order to ensure the content validity, the researcher seeks other opinions about the operationalized items. The opinions can be those of expert judges (experts in the development scales or experts in the target construct) or target population judges (potential users of the scale), enabling the researcher to ensure that the hypothesis elaborated in the research appropriately represents the construct of interest" (Morgado et al. 2017, p. 2). Therefore, a pilot study was conducted among 66 researchers. This allowed to identify and rigorous initial testing of a scale to measure the antecedents of crowdsourcing in science.

2. Scale development methodology

Given that the purpose of this article is to construct a measurement tool, the widely used scale development paradigm by Churchill (1979) was adopted, which was then extended by many researchers (Bagozzi et al. 1991). The procedure consists of five steps divided into two phases: (1) item generation (literature review and focus research group), (2) content validation. In the second phase, scale development was started, in particular: (3) sampling procedure and data collection, and (5) item purification.

3. Results

The estimation of the reliability of the scale was based on the analysis of the correlation between the questions in relation to the variance of the question (Tavakol &Dennick 2011). The entire research tool has a value of 0.939, which proves a very high reliability of the tool. The reliability of the tool can also be said to be very high. Neither of the items needed to be eliminated as they all loaded high between 0.924 and 0.979. Additionally, the raw data was tested for the risk of common method bias (Podsakoff et al. 2003) using the main Harman factors test (Podsakoff et al. 2003). The results showed that the variance of the one-way solution did not exceed 70% (Fuller et al. 2016) and was 61.80% – so there was no error of common variance of the methods.

The design validity was assessed using EFA. To verify the basic requirements of the EFA, the adequacy and sphericality of the sample were verified prior to the analysis. KMO measures indicate that it is possible to use EFA to analyse and isolate the main components (Tabachnick & Fidell 2014). The KMO result was 0.915 and the Bartlett sphericity test was 657,846, with a significance level of <0.001, indicating the validity of the use of factor analysis.

The EFA method of principal components was used to determine the internal structure of the scale. As a result, eight factors with the eigenvalue above 1 were obtained, explaining in total 82.980% of the variance. Taking into account the recent concerns about the reliability of the Fornell and Larcker criteria (1981), an additional analysis of the correlation with Pearson's coefficients was performed. The results show that the strongest relationship with behavioural intention is shown by normative pressure ($r = 0.536$; $p < 0.01$), then subjective norms ($r = 0.468$; $p < 0.01$), perceived organizational support ($r = 0.425$; $p < 0, 01$) and attitudes (0.440; $p < 0.01$). There is a weaker relationship between behavioural intentions and motivation ($r = 0.397$; $p < 0.01$), perceived usefulness ($r = 0.377$; $p < 0.05$), descriptive norms ($r = 0.374$; $p < 0.05$), perceived behavioural control ($r = 0.335$; $p = 0.006$), perceived risk ($r = 0.329$; $p = 0.007$) and perceived utility ($r = 0.309$; $p < 0.012$). Additionally, on the basis of significant correlation results at the level of 0.000 <0.05, it can be concluded that there is no relationship between the perceived effort and institutional pressure.

Taking into account the results of the analysis of the correlation between antecedents and behavioural intention, as well as the results of EFA, CR and AVE, as well as recommendations of other researchers regarding the reasons for the participation of researchers in initiatives based on public involvement (Poliakoff & Webb 2007), the following antecedents that showed the strongest relationship with behavioural intention, i.e. normative pressure, subjective norms, perceived organizational support, perceived usefulness, and attitudes. The final version of the crowdsourcing in science antecedents tool consists of 38 items including 5 antecedents of crowdsourcing in science.

4. Discussion and conclusion

This publication responds to the calls and extends the findings of other researchers regarding the explanation of why academics use crowdsourcing in science. Therefore, the publication, based on the results of a systematic literature review and focus group interviews, provides a list of 12 antecedents of crowdsourcing in science. Moreover, due to the fact that there is currently no reliable and up-to-date tool for measuring antecedents of crowdsourcing in science in the literature (Beck et al. 2020; Franzoni et al. 2021), a pilot study was conducted. Their results suggest that the 5 antecedents have a moderate impact on the intention of scientists to use crowdsourcing in science (normative pressure, subjective norms, perceived organizational support perceived usefulness, and attitudes). In contrast, a weaker relationship of behavioural intentions is observed in relation to motivation, descriptive norms, perceived behavioural control, perceived risk, and perceived utility. Additionally, there is no relationship between behavioural intention and perceived effort and institutional pressure. The obtained results are consistent with the findings of other researchers to date. They also fit into planned theory of planned behaviour (TPB) (Ajzen 1991) and institutional theory (DiMaggio & Powell 1983). Therefore, this article proposes a tool for measuring crowdsourcing in science antecedence consisting of 38 items (using the 7-

point Likert scale) taking into account 5 antecedents of crowdsourcing in science. These tools were provided and tested using the approach proposed by Churchill (1979).

Acknowledgments

This project was financed from the funds provided by the National Science Centre, Poland awarded on the basis of decision number DEC-2019/35/B/HS4/01446.

References

Beck, S., Bergenholtz, C., Bogers, M., Brasseur, T.-M., Conradsen, M. L., Di Marco, D., Distel, A. P., Dobusch, L., Dörler, D., Effert, A., Fecher, B., Filiou, D., Frederiksen, L., Gillier, T., Grimpe, C., Gruber, M., Haeussler, C., Heigl, F., Hoisl, K., Hyslop, K., Kokshagina, O., LaFlamme, M., Lawson, C., Lifshitz-Assaf, H., Lukas, W., Nordberg, M., Norn, M. T., Poetz, M., Ponti, M., Pruschak, G., Pujol Priego, L., Radziwon, A., Rafner, J., Romanova, G., Ruser, A., Sauermann, H., Shah, S. K., Sherson, J. F., Suess-Reyes, J., Tucci, C. L., Tuertscher, P., Vedel, J. B., Velden, T., Verganti, R., Wareham, J., Wiggins, A. and Xu, S. M. (2020) "The Open Innovation in Science research field: a collaborative conceptualisation approach", *Industry and Innovation*, pp. 1-50.

Behrend, T. S., Sharek, D. J., Meade, A. W., & Wiebe, E. N. (2011) "The viability of crowdsourcing for survey research", *Behavior Research Methods*, Vol. 43, No. 3, pp. 800-813.

Boell, S., Cecez-Kecmanovic, D. (2014) "A hermeneutic approach for conducting literature reviews and literature searches", *Communications of the Association for Information Systems*, Vol. 34, No. 1, pp. 257–286.

Correia, A., Schneider, D., Fonseca, B. and Paredes, H. (2018) *Crowdsourcing and Massively Collaborative Science: A Systematic Literature Review and Mapping Study. Collaboration and Technology*, Springer International Publishing, Cham.

Eklund, L., Stamm, I. and Liebermann, W. K. (2019) "The crowd in crowdsourcing: Crowdsourcing as a pragmatic research method", *First Monday*, Vol. 24, No. 10, pp. 1-14.

English, P. B., Richardson, M. J. and Garzón-Galvis, C. (2018) "From Crowdsourcing to Extreme Citizen Science: Participatory Research for Environmental Health', *Annual Review of Public Health*, Vol. 39, pp. 335-350.

Franzoni, C. and Sauermann, H. (2014) "Crowd science: The organization of scientific research in open collaborative projects", *Research Policy*, Vol. 43, No. 1, pp. 1-20.

Hedges, M. and Dunn, S. (2018) *Introduction: Academic crowdsourcing from the periphery to the centre*, in Hedges, M. and Dunn, S. (eds.) *Academic Crowdsourcing in the Humanities*, Chandos Publishing, UK, pp. 1-12.

Howe, J. (2006) "The Rise of Crowdsourcing", *Wired*, Vol. 14, No. 6, pp.1-4.

Law, E., Gajos, K. Z., Wiggins, A., Gray, M. L. and Williams, A. (2017) "Crowdsourcing as a Tool for Research: Implications of Uncertainty'", Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, Portland, Oregon, USA.

Levy, M. and Germonprez, M. (2017) "The Potential for Citizen Science in Information Systems Research", *Communications of the Association for Information Systems*, Vol. 40, pp. 22-39.

Lukyanenko, R., Parsons, J., Wiersma, Y. F. and Maddah, M. (2019) "Expecting the unexpected: effects of data collection design choices on the quality of crowdsourced user-generated content", *MIS Quarterly*, Vol. 43, No. 2, pp. 623–648.

Riesch, H. and Potter, C. (2014) "Citizen science as seen by scientists: Methodological, epistemological and ethical dimensions", *Public Understanding of Science*, Vol. 23, No. 1, pp. 107-120.

Sari, A., Tosun, A. and Alptekin, G. (2019) "A Systematic Literature Review on Crowdsourcing in Software Engineering", *Journal of Systems and Software*, Vol. 153, pp. 200-219.

Scheliga, K., Friesike, S., Puschmann, C. and Fecher, B. (2016) "Setting up crowd science projects", *Public Understanding of Science*, Vol. 27, No. 5, pp. 515-534.

Schildhauer, T. and Voss, H. (2014) *Open Innovation and Crowdsourcing in the Sciences*, in Bartling, S. and Friesike, S. (eds.) *Opening Science: The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing*, Springer International Publishing,Cham: pp. 255-269.

Schlagwein, D., and Farhad, D. (2014) "User Requirements of a Crowdsourcing Platform for researchers: Findings from a Series of Focus Groups", PACIS.