What Can NeuroIS Learn from the Replication Crisis in Psychological Science?

Colin Conrad and Lyam Bailey[1]

[1] Dalhousie University, Halifax, Canada
{colin.conrad, lyam.bailey}@dal.ca

Abstract. The Reproducibility Crisis is a phenomenon that has gained considerable attention in the psychological sciences. Scholars in these fileds have found that many high profile findings are either difficult to reproduce or could not be replicated. These findings have ultimately encouraged researchers to adopt pre-registered results, replication in study design and open data. As an emerging field, NeuroIS has an opportunity to learn from this crisis and adopt new practices based on the lessons learned in the psychological sciences. We explored the current state of NeuroIS research from the perspective of reproducibility by conducting a survey of the extant NeuroIS literature. We conclude by suggesting two practices that the NeuroIS community can undertake to help address the replication problem.

Keywords: NeuroIS · replication · research methods · construct validity

1 Introduction

Reproducibility is often regarded to be one of the defining characteristics of hypothesis-driven science. When empirical observations are reproduced through experimentation, observers gain evidence for the relationships between observed phenomena. Additional empirical observations decrease the likelihood of a false positive or false negative result, allowing scientists to be more confident about the relationship being observed. It is through this process that scientists often induce causal relationships [1]. The managerial sciences are no exception; the logical validity of our work as Information Systems scholars is improved by its reproducibility. Unfortunately, the benefits of replication are often overshadowed by motivations to publish novel, high-impact work. Many scholars will therefore forego replication of a previously observed phenomenon in favor of new, more exciting research. We posit that this may become a non-trivial concern for research in NeuroIS. As an emergent field, NeuroIS holds the potential for large or surprising gains in understanding of information systems and information technology phenomena. Scholars in this field, therefore, would be well-served to ensure that any such gains in understanding are replicable from the outset.

To tangibly demonstrate the importance of this issue, and why it should not be left unchecked, we point to the ongoing Replication Crisis in psychological science. In

August 2015, The Open Science Collaboration published a research article demonstrating that out of a selection of 100 experiments published in prominent psychological journals, 97 of which originally yielded significant effects, only 35 of the original observations (i.e. less than half) were replicable [2]. This paper, among others, sparked a wider conversation about reproducibility and openness in science broadly [3]. Fortunately, scholarly responses to these concerns have been largely positive. For example, one of the largest barriers to replication-oriented research is publication biases favoring novel and significant results. In simple terms, a replication experiment is considerably less likely to be published if it fails to reproduce a previously observed finding (i.e. obtains a null result) [4,5]. Many psychological journals have taken steps to combat this issue by endorsing so called "pre-registered reports" – dedicated article types for which the decision to accept or reject is made prior to data collection and is based entirely on the merits of the proposed methods [6]. In this vein, some journals even endorse pre-registered *replication* articles, which carry the benefits of 'regular' pre-registered reports, but which specifically foster replication-oriented research. Moreover, recent years have seen movements towards greater transparency in science. Most notably, the recently established Open Science Framework provides an online platform for researchers to pre-register experimental designs and hypotheses, and to store and share their data [7].

The field of neuroimaging faces its own unique concerns surrounding replicability [8,9]. Neuroimaging experiments require large financial investment (owing to the costs of equipment maintenance and technical staff), meaning that the feasibility of replication is constrained by available funds. Moreover, neuroimaging studies often make large numbers of comparisons. Despite these issues, replication in neuroimaging is by no means a lost cause. The recent emergence of publicly accessible neuroimaging datasets such as the Human Connectome Project [10] the Cambridge Centre for Ageing and Neuroscience database [11], and the Alzheimer's Disease Neuroimaging Initiative [12] has allowed researchers to explore new lines of research *and* explore the replicability of old ones without incurring prohibitive financial costs.

We posit that the steps taken in psychological science to counter the replicability crisis may also be applied to the emergent field of NeuroIS. NeuroIS researchers have recognized the need to adopt best practices and use multiple methods to understand the relationship between neurophysiological observations and IS constructs [13,14]. However, to the best of our knowledge there has been limited discussion about reproducibility in NeuroIS specifically. Though there has been some recognition of the importance of replication in the broader Information Systems community [15], we believe that a conversation about the value of replication in NeuroIS research is warranted. In this paper, we briefly explore the extent that replication has been incorporated in past NeuroIS research by observing the methods employed by published NeuroIS papers originally described by Riedl et al. (2017) [16]. After assessing the findings of the review, we conclude by outlining two potential practices that can be adopted in future NeuroIS research.

2 Replication and NeuroIS

In order to assess how NeuroIS researchers have incorporated replication in the past, we manually reviewed the 164 papers specified by Riedl et al. (2017) [16]. We opted to observe the previously reported works because, to the best of our knowledge, this work represents the only comprehensive literature review on the NeuroIS discipline and extending a comprehensive literature review is outside of the scope of this short paper. Given that new NeuroIS research involves either novel measures or novel applications, we opted to investigate studies that incorporated elements of replication in their research design. We identified papers based on whether they conducted empirical studies, which methods were used, whether they incorporated multiple experiments in their research, and whether the authors reported replicating their neurophysiological findings. Our analysis similarly found 103 empirical NeuroIS papers, summarized in Table 1.[1]

Table 1. NeuroIS studies identified which reported multiple experiments or replication.

Primary tool	Number	No. reporting multiple experiments	No. reporting measure replication
Eye tracking	46	7	5
EEG	22	5	5
fMRI	9	2	1
Other than above	26	5	2
Total	**103**	**20**	**13**

Though we identified 20 studies which incorporated multiple experiments, 9 of the studies described multiple experiments using different methods (e.g. neurophysiological, self-report) in an effort to triangulate findings with prior IS constructs [17-25]. We identified 10 studies which employed multiple experiments and reported successfully replicating observed results through a later experiment [26-36]. We also observed 3 papers which employed a single study but directly replicated an experiment previously reported in another paper [37-39]. Furthermore, we observed a trend that research published in the AIS Senior Scholars Basket [40] or highly recognized conferences were more likely to report multiple experiments or replicate NeuroIS phenomena. Table 2 summarizes findings from the AIS Senior Scholar’s Basket, a widely recognized basket of quality journals in the IS discipline. We observe that studies published in the basket are more likely to report multiple experiments (though not necessarily multiple neurophysiological experiments).

[1] Discrepancies between the findings of this paper and Riedl et al. [16] can be attributed to differences in how authors interpreted studies as complete or empirical, and the subjective judgement employed by us when identifying a primary research method for each study.

Does NeuroIS have a reproducibility problem?

The results of our investigation suggest that some NeuroIS researchers have taken steps to include multiple experiments or replication in their designs when appropriate, especially when publishing in highly recognized venues. Notably, we did not observe any fMRI studies which replicate their neurophysiological findings, as the one study identified replicated behavioural results [38]. This observation is likely attributable to high cost of running fMRI experiments – often around $500 USD per hour – alluded to earlier. Given that our investigation concerned publications published prior to 2017 however, we are led to conclude that NeuroIS has largely conformed to the standards in Neuroimaging, which similarly did not concern itself with replication until recently. NeuroIS would nonetheless benefit from similarly addressing the problem.

Table 2. Summary of empirical studies published in the AIS Senior Scholar's Basket reporting multiple experiments or replication

Primary tool	Number	No. reporting multiple experiments	No. reporting measure replication
Eye tracking	4	1	1
EEG	7	2	1
fMRI	6	2	0
Other than above	5	3	2
Total	**22**	**8**	**4**

3 Recommendations for Future NeuroIS Research

Consider incorporating replication in study designs

We encourage NeuroIS researchers to test the replicability of their observed findings wherever possible. Such efforts might not necessarily take the form of a so-called "straight replication" – that is, repeating a study using precisely the same methods, on a new sample of participants. Replication may instead be achieved by further exploring an observed finding. For example, a research team might observe that a particular stimulus elicits a particular neural response, as measured with non-invasive neuroimaging. The team could then explore whether this neural response is modulated by certain experimental parameters, by the demographic/cognitive characteristics of participants, or is corroborated by psychometric measures (when possible). Such a follow-up experiment would not only serve to replicate the first, but would provide a novel advancement in understanding, with respect to the neural response in question.

Such efforts might also be incentivized by prominent IS journals and conferences, which could consider special treatment for research articles aiming to test the replicability of previously observed findings. For example, having a dedicated article type for replication or for the acceptance of pre-registered reports, as is becoming more common practice in psychological science. Furthermore, NeuroIS researchers could consider disseminating replication experiments in the *AIS Transactions on Replication Research* [15] or similar venues dedicated to supporting replication.

Towards an open NeuroIS data repository

We also encourage NeuroIS researchers to explore collaborative data-sharing initiatives. In the field of Neuroimaging, publicly available repositories such as the Human Connectome Project [10], amongst others [11,12] enable researchers to test the replicability of previously reported effects, often with considerably larger samples. NeuroIS may similarly benefit from the creation of a repository of publicly accessible experiment data or a publicly available dataset. Such a resource might be achieved by means of a dedicated project (as is the case for the databases cited previously), from multiple collaborators pooling their anonymized data collected for the purposes of ongoing research studies, or even simply by making past data accessible to the NeuroIS community through the web. Calls for open software and data have been made in the broader IS community [41]; in the NeuroIS context, such an undertaking has the further potential to foster sharing of best practices and provide a resource for training research students. The final result of such an undertaking would not merely be increased transparency of NeuroIS research, but also the wider dissemination of NeuroIS research to other academic communities.

References

1. Hempel, C. G.: Maximal Specificity and Lawlikeness in Probabilistic Explanation. In: Philosophy of Science, vol. 35, iss. 2, pp. 116-133 (1968)
2. Open Science Collaboration: Estimating the reproducibility of psychological science. In: Science, vol. 349, iss. 6251, aac4716 (2015)
3. Baker, M.: Is There a Reproducibility Crisis? In Nature, vol. 533 (2016)
4. Francis, G.: Too good to be true: Publication bias in two prominent studies from experimental psychology. In: Psychonomic Bulletin & Review, vol. 19, iss. 2, pp. 151-156 (2012)
5. Francis, G.: Publication bias and the failure of replication in experimental psychology. In: Psychonomic Bulletin & Review, vol. 19, iss. 6, pp. 975-991 (2012)
6. Gonzales, J.E., and Cunningham, C.A.: The promise of pre-registration in psychological research. `https://www.apa.org/science/about/psa/2015/08/pre-registration`
7. Foster, E. D., and Deardorff, A.: Open science framework (OSF). In: Journal of the Medical Library Association, vol. 105, iss. 2, pp. 203 (2017)
8. Luck, S. J. and Gaspelin, N.: How to get statistically significant effects in any ERP experiment (and why you shouldn't). In: Psychophysiology, vol. 54, pp. 146-157 (2017)
9. Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafo, M. R., Nichols, T. E., Poline, J.-B., Vul, E., Yarkoni, T.: Scanning the horizon: towards transparent and reproducible neuroimaging research. In: Nature Reviews, vol. 18 (2017)
10. Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T. E. J., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtis, S. W., and Della Penna, S: The Human Connec-

tome Project: a data acquisition perspective. In: Neuroimage, vol. 62 iss. 4, pp. 2222-2231 (2012)
11. Taylor, J.R., Williams, N., Cusack, R., Auer, T., Shafto, M. A., Dixon, M., Tyler, L. K., Henson, R. N.: The Cambridge Centre for Ageing and Neuroscience (CamCAN) data repository: Structural and functional MRI, MEG, and cognitive data from a cross-sectional adult lifespan sample. In: NeuroImage, vol. 144, pp. 262-269 (2017)
12. Mueller, S. G., Weiner, M. W., Thal, L. J., Petersen, R. C., Jack, C. R., Jagust, W.,Trojanowski, J. Q., Toga, A. W. and Beckett, L.: Ways toward an early diagnosis in Alzheimer’s disease: the Alzheimer’s Disease Neuroimaging Initiative (ADNI). In: Alzheimer's and Dementia, vol. 1 iss. 1, pp. 55-66 (2005)
13. Dimoka, A., Davis, F. D., Gupta, A., Pavlou, P. A., Banker, R. D., Dennis, A. R., Ischebeck, A., Müller-Putz, G., Benbasat, I., Gefen, D., Kenning, P. H., Riedl, R., vom Brocke, J., and Weber, B.: On the Use of Neurophysiological Tools in IS Research: Developing a Research Agenda for NeuroIS. In: MIS Quarterly, vol. 36, iss. 3, pp. 679-702 (2012)
14. vom Brocke, J., and Liang, T.-P.: Guidelines for Neuroscience Studies in Information Systems Research. In: Journal of Management Information Systems (2014)
15. Dennis, A. R. and Valacich, J. S.: A Replication Manifesto. In: AIS Transactions on Replication Research, vol. 1, pp. 1-4 (2014)
16. Riedl, R., Fischer, T., and Léger, P.-M.: A Decade of NeuroIS Research: Status Quo, Challenges, and Future Directions. In: Proceedings of the 38th International Conference on Information Systems, Seoul (2017)
17. Dimoka, A.: What does the brain tell us about trust and distrust? Evidence from a functional neuroimaging study. In: MIS Quarterly, vol. 34, iss. 2, pp. 373-396 (2010)
18. De Guinea, A. O., and Webster, J.: An investigation of information systems use patterns: Technological events as triggers, the effect of time, and consequences for performance. In MIS Quarterly, vol. 38, iss. 4, pp. 1165-1188 (2013)
19. Goggins, S. P., Schmidt, M., Guajardo, J., and Moore, J.: Assessing multiple perspectives in three dimensional virtual worlds: eye tracking and all views qualitative analysis (AVQA). In: 2010 43rd Hawaii International Conference on System Sciences, pp. 1-10. IEEE (2010)
20. Gregor, S., Lin, A. C., Gedeon, T., Riaz, A., and Zhu, D.: Neuroscience and a nomological network for the understanding and assessment of emotions in information systems research. In: Journal of Management Information Systems, vol. 30, iss. 4, pp. 13-48 (2014)
21. Perrin, J. L., Paillé, D., and Baccino, T.: Reading tilted: Does the use of tablets impact performance? An oculometric study. In: Computers in Human Behavior, vol. 39, pp. 339-345 (2014)
22. Clayton, R. B., Leshner, G., and Almond, A.: The extended iSelf: The impact of iPhone separation on cognition, emotion, and physiology. In Journal of Computer-Mediated Communication, vol. 20, iss. 2, pp. 119-135 (2015)
23. Cole, M. J., Hendahewa, C., Belkin, N. J., and Shah, C.: User activity patterns during information search. In: ACM Transactions on Information Systems (TOIS), vol. 33, iss. 1, pp. 1-39 (2015)
24. Jenkins, J. L., Anderson, B. B., Vance, A., Kirwan, C. B., and Eargle, D.: More harm than good? How messages that interrupt can make us vulnerable. In Information Systems Research, vol. 27, iss. 4, pp. 880-896 (2016)
25. Luan, J., Yao, Z., Zhao, F., and Liu, H.: Search product and experience product online reviews: an eye-tracking study on consumers' review search behavior. In Computers in Human Behavior, vol. 65, pp. 420-430 (2016)

26. Bahr, G. S., and Ford, R. A.: How and why pop-ups don’t work: Pop-up prompted eye movements, user affect and decision making. In: Computers in Human Behavior, vol. 27, iss. 2, pp. 776-783 (2011)
27. Nunamaker, J. F., Derrick, D. C., Elkins, A. C., Burgoon, J. K., and Patton, M. W.: Embodied conversational agent-based kiosk for automated interviewing. In: Journal of Management Information Systems, vol. 28, iss. 1, pp. 17-48 (2011).
28. Astor, P. J., Adam, M. T., Jerčić, P., Schaaff, K., and Weinhardt, C.: Integrating biosignals into information systems: A NeuroIS tool for improving emotion regulation. In: Journal of Management Information Systems, vol. 30, iss. 3, pp. 247-278 (2013)
29. De Guinea, A. O., Titah, R., and Léger, P. M.: Measure for measure: A two study multi-trait multi-method investigation of construct validity in IS research. In: Computers in Human Behavior, vol. 29, iss. 3, pp. 833-844 (2013).
30. Jay, C., Brown, A., and Harper, S.: Predicting whether users view dynamic content on the world wide web. In: ACM Transactions on Computer-Human Interaction (TOCHI), vol. 20, iss. 2, pp. 1-33 (2013)
31. Dogusoy-Taylan, B., and Cagiltay, K.: Cognitive analysis of experts’ and novices’ concept mapping processes: An eye tracking study. In: Computers in human behavior, vol. 36, pp. 82-93 (2014)
32. Molina, A. I., Redondo, M. A., Lacave, C., and Ortega, M.: Assessing the effectiveness of new devices for accessing learning materials: An empirical analysis based on eye tracking and learner subjective perception. In: *Computers in Human Behavior*, vol. *31*, pp. 475-490 (2014)
33. Wang, C. C., and Hsu, M. C.: An exploratory study using inexpensive electroencephalography (EEG) to understand flow experience in computer-based instruction. In: *Information & Management*, vol. *51, iss.* 7, pp. 912-923 (2014)
34. Galluch, P. S., Grover, V., and Thatcher, J. B.: Interrupting the workplace: Examining stressors in an information technology context. In: *Journal of the Association for Information Systems*, vol. *16*, *iss.* 1, pp. 1-47 (2015)
35. Huang, Y. F., Kuo, F. Y., Luu, P., Tucker, D., and Hsieh, P. J.: Hedonic evaluation can be automatically performed: An electroencephalography study of website impression across two cultures. In: Computers in Human Behavior, vol. 49, pp. 138-146 (2015).
36. Hu, Q., West, R., and Smarandescu, L.: The role of self-control in information security violations: Insights from a cognitive neuroscience perspective. In: Journal of Management Information Systems, vol. 31, iss. 4, pp. 6-48 (2015).
37. Adam, M. T., Krämer, J., and Weinhardt, C.: Excitement up! Price down! Measuring emotions in Dutch auctions. In: International Journal of Electronic Commerce, vol. 17, iss. 2, pp. 7-40 (2012).
38. Riedl, R., Mohr, P. N., Kenning, P. H., Davis, F. D., and Heekeren, H. R.: Trusting humans and avatars: A brain imaging study based on evolution theory. In: Journal of Management Information Systems, vol. 30, iss. 4, pp. 83-114 (2014).
39. Labonté-LeMoyne, É., Santhanam, R., Léger, P. M., Courtemanche, F., Fredette, M., and Sénécal, S.: The delayed effect of treadmill desk usage on recall and attention. In: Computers in Human Behavior, vol. 46, pp. 1-5 (2015).
40. Association for Information Systems: Senior Scholars’ Basket of Journals. `https://aisnet.org/page/SeniorScholarBasket`
41. van der Aalst, W., Bichler, M., & Heinzl, A. Open research in business and information systems engineering. In: Business & Information Systems Engineering, vol. 58, uss. 6, pp. 375-379 (2016).