Promoting replications through positive incentives

Krzysztof Gorgolewski[1], Thomas Nichols[234], David N. Kennedy[5], Jean-Baptiste Poline[67], and Russell A. Poldrack[1]

1 Department of Psychology, Stanford University, Stanford, USA
2 Oxford Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Population Health, University of Oxford, UK
3 Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford, UK
4 Department of Statistics, University of Warwick, UK
5 Department of Psychiatry, University of Massachusetts Medical School, Worcester, MA, USA
6 Montreal Neurological Institute, McGill University, Quebec, Canada,
7 Helen Wills Neuroscience Institute, University of California Berkeley, California, USA

Abstract

Despite their unquestionable value replication studies are hard to find in the literature. We propose using academic awards to improve this situation and show, using a pilot award in the field of human neuroimaging, that they can be an effective way to promote and incentivize replication studies.

In recent years the academic community has been stunned by reports of failed attempts to replicate a number of experiments in behavioural (Open Science Collaboration 2015) and clinical sciences (Nosek and Errington 2017). Coupled with reports of insufficient statistical power (Button et al. 2013), biases (Fanelli, Costas, and Ioannidis 2017) and questionable research practices (John, Loewenstein, and Prelec 2012; Fiedler and Schwarz 2015) this has led to a spread of negative and alarmist press coverage.

The degree to which the lack of replicability was surprising may reflect the fact that pure replication attempts are relatively uncommon in many areas of basic science. Despite the fact that knowing whether previously reported effects can be replicated by independent scientists on a newly acquired set of data is an essential part of the scientific process, the incentives for scientists to perform and publish replications are lacking. Different approaches to incentivizing replications has been explored in the past such as editorial policies (Fletcher and Grafton 2013) and special funding programs dedicated to replications[1]. Here we investigate another approach - scientific awards.

[1] https://grants.nih.gov/grants/guide/rfa-files/RFA-CA-09-003.html

Science praises ingenuity - often featuring an individual scientist, singled out of a crowd of peers. Numerous awards such as the Nobel Prize or the Fields Medal are the most highly praised academic achievements, and universities compete between each other in terms of number of scientists awarded those prizes. Academic awards are proxies of excellence used in hiring and job evaluation procedures. Specialized awards exist in almost every academic field - usually awarded during a large annual meeting. For example: the Society for Neuroscience awards the The Gruber Foundation Neuroscience Prize, The American Chemical Society awards the Priestley Medal, and the American Astronomical Society awards the Helen B. Warner Prize for Astronomy.

Awards can be used to send a powerful message about the desired achievements and traits of awarded scientists. In order to incentivize and elevate the value of replication studies within the neuroimaging community, we proposed a Replication Award honoring the best replication study in the field of human neuroimaging. The award has been embraced by the leading conference in the field - Organization for Human Brain Mapping[2] and funded by the Laura and John Arnold Foundation[3]. We have designed a detailed and transparent review protocol (each submission was evaluated along two dimensions - the quality of the replication study and the impact of the study being replicated) and made sure that all of the elements of prestige associated with a scientific award were there (i.e. presentation during the opening ceremony, press coverage etc.).

In terms of number, quality, and diversity of submissions the initial edition of the OHBM Replication Award was definitely a success (see Box 1 and 2). Twenty two eligible submissions were received which included large and small studies, meta-analyses, successful and failed replications, replications of empirical studies as well as methodological replications. The average score was high along both dimensions. It's also worth mentioning that over 30% of submissions were preprints - highlighting the fact that this valuable work is often not published in academic journals.

Box 1 - Finalists of the 2017 OHBM Replication Award

In the first round of the selection process each submission was evaluated by at least 3 reviewers. The submissions were scored along two dimensions: quality and impact. The following five submissions earned the highest average scores:

Altered Brain Activity in Unipolar Depression Revisited: Meta-Analyses of Neuroimaging Studies by Muller et al. performed a meta-analysis of 57 brain imaging studies of depression and found no consistent pattern of activation thus failing to replicate previous meta-analyses (Müller et al. 2017).

[2] https://www.humanbrainmapping.org/
[3] http://www.arnoldfoundation.org/

Neuroanatomical Diversity of Corpus Callosum and Brain Volume in Autism: Meta-analysis, Analysis of the Autism Brain Imaging Data Exchange Project, and Simulation by Lefebvre et al. in a highly powered sample of 694 participants failed to replicate previous claims of smaller Corpus Callosum size in patients diagnosed with Autism (Lefebvre et al. 2015).

A Purely Confirmatory Replication Study of Structural Brain-Behavior Correlations by Boekel et al. performed a replication 17 previously reported brain-behaviour relationships, but found enough evidence to confirm only one of them - grey matter volume in amygdala positively correlated with social network size (Boekel et al. 2015; Keuken et al. 2017).

An Attempt to Replicate a Dissociation between Syntax and Semantics during Sentence Comprehension Reported by Dapretto & Bookheimer (1999, Neuron) by Siegelman et al. failed to replicate previous reports of processing the meanings of individual words vs. assembling words into phrases and sentences relying on distinct pools of cognitive and neural resources (Siegelman et al. 2017).

Personality Reflection in the Brain's Intrinsic Functional Architecture Remains Elusive by Tomeček et al. failed to replicate previous reports of relation between resting state functional connectivity and personality traits (Tomeček et al. 2017).

It is worth noting that the outcome of the replication attempt was not one of the directly evaluated properties of submissions. Thus the award is not an ultimate endorsement of the replication outcome, and discordant results are not favored either. Many research questions will require more replications under different conditions.

The full list of all submissions and scores can be found in (Gorgolewski et al. 2017).

Box 2 - The Winner of the OHBM Replication Award

In the second phase of the selection process all reviewers were asked to rank the finalists according to their preference. The individual rankings were compiled using an instant runner-off method. The study **A Purely Confirmatory Replication Study of Structural Brain-Behavior Correlations** by Wouter Boekel, Eric-Jan Wagenmakers, Luam Belay, Josine Verhagen, Scott Brown and Birte U. Forstmann was ranked first by more than 50% of reviewers and thus won the award. This study investigated 17 brain behaviour relationships out of which the authors found evidence confirming one of the relation (correlation between real-world social network size and grey matter volume in the amygdala). Using Bayesian statistics authors found evidence in favor of the null-hypothesis (thus contradicting previous results) for the following 8 relationships:

- Efficacy of changing response caution correlated with tract strength between pre-supplementary motor area and striatum.
- Facebook network size correlated with grey matter volume in
    - right entorhinal cortex,
    - left amygdala, and
    - right amygdala.
- Behavioral Activation System (BAS) scale correlated with parallel eigenvalue in left corona radiata and left superior longitudinal fasciculus.
- Fun-Seeking subscale of BAS correlated with fractional anisotropy in left corona radiata and left superior longitudinal fasciculus.
- Fun-Seeking subscale of BAS correlated with parallel eigenvalue in left corona radiata and left superior longitudinal fasciculus.
- Alerting components of attention correlated with cortical thickness in left superior parietal lobe.

The results of the remaining 8 replications remain inconclusive (possibly due to insufficient statistical power).

The study was praised by the reviewers for its methods (preregistration and the use of Bayesian statistics) although it has been criticized for its low sample size (n=36 see Kanai 2016; Muhlert and Ridgway 2016; Boekel, Forstmann, and Wagenmakers 2016).

To further investigate if the award had a positive impact on the field we conducted a survey among the human neuroimaging community. Out of 226 respondents familiar with the award 49% declared it made them more likely to perform replications in the future and 41% declared it made them more likely to publish replication studies (see Figure 1).

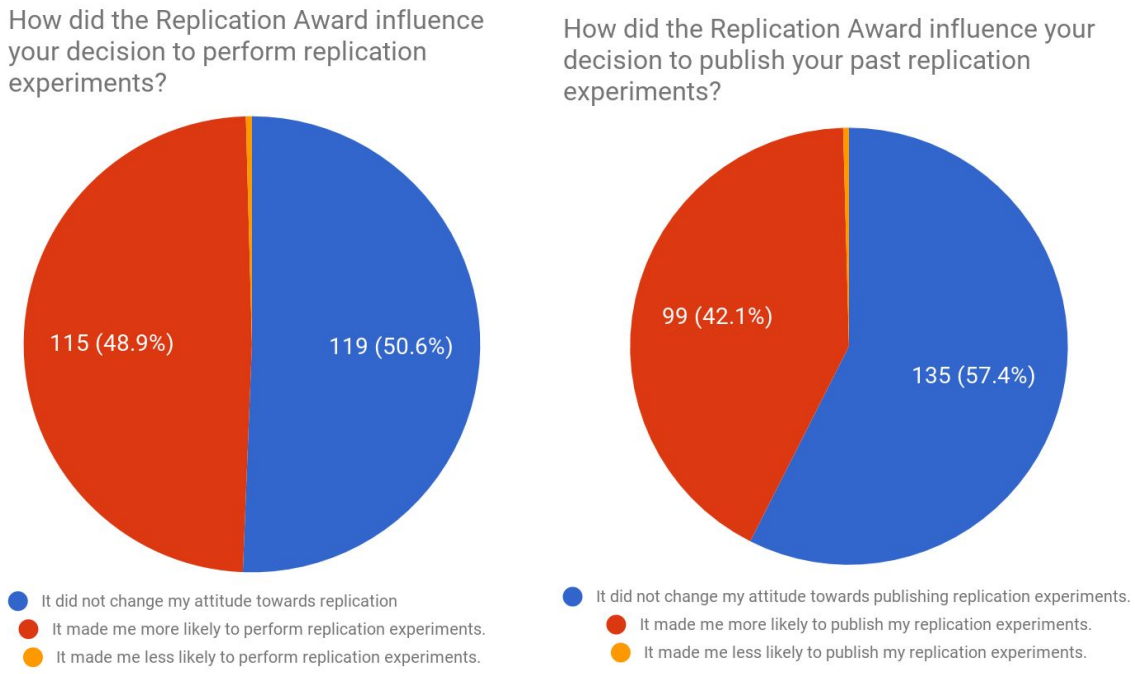

Figure 1. Survey results. 234 out of 296 (79%) of respondents that previously heard about the Replication Award were asked how did it influence their attitude towards performing (left) and publishing (right) replication studies.

An award can play a significant role in shaping a field and incentivizing underrepresented research. We believe that the OHBM Replication Award has sent a clear message about the importance of replication in the field of human neuroimaging and thereby significantly increased the probability of such studies being performed and published. We hope that other fields will also adopt the approach of positive reinforcement of good research practices and promote replications amongst their ranks.

References

Boekel, Wouter, Birte U. Forstmann, and Eric-Jan Wagenmakers. 2016. “Challenges in Replicating Brain-Behavior Correlations: Rejoinder to Kanai (2015) and Muhlert and Ridgway (2015).” *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* 74 (January): 348–52. doi:10.1016/j.cortex.2015.06.018.

Boekel, Wouter, Eric-Jan Wagenmakers, Luam Belay, Josine Verhagen, Scott Brown, and Birte U. Forstmann. 2015. “A Purely Confirmatory Replication Study of Structural Brain-Behavior Correlations.” *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* 66 (May): 115–33. doi:10.1016/j.cortex.2014.11.019.

Button, Katherine S., John P. A. Ioannidis, Claire Mokrysz, Brian A. Nosek, Jonathan Flint, Emma S. J. Robinson, and Marcus R. Munafò. 2013. “Power Failure: Why Small Sample Size Undermines the Reliability of Neuroscience.” *Nature Reviews. Neuroscience* 14 (5): 365–76. doi:10.1038/nrn3475.

Fanelli, Daniele, Rodrigo Costas, and John P. A. Ioannidis. 2017. “Meta-Assessment of Bias in Science.” *Proceedings of the National Academy of Sciences of the United States of*

America 114 (14): 3714–19. doi:10.1073/pnas.1618569114.
Fiedler, Klaus, and Norbert Schwarz. 2015. “Questionable Research Practices Revisited.” *Social Psychological and Personality Science* 7 (1). SAGE Publications: 45–52. doi:10.1177/1948550615612150.
Fletcher, Paul C., and Scott T. Grafton. 2013. “Repeat after Me: Replication in Clinical Neuroimaging Is Critical.” *NeuroImage. Clinical* 2 (January): 247–48. doi:10.1016/j.nicl.2013.01.007.
Gorgolewski, Krzysztof, Thomas Nichols, David Kennedy, Jean-Babtiste Poline, and Russell A Poldrack. 2017. “The 2017 OHBM Replication Award.” *Figshare*. doi:10.6084/m9.figshare.5140294.v1.
John, Leslie K., George Loewenstein, and Drazen Prelec. 2012. “Measuring the Prevalence of Questionable Research Practices with Incentives for Truth Telling.” *Psychological Science* 23 (5): 524–32. doi:10.1177/0956797611430953.
Kanai, Ryota. 2016. “Open Questions in Conducting Confirmatory Replication Studies: Commentary on Boekel et Al., 2015.” *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* 74 (January): 343–47. doi:10.1016/j.cortex.2015.02.020.
Keuken, Max C., Alexander Ly, Wouter Boekel, Eric-Jan Wagenmakers, Luam Belay, Josine Verhagen, Scott D. Brown, and Birte U. Forstmann. 2017. “Corrigendum to ‘A Purely Confirmatory Replication Study of Structural Brain-Behavior Correlations’ [Cortex 66 (2015) 115-133].” *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* 93 (August): 229–33. doi:10.1016/j.cortex.2017.03.007.
Lefebvre, Aline, Anita Beggiato, Thomas Bourgeron, and Roberto Toro. 2015. “Neuroanatomical Diversity of Corpus Callosum and Brain Volume in Autism: Meta-Analysis, Analysis of the Autism Brain Imaging Data Exchange Project, and Simulation.” *Biological Psychiatry* 78 (2): 126–34. doi:10.1016/j.biopsych.2015.02.010.
Muhlert, Nils, and Gerard R. Ridgway. 2016. “Failed Replications, Contributing Factors and Careful Interpretations: Commentary on Boekel et Al., 2015.” *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* 74 (January): 338–42. doi:10.1016/j.cortex.2015.02.019.
Müller, Veronika I., Edna C. Cieslik, Ilinca Serbanescu, Angela R. Laird, Peter T. Fox, and Simon B. Eickhoff. 2017. “Altered Brain Activity in Unipolar Depression Revisited: Meta-Analyses of Neuroimaging Studies.” *JAMA Psychiatry* 74 (1): 47–55. doi:10.1001/jamapsychiatry.2016.2783.
Nosek, Brian A., and Timothy M. Errington. 2017. “Making Sense of Replications.” *eLife* 6 (January). doi:10.7554/eLife.23383.
Open Science Collaboration. 2015. “Estimating the Reproducibility of Psychological Science.” *Science* 349 (6251). doi:10.1126/science.aac4716.
Siegelman, Matthew, Zachary Mineroff, Idan Blank, and Evelina Fedorenko. 2017. “An Attempt to Replicate a Dissociation between Syntax and Semantics during Sentence Comprehension Reported by Dapretto & Bookheimer (1999, Neuron).” *bioRxiv*. doi:10.1101/110791.
Tomeček, David, Renata Androvičová, Iveta Fajnerová, Filip Děchtěrenko, Jan Rydlo, Jiří Horáček, Jiří Lukavský, Jaroslav Tintěra, and Jaroslav Hlinka. 2017. “Personality Reflection in the Brain’s Intrinsic Functional Architecture Remains Elusive.” *PsyArXiv*, May. doi:10.17605/OSF.IO/EMMXT.