

Transparent Psi Project

The Transparent Psi Project, led by psychologists at Budapest's Eötvös Loránd University, failed in 2022 to replicate a precognitive effect reported by Daryl Bem in 2011, while using the most stringent methods to reduce the likelihood of methodological and statistical errors.

Background

The 2010s saw the emergence of a [replication crisis](#) in the social sciences, with the realization that positive experimental results might in fact be due to statistical and methodological errors, along with artificially inflated success rates – so-called questionable research practices (QRPs). The crisis was fuelled by the publication in 2011 of [Daryl Bem's](#) report of [experiments yielding evidence of precognitive effects](#),^[1] which were widely disbelieved within the science community. Subsequent surveys indicated that many psychologists knowingly commit QRPs.

Concerned by the survey results and recently uncovered fraud by social psychologist [Diederik Stapel](#), the [Center for Open Science](#) and its co-founder [Brian Nosek](#)^[2] launched the [Reproducibility Project](#), a multi-centre attempt to replicate a hundred studies originally published in 2008. Only 36% of the studies were successfully replicated.^[3] The results were published in 2015, and further large-scale efforts^[4] since then have met with similarly low replication rates.

Transparent Psi Project

The [Transparent Psi Project](#) (TPP) is a consensus-driven attempt by psychologists to apply stringent state-of-the-art protocols aimed at reducing the likelihood of QRPs in the social sciences, adopting methods such as preregistration, consensus protocols, pilot testing and independent auditing.

The project, led by Zoltán Kekecs and colleagues at the Eötvös Loránd University in Budapest, first aimed to replicate the precognitive effects claimed by Daryl Bem. It chose one of Bem's nine reported experiments, involving the detection of erotic stimuli, for its relative simplicity of design and because it had produced the highest effect size. Participants seated in front of a computer screen were presented with two curtains and asked to choose which one concealed a picture. After making a selection, the curtain opened to reveal either a blank slate or an erotic picture – the choice being determined randomly by the computer *after* the participant's decision was indicated. In Bem's experiment, participants uncovered the erotic picture 53.1% of the time, a statistically significant departure from the expected 50% chance mean ($p = 0.01$).

The TPP ran from January 2020 to April 2022. Ten laboratories tested 2,220 subjects, in person not online, producing a total of 37,836 trials.^[5] An erotic image was predicted 49.89% of the time, almost exactly the 50% chance level, meaning that Bem's experiment was not replicated. Additionally, there was no evidence of exceptionally lucky (or unlucky) sub-groups of subjects cancelling each other's

scoring rates, as correct guesses were distributed normally, with no evidence of outliers. Examination of the effect of experimenter and subject psi belief on the results also found no influence on scoring rates. The authors write:

The failure to replicate previous positive findings with this strict methodology indicates that it is likely that the overall positive effect in the literature might be the result of recognized methodological biases rather than ESP. However, the occurrence of ESP effects could depend on some unrecognized moderating variables that were not adequately controlled in this study, or ESP could be very rare or extremely small, and thus undetectable with this study design. Nevertheless, even if ESP would exist, our findings strongly indicate that this particular paradigm, used in the way we did, is unlikely to yield evidence for its existence.^[6]

Criticism

The Transparent Psi Project has received criticism from within the parapsychological community.

In a [preprint report](#), Italian psychologist [Patrizio Tressoldi](#) discusses the limitations of using non-selected subjects tested in a normal state of consciousness. He argues that meta-analyses demonstrate the importance of employing the selected subjects in psi-enhancing states of consciousness to maximise the probability of finding psi.^[7]

[Dean Radin](#) has argued that the kind of ‘hyper-objective’ approaches used in the TPP are counterproductive.

Obsessively strict pre-registration and detailed recording of all aspects of the study are predicated on two assumptions: (1) observation of a system under study does not affect that system, and (2) exact repeatability establishes if a phenomenon is scientifically "real."

These assumptions work quite well for many macroscopic physical systems. They do not work quite as well for living systems. And they work poorly when it comes to the study of subtle psychological effects, including psi. They also sidestep the QM uncertainty principle and QM observational effects, which tells me that while those two assumptions are considered sacrosanct requirements within science today, they are not appropriate for the study of all possible things.^[8]

Andrew Gelman, a professor of statistics and political science at Columbia University, declined Kekecs's invitation to take part in the TPP on the grounds that ‘preregistration would not solve the problem of small effect sizes and poor measurements ... and there’s a 5% chance you’ll see something statistically significant just by chance, leading to further confusion’. He argued that the project would be better directed at established psychological phenomena such as social priming than ESP, which might not exist at all.^[9]

Kekecs replied

Bias can always be introduced by different questionable research practices (QRPs), but if we are able to design a study completely immune to the QRPs, there is no real possibility for bias toward type I error. Of course, if the effect really exists, all the usual threats to validity can have an influence (for example, it is possible that people can get “psi fatigue” if they perform a lot of trials, or that events and contextual features, or even expectancy can have an effect on performance), but we cannot make a type I error in that case, because the effect exists, we can only make errors in estimating the size of the effect, or a type II error.

So understanding what is underlying the dominance of positive effects in ESP research is very important. If there is no effect, psi literature can serve as a case study for bias in its purest form, which can help us understand it in other research fields. On the other hand, if we find an effect when all QRPs are controlled for, we may need to really rethink our current paradigm.^[10]

Further Testing

The Advanced Meta-Experimental Protocol (AMP)^[11] was developed by Jan Walleczek at [Phenosience Laboratories](#) (a research program investigating the links between physics, biology, and consciousness research) with the aim of introducing robust controls into scientific testing. A particular focus of the AMP is testing for false positives - a result that indicates the presence of an effect that does not actually exist.

In 2019 an attempted replication^[12] of Dean Radin’s double slit experiment using the AMP method produced evidence of false positives, but these conclusions were emphatically disputed by Radin and his colleagues.^[13]

In 2020, [Zoltán Kekecs](#), one of the architects of the Transparent Psi Project, collaborated with Walleczek, Nikolaus von Stillfried and [Stefan Schmidt](#) to run the AMP alongside the TPP, as a sister-project and control check. Unlike the TPP, the first AMP replication attempt (referred to as AMP-TPP1) used online testing only. Against prediction, subjects tested under AMP-TPP1 scored significantly *below* chance (37,836 trials, 49.47 percent correct responses, where 50 percent expected by chance, $p < 0.05$).

This unexpected post-hoc effect was tested in a second formal preregistered replication referred to as AMP-TPP2.^[14] Analysis of 127,000 trials confirmed the predicted negative scoring (49.65 percent, $p = 0.005$). In parapsychology it is fairly uncommon to replicate a preregistered effect under such exacting conditions – a notable finding for the field.

A [third study \(AMP-TPP3\)](#) ran the same protocol in the absence of subjects, in order to test the possibility that the earlier significant results were the result of false positives. These sham trials found nothing out-of-the-ordinary, therefore confirming the AMP-TPP2 result.^[15]

Michael Duggan

Literature

Baptista, J., Derakhshani, M., Tressoldi, P. E. (2015). Explicit anomalous cognition: A review of the best evidence in ganzfeld, forced choice, remote viewing and dream studies. In *Parapsychology: A Handbook for the 21st Century* (192-214), ed. by E. Cardeña, J. Palmer, & D. Marcusson-Clavertz. Jefferson, NC: McFarland.

Bem, D.J. (2011a). [Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect](#). *Journal of Personality and Social Psychology* 100, 407-25.

Bem, D., Tressoldi, P., Rabeyron, T., & Duggan, M. (2016). Feeling the future: A meta-analysis of 90 experiments on the anomalous anticipation of random future events. *F1000Research* 4. <https://doi.org/10.12688/f1000research.7177.2>

Bem, D., Utts, J., & Johnson, W. O. (2011). Must psychologists change the way they analyze their data? *Journal of Personality and Social Psychology* 101/ 4, 716–19. <https://doi.org/10.1037/a0024777>

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B. A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., ... Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour* 2/9, 637. <https://doi.org/10.1038/s41562-018-0399-z>

Gelman, A. (2017). [Why I'm not participating in the Transparent Psi Project](#). [Web page]

Kekecs, Z., Palfi, B., Szaszi, B., Szecsi, P., Zrubka, M., Kovacs, M., et al.(2022). Royal Society Publishing. [Raising the value of research studies in psychological science by increasing the credibility of research reports: the transparent Psi project](#). <https://doi.org/10.31234/osf.io/uwk7y>

Nosek, B.A. (2015). *Open Science Collaboration*. "Estimating the reproducibility of psychological science". *Science*. 349/6251.

Nosek, B.A., Alter, G., Banks, G.C., Borsboom, D., Bowman, S.D., Breckler, S.J., Buck, S., Chambers, C.D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M. Yarkoni, T. (2015). Promoting an open research culture. *Science* 348/6242, 1422-25. <https://doi.org/10.1126/science.aab2374>

Radin, D., Wahbeh, H., Michel, L., & Delorme, A. (2020). [Commentary: False-Positive Effect in the Radin Double-Slit Experiment on Observer Consciousness as Determined With the Advanced Meta-Experimental Protocol](#). *Frontiers in Psychology* 11, 15 April.

Targ, R., Puthoff, H. (1974). Information transmission under conditions of sensory shielding. *Nature* 251, 602-607. <https://doi.org/10.1038/251602a0>

Walleczek J, von Stillfried N. (2019). [False-Positive Effect in the Radin Double-Slit Experiment on Observer Consciousness as Determined With the Advanced Meta-Experimental Protocol](#). *Frontiers in Psychology*, 22 Aug. 10:1891. doi: 10.3389/fpsyg.2019.01891.

Walleczek J, von Stillfried N. (2020). [Response: Commentary: False-Positive Effect in the Radin Double-Slit Experiment on Observer Consciousness as Determined With the Advanced Meta-Experimental Protocol](#). *Frontiers in Psychology*, 3 Dec. 11:596125. doi: 10.3389/fpsyg.2020.596125. PMID: 33343467

Endnotes

Footnotes

1. ^ Bem (2011).
2. ^ Nosek et al (2015).
3. ^ Nosek (2015).
4. ^ Camerer et al (2018).
5. ^ Kekecs et al (2022).
6. ^ Kekecs et al (2022).
7. ^ Tressoldi (2022).
8. ^ Private forum, January 2023, quoted with permission.
9. ^ Gelman (2017).
10. ^ Gelman (2017).
11. ^ Walleczek, J., Stillfried, N, v. (2019).
12. ^ Walleczek, J., Stillfried, N, v. (2020).
13. ^ Radin et al (2020).
14. ^ Walleczek, J., Stillfried, N.v., Schmidt,S., Kekecs, Z. (2022).
15. ^ Schmidt, S. Personal communication, 20 September, 2022.