Open Peer Comment to

"Anomalous Cognition: An Umbrella Review of the Meta-Analytic Evidence"

Stefan Schmidt[1],
Institute for Frontier Areas of Psychology and Mental Health

[Note from the Editor: JAEX uses an anonymous system for both authors and reviewers because in a small area of inquiry identifying either of them would make it more likely that bias for or against, or future reciprocation would creep into the process. In this case, both the authors of the paper and one of the reviewers recognized each other nonetheless, and the authors of the paper recommended that the reviewer write a comment to which they would respond. I should mention that the statistician to whom Professor Schmidt alludes is someone I first approached for an opinion, besides requesting the opinion of two other eminent authors, trying to clarify the issues raised in this comment. Also, I did not accept the second revision "straightaway" but requested some additional changes in response to comments by additional consultants.]

Background: The approach of the authors to systemically compile meta-analyses from psi research is a very fruitful one and I was happy to get selected as one of two reviewers of the manuscript. The first version of this publication was quite short (2,400 words) and lacked detailed information in several places. Thus, it was difficult to follow the methodological approach of the authors. The first revision provided some more details on procedures. To some of the comments of the reviewers the authors were not very responsive. Furthermore, the new information on the methodological details raised more concerns regarding the methodological adequacy of some approaches. I pointed them out in my second review and informed the editor accordingly. When the second revision came in the editor accepted it straightaway. After that he sent me the manuscript and invited me to write an accompanying letter should I not be satisfied. In what follows, I will raise some critical issues regarding the publication.

Inclusion Criteria: The authors performed a systematic research in order to find anomalous cognition meta-analyses. However, the inclusion/exclusion of studies

1 Address correspondence to: Stefan Schmidt, Ph. D., Department of Psychosomatic Medicine and Psychotherapy. Faculty of Medicine, Albert-Ludwigs-Universität Freiburg, stefan.schmidt@uniklinik-freiburg.de

2021, Vol. 1, No. 1-2, pp. 73-75
https://doi.org/10.31156/jaex.23439

Journal of Anomalous Experience and Cognition (JAEX) 2021, Vol. 1, No. 1-2, pp. 73-77

seems to be arbitrary. Both reviewers had the feeling that the exclusion of Milton and Wiseman (1999) was not systematic. The authors argued "We also excluded the Milton and Wiseman (1999a) meta-analysis because it was related to mass participation without any control over recruitment and motivation of participants who were requested to predict masked targets, similar to the lottery guessing tasks." After this critique, the authors added the inclusion criteria to the manuscript which had not been mentioned before. However, it is still puzzling that, for instance, the presentiment meta-analysis is included while all the DMILS/remote staring meta-analyses are not. Regarding this issue the authors replied: "DMILS is not generally argued to be a form of ESP (anomalous perception)." This statement is a mystery to me. Why should sensing that somebody is staring at me from behind not be some form of ESP or anomalous perception?

Heterogeneity: The quality of the results of a meta-analysis can be assessed by several criteria. There are, for instance, the mean effect size and the significance level. Another crucial indicator is heterogeneity. It assesses whether within the sample of the effect-sizes there is more variance than would be expected by sampling error. Since some of these meta-analyses are quite old and some have complex combinations of studies, heterogeneity is an important issue to report. It is a standard in mainstream reviews. The authors refused to report this information.

Quality of Meta-Analyses. In a similar fashion, I suggested to them to include a quality rating of the different meta-analyses since in the last 30 years of meta-analyses the methodology and reporting standards have made some advances. My idea was to select variables referring to quality issues that vary throughout the sample. The authors insisted to remain with the MARS criteria. I consider them useless because they are so simple that the sample has hardly any variance. Thus, I also suggested deleting table 2. Since almost all studies but two report the exactly same features the information contained in this table is close to zero and could be easily summarized in one sentence.

Effect Size Metric: The authors combine and compare different effect types of effect sizes in their approach. They also compute correlations with these values. This requires that they be mathematically equivalent in the sense that, for instance, the double in size refers to an effect twice as large. Whether this is true for the different types of effect sizes in this review and whether all these effect sizes in the paper share the same metric is difficult to say. My request to give more background information on effect size types, formulas, and converting procedures were unfortunately declined. I made a few tests with my own data, read the literature, and debated with a statistician about this issue. I remain uncertain whether all these effect sizes share the same metric as the authors claim. This is especially problematic regarding the ES = $z/\sqrt{n}$, where so far nobody found a reference stating that this belongs to the *d*-type effect size family. It might be that such an ES underestimates the size of effect if taken as a *d*-type effect size. So, all comparisons and computations with this set of effect sizes reported here need be interpreted in a very cautious way. This is even more true since some of the primary studies are contained in several meta-analyses.

Line in the Graph: In figure 1, the mean effect sizes of the different meta-analyses are depicted graphically. A line connects them and a line in a graph indicates that several observations reflect moments in a continuous variable such as temperature in a weather chart. In this case, the graph displays single independent observations. Furthermore, the sequence of these observation was chosen by the authors based on the size. In such a graph, single observations should not be connected by a line since this is misleading, which is basic knowledge in statistics. The authors did not agree to this reasoning and insisted on the connecting line.

Moderator Analyses: The authors chose two moderators, one is response type and the other state of consciousness. They gave point ratings for each of the moderators. For instance, a study with a physiological dependent variable received three times as many points as one with applying a forced choice variable. In order to assess the effects of the moderators on effect size they combined these two moderators into one score by adding points. This correlation turned out to be significant. The authors state in their reply to my comments that the decision to combine these two moderators "did not derive from theoretical reasoning, but was empirically tested." This sounds a bit like *p*-hacking. An appropriate approach would be to report the results of the respective moderators separately and to combine them in a simple meta-regression.

There a couple of more issues I disagree with (e.g., the arguments regarding QRPs) but this comment should not become too long. As a scientist, my approach to the peer review process is to make contributions better in a joint effort. I have learnt many important things from reviewers of my manuscripts and I think the peer review process is one of the strengths of science. What happened here is that the authors chose to ignore almost all suggestions and insisted on their approach. This has never happened to me before. I am thus, thankful to the editor for giving me a chance to publish my deviating opinions together with the original publication.