Expert Consensus Procedure (ECO): Facilitating Robust Scientific Outputs

Zoltan Kekecs[1,2*], Barnabas Szaszi[1], Balazs Aczel[1]

[1]Institute of Psychology, ELTE Eotvos Lorand University

[2]Department of Psychology, Lund University

This manuscript has been submitted for peer review to a scientific journal. This is the version at the time of submission

Corresponding author: Zoltan Kekecs, Department of Affective Psychology, Eotvos Lorand University, Izabella u. 46. 1064 Budapest, Hungary. E-mail: kekecs.zoltan@gmail.com

Abstract

This tutorial provides a step-by-step guide for the implementation of the Expert Consensus procedure (ECO) for scientific projects. The ECO aims to facilitate consensus among a panel of experts about a target scientific output. Consensus is achieved through an iterative process including surveys and feedback to the panel members. Following the procedure can achieve two main aims: (1) to decrease the chance that the target output will be subject to conceptual or methodological mistakes, and (2) to increase the chance that the target output would be acceptable by the stakeholders on the field.

Keywords: consensus, research design, meta-science

Glossary

ECO - Expert Consensus Procedure

Previous Round Summary - A personalized feedback for each panel member about the results of the previous round of the iterative survey.

Target Output - The output that the consensus procedure aims to produce.

Survey Package - A collection of materials that is sent to the panel members in the rounds of the iterative survey. It can include data collection survey, proposed or amended version of the Target Output, and the Previous Round Summary.

Expert Consensus Procedure (ECO): Facilitating Robust Scientific Outputs

In this tutorial, we provide a guideline to conduct an Expert Consensus procedure (ECO) in scientific projects. Applying the ECO can help achieve two main aims: (1) to decrease the chance that the product will be subject to conceptual or methodological mistakes, and (2) to increase the chance that the product of the project would be acceptable by the stakeholders on the field.

While individual researchers or their teams strive to provide value to their field, their success is often delayed or halted by the post-hoc criticism of the field. Projects on the cutting edge or outside of the mainstream are most vulnerable to post-hoc criticism since there are fewer standard protocols to follow, and less prior data to build on, adding an extra layer of risk to such projects. Such vulnerability for projects to be dismissed after resources have already been spent makes the scientific endeavour more risky and more inefficient than it needs to be. Criticism and falsification is central to the scientific process, so the solution is not to decrease criticism. Instead, we should strive to push the scientific discourse as early as possible in the project lifecycle to increase the project's potential impact. Exposing projects to early scrutiny not only decreases waste of valuable resources, but also improves the signal to noise ratio in the scientific literature overall.

Early critical discourse about the research output already manifests to some extent in adversarial collaborations (e.g., Latham et al., 1988), Large-scale collaborations and

crowd-sourced research projects (e.g., Klein et al., 2019; Moshontz et al., 2018), and Registered reports (see Center for Open Science, n.d., https://cos.io/rr/). Critical discussion is also possible in the recently suggested adversarial commentaries (Heyman., Moors & Rabagliati, 2020). However, these methods are not designed to build consensus in a representative sample of stakeholders on the field. Furthermore, these methods are mainly applied in empirical research. Options for improving robustness are especially limited for theoretical and methodological work. Thus, there is a need for new approaches for maximizing the robustness of research projects that are applicable for a wide range of scientific projects.

To address this need, we turn to consensus procedures. In medical research, consensus procedures are applied mainly when there is a lack of empirical data to build guidelines on, and thus there is a lack of general consensus on the field on criteria for diagnosis or treatment approaches. In these cases a panel is assembled of widely recognized experts on the specific subject, and they work together to formulate a "consensus statement". This consensus statement is then used as a basis for decision making in the front line until more empirical data is accumulated to support new guidelines. There are multiple formal procedures that are used in the clinical sciences to elicit consensus, including the Delphi method, the Nominal Group Technique, the RAND/UCLA Appropriateness Method, and the National Institutes of Health consensus development conference (Nair et al., 2011). A common feature of these methods is that a group of experts is subjected to a formalized process to guide the panel members toward consensus on the topic of interest. We have adopted the most useful features of these procedures to design a method that can be used to formulate consensus on research design, theoretical or

methodological issues. With this approach our aim was to maximize the robustness and impact of research products.

Expert Consensus procedure (ECO)

The Expert Consensus procedure (ECO) aims to facilitate consensus among a panel of experts on a Target Output. This is achieved through an iterative process including surveys and feedback to the panel members, at the end of which a consensus-based Target Output is created, or lack of consensus is declared. The Target Output can vary between projects. The procedure is very versatile and the Target Output can be any form of scientific product. Table 1. contains some example projects where the ECO could be applied.

Table 1 - Example use cases for the application of the ECO

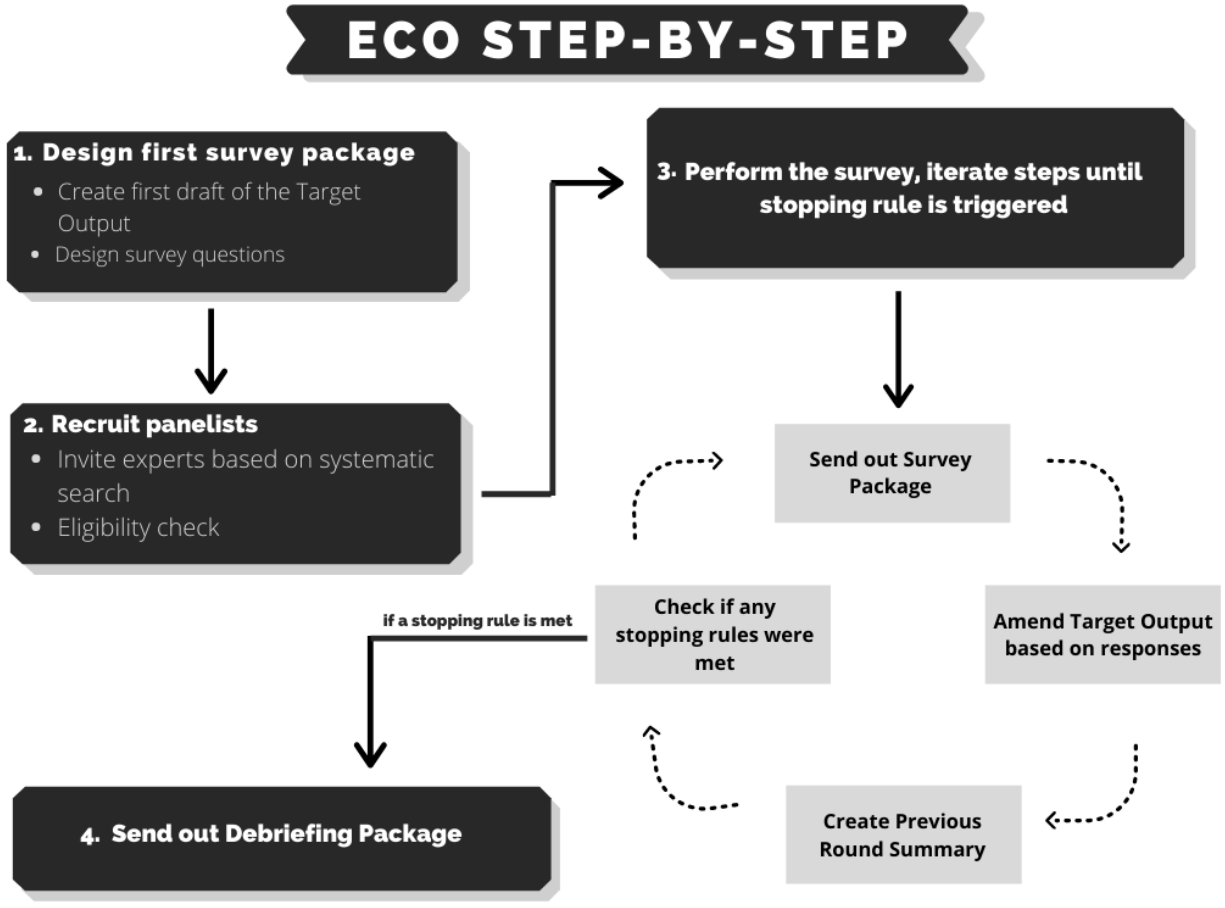
Type of research project	Formulate consensus/agreement on
empirical research study	research protocol
	analysis strategy and inference thresholds
	sampling and recruitment details
	treating missing data
	range of possible conclusions
systematic review of the literature	literature search strategy

	study eligibility criteria
	hypotheses
	what to take into account when assessing quality of evidence
	range of possible conclusions for specific findings
conceptual replication study	appropriate study design
	expected outcomes
	range of possible conclusions for specific findings
developing a multi-item measure or other multi-item research tool	items to include or exclude from a list of possible items
	suggested new/additional items
developing guidelines, recommendations	items to include or exclude from a list of possible recommendations/guidelines
	content of the recommendations/guidelines
presenting a theory or model	testable predictions of the theory/model

Below we provide a step-by-step guide for the implementation of the ECO. An outline is shown in Figure 1. We have used the ECO in three cases so far to produce three different Target

Outputs: a research protocol of the Transparent Psi Project (Kekecs et al., 2019 in principle acceptance), the Transparency Checklist (Aczel, Szaszi, et al., 2020), and the Thinking Guideline for Bayesian Inference (Aczel, Hoekstra, et al., 2020). To facilitate the understanding of the steps of the ECO, we provide concrete examples for each step implemented across these three projects (see Appendix, Table A1).

Figure 1. **The ECO step-by-step**



1. Design first Survey Package

At first, the coordinators of the ECO design a Survey Package which has two main components: 1) the first draft of the Target Output, and 2) an iterative survey.

1.1. The first draft of the Target Output

First, the initial version of the Target Output is created. This first draft of the Target Output aims to structure and direct the discussion of the experts on the given topic. Depending on the purpose of the consensus-building process, this could be a very detailed draft where the panel is expected to fine-tune the contents to make it acceptable, or it could be a rough outline where the panel is expected to suggest new content.

1.2. The iterative survey

This is followed by the design of an iterative survey. The goal of this survey is twofold. First, it aims to assess the level of agreement with the overall state or specific components of the Target Output. For example, if the Target Output is a list of guidelines, the survey can ask the panelists to rate their agreement with each item on the list separately on a Likert-type rating scale; or if the Target Output is a research protocol, the survey can assess the acceptability of certain protocol elements. Second, the survey collects suggestions for improvement of the Target Output. Especially if the respondent does not indicate perfect agreement with the content, the survey should ask for suggestions on how the Target Output can be improved, usually via a free text field.

The Survey Package should be pilot tested with a group of individuals who have similar domain or methods expertise as the panelists, but who will not be members of the expert panel.

Piloting can help in improving the clarity of the survey questions and can help avoid other ambiguities in the Survey Package or the ECO process itself.

2. Recruit panelists

After the Survey Package is finalized, participating experts need to be recruited.

Systematic search for potential panel members

One of the key distinguishing feature of the ECO is that instead of asking for feedback from people based on who is conveniently available (this is analogous to convenience sampling in a research study), the coordinating researchers make an effort to give an equal chance for all of the relevant experts or qualified stakeholders to be involved in the consensus panel. The main principle here should be that the sampling procedure should be able to reach a representative sample of the population of experts or qualified stakeholders. In some cases, the sampling procedure can be extended to include all eligible individuals, while in other cases some restrictions might be applied to keep the number of panelists manageable. If sampling needs to be restricted, random sampling is preferable to avoid systematically excluding relevant subgroups of the population of interest.

Invitation and eligibility check

The coordinating researchers should pre-specify eligibility criteria for participation in the panel. These criteria can be put in place to make sure that people in the panel are really part of the target population, and that they have the necessary expertise or that their voice is relevant to be included in the consensus. Depending on the Target Output panelists do not necessarily have

to be “experts” in the classical sense, sometimes the goal might be to extend the consensus to non-expert stakeholders. Preregistration of these criteria could prevent the coordinating researchers from manipulating the outcome of the ECO by excluding some participants post-hoc to trigger stopping rules or to otherwise manipulate the Target Output. As stated above, a systematic search is conducted to identify the list of experts who potentially meet the eligibility criteria. Once the target list of experts is created, they are invited to participate in the study. The invitation should contain a short description of the aims and process of the ECO, and of potential compensation or incentives for participation (see also the section Motivation, retention, compensation, and crediting of panel members below). Invitees are asked for intent for participation as well as information through which eligibility to participate in the panel can be verified. (In some cases the eligibility check may be done before the invitation using publicly available information).

Performing the iterative survey

Once the Survey Package has been finalized, the iterative survey process is conducted in order to assess the level of consensus and to improve the Target Output based on the collected feedback. The following steps are iterated until a stopping rule is triggered:

1) The Survey Package is sent out. In order to increase effectiveness, the message containing the survey package should restate the aims of the project together with a clear deadline for completion of the survey, and publication plans for the Target Output. Regular reminders should be sent out to ensure timely completion of the surveys.

2) Once all the responses are collected, the ECO coordinators amend the Target Output based on the collected responses to be more agreeable for the majority of the panelists. Changes should be tracked so that the panelists can easily see what changed between the survey rounds.
3) After the changes are made, a personalized feedback is created for each panel member (Previous Round Summary). This Previous Round Summary is designed to allow the panelists to see how their own opinions relate to others panelists' opinions, and the changes made to the Target Output compared to the previous round (if this is not the first round of the survey).
4) Then, the stopping rules are checked. If the stopping rules are not met, the process re-starts again by sending out the Survey Package containing the Previous Round Summary, the amended version of the Target Output, together with a new survey to re-evaluation the Target Output. If one of the stopping rules are met, the iterative survey process can be finished by sending out the Previous Round Summary accompanied by the finale version of the Target Output.

The key elements of this iterative process are discussed in detail below:

Means of communication with the panel members

The main source of communication with the panelists is via an iterative survey. The structured format provided by the survey allows the ECO coordinators to get relevant answers on key indicators of agreeability of the Target Output, and the panelists can get clear feedback about how their own responses compare to the panel's overall opinion. Nevertheless, this does not

mean that the survey is the only medium through which the coordinators can communicate with the panelists. In some situations back-and-forth communication might be necessary to clarify the details of what would make the Target Output more agreeable for a certain panel member, which is not possible within the confines of the formal survey. Thus, one-on-one communication with individual panel members can be an integral part of the ECO.

Previous Rounds Summary

A key method for facilitating consensus is to give structured feedback about the panelists' own previous responses and the panel's overall responses. Thus, the coordinators should prepare a Previous Round Summary. This should include the following components:

1) The respondent's own numerical responses to the survey questions. Since several weeks can pass between the survey rounds, this serves as a reminder of the participant's previous opinion.
2) Summary results of numerical responses of all panelists. This usually consists of the mean, median, or mode and some spread (sd, range, or distribution) characteristic of the numerical responses of all panelists. By being reminded of their own responses and the panel's overall ratings, the panelists get detailed feedback allowing them to assess their previous ratings in contrast to other panelists' opinions, and may help them to better evaluate how the changes to the Target Output should affect their new rating.
3) List of changes to the Target Output or other actions taken in response to the outcome of the survey round

4) List of all suggestions and comments by all panelists. The panelists may get access to the notes and suggestions of all panelists with regard to how the Target Output can be improved. This improves the transparency of the ECO process and helps to put the list of changes to the Target Output into context. Importantly, the comments should be anonymized to avoid common sources of bias like prestige or known affiliation with certain groups or theoretical alignment. Anonymity of comments is also thought to prevent polarization of the panel. Note, that in very large projects sharing free text notes in their raw form may not be an effective way of sharing information. Instead, the notes might have to be processed, grouped, or summarized in ways that makes it easier for the panel members to process this. The comments can be grouped based on multiple aspects such as type of the content, the (non-identifiable) ID of the person making the comment, the section of the survey where the comment was made, or the component of the Target Output referenced in the comment, etc. This can be made user friendly by a searchable spreadsheet or Shiny app allowing for filtering based on multiple categories.

Stopping rules

After each round, stopping rules for the ECO should be checked. The stopping rules can depend on the resources and time available for the procedure and the goal, but usually they are linked to some main indicator of consensus or agreeability of the Target Output or some of its components. In some cases consensus might not be attainable either because of the polarization of the field or because of time constraints. The project coordinators should consider this

possibility and keep this in mind while planning the ECO and while formulating the stopping rules.

Restarting the iterative survey

If none of the stopping rules are triggered, a new Survey Package is created which includes the Previous Round Summary, the amended Target Output, and the survey questions, and the process is repeated. The survey questions might change from round to round depending on the type of information needed to improve the Target Output at that stage. In some cases it might not be necessary to include the Target Output in its entirety in all iterations of the Survey Package, rather, it is possible to only include the components for which consensus is not yet achieved.

Finishing the iterative survey

If any of the stopping rules is triggered, a Debriefing Package is created, including the Previous Round Summary, and the amended (final) Target Output. Usually no more survey questions are required at this point, although feedback could be solicited from the panel on finalizing the Target Output. Most iterative consensus projects reviewed by us are finished within 2-3 rounds either because they are pre-planned to stop at 2-3 rounds or because the stopping rules are triggered by then (Ager et al., 2007; Arthur et al., 2013; Berk et al., 2011; Lindqvist et al., 2013; Loblaw et al., 2012; Murphy et al., 1998; Snape et al., 2014; Sunderji et al., 2016).

Additional considerations

Remote usage

The iterative survey is designed in a way that it can be easily used remotely. The remote format is advantageous because it helps to avoid group formation around influential leaders and polarization of the panel. However, limitations to this remote format include long response times and attrition. Thus, the coordinators of the ECO should be prepared to address these issues.

Motivation, retention, compensation, and crediting of panel members

The success of the ECO depends on the motivation and buy-in of the panel members since they need to devote time and effort to seriously consider the Target Output and the survey questions. The project coordinators should consider before starting the project what the main drivers of panel members will be and ways in which motivation can be maintained throughout the process. One consideration related to motivation is how the panel members will be credited for their contributions. It is important that this is clearly communicated with the panel members from the beginning of the process. Avoiding attrition should be a primary goal since dropout between survey rounds might mean that the concerns or viewpoints of some panel members is not taken into consideration in the Target Output. Even if retention fails, the coordinators of the project should make every effort to follow up with the panel members who did not take part in a survey round to get information about the reason for them dropping out.

How to quantify the level of consensus?

Based on the literature covering other consensus methodologies, a good criteria for declaring consensus needs to take into account both the size or value and the frequency or variability of the ratings (Jorm, 2015; Murphy et al., 1998). In such procedures it is common for the distribution of the numerical ratings to be non-normally distributed, so using the mean and standard deviation alone might be misleading. Thus, if a Likert-type scale is used for rating acceptability of the elements of the Target Output, we recommend predefining a certain median rating threshold with restrictions on the maximum size of the interquartile range. If this criteria is satisfied, that particular element or characteristic of the Target Output can be considered to be acceptable to the consensus panel.

What is the ideal number of panelists?

In a review by Murphy and colleagues (1998) the question of the ideal number of panelists in consensus panels is investigated in detail. The authors conclude that reliability of group judgement increases substantially with every additional panel member up to 6 members. Reliability still increases after that with every additional member, but after reaching 12 members the return of adding new members in terms of reliability is minimal. Nevertheless, adding more members can increase the diversity of the group and make the Target Output more robust, so 12 members should not be considered an upper limit for consensus panels. As a downside of a larger panel, time required for coordination and potentially increased waiting times between rounds should be considered as a cost for adding more members.

How to report the composition of the panel?

The readers or users of the Target Output may be interested in the composition of the consensus panel. Importantly, the readers should be able to get accurate information about the expertise of the panel members to be able to judge the reliability of the Target Output. In some cases, it might be possible to reveal names and affiliations with the consent of the panel members, while in others, summary information is more appropriate. If there was considerable attrition, it might also be relevant to report the panel composition at different stages of the survey. Consent to reveal information about the panelists (if relevant) should be acquired at the beginning of the ECO anticipating dropout of panel members. The ECO coordinators should also consider reporting information about the panel composition in terms of theoretical alignment if relevant to the project.

How can post-consensus-procedure amendments be made to the Target Output?

It is possible that the lead researchers or coordinators of the project encounter a need to change the Target Output after the ECO. For example, changes might be suggested by reviewers of the manuscript reporting on the Target Output, or with time new information comes to light that necessitates change. So the project coordinators should pre-specify under which circumstances can the Target Output be changed and still be called ‘consensus-based’. Depending on the extent of the change this might require re-conferring the consensus panel and initiating a new survey round, just an update to the panel members asking for comments or suggestions, or no involvement of the panel at all. The important thing is to maintain the acceptability of the Target Output for the stakeholders, that the process for updating the Target

Output is made clear to the panelists during the ECO, and that it is made transparent in the final Target Output what (if anything) was changed after the ECO.

Preregistration of the consensus procedure

We recommend that the ECO should be preregistered. The registration should include the aims of the project, description of the desired Target Output, the sampling strategy to invite panelists, the eligibility criteria for participation, and the stopping rules of the consensus process. It should also be pre-determined whether people who do not take part in a survey round can still participate in subsequent survey rounds, and what is the process (if any) for making changes to the Target Output after the ECO has finished. When considering what should be preregistered the general rule is to make the process of "declaring consensus" as objective as possible and to limit degrees of freedom in cherry picking experts and comments to take into account when amending the Target Output.

A cooperative mindset

The ECO inevitably involves conflicting opinions. In polarized topics the tone of the conversation (even in survey responses) might be competitive. Nevertheless, the coordinating team should keep a cooperative mindset where the goal is a net gain for the field, not to bring out one or the other theoretical camp as the winner. In many cases it is possible to incorporate seemingly conflicting views into the same Target Output. For example during the consensus design of the Transparent Psi Project study protocol, one of our initial goals was to select the method used to make statistical inference. This traditionally involves the use of a single

statistical test. However, some members of the expert panel had strong preferences for different statistical tests. Instead of choosing one of the tests which would have made only a portion of the panelists content, we finally chose to propose a hybrid approach, where we would use multiple statistical tests that all were reasonable to test the hypothesis, and only make a statistical inference if all of the tests supported the same conclusion.

Summary

In this tutorial we described the ECO which is a tool to improve the acceptability and robustness of research products. We showed how to implement the ECO and what aspects need to be considered during the planning of the procedure.

In the paper, we advocate for the use of the ECO as a formalized process to improve research outputs, but we are not claiming that this is the only or the best method to achieve consensus among researchers or to improve robustness of research outputs. There are alternatives such as the registered report peer-review process or classical adversarial collaborations that can produce similar results. The ECO has several advantages compared to these alternatives in the number and diversity of opinions it can integrate. We recommend the use of this procedure when the researchers want to be comprehensive in integrating opinions from a significant sample of the stakeholders. The systematic sampling procedure ensures not only a large panel size, but also the inclusiveness of the process. This way, smaller subgroups of the stakeholders that otherwise would be left out can voice their opinions. Another advantage of the ECO is that many of the major stakeholders on the field can impact the project at its planning stage, eliciting buy-in,

which greatly reduces the potential sources of significant post-hoc criticism compared to the less inclusive alternatives.

We consider the ECO to be most beneficial for resource-intensive projects where there are no widely-agreed standards or reliable prior data to build on, and where the impact of the project is highly reliant on stakeholder consensus. It is not a tool for seeking affirmation from peers about a well developed theory or procedure in which there is no room for significant change. The ECO requires openness to compromise and flexibility in the structure and content of the Target Output for it to produce the best outcomes.

Disclosures

Authors contributions

Conceptualization: Zoltan Kekecs, Barnabas Szaszi and Balazs Aczel.

Funding Acquisition: Zoltan Kekecs.

Investigation: Zoltan Kekecs.

Methodology: Zoltan Kekecs, Barnabas Szaszi and Balazs Aczel.

Writing - Original Draft Preparation: Zoltan Kekecs, Barnabas Szaszi and Balazs Aczel.

Writing - Review & Editing: Zoltan Kekecs, Barnabas Szaszi and Balazs Aczel.

(created with Tenzing: https://martonbalazskovacs.shinyapps.io/tenzing/)

Conflict of interest

The authors have conflicts of interest: the acceptance of this manuscript and the citations of the accepted paper partially determine the authors' career prospects and chances for getting future research funding.

Funding

The project was funded by the Bial Foundation via Grant nº 122/16.

Zoltan Kekecs was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

Zoltan Kekecs was supported by the ÚNKP-20-5 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund

Acknowledgement

We would like to thank Aba Szollosi, Bence Palfi, Melvin Marsh, and Pietro Rizzo for their help in conducting the literature review of expert consensus methods and other relevant literature, and Anna Lutoskin for her help in formatting the manuscript.

Supplementary Material

Table A1. Exemplary operational documents from previous studies

Prior versions

A previous version of the manuscript was posted on PsyArXiv: https://psyarxiv.com/9gqru/

References

Aczel, B., Hoekstra, R., Gelman, A., Wagenmakers, E.-J., Klugkist, I. G., Rouder, J. N., Vandekerckhove, J., Lee, M. D., Morey, R. D., Vanpaemel, W., Dienes, Z., & van Ravenzwaaij, D. (2020). Discussion points for Bayesian inference. *Nature Human Behaviour*, *4*(6), 561–563. https://doi.org/10.1038/s41562-019-0807-z

Aczel, B., Szaszi, B., Sarafoglou, A., Kekecs, Z., Kucharský, Š., Benjamin, D., Chambers, C. D., Fisher, A., Gelman, A., Gernsbacher, M. A., Ioannidis, J. P., Johnson, E., Jonas, K., Kousta, S., Lilienfeld, S. O., Lindsay, D. S., Morey, C. C., Munafò, M., Newell, B. R., … Wagenmakers, E.-J. (2020). A consensus-based transparency checklist. *Nature Human Behaviour*, *4*(1), 4–6. https://doi.org/10.1038/s41562-019-0772-6

Ager, A., Boothby, N., & Wessells, M. (2007). The use of consensus methodology in determining key research and practice development questions in the field of intervention with children associated with fighting forces. *Intervention: International Journal of Mental Health, Psychosocial Work and Counselling in Areas of Armed Conflict*, *5*(2), 124–129.

Arthur, C., Levett-Jones, T., & Kable, A. (2013). Quality indicators for the design and implementation of simulation experiences: A Delphi study. *Nurse Education Today*, *33*(11), 1357–1361. https://doi.org/10.1016/j.nedt.2012.07.012

Berk, L., Jorm, A. F., Kelly, C. M., Dodd, S., & Berk, M. (2011). Development of guidelines for caregivers of people with bipolar disorder: A Delphi expert consensus study: Caregiver guidelines. *Bipolar Disorders*, *13*(5–6), 556–570.

https://doi.org/10.1111/j.1399-5618.2011.00942.x

Center for Open Science. (n.d.). *Registered Reports*. Retrieved October 25, 2020, from https://www.cos.io/initiatives/registered-reports

Heyman, T., Moors, P., & Rabagliati, H. (2020). The benefits of adversarial collaboration for commentaries. *Nature Human Behaviour,* 1-1.

Jorm, A. F. (2015). Using the Delphi expert consensus method in mental health research. *Australian & New Zealand Journal of Psychiatry*, *49*(10), 887–897. https://doi.org/10.1177/0004867415600891

Kekecs, Z., Aczel, B., Palfi, B., Szaszi, B., Szecsi, P., Zrubka, M., Kovacs, M., Bakos, B. E., Cousineau, D., Tressoldi, P., Grassi, M., Arnold, D., Evans, T. R., Yamada, Y., Miller, J. K., Liu, H., Yonemitsu, F., & Dubrov, D. (2019). *Raising the value of research studies in psychological science by increasing the credibility of research reports: The Transparent Psi Project - Preprint* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/uwk7y

Klein, R. A., Cook, C. L., Ebersole, C. R., Vitiello, C. A., Nosek, B. A., Chartier, C. R., Christopherson, C. D., Clay, S., Collisson, B., Crawford, J., Cromar, R., Vidamuerte, D., Gardiner, G., Gosnell, C., Grahe, J. E., Hall, C., Joy-Gaba, J. A., Legg, A. M., Levitan, C., … Ratliff, K. A. (2019). *Many Labs 4: Failure to Replicate Mortality Salience Effect With and Without Original Author Involvement* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/vef2c

Latham, G. P., Erez, M., & Locke, E. A. (1988). Resolving scientific disputes by the joint design of crucial experiments by the antagonists: Application to the Erez–Latham dispute regarding participation in goal setting. *Journal of Applied Psychology*, *73*(4), 753–772.

https://doi.org/10.1037/0021-9010.73.4.753

Lindqvist, O., Lundquist, G., Dickman, A., Bükki, J., Lunder, U., Hagelin, C. L., Rasmussen, B. H., Sauter, S., Tishelman, C., & Fürst, on behalf of OPCARE9, C. J. (2013). Four Essential Drugs Needed for Quality Care of the Dying: A Delphi-Study Based International Expert Consensus Opinion. *Journal of Palliative Medicine*, *16*(1), 38–43. https://doi.org/10.1089/jpm.2012.0205

Loblaw, D. A., Prestrud, A. A., Somerfield, M. R., Oliver, T. K., Brouwers, M. C., Nam, R. K., Lyman, G. H., & Basch, E. (2012). American Society of Clinical Oncology Clinical Practice Guidelines: Formal Systematic Review–Based Consensus Methodology. *Journal of Clinical Oncology*, *30*(25), 3136–3140. https://doi.org/10.1200/JCO.2012.42.0489

Moshontz, H., Campbell, L., Ebersole, C. R., IJzerman, H., Urry, H. L., Forscher, P. S., Grahe, J. E., McCarthy, R. J., Musser, E. D., Antfolk, J., Castille, C. M., Evans, T. R., Fiedler, S., Flake, J. K., Forero, D. A., Janssen, S. M. J., Keene, J. R., Protzko, J., Aczel, B., … Chartier, C. R. (2018). The Psychological Science Accelerator: Advancing Psychology Through a Distributed Collaborative Network. *Advances in Methods and Practices in Psychological Science*, *1*(4), 501–515. https://doi.org/10.1177/2515245918797607

Murphy, Black, Lamping, McKee, Sanderson, Askham, & Marteau. (1998). Consensus development methods, and their use in clinical guideline development. *Health Technology Assessment*, *2*(3). https://doi.org/10.3310/hta2030

Nair, R., Aggarwal, R., & Khanna, D. (2011). Methods of Formal Consensus in Classification/Diagnostic Criteria and Guideline Development. *Seminars in Arthritis and Rheumatism*, *41*(2), 95–105. https://doi.org/10.1016/j.semarthrit.2010.12.001

Snape, D., Kirkham, J., Preston, J., Popay, J., Britten, N., Collins, M., Froggatt, K., Gibson, A., Lobban, F., Wyatt, K., & Jacoby, A. (2014). Exploring areas of consensus and conflict around values underpinning public involvement in health and social care research: A modified Delphi study. *BMJ Open*, *4*(1), e004217. https://doi.org/10.1136/bmjopen-2013-004217

Sunderji, N., Waddell, A., Gupta, M., Soklaridis, S., & Steinberg, R. (2016). An expert consensus on core competencies in integrated care for psychiatrists. *General Hospital Psychiatry*, *41*, 45–52. https://doi.org/10.1016/j.genhosppsych.2016.05.003

Appendix

Table A1. Exemplary operational documents from previous studies

First draft of the Target Output	Transparent Psi Project: https://osf.io/rh5hb/wiki/Initial%20research%20plan/ Transparency Checklist: https://osf.io/x3g2t/
Invitation of experts	Transparency Checklist, email to target journal editors: https://osf.io/8b52k/ Transparency Checklist, eligibility criteria (in the preregistration)**:** https://osf.io/3tkhn Transparency Checklist, list of target journals: https://osf.io/54pzb/ Transparent Psi Project, eligibility criteria: https://osf.io/he8sb/wiki/Eligibility%20criteria/
Initial round of survey questions	Transparent Psi Project: https://osf.io/2mrn9/ Transparency Checklist: https://osf.io/ptjcz/ Bayesian Inference Study: https://osf.io/u7fxv/
Second round of survey questions including previous rounds summary	Transparent Psi Project: https://osf.io/tsqwa/ Transparency Checklist: https://osf.io/5xh7d/
Stopping rules	Transparent Psi Project: https://osf.io/he8sb/wiki/Stopping%20rules/ Transparent Psi Project (in the preregistration): https://osf.io/3tkhn Bayesian Inference Study (in the preregistration): https://osf.io/q37as
Debriefing	Transparent Psi Project: https://osf.io/qp5rc/