# Introducing synchronous robustness reports

Bartoš, F.; Sarafoglou, A.; Aczel, B.; Hoogeveen, S.; Chambers, C.D.; Wagenmakers, E.-J.

**DOI**
10.1038/s41562-025-02129-1

**Publication date**
2025

**Document Version**
Final published version

**Published in**
Nature Human Behaviour

https://doi.org/10.1038/s41562-025-02129-1

# Introducing synchronous robustness reports

František Bartoš, Alexandra Sarafoglou, Balazs Aczel, Suzanne Hoogeveen, Christopher D. Chambers & Eric-Jan Wagenmakers



Most empirical research articles feature a single primary analysis that is conducted by the authors. However, different analysis teams usually adopt different analytical approaches and frequently reach varied conclusions. We propose synchronous robustness reports — brief reports that summarize the results of alternative analyses by independent experts — to strengthen the credibility of science.

In recent years, several large-scale research projects have highlighted the advantages of 'many-analysts' reports and the pitfalls associated with traditional 'single-analyst' reporting practices. For instance, a 2020 effort to learn the reproduction number, *R*, for COVID-19 infections in the UK involved nine expert teams, each of which independently estimated *R* (ref. 1). The results, shown in Fig. 1, reveal considerable heterogeneity across teams in the reported outcome — this heterogeneity exists despite the clarity of the research question and despite the fact that the expert teams had access to the same abundant information (deaths, hospital admissions and testing rates). Had the estimate of *R* been informed by any single team, the associated narrow uncertainty intervals would have instilled a false sense of confidence in the conclusion. By involving multiple teams, the UK group was able to produce a consensus report that accounted for the kind of statistical uncertainty that usually remains hidden beneath the surface: the uncertainty associated with different analysis pipelines and different modelling choices[2].

This between-analyst variability in conclusions arises across a wide range of empirical disciplines (for example, neuroscience[3], psychology[4], social science[5], medical sciences and epidemiology[6], and economics and finance[7]). Moreover, the variability does not appear to be a result of suboptimal analytic choices[5]. This means that for a traditional single-analyst research project, it can be difficult to predict whether or not the conclusions are robust under alternative plausible analysis procedures.

There is, of course, a straightforward method to gauge the extent to which a scientific conclusion is either fragile or robust: conduct a many-analysts study and assess the heterogeneity across the analysis teams. Unfortunately, many-analysts studies come with considerable logistical challenges and their routine application is therefore unattractive and, arguably, unfeasible. As a solution, Botvinik-Nezer and colleagues[3] and Wagenmakers and colleagues[2] suggested that results of particular scientific interest or societal importance could be accompanied by two or three short reports that summarize the results of alternative analyses — what we term synchronous robustness reports (SRRs).

Here, we provide journal guidelines for seamlessly integrating SRRs into the current publication workflow. We then briefly discuss the direct and indirect advantages of SRRs, as well as anticipated

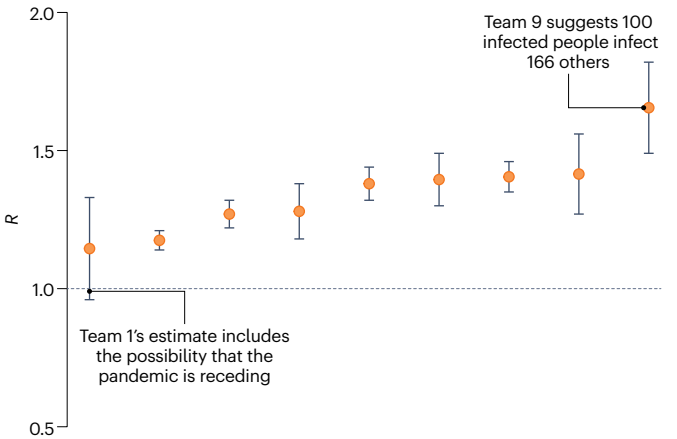


**Fig. 1 | COVID-19 *R* estimates from nine expert teams.** Nine expert teams independently estimated *R* for the spread of COVID-19 in the UK in October 2020. The uncertainty across teams is substantial. Adapted from ref. 2, Springer Nature Ltd, on the basis of data reported in ref. 1. Error bars denote 95% confidence intervals.

questions (Box 1). Because SSRs have not yet been introduced into academic practice, a discussion on their informativeness and feasibility is necessarily speculative. We emphasize that the article-by-article deployment of SRRs is at the editor's discretion, and that a journal can experiment with the format to ascertain its value. We hope that this Comment will be useful in designing such an experiment.

## Proposed format

The goal of SRRs is to offer a concise and immediate reanalysis of the key findings of a recently accepted empirical article. The concise nature of the SRR is essential for the review process to be efficient enough for the reports to be published on the heels of the empirical article, a synchronicity that we regard as essential. The proposed SRR format consists of five short sections:

1. A goal section that outlines the rationale of the reanalysis (for example, the question that the reanalysis is trying to address);
2. A methods section that provides the necessary background information about the reanalysis;
3. A results section that describes the main outcomes of the reanalysis;
4. A conclusion section that compares the results from the reanalysis to those from the original analysis and assesses the degree to which the reanalysis corroborates or undercuts the conclusions from the original analysis; and
5. A code and literature section that must, at a minimum, include a link to a public repository that contains the code used for the reanalysis; this public repository might also contain a more extensive version of the SRR, and can of course also be used to refer to relevant methodological literature.

## BOX 1

# Frequently asked questions

The application of the SRR is always under editorial control; journals will therefore not be swamped by SRRs, and SRRs will not be required to follow every single empirical article. We anticipate the following questions:

**Will SRRs overload the review process?**
It is true that the present system is already overloaded, and adding more articles does nothing to alleviate the pressure. Publishing SRRs does take time and effort. Ultimately, every journal has to negotiate its own cost–benefit trade-off. We believe that the time investment can be minimized by adopting the methodology outlined in this Comment. Moreover, the experience from previous multi-experiment articles is that methodologists are usually enthusiastic about reanalysing topical data with their favourite state-of-the-art statistical procedures. It might even be that the prospect of a potential SRR makes it easier to find statistical experts who are willing to accept review assignments.

**Will SRRs add only limited insight?**
Existing many-analysts studies suggest that SSRs will add considerable insight. The SRR contribution can be maximized when a knowledgeable methods editor selects from a broad and diverse range of methodological and statistical expertise.

**Why publish an SRR when space in journals is at a premium?**
Each SRR is 500 words long and provides a complementary perspective on an important or groundbreaking paper. SRRs are no longer than the letter or correspondence formats that several journals already publish.

**Will SRRs affect the journal impact factor?**
It is not clear what effect SRRs will have on journal impact, including a journal's impact factor. SRRs might enhance the visibility of the associated empirical research, and they might provide methodological inspiration for future empirical work as well.

**Will analysts be interested in contributing SRRs at short notice?**
Recent experience from many-analysts projects[12] suggests that researchers are enthusiastic about participating in reanalysis efforts.

We suggest that, excluding the code and literature section, an SRR be limited to 500 words and a single display element (that is, a figure or table); refs. 8–10 provide an example of a manuscript[8] accompanied by two SRRs[9,10].

## Workflow

To integrate SRRs seamlessly into the publication process, we suggest the framework outlined as a flowchart in Fig. 2. As the flowchart shows, the SRRs form a natural extension to the standard review process.

After an empirical manuscript has been accepted for publication, the journal editor or action editor makes the executive decision whether or not to initiate an SRR. This decision forms a vital component of our proposal, as it means that journals retain complete editorial control over how many SRRs are published and for which articles (note the flow of 'Editorial control' in Fig. 2). Of course, both authors and reviewers could alert the editor to the value of initiating an SRR.

After deciding that an article merits SRRs, the editor needs to invite expert analysis teams to submit an SRR. As the number of SRRs that accompany any particular article is limited, coordinating the contributions is necessary to safeguard methodological diversity, efficiency and quality. Ideally, a methods editor would be assigned with the tasks of inviting teams, coordinating their contributions and reviewing their analyses. It is up to the journal to decide whether the methods editor acts as a reviewer of the analyses or whether they only coordinate the peer review of the analyses. The SRR teams can be invited as soon as the original article has been accepted for publication (or perhaps even earlier, when a provisional decision for acceptance has been made).

We propose that a reanalysis and 500-word SRR need not take more than three weeks before it can be submitted. The SRR review process should be constructive — the SRR teams are experts in the relevant analysis domain, and the SRR is mostly a factual description of an alternative analysis procedure. From invitation to final acceptance, the SRR publication process should to take no more than six weeks. This implies that the gap between the potential publication of the target article and the accompanying SRRs is, at most, six weeks. A journal can then choose either to publish the target article first or to delay the publication of the target article until the SRRs are ready to be published alongside it.

The process can be made even more efficient when potential SRR contributors are involved in the review process of the empirical article. Most journals already accommodate letters or comments — the SRR deviates from these only in terms of the degree of synchronicity and in its focus on reanalysis.

## TOP guidelines

The TOP ('Transparency and Openness Promotion') guidelines provide journals with a concrete list of actions (for example, concerning replication, data sharing, code sharing and preregistration) that they can engage in to encourage scientific transparency[11]. The guidelines have been signed by over 5,000 journals, funders and institutions. It has arguably been critical to the success of the TOP guidelines that they encompass different levels of compliance with best-practice recommendations.

Similar to the TOP guidelines for other transparency-related practices, we therefore propose the following TOP levels for the implementation of SRRs:

- Level 0: do nothing;
- Level 1: invite SRRs for a select subset of articles;
- Level 2: invite SRRs for a select subset of articles, with a methods editor coordinating the process; and

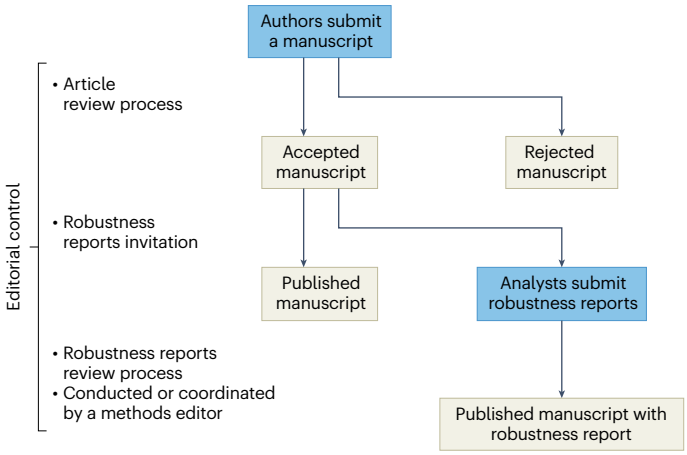


**Fig. 2 | Integrating SRRs into the publication process.** This flowchart outlines the necessary steps for adding SRRs to current publication practices.

- Level 3: invite SSRs for all empirical articles published in the journal, with a methods editor coordinating the process.

## Advantages

The SRR format brings both direct and indirect advantages. The direct advantage is that the SRR enables some inspection of what lurks beneath the surface. This is similar to the recently initiated 'Reproduction and replication at scale' project[12,13] – the main difference is that SRRs are synchronous and follow strict reporting guidelines. Readers are thus alerted to the kind of statistical uncertainty that usually remains hidden from view, and thereby enable an interpretation of the empirical data that is more comprehensive, and fairer. This is particularly important when the original findings have ramifications for scientific practice, public policy or health care.

Among the indirect advantages of the SRR format, we call attention to the following:

- Data can only be reanalysed when they have been properly archived; the adoption of SRRs therefore provides concrete and practical encouragement for authors to adhere to FAIR (findable, accessible, interoperable and reusable[14]) principles of data archiving.
- Authors who anticipate that their work might be selected for an SRR might be particularly motivated to showcase the robustness of their conclusions in the original article, or draw conclusions that are more modest than they would otherwise have been.
- The SRR might facilitate the broader adoption of state-of-the-art statistical methodology, and might boost collaboration between methodologists and substantive researchers.
- The realization that data rarely provide a conclusive story might calibrate expectations from editors and reviewers, and consequently lessen the pressure on authors to present their data in the most favourable light. At the same time, this realization might lead researchers to consider robustness reports for previously published findings as well.

## Concluding comments

We believe that SRRs hold considerable promise as a method to gauge robustness and encourage a more diverse statistical perspective on research with profound scientific or societal ramifications.

To those journal editors who worry that SRRs might be impractical, uninformative or otherwise inadvisable, we issue a challenge: assign a methods editor and task them to arrange SRRs for, say, five empirical articles that will be published in your journal. On the basis of these concrete outcomes you can then make an evidence-based decision on whether or not your journal should to be open for SRRs on a more permanent basis.

To those journal editors who believe that SRRs fall short of best practices and wish to raise the bar even further, note that SRRs can be seamlessly incorporated with preregistration[15], registered reports[16] or analysis blinding[17], or can be extended into the many-analysts format[18].

In sum, SRRs are a straightforward method to test robustness and reveal an important source of uncertainty that usually remains hidden. Their added value can only be assessed by practical implementation, but prior knowledge and limited experience suggest that SRRs are both feasible and informative.

**František Bartoš [1] ✉, Alexandra Sarafoglou [1], Balazs Aczel [2], Suzanne Hoogeveen [3], Christopher D. Chambers [4] & Eric-Jan Wagenmakers [1]**

[1]Psychological Methods Unit, University of Amsterdam, Amsterdam, the Netherlands. [2]Institute of Psychology, ELTE Eotvos Lorand University, Budapest, Hungary. [3]Department of Methodology and Statistics, Utrecht University, Utrecht, the Netherlands. [4]School of Psychology, Cardiff University, Cardiff, UK.
✉e-mail: f.bartos96@gmail.com

Published online: 7 March 2025

## References

1. Scientific Pandemic Influenza Group on Modelling. *SPI-M-O: Consensus Statement on COVID-19, 8 October 2020* (2020).
2. Wagenmakers, E.-J., Sarafoglou, A. & Aczel, B. *Nature* **605**, 423–425 (2022).
3. Botvinik-Nezer, R. et al. *Nature* **582**, 84–88 (2020).
4. Silberzahn, R. et al. *Adv. Methods Pract. Psychol. Sci.* **1**, 337–356 (2018).
5. Breznau, N. et al. *Proc. Natl Acad. Sci. USA* **119**, e2203150119 (2022).
6. Bastiaansen, J. A. et al. *J. Psychosom. Res.* **137**, 110211 (2020).
7. Menkveld, A. J. et al. *J. Finance* **79**, 2339–2390 (2023).
8. Wagenmakers, E.-J. & Bartoš, F. Preprint at *PsyArXiv* https://doi.org/10.31234/osf.io/2u84v (2025).
9. Sarafoglou, A. Preprint at *PsyArXiv* https://doi.org/10.31234/osf.io/6je5r (2025).
10. Hoogeveen, S. Preprint at *PsyArXiv* https://doi.org/10.31234/osf.io/sy5n8 (2025).
11. Nosek, B. A. et al. *Science* **348**, 1422–1425 (2015).
12. Brodeur, A., Dreber, A., Hoces de la Guardia, F. & Miguel, E. *Nat. Hum. Behav.* **8**, 2–3 (2024).
13. *Nat. Hum. Behav.* **8**, 1 (2024).
14. Wilkinson, M. D. et al. *Sci. Data* **3**, 160018 (2016).
15. Hardwicke, T. E. & Wagenmakers, E.-J. *Nat. Hum. Behav.* **7**, 15–26 (2023).
16. Chambers, C. D. *Cortex* **49**, 609–610 (2013).
17. Dutilh, G., Sarafoglou, A. & Wagenmakers, E.-J. *Synthese* **198** (Suppl. 23), 5745–5772 (2021).
18. Aczel, B. et al. *eLife* **10**, e72185 (2021).

## Acknowledgements

The authors were supported in part by an NWO Open Science grant (OSF23.1.030) and a 2024 Ammodo Science Award.

## Competing interests

F.B., A.S. and E.-J.W. declare their involvement in the open-source software package JASP (jasp-stats.org), a non-commercial, publicly funded effort to make both Bayesian and classical statistics accessible to a broader group of researchers and students. The other authors declare no competing interests.