SIPS 2021 Unconference:
'Improving Registered Reports for authors, reviewers and editors'

Session Leaders:

Loukia Tzavella, Ben Meghreblian, Aoife O'Mahony

We have kept anonymity for session participants unless otherwise stated.

Contributors:

David Moreau; d.moreau@auckland.ac.nz, Veli-Matti Karhulahti, vmmkar@utu.fi, Hans IJzerman,

h.ijzerman@gmail.com, Jacqueline Thompson, jackie.thompson@aya.yale.edu

Session Details

For this SIPS unconference session we invited researchers from all career levels to discuss how we can improve Registered Reports (RRs) for authors, reviewers, and editors. Discussion themes included addressing the key challenges surrounding the adoption and implementation of RRs, expanding the format for more types of research, and increasing their overall accessibility. The insights gained from this discussion will be used to guide improvements to the RR process and quality of RRs being published (e.g. RRs study design template, community feedback, quality monitoring). For more details please see our landing page.

Direct Replications & Registered Reports

Registered Reports (RRs) often involve direct replications of original studies that are peer-reviewed before the research is conducted and are published irrespective of the results if they receive in-principle acceptance (IPA). A key issue was identified with the RR process that may require clarification in terms of journal guidelines and improvement from the authors' perspective. Session participants shared their experiences on challenges they encountered when the original author(s) of the study were involved in the peer-review process and requested changes to methods and/or analyses.

We agreed that the original authors' involvement may well depend on the purpose of the replication, as for example whether it is only a specific study protocol or the generalisability of the claims the authors made in their original research. In the latter scenario, the input of the original

authors may be paramount for clarifying which parameters have to be considered for generalisation (e.g. sample or specific stimuli used in an experiment). In theory, such generalisability clarifications should be added to original research to allow for well-constructed and clearly interpreted direct replications.

To avoid such problems in the future there could be clear guidelines that the editorial teams have to follow regarding the **involvement of the original authors in the peer-review process**. For example, should they be invited to provide a full review of the Stage 1 proposal or simply be invited to check the RRR protocol for misinterpretations of the methods and/or claims? Such guidelines should ideally involve editorial actions that can prevent or resolve the influence of bias and potential conflicts of interest in the peer-review process. Editorial guidelines could also emphasize that exclusion requests by the authors should be honoured when appropriate - e.g. accept an exclusion request when there is known conflict between the authors of the original study and the replication study but not when two teams simply have competing theories and engage in *scientific* discourse). Another idea for improving this process was **publishing the reviews openly** as part of the final manuscript, which could deter original authors and other reviewers from requesting too many deviations from the original protocol or negatively bias the RR process.

Reducing the peer-review time at Stage 1

The time for a Stage 1 RR proposal to be thoroughly peer-reviewed and receive IPA may vary across journals and types of research. It is also a very general issue that was mentioned throughout different topics (e.g. from a reviewers' perspective) including the overall accessibility of RRs. For ECRs and other researchers that are employed on short-term contracts and have limited time and/or funding to collect data the total peer-review time can add a significant burden on their overall progress. The same applies to clinical areas of research where there are already potentially long time demands not only for data collection but obtaining ethics approval at Stage 1.

Session participants shared experiences and/or opinions about cases of 'bureaucratic tennis' between the editorial team (including reviewers) and an ethics committee. The requirement for RRs to have ethics approval when a Stage 1 proposal is first submitted may very well be problematic for clinical research, where ethics approval can sometimes take up to 6 months, and for multi-lab studies which require ethics from more than one institution and other cases that are not listed here. If Stage 1 peer-review results in any major revisions to the methods, authors may still need to submit amendments to their approved Ethics. At least as a first step, journal guidelines should accommodate for such cases and provide the authors with the **flexibility to provide a statement regarding ethics approval when IPA is received** (e.g. see policy in PCI RR). If peer-review can

last more than 4-5 months for certain manuscripts, this should not be a problem; by the time the research can actually commence the authors can receive both IPA and ethics approvals.

A very interesting solution for reducing the time of Stage 1 peer-review could be **submitting 'one-pagers' to journals before developing a full protocol** that can often require time in itself (e.g. simulating data, learning about power analyses, coding checks, piloting, internal reviews with many co-authors). Similar to enquiries for submitted commentaries, reviews and opinion pieces, journals could solicit short RR proposals. This idea was discussed with regards to the implementation of a 'custom' RR workflow at *International Review of Social Psychology* (*IRSP*; see https://www.rips-irsp.com/about/confirmatory-reports/). This is consistent with the idea of 'scheduled reviews' (see https://osf.io/preprints/metaarxiv/43298/), where authors can prepare their complete Stage 1 proposal in a pre-specified time (e.g. 6 weeks) and the reviewers' time is booked for that time - i.e. they agree to provide a review in advance. This may mean that the time for peer-review is considerably reduced.

In some studies, like those of qualitative research where interview participants are recruited from special groups, data collection can take a long time and it may be impossible for researchers to wait for the IPA before starting to recruit participants in projects with time restrictions. For these scenarios, journals could provide the option for authors to **apply for a data collection permission before the IPA**. To receive such permission, the editorial team would have to carry out a partial review on the data collection section of the Stage 1 RR and accept the collection procedure (but leave analysis, introduction, and other parts for further evaluation).

Recognising the reviewers' contributions

A prominent discussion issue during the session was the role of the reviewers at Stage 1 and how their contributions may not be recognised in any formal manner when the final manuscript is published. The question of whether to **credit reviewers' contributions using the CRediT Taxonomy** (https://casrai.org/credit/) was difficult to answer as on the one hand 'rewarding' reviewers for their review by adding them as co-authors in the final output can violate existing guidelines (e.g. COPE - unless they are then replaced by other reviewers) and lead to biased reviews, but conversely not recognising their contributions when they have influenced the overall study protocol (theory, methods and/or analyses) is not ideal to say the least. Publishing the reviewers' comments alongside the final manuscript and acknowledging their contributions can easily be implemented without presenting any conflicts of interest. At PCI RR, all reviews are published and reviewers have the option to sign, and their review gets a DOI so it is a citable object. This provides a way for reviewers to demonstrate a scholarly contribution independently of

authorship. Another approach which has been proposed is to list Stage 1 reviewers who make a significant contribution to the design of a study as **"Reviewer contributors"**, separately from the authors. The recognition of reviewers' contributions on published RRs could also be embedded into promotion/tenure decisions. However, although not directly linked to this idea, we highlighted the **potential pitfalls of favouring RRs over standard articles** with the assumption that they can demonstrate greater quality and/or quantity of work when:

(i) Some types of research, such as theoretical studies (humanities) and ethnographic fieldwork (anthropology), are not currently published as RRs. Moreover, while there is a growing number of qualitative RR submissions (QRRs), it is still unknown to what extent will journals publish QRRs: because the methods of analysis and reporting in many qualitative formats involve creativity and uncertainty that cannot be verified beforehand, there may be an increased risk for Stage 2 rejects, ultimately resulting in many qualitative studies not being submitted as QRRs at all. In the future, QRRs may benefit from a special Stage 2 review procedure where the analysis is evaluated by case-specific criteria set during Stage 1.

(ii) Many scholars are incentivised or even required to publish in specific journals. These journals may not support RRs, moreover, some RR positive journals have publication fees that make them out of reach for some teams. These issues may resolve along with the increasing number of RR positive journals, however, they currently set limits for many authors.

Accessibility of Registered Reports

Discussion on how to define the accessibility of RRs resulted in identifying several challenges with the process that could be improved in future developments. First, we should consider whether journal-specific implementations of RR guidelines are helpful to authors who are neurodivergent and how the process can be simplified in terms of manuscript preparation. It may be helpful to assess which workflows for assessing RR criteria are most efficient for authors (e.g. comprehensive checklist before submission versus generic guidelines). Accessibility in this context is complex but one question to ask first is whether there are any RR-specific challenges that may not apply to standard publication formats. The RR process in general tends to be longer, from preparation to final publication, and the procedures at different journals may not always be consistent - with author guidelines not being flexible or specific enough, deeming them difficult to navigate at times (also see Journal-specific RR criteria & implementation). However, problems could occur when guidelines are *too* specific and strict with regards to their actual implementation. For example the requirement to add statements in an RR manuscript that read more like 'legal' rather than 'academic' narratives may be challenging for some authors. Another potential issue with the RR format would be how much it actually depends on writing the Stage 1 carefully so that there are no deviations from the protocol after data collection. Assessment for RR criteria often involves a detailed examination of how specific and consistent the different protocol elements are (hypotheses, sampling plan, statistical tests and contingency plans - if any). For many of these elements language matters (e.g. exactly how you define the expected effect or relationship between variables) and this could add more pressure to authors when preparing the Stage 1 manuscript. It may be too early to tell if RR guidelines present this problem for authors that do not identify as neurotypical, but in the session we agreed that we would benefit from **making the guidelines clear, enforced, but flexible** when possible - with clear indications of where that flexibility lies and what actions an author can take to discuss any uncertainties with the Editor(s).

Modularised reviews for different areas of expertise

Session participants shared a specific approach about RR peer-review which may not be widely known or promoted for journals that currently offer this article format or are prospective adopters. For example, Michael Matthews is an editor at a journal that accepts RRs (*Gifted Child Quarterly*) and described the considerations that the editorial team follows with regards to reviewer selection. The reviewers are familiar and/or supportive of RRs in general and they also get reminders about how the RR process works. More importantly, the **review at Stage 1 can be modularised according to the reviewers' areas of expertise**. For example, manuscripts that present more

complex or advanced statistical analyses will be assigned to a reviewer that has the expertise to review them appropriately. Another reviewer will be responsible for the theory and methods that are specific to their research area/topic. With a total of three-to-four reviewers there can be adequate expertise to cover different manuscript components (e.g. methods, analyses, theory/interpretation). This can also potentially help reduce the workload and time associated with providing an in-depth review for a Stage 1 RR, which would benefit authors, reviewers and editors. This approach is also followed by *IRSP* as indicated by Hans IJzerman. To further reduce the editorial workload associated with identifying areas of expertise and review requirements for a specific manuscript, the authors could complete a short checklist where they identify key areas of their research that may require expert review (e.g. method used, specific theory, statistics, sample/culture discussed). Reviewers could then fill in which areas are applicable to them and any gaps would be identified - this idea was developed in another SIPS 2021 unconference session ("Improving Interdisciplinary Review").

Note. Although checklists can be helpful some aspects of the research are often not explicitly stated in the actual manuscript and this could potentially hinder the process of identifying such areas of expertise for peer-review. Authors could include **constraints on generality** statements/sections in the manuscript and also specify the research sample in the manuscript titles and/or abstracts (e.g. see this paper where the issues of generalizability are discussed with regards to WEIRD samples).

Journal-specific RR criteria & implementation

There may be cases where the official RR criteria and guidelines that are advertised on a journal's website are not the only ones being implemented behind the scenes. Without disclosing further information that could identify session participants, this can be described as an Editor-specific set of rules that are only mentioned after the Stage 1 manuscript has been submitted or new guidelines and checklists that are not included in the instructions for authors. There can also be cases where RRs are desk-rejected based on the Editor's discretion for reasons that are not necessarily compatible with the 'official' criteria[1].

This issue is related to an upcoming development for RRs, which is a **community monitoring mechanism and feedback website for journal submissions** (PhD project by Ben Meghreblian). Journal-specific issues with RRs may not always be communicated by authors to the research community and the level of familiarity with specific editorial teams and procedures may differ according to individual career levels. For example, in the session we discussed how

[1] This issue would be resolved when the peer-review at Stage 1 is conducted independently of a specific journal; for more details see the Peer Community in Registered Reports (PCI RRs; https://rr.peercommunityin.org/about/about).

sometimes early career researchers (self-identified; ECRs) are not adequately informed and/or trained about article submissions and editorial handling procedures. They may be reluctant to contact the Editor directly for clarification or to raise concerns about a manuscript submission.

A related idea for improving the RR peer-review process (not only RRs however) is having more **webinars or drop-in sessions with publishers where authors can receive additional guidance** about prospective submissions, especially when journal guidelines are difficult to navigate. The instructions for authors could further emphasize that authors can send pre-submission enquiries and specifically questions regarding RR criteria before preparing a full Stage 1 proposal. On a last note, the reasons that peer-review can include non-official RR criteria as part of the assessment process may not always be explicit. For instance, as the RR adoption rate increases, many journals may use the generic RR guidelines without further considerations for the demands of the specific journal submissions (e.g. research area, type of analyses, practical limitations for power requirements). Although generic RR guidelines can be a helpful starting point and many variations may potentially cause more problems in the long-term, it may be best to go through different options in the guidelines and **tailor them to journal-specific needs** before this article format is launched.

Other Considerations

The following points were not covered in the session itself. Most were intended as prompts if needed, but did not actually come up in the discussion during the session. However, they are still important aspects to consider and so have been included here as additional points.

RRs using secondary data analysis

While not discussed during this unconference session, another unconference session hosted by Olivia Kirtley and Ginette Lafit raised a number of useful considerations. Specifically, there appears to be a considerable lack of clarity about RRs that use existing data. While examples exist of such RRs, it is not clear how authors can actually demonstrate that they did not access or analyse the data prior to submitting their Stage 1 protocol. As not having seen the data would be a condition of the acceptance decision, any criteria or actions that authors need to take in order to demonstrate this must be made clear by the journal(s). Where researchers have access to a data manager, this person could potentially provide a statement saying that the author has not seen the data.

Other guidelines in relation to writing a stage 1 RR for secondary/existing data could include whether sampling plans/power calculations are necessary to report at stage 1 when the data has already been collected from another source. It seems likely that in the absence of any strong and clear guidelines regarding such expectations, there is potential for considerable variation in how

journals approach and implement RRs that rely on secondary data. Clarifying these expectations and identifying any best practices that could be implemented across journals would be highly beneficial, particularly for promoting re-use of data, and maximising the value of datasets from hard-to reach populations.

At PCI RR this issue can potentially be resolved using the 6-level bias control taxonomy where studies can be assigned to one of 6 levels, ranging from Level 6 (data does not and will not exist prior to IPA) to Levels 5 down to 1 (various levels of bias control for existing data). To complement such taxonomies however data sharing platforms/repositories could have clear sharing regulations where data entries are controlled and authors can demonstrate whether the data were accessed before submission (e.g. Finnish Social Science Data Archive).

How do we increase adoption of RRs?

How can we increase adoption of RRs (at journals), particularly for disciplines/sub-disciplines/approaches where RRs are currently not widely represented. What are the unique challenges associated with this compared to areas where RRs are more common? How can we overcome these challenges?

E.g. Editors may not be convinced that the 'quality' of Stage 2 manuscripts will be high when the review after IPA is constrained only to specific elements (Results, Discussion) and other issues are identified at Stage 2.

Variant formats

Promoting the adoption of RR variants for more flexibility (e.g. results blind review, Exploratory Reports, accountable replications, verification reports)

Exploratory reports

- What are the challenges in adoption and implementation at journals (in common with, and in addition to the challenges experienced in relation to RRs)?
- Whether these formats might help to increase adoption of such approaches in other disciplines/approaches where exploratory research is more widespread/common?
- Example guidelines from *IRSP*: https://www.rips-irsp.com/about/exploratory-reports/