Thoroughly thought through? Experimenting with registered reports

Article  (Accepted Version)


This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/106388/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher’s version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.
We are happy to present our special section on Registered Reports, which includes five papers focusing on key tenets of Social Identity Theory. Rather than reviewing the content of these papers, in this introduction we aim to share what we have learnt from the process of handling RRs and our recommendations for editors as well as authors considering this format.

Registered Reports (RRs) are research plans that are peer-reviewed before the data are collected. As such, the review process focusses on the originality of the research question and the soundness of the employed methods (both under control of the authors) rather than the outcome, the data (that should not be under control of the authors). Despite general enthusiasm, we had some hesitations. One was the editorial team’s almost complete lack of experience with this format. Another concern was whether the already challenging task to find reviewers would be further aggravated by this format and some reviewers’ hesitancy to enter this terra incognita. We proposed a focused special section that would allow the editorial team to trial the process without fully committing to the format on a topic. We invited RRs in the research tradition of Social Identity Theory (SIT), as on the one hand, EJSP has had an intimate connection with this theory SIT since its beginning, and on the other hand, SIT seemed like a research area where the adoption of RRs was not well advanced, but where our call might actually ignite development. What we envisioned was a collection of articles that would each address a basic tenet of Social Identity Theory (or Self-Categorization Theory) and put it to the empirical test according to current standards. Ideally, this would not only help the field to formalize the central tenets of the theor(ies) but also test their empirical robustness.

The Submission Experience

In response to our call, we received sixteen submissions. Two papers were desk-rejected for lack of fit to the call. After adjustments (powering for 90% or even 95% instead of 80%), all remaining papers were assigned to a handling editor (one of the authors of this editorial) who then invited reviewers to evaluate the Stage 1 reports. Needless to say, for many of the reviewers it was also a new experience to judge the merits and the added value of a manuscript that describes an idea but provides no data – and not everybody seemed fond of it (although many reviewers were enthusiastic to learn about and experiment with this format).

As became clear, our initial idea to re-scrutinize a theoretical framework was too ambitious. Arguably, it was a goal difficult to achieve, as the format of an RR Special Section is ill-suited to encourage such submissions for several reasons. First, we imposed time constraints that made it unrealistic for people to actually do the necessary groundwork of demarcating the theory, distill its central tenets and then design a critical study. Second, particularly for RRs it may not even be a good idea to step outside of one’s expertise. Many of the submissions we received felt like research lines that authors were already invested in that had some connection to SIT, rather than research ideas developed in direct response to our call. Admittedly disappointed at first, we quickly realized that this is probably for the best. Registered Reports seemed to work best when authors know their way around in the subject, have experience with manipulations and measures, and ideally have pilot data
or have conducted first studies. All this makes planning for a compelling experiment easier and more convincing.

A common experience for us as editors but also the reviewers who gave feedback was that Registered Reports prompt a different mindset in approaching others’ research. Almost in-built is a more constructive than judgmental approach. Many of us reframed our role into the direction of helping authors to make their research stronger rather than gatekeeper to the holy halls of publication. Although this was a good experience for most of us, there were differing boundaries to what seemed like an adequate contribution to a paper by both editors and reviewers. In other words: How much feedback can you give before you start to do the authors’ work? In a standard submission, an unconvincing theorizing or inadequate research design leave few alternatives to rejection, whereas in Registered Reports, nothing is lost yet at Stage 1 evaluation. In principle, everything can be “repaired”, making it a tough decision to reject a submission.

**Lessons learned**

One lesson we learned throughout the process is that not every pre-registerable study makes sense as an RR. One of the many benefits of RRs is that it facilitates are more balanced representation of the empirical world by getting journals to accept null results, but it is debatable whether null results have the same informational value for completely new hypotheses (as they have for replications of established effects). A paradigm, an idea, an experimental setup that does not “deliver” with the first study might need more refinement, more thought, more experience before officially declaring its wrongness. As an additional consideration, most novel research question cannot be answered in just one study. Admittedly, the encouragement to invest in novel ideas even if the first results are unpromising may sound like a call to torture data and paradigm until they confess. We would argue that this not only about refinement of paradigms, though. In the first, researchers develop a novel paradigm to test a research idea and the analyses yield support in the first nine studies but not in the last. Cumulatively, this still seems like strong support, particularly in light of the fact that one non-significant finding is to be expected if the studies have 90% power. Now, randomness has no particular obligation to create event in this order. It might very well be that the first study the researchers run is the non-significant one. If that is published as an RR and the null effect discourages them and others to further pursue the idea, a false negative emerges. Obviously, being discouraged if the first study fails is not restricted to RRs, but this may serve as an example why we got the impression that null effects – and hence RRs - are not always equally informative. In our experience, the most promising proposals were the ones this addressed unresolved or contested issues within the field. In an ideal case, the review process can then involve reviewers from both position to ensure that any design is a fair test of competing hypotheses.

Our experience also somewhat shifted our perspective of who benefits most from such a format. Our a priori assumption might have been that RRs are particularly attractive to junior scholars: They get a guaranteed publication and often-unpredictable navigation of academic publishing becomes a little more foreseeable for those who need to demonstrate their productivity the most. While all these arguments are still valid, it is our impression that this is only a viable path if junior scholars receive strong support and guidance. In our impression, RRs are a good idea when there is sufficient expertise for both the theoretical background and the employed paradigms. A broad overview over what the relevant and pressing research questions are as well as experience with the planned methods helps to make a stronger case. Even then though, the length of the waiting period between having your study ready to run and waiting for green light from anonymous reviewers can also be seen as a disadvantage for early career researchers under time pressure. Finally, the RR process requires open-minded participants on all fronts. Not only reviewers and editors, but also authors should be genuinely open to being wrong. To the extent that RRs are seen
to authoritatively prove one’s point, we feel it is a wasted chance. RR should be conducted in the spirit of a genuine curiosity and openness to whatever the data will tell.

**Tips for editors**

Based on our experience, we would probably not recommend following our example and having an RR Special Issue. The fixed time schedule typical for a special issue and the need to have all contributions finalized in synchrony does not seem to be well suited for RRs where each project is different and will require different degrees of revision both before and after data collection.

Be clear about how much you want to invest. Clearly, RRs can provide with an opportunity to invest a lot of time by checking every item formulation, every line of code, everything. If that leaves you burned and bitter because your massive contribution is not even acknowledged in any way, adapt your decision criteria. It is fine to demand that RRs Stage 1 report need to meet a higher bar. After all, issuing an IPA is an acceptance. Considering in what shape a paper must be to receive immediate acceptance might serve as a valid comparison here. You can of course decide that you do want to invest and water the plant. Being enthusiastic about the project and giving everything you can to make the project stronger is of course an equally valid option. No matter how you decide, be clear in your communication to the reviewers. It will make your life dramatically easier if reviewers know what you are expecting – a highly detailed list of suggested alteration or a more abstract statement about the theoretical and methodological contribution the paper makes.

Speaking of reviewers: Choose wisely. All of us handled regular papers as well as RRs simultaneously and we as special section editors advised our handling editors to reserve those reviewers for the RR papers with whom they had the best experience. Helpful reviews make an editor’s life easier for any submission, but we felt that this was particularly true for RRs.

Do not lose sight of the forest for the trees. It is very easy to be drawn into discussing methodological details and exact power calculations, but it is arguably (at least for our special section) equally important to pay attention to the theoretical contribution. If a study is methodologically adequate, there might be the instinct to invite a revision or IPA without thinking about what it adds to our understanding of theory.

**Tips for authors**

In light of our experience with the papers we received, it seems the most straight-forward encouragement to authors would be: Dare to be humble! Constrain your thinking to one critical informative test rather than a complex interaction pattern. Aside from greater theoretical clarity - remember that interactions are fiendishly difficult to find and replicate! It helps a lot to simplify designs and avoid designs that require finding a significant 3 x 3 interaction. Such designs are tricky to power adequately, it is tricky to come up with really precise predictions and your hypothesis easily get a bit messy or even arbitrary even if you have strong theories. Keep it simple.

Think clearly about what study you want to submit as an RR. In our perspective, studies that were specifically designed with an RR in mind were more convincing that writing a Stage 1 for a study that was not designed as a focused test of one specific research question. Experience helps, but pilot data and pre-testing also helps convince reviewers that you have experimental control over the phenomenon you are targeting. Always include a positive control, that is a manipulation check that shows that your explanatory construct has indeed been affected by your manipulation.

In Stage 2 report, don’t be tempted to go beyond your data. In our impression, some authors still fell into some of the same traps in the discussion as we see in non-RRs. Accept the RR for what it is without the need to make a stronger case than is commensurate with the data.
Open questions

Having made the experience we are collectively convinced that it is a valuable format, but likely not the only one to improve our science. It is and should remain part of our toolkit, but it is not an all-purpose weapon to overcome the challenges to field faces. We also realized that we have questions about the format that seem unanswered currently.

The overarching question is: What is the function of Registered Reports? Are they seen as an improved way to do what we are doing anyway, the future of our field? Or are they particularly useful for some kinds of research questions but arguably not for other? Might they be ideal to revisit textbook theories, but less ideal to develop a novel research program? Might they work well for clear quantitative predictions derived from a highly formalized theory but less so in fuzzier areas like qualitative research? Every tool is tremendously more useful when we know what it is for.

Are there any opportunity costs? What do we lose by doing research this way? As it is extremely hard to pre-register a series of studies (and there is no reward for series of studies if a single study suffices to receive an IPA), we might end up with single short papers in this format. Are RRs thus a return of single study papers though the back door? And is that what we want in order to improve our science?

Is there an expiration date on IPAs? For this issue, there was a long gap between preregistration and actual data collection and final paper production. Within what period does data collection and final paper production have to be completed, given that, in some areas, timeliness and topicality may feature in editorial decision-making?

How do we incentivize and reward the tremendous work that reviewers (a big shout-out to the amazingly helpful reviewers that helped us in this endeavor) and editors invest in a task that traditionally is the authors’ responsibility (delineating and designing a convincing study)? How much are reviewers and editors expected to contribute? How do we negotiate “authorship” in a collective effort, as many RRs seem to be?

We are curious how this format will further develop and whether we as a field can find answers to these and other unresolved issues on our way towards a more robust and solid social psychological science. Despite all the challenges and setbacks encountered and the lingering questions outlined above, we feel that all parties involved (authors, reviewers, editors...) learned a lot from this experience. We especially feel that we grew as an editorial team, and are thus proud to present this special section composed of original and innovative contributions.