Registered Reports Literature Review

By Aleksandar Bogdanoski, Berkeley Initiative for Transparency in the Social Sciences (BITSS)

Contents

Motivating issues ....................................................................................................................... 2
  Publication bias .................................................................................................................. 2
  Questionable research practices ....................................................................................... 3
Definition of Registered Reports .............................................................................................. 3
  Study eligibility .................................................................................................................. 4
Distinguishing features of Registered Reports ........................................................................ 4
  a. Registered Reports and study pre-registration ............................................................. 4
  b. Registered Reports and results-blind review ............................................................... 5
  c. Registered Reports and grant proposal review ............................................................. 5
Advantages of Registered Reports ............................................................................................. 5
  Publication bias and “the file drawer problem” ................................................................. 5
  Addressing specification searching .................................................................................. 6
Challenges and limitations ......................................................................................................... 6
  Challenges associated with the journal publishing process .............................................. 6
  Challenges associated with pre-specification .................................................................... 7
Bibliography ............................................................................................................................... 8
Appendix 1: Flowchart of the editorial process in Registered Reports ....................................13
Motivating issues

Registered Reports (RRs) are one in a line of innovations in scholarly communication and peer review that look to reform the way scientists share findings from their work with the rest of the scientific community. Other innovations in this realm include results-blind peer review, open (and continuous) peer review, open access, preprints, reporting standards, and other. These are motivated by a number of challenges associated with conventional scholarly communication, including concerns about efficiency, transparency, equity and impartiality.

Nosek and Bar-Anan (2012) propose a useful catalogue of problems in scholarly communication:

a. *No communication*, where studies with null findings are not published or written up by scientists;

b. *Slow communication*, where inefficiencies in the publishing process in academic journals slow down collective learning in a given discipline;

c. *Incomplete communication*, where key parts of the methodological framework of studies are missing;

d. *Inaccurate communication*, where error-detection mechanisms in conventional peer review are hampered by lack of transparency in what is reported, and insufficient resources for thorough peer-review; and

e. *Unmodifiable communication*, where published articles are static and fail to reflect the evolving knowledge in a given topic, outside of the possibility of retraction, publication of errata, or comments on the original article.

The RRs format is motivated by concerns about biases in the way scientific findings are reported, evaluated and published elaborated in categories a) and c) in the catalogue above. In the meta-science literature problems related to no scholarly communication are commonly referred to as “publication bias” (Franco, Malhotra, and Simonovits 2014) and “the file drawer problem” (Rosenthal 1973). Incomplete scholarly communication is generally related to lack of transparency in the way researchers report steps in conducting research, particularly in terms of sample size, analytical models.

At a macro level, it is likely that publication bias and questionable research practices (QRPs) are a result of a mix of perverse incentives in the wider academic research environment, including hyper-competition for diminishing research funding, the emphasis on output quantity in performance measurement, and the changing business model for higher education (Edwards and Roy 2017). However, without discounting the impact of all of those factors, it is necessary to recognize the centrality of the peer-reviewed journal article as the primary means of communication of research methods, data and findings in the scientific community (B. A. Nosek et al. 2015), and look to address publication bias and QRPs in this realm by disrupting incentives for the main parties involved in the publishing process, including researchers, editors and reviewers.

Publication bias

Numerous studies have provided strong evidence of a pervasive results-based bias in academic journal publishing, suggesting that the decision to publish a given study is often influenced by whether its results are novel, theoretically tidy and statistically significant (Dickersin 1990; Franco, Malhotra, and Simonovits 2014; Ioannidis, Stanley, and Doucouliagos 2017; Brodeur et al. 2016). The propensity of academic journals to publish such studies has correspondingly discouraged authors from reporting analyses which may make their studies less “publishable”, or even writing up papers when studies fail to find “publishable” results (Franco, Malhotra, and Simonovits 2014).
Fundamentally, publication bias exists when the published literature is not representative of all available or possible studies (Findley et al. 2016). As a result, it is very difficult to assess the actual state of knowledge in a given scientific field, given that the published literature is inflated with statistically significant results which were most likely results of chance (Brodeur et al. 2016; Gerber and Malhotra 2008; Ioannidis, Stanley, and Doucouliagos 2017). Publication bias also contributes to considerable time and resources being wasted on studies which were already conducted but were never published (Bruckner 2018).

**Questionable research practices**

Given that chances for a study to get published are influenced by the statistical significance and the direction of its results, and publication record clearly plays a role in career advancement in academia (Nosek, Spies, and Motyl 2012), any behavior which maximizes the chances for publication may be considered rational at the level of the individual researcher.

By relying on undisclosed flexibility around excluding certain observations, conducting subgroup analyses, combining different treatment or control variables, transforming data and many other techniques of “torturing data until they confess”, researchers can make virtually any result statistically significant (see Simmons, Nelson, and Simonsohn (2011); Wagenmakers et al. (2012); Leamer (1983); Humphreys, Sierra, and Windt (2013) and Casey, Glennerster, and Miguel (2012)). Such questionable research practices (also referred to as “specification searching”, “p-hacking”, “fishing”, “data mining”, “data dredging”, etc.) deprive hypothesis-testing from its purpose, as they conflate the probability of finding a statistically significant relationship even if none exists as such (Wagenmakers et al. 2012).

Specification searching practices also raise ethical concerns in the form of falsification, which according to the National Science Foundation is defined as intentional “manipulating research materials, equipment, or processes, or changing or omitting data or results such that the research is not accurately represented in the research record” (Title 45, Code of Federal Regulations, Ch. VI).

**Definition of Registered Reports**

The idea of registered reports can be ascribed to Robert Rosenthal, who in 1966 proposed a system for evaluating research where papers would be evaluated and accepted based on the appropriateness and rigor of proposed procedures, either through results-blind review, or even before the study is conducted (Rosenthal 1966, 35–37). In the current form, RRs are empirical articles which are peer-reviewed and receive an editorial decision for publication based on their pre-registered research design, before the actual results of the study are known. The decision to publish the study is based on the importance of its research question, the soundness of its theoretical reasoning, and the credibility and feasibility of its research design.

Registered Reports aim to change the way research is designed, conducted, and published. This publishing format is based on a two-stage peer review process during which authors develop and write a single study, but receive feedback from reviewers on both its research design (Stage 1) and its implementation (Stage 2). Stage 1 submissions include a description of the key background literature and motivation for the study, hypotheses, study procedures, statistical analysis plan, a statistical power analysis, and pilot data (wherever applicable) (Center for Open Science 2017d). Following Stage 1 peer review, high-quality submissions are granted in-principle acceptance for publication (IPA), which constitutes a commitment on the side of the journal to publish the paper, provided that the study is implemented according to the research design accepted at Stage 1. Following IPA, authors implement the study and submit a full manuscript.
(see Appendix 1). The Stage 2 submission undergoes peer review to ensure the study was implemented based on the pre-specified research design and the implementation process meets standards of quality, including pre-specified data quality checks. In cases of deviations from the pre-specified research design, Stage 2 referees also assess whether deviations were theoretically and methodologically justifiable.

Study eligibility

The RRs format is best suited for hypothesis-testing research where study procedures can be pre-specified before the results are known.

In terms of study design, this includes experiments (and quasi-experiments), but also certain observational research designs, particularly cases where researchers use existing data to make a prediction about prospective data (e.g. Neumark (2001), the 2016 Election Research Pre-acceptance Competition (ERPC), the 2016 Italian Constitutional Referendum Research Pre-acceptance Competition (ITANES), and a special issue of Caucasus Survey (see Erlich (2018)).

In terms of data availability, registered reports are best suited for instances where data is yet to be collected at the time of the first submission. Even though journals increasingly accept research with existing data sets, such cases raise questions about how researchers can verifiably demonstrate they had not accessed the data before submitting and give assurance that the proposed analyses is not biased by the data. This may be possible for analyses using non-publicly available administrative or survey data (e.g., data from the Demographic and Health Surveys Program) or administrative data that are yet to be released at the when the first submission is filed.

Distinguishing features of Registered Reports

In its essence, the Registered Reports format synthesizes two distinct research transparency tools and practices: i) pre-registration and pre-analysis plans (PAPs), and ii) results-blind peer review and editorial decision-making. Given these similarities, it is worth to the distinctive features of RRs.

a. Registered Reports and study pre-registration

A key distinction from pre-registration and PAPs is that pre-specified research designs undergo peer review as part of a formalized process within an academic journal. The benefits of this approach extend beyond the traditional function of peer review as a quality vetting process in academic publishing, as the Stage 1 peer review process is also intended to provide authors with constructive feedback early in the lifecycle of a research project (Chambers et al. 2014; Royal Society Open Science 2017; Graf 2017). Moreover, Stage 2 of peer review is meant to help ensure that the pre-specified research design is consistently implemented, and analyses are clearly reported in the final publication (e.g. all pre-specified hypotheses are reported, and exploratory analyses are properly labelled). Even in the life sciences, where clinical trials pre-registration is formally required, multiple studies have shown that without peer review of the pre-specified research design, publications fail to fully report pre-registered analyses (Hartung et al. 2014; Howard et al. 2017) or even pre-register, in the first place (Rayhill et al. 2015). Finally, whereas pre-registration alone may address undisclosed analytical flexibility on the side of the researcher, it does not eliminate the results-based bias from the editorial and peer review process, nor does it guarantee publication. In RRs, editors have no means of knowing the research results when making the decision to grant IPA, and publication is “virtually guaranteed" for articles that reach Stage 2 (Chambers 2014; Brian A. Nosek and Lakens 2014).
b. Registered Reports and results-blind review

RRs have developed into a wider-encompassing publishing model than results-blind review or “sealed-envelope” submissions. Whereas the latter may involve peer review of studies which were already completed (e.g. the special issue of *Comparative Political Studies* - see Findley et al. (2016)), peer review in RRs normally takes place before data collection and analysis. Thus, whereas results-blind review may limit results-related bias in the editorial process (Dufwenberg and Martinsson 2014; Findley et al. 2016), it does not safeguard against QRPs during the study implementation process. By virtually guaranteeing the publication of studies regardless of their direction or statistical significance, RRs promise to limit incentives for researchers to engage in selective reporting or undisclosed specification searching (Chambers et al. 2014; Chambers 2014; Brian A. Nosek and Lakens 2014).

c. Registered Reports and grant proposal review

RRs also bear similarities to research proposals at grant-making agencies, where the substance of the research design is reviewed before the results are available. However, Stage 1 of RRs peer review puts a much greater emphasis on the study design and analysis plan, and Stage 2 ensures that research is consistently implemented (Chambers et al. 2014). Finally, whereas many studies supported by grants are not published, or even written up (Franco, Malhotra, and Simonovits 2014), registered reports have a much higher likelihood of being published.

Advantages of Registered Reports

*Publication bias and “the file drawer problem”*

In RRs, the editorial decision to publish a given study is made based explicitly on its methodological and theoretical merit as proposed at Stage 1 – before the results are known to all parties involved in the process. This is important because null or statistically insignificant results have traditionally faced a higher burden of proof in the refereeing process, as reviewers need to verify if such results were due to incorrect theory or poor implementation (Findley et al. 2016), or their decisions were influenced by their own preexisting beliefs (Bastardi, Uhlmann, and Ross 2011).

In-principle acceptance is a key segment of the RRs publishing format, and most journals explicitly state that “editorial decisions will not be based on the perceived importance, novelty or conclusiveness of the results” (Center for Open Science 2017d). In theory, this should help reduce the cost of producing papers with results that are not theoretically tidy, but are based on rigorous research design (Goodchild van Hilten 2015). An even more modest intervention than results-blind reviewing at eight health economics journals reminding referees not to dismiss submissions with negative findings arguably reduced the incentives for selective reporting and increased the share of publications with null findings (Brodeur and Blanco-Perez 2017).

In fact, an author’s attempt to steer the analysis towards certain type of results may come at a cost during Stage 2 of peer review, because authors are expected to disclose and justify deviations from the pre-specified research design. Any undisclosed deviation may result in retraction of the in-principle acceptance granted earlier.

A highly-detailed registration of the research design prior to data analysis has been often proposed as a possible way to address publication bias and the associated file drawer problem (Christensen and Miguel 2016; Maniadis, Tufano, and List 2015; Miguel et al. 2014). This solution is central to the RR format, as authors submit a detailed methodological and analysis framework as part of their Stage 1 submissions. Seventy-three of the 165 journals which have implemented RR to date also require authors to submit protocols to public registries following Stage 1
provisional acceptance, and practically all of them encourage authors to do so (Center for Open Science 2019). Mandating registration of clinical researchers has made it possible to review a significantly larger body of literature within a given discipline, regardless of publication status (Miguel et al. 2014). At the very minimum, pre-registration can reveal selective reporting practices in a given discipline and point out the challenges in the practice of current reporting standards (Mathieu et al. 2009).

**Addressing specification searching**

Pre-registration and pre-specification of research designs through PAPs makes it easier to differentiate hypothesis-testing (confirmatory) research from hypothesis-generating (exploratory) research, without favoring one over the other (Brian A. Nosek et al. 2017, 6; Humphreys, Sierra, and Windt 2013; Olken 2015; Rasmussen, Malchow-Møller, and Andersen 2011). In RRs pre-registrations are reviewed before and after fieldwork, which is meant to reduce the possibility of QRPs to a minimum. Moreover, by mandating statistical power analyses before data collection (often times with a minimum threshold for statistical power of 90%), RRs can reduce false negatives in comparison to standard articles (Fiedler, Kutzner, and Krueger 2012). Beyond its potential to mitigate specification searching and false positives, pre-specification helps researchers consider the feasibility of their data collection and analysis strategy before its implementation (Olken 2015), and develop more efficient research designs (Daniël Lakens 2014).

**Challenges and limitations**

According to the Center for Open Science (COS), 215 journals have adopted RRs in some form (2017b), and 124 articles have been published as registered reports to date (2017c). Even though these numbers are steadily rising, RRs are still a novelty in academic journals, and their capacity to strengthen the credibility and integrity of research is yet to be confirmed through research. Nevertheless, it is important to delineate some of its limitations and the challenges faced by implementing journals, and consider some of the possible ways to mitigate them.

**Challenges associated with the journal publishing process**

One of the biggest concerns among journal editors about RRs is that null results are less interesting, and therefore may lead to a fewer citations and lower Journal Impact Factor (JIF). A study is currently underway to compare the citation counts of articles published as RRs in comparison to comparable publications in the traditional format (Hummer, Nosek, and Errington 2017), however preliminary results suggest that articles published as RRs have an average citation ratio of 2.07 on Google Scholar relative to the impact factor of the journal in which they were published (Center for Open Science 2017c). According to COS (2017a), this is because the RRs review process ensures that findings are credible and replicable, and editors can still select the studies that address the most significant research questions through credible, innovative and feasible research designs.

A related concern from the editorial perspective is that due to its tolerance for null results, the RRs format may turn certain journals into ‘dumping grounds’ for studies with null results. Whereas an early analysis of registered reports confirms that papers in this format are much more likely to report null findings in comparison to the general literature (Allen and Mehler 2018), such concerns are not justified on the merits of the scientific method. If the overarching goal of science is to build a shared body of knowledge through structured testing of assumptions against empirical data, then “well-designed, well-executed empirical studies that address interesting and important problems (…), utilize appropriate data in a sound and creative manner, and deploy innovative conceptual and methodological approaches (…) have potential scientific and publication merit
regardless of whether such studies’ empirical findings do or do not reject null hypotheses that may be specified” (“Editorial Statement on Negative Findings” 2015). RRs ensure that the scientific contribution of a research endeavor is evaluated based on the soundness of its theoretical reasoning, and the credibility and feasibility of its research design – regardless of whether its results conform with anyone’s pre-existing beliefs and preferences.

**Challenges associated with pre-specification**

Because RRs incorporate pre-registration and PAPs, many lines of criticism against these research transparency tools can be also applied against RRs. Those can be summarized in the arguments that pre-registration incurs **unnecessary additional costs for researchers** (Olken 2015; Mathôt 2013) and may **restrict creativity** in terms of research design and limit data exploration (Ansell and Samuels 2016).

Pre-registration **adds costs in effort and time** at the beginning of the research project, as it requires that authors spell out the analysis and pre-specify dependencies in full before collecting and analyzing the data (Mathôt 2013; Lin and Green 2016). This is particularly true in economics, where studies are based on increasingly complex designs and typically dozens of hypotheses (Hamermesh 2013). RRs proponents point out that the workload to prepare a RR is comparable to that of a conventional article, because decisions about research design and analysis strategies have to be done nevertheless – the only difference is when (Chambers et al. 2014). Most importantly, full research design specification with RRs offers the reward of **acceptance** before the project is actually implemented. Pre-registration also allows authors to develop efficient research designs while ensuring the robustness of their findings (Olken 2015; Daniel Lakens 2017) and to clarify in advance the terms and conditions with project partners and sponsors (Casey, Glennerster, and Miguel 2012; Olken 2015).

Most RRs journals at Stage 2 sanction **failure to disclose deviations** from the pre-specified research design, however tolerate deviations (Center for Open Science 2017d). To maintain IPA, authors may be required to consult the Editor whenever they cannot complete the study as pre-specified (a solution practiced by most psychology journals, including Cortex, Collabra, Cognitive Research, BMJ Open Science, Royal Society Open Science) or simply disclose deviations and provide theoretical and/or methodological justification in their Stage 2 submission (e.g. the Election Research Pre-acceptance Competition and the Comparative Political Studies results-blind review special issue). Beyond tolerance for deviations in RRs author guidelines, journals may also encourage authors to use Standard Operation Procedures (SOPs), which may help them decisions about the implementation of their research design when facing objective challenges they could not have foreseen in their PAPs (see Lin and Green 2016).

A related concern is that RRs may **limit creativity and exploratory analysis**, thereby impeding the discovery of new and unexpected results. However, RRs do not restrict authors from reporting exploratory analyses, but only mandate that those are clearly delineated from the analyses pre-specified at Stage 1 (Chambers et al. 2014). Rather than a limitation of the RRs publishing model in particular, this argument is criticism of pre-registration and PAPs, which place such restrictions in order to preserve the integrity of results of hypothesis-testing (Olken 2015).
Bibliography


———. 2017c. “Registered Reports - Citations.” Center for Open Science. https://docs.google.com/spreadsheets/d/1gDk6bQLT9fH-J67uP0pRnh250eK2ZocyfGzZKlmkK4/edit#gid=0.


Appendix 1: Flowchart of the editorial process in Registered Reports

1. **Submit Proposal**
   - Desk rejection

2. **Editorial pre-screening**
   - Proposal rejected

3. **Peer review (research design)**
   - Study accepted based on pre-results review
   - Author(s) invited to revise and resubmit

4. **Stage 1**
   - Author(s) conduct study based on pre-specified research design
   - Author(s) free to submit to another journal or withdraw

5. **Submit Full Article (with results)**
   - Peer review (research implementation and quality assurance)
   - Article rejected
   - Article published

- Submission advances in the publishing process
- Submission undergoes revision
- Submission rejected