Moving Sport and Exercise Science Forward: A Call for the Adoption of More Transparent Research Practices

Aaron R. Caldwell[1], Andrew D. Vigotsky*[2], Greg Nuckols[3], Ian D. Boardley[4], Julia Schmidt[5,6], Matthew S. Tenan[7], Jakob Škarabot[8], Rémi Radel[9], Mitchell Naughton[10], Brad J. Schoenfeld[11], Johan Lahti[9], Rosie Twomey[12], Keith R. Lohse[13], David T. Mellor[14], Andreas Kreutzer[15], Ian M. Lahart[16], John P. Mills[17], Jean-Benoit Morin[9], Brad P. Dieter[18,19], Zad R. Chow[20], Brooke Bouza[21], Zachary Zenko[22], Jason L. Neva[5], Richard S. Severin[23,24], Matthieu P. Boisgontier[5], James Steele[25,26], David Nunan[27], Bret Contreras[28], Cody T. Haun[29], Megan E. Rosa-Caldwell[1], Brandon M. Roberts[30], Boris Cheval[31], Matthew W. Miller[32], Duane Knudson[33], Sue Peters[5], and Israel Halperin[34,35]

[1]Exercise Science Research Center, University of Arkansas-Fayetteville, Fayetteville, AR, USA, [2]Department of Biomedical Engineering, Northwestern University, Evanston, IL, USA, [3]Sport and Exercise Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, [4]School of Sport, Exercise, & Rehabilitation Sciences, University of Birmingham, Birmingham, UK, [5]Department of Physical Therapy, University of British Columbia, Vancouver, British Columbia, Canada, [6]Department of Occupational Therapy, La Trobe University, Melbourne Australia, [7]Department of Biomedical Engineering, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, [8]Faculty of Health and Life Sciences, Northumbria University, Newcastle upon Tyne, UK, [9]Laboratoire LAMHESS, Université Côte d'Azur, Nice, France, [10]University of New England, Armidale, New South Wales, Australia, [11]Health Sciences Department, CUNY Lehman College, Bronx, NY, USA, [12]Faculty of Kinesiology, University of Calgary, Calgary, Alberta, Canada, [13]Department of Health, Kinesiology, & Recreation; Department of Physical Therapy and Athletic Training; University of Utah, Salt Lake City, UT, USA, [14]Center for Open Science, Charlottesville, VA, USA, [15]Department of Kinesiology, Texas Christian University, Fort Worth, TX, USA, [16]Institute of Human Sciences, University of Wolverhampton, Wolverhampton, UK, [17]School of Sport, Rehabilitation and Exercise Sciences, University of Essex, Colchester, UK, [18]Washington State University, Pullman, WA, USA, [19]Providence Medical Research Center, Providence Health Care, Spokane, WA, USA, [20]Department of Population Health, NYU Langone Medical Center, New York, NY, [21]Department of Health, Human Performance, and Recreation, University of Arkansas-Fayetteville, Fayetteville, AR, USA, [22]Department of Kinesiology, California State University Bakersfield, Bakersfield, CA, USA, [23]Department of Physical Therapy, The University of Illinois at Chicago, Chicago, IL, USA, [24]Doctor of Physical Therapy Program, Baylor University, Waco, TX, USA, [25]ukactive Research Institute, London, UK, [26]School of Sport, Health, and Social Sciences, Solent University, Southampton, UK, [27]Centre for Evidence-Based Medicine, University of Oxford, Oxford, UK, [28]Sport Performance Research Institute, Auckland University of Technology, Auckland, NZ, [29]Department of Exercise Science, LaGrange College, LaGrange, GA, USA, [30]Department of Cell, Developmental and Integrative Biology, University of Birmingham at Alabama, Birmingham, AL, USA, [31]Department of Psychology, University of Geneva, Geneva, Switzerland, [32]School of Kinesiology and Center for Neuroscience, Auburn University, Auburn, AL, USA, [33]Department of Health and Human Performance, Texas State University, San Marcos, TX, USA, [34]School of Public Health, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel, [35]Sylvan Adams Sports Institute, Tel Aviv University, Tel Aviv, Israel

January 12, 2019

Abstract

The primary means for disseminating sport and exercise science research is currently through journal articles. However, not all studies, especially those with null findings, make it to formal publication. This publication bias towards positive findings may contribute to questionable research practices. Preregistration is a solution to prevent the publication of distorted evidence resulting from this system. This process asks authors to register their hypotheses and methods *before* data collection on a publicly available repository or by submitting a Registered Report. In the Registered Reports format, authors submit a Stage 1 manuscript to a participating journal that includes an introduction, methods, and any pilot data indicating the exploratory or confirmatory nature of the study. After a Stage 1 peer review, the manuscript can then be offered *in-principle acceptance*, rejected, or sent back for revisions to improve the quality of the study. If accepted, the project is guaranteed publication, assuming the authors follow the data collection and analysis protocol. After data collection, authors re-submit a Stage 2 manuscript that includes the results and discussion, and the study is evaluated on clarity and conformity with the planned analysis. In its final form, Registered Reports appear almost identical to a typical publication, but give readers confidence that the hypotheses and main analyses are less susceptible to bias from questionable research practices. From this perspective, we argue that inclusion of Registered Reports by researchers and journals will improve the transparency, replicability, and trust in sport and exercise science research.

*Corresponding Author: Andrew D. Vigotsky (avigotsky@gmail.com)
This work is a preprint and has not yet been peer-reviewed.
All authors have read and approved this version of the manuscript for pre-print.
This work can be cited as: Caldwell et al. (2018). Moving Sport and Exercise Science Forward: A Call for the Adoption of More Transparent Research Practices. SportRχiv Preprint. doi:10.31236/osf.io/fxe7a

1 Introduction

Reproducibility and replicability are defining features of science [1]. Many researchers publish studies that fail to meet the criteria of reproducibility ("the ability of a researcher to duplicate the results of a prior study using the same materials as were used by the original investigator" [2]) and replicability ("the ability of a researcher to duplicate the results of a prior study if the same procedures are followed but new data are collected" [2]) [3, 4, 5]. This may be due, in part, to the widespread adoption of questionable research practices (QRPs) [6, 7], which represent a major obstacle for reducing uncertainty in scientific research. QRPs can take various forms, such as the post-hoc manipulation of hypotheses after the results are known (i.e., HARKing), manipulating data to meet the conventional alpha-level (i.e., p-hacking), selectively discarding non-significant results (i.e., cherry picking), only publishing 'statistically significant' findings (i.e., the file drawer Problem), conducting underpowered research, primary outcome switching, or fraudulently fabricating data [8, 9]. Current evidence suggests that while QRP is widespread, it may not represent the majority of research [6, 10]. For instance, about 2% of social scientists admitted to fabricating, falsifying or modifying data or results, and approximately one-third have admitted to employing other questionable research practices [11]. In nutrition, a field adjacent to sport and exercise science, recent investigations of questionable research practices have led to the retraction of numerous high profile research articles [12].

Although the prevalence of such QRPs is yet to be established within sport and exercise science, given the interdisciplinary nature of this field and the direct overlaps with both the psychological and biomedical sciences, there is little reason to believe that this field is immune to the issues other fields are facing [13, 14]. For example, the very public mistakes found within the "Pacing, graded Activity, and Cognitive behaviour therapy; a randomised Evaluation" (PACE) [15] trial are likely the result of QRPs and undisclosed analytical flexibility. Sampling and statistical analyses within sport and exercise science have long been known to be underpowered and produce biased effect sizes [16]. We suggest there is an urgent need for improved scientific practice and transparency within sport and exercise science to avoid attempts to build upon a fragile scientific foundation. Here, we outline how several QRPs infect scientific practices and suggest a few potential cures for sport and exercise science. This article focuses primarily upon sport and exercise science, which is synonymous with kinesiology though it is likely that our discussion here will relate to fields like athletic training, ergonomics, rehabilitation, and sports and exercise medicine.

2 Common Questionable Research Practices

2.1 HARKing

In confirmatory research, hypotheses and research questions should be clear from the outset of the experiment. However, too often hypotheses and research questions are unspecified prior to data collection and analysis, are occasionally formulated to fit the observed data, and are subsequently reported without indication of *post hoc* conceptualization. Kerr [17] referred to this as "hypothesizing after the results are known," or simply HARKing. Whilst problematic, HARKing may result from hindsight bias or a poor understanding of scientific research practices, rather than by intentional deception [17]. Regardless, this practice distorts scientific understanding by creating the perception that a study's results were more certain—or predictable—than they were in reality [18]. While researchers should be open to serendipitous findings, they should be careful to avoid overinterpreting statistical noise [19, 20].

2.2 P-Hacking and Data Dredging

Even the most rigorous researchers can overinterpret data due to the ease of modern data analysis [21] increasing the risk of apophenia—the tendency to see patterns in random data [22]. For a single dataset, there may be hundreds or thousands of analysis options [23, 24, 25], which creates a "garden of forking paths" [26, 27, 28], and thus enables the overinterpretation of data. For instance, the average sport and exercise scientist can easily open point-and-click software and produce dozens of analyses of the same data within minutes (e.g., by adding or removing covariates, considering various means of operationalizing an outcome measure, or adding or removing sup-populations).

When the analysis plan has not been registered in advance, researchers may attempt multiple statistical analyses or data transformations, but then only report the analysis which best fits their biases or hypotheses. It is likely that many exercise scientists (particularly early career scientists) are unaware that this is poor practice, and may be encouraged to engage in such practices under the guidance of equally naïve senior colleagues [29]. Analytical flexibility may entice "p-hacking," or the re-analyzation of data until a "statistically significant" p-value is observed when no effect truly exists [30, 31, 32]. With a multitude of analysis options, researchers can easily find a desirable, likely significant, result, and this analytic flexibility occurs unbeknownst to the reader. With the alpha level fixed at 5% and a multitude of analysis options, a statistically significant result can almost always be found if nothing is planned to correct for the multiplicity of tests or the optional cessation of data collection [33, 34].

2.3 Cherry Picking and the File Drawer Problem

There is good evidence to suggest that the scientific literature in most fields is biased toward reporting statistically significant results, which has created a distorted view of reality (Figure 1) [35]. This is, in part, caused by publication bias or a "file drawer problem," where negative results from original studies and meta-analyses are less likely to be published than those reporting statistically significant results [5, 20]. Moreover, publication bias extends to situations wherein positive or novel results are more likely to be published than those that make incremental advancements in knowledge. Although there now exists a number of journals that publish

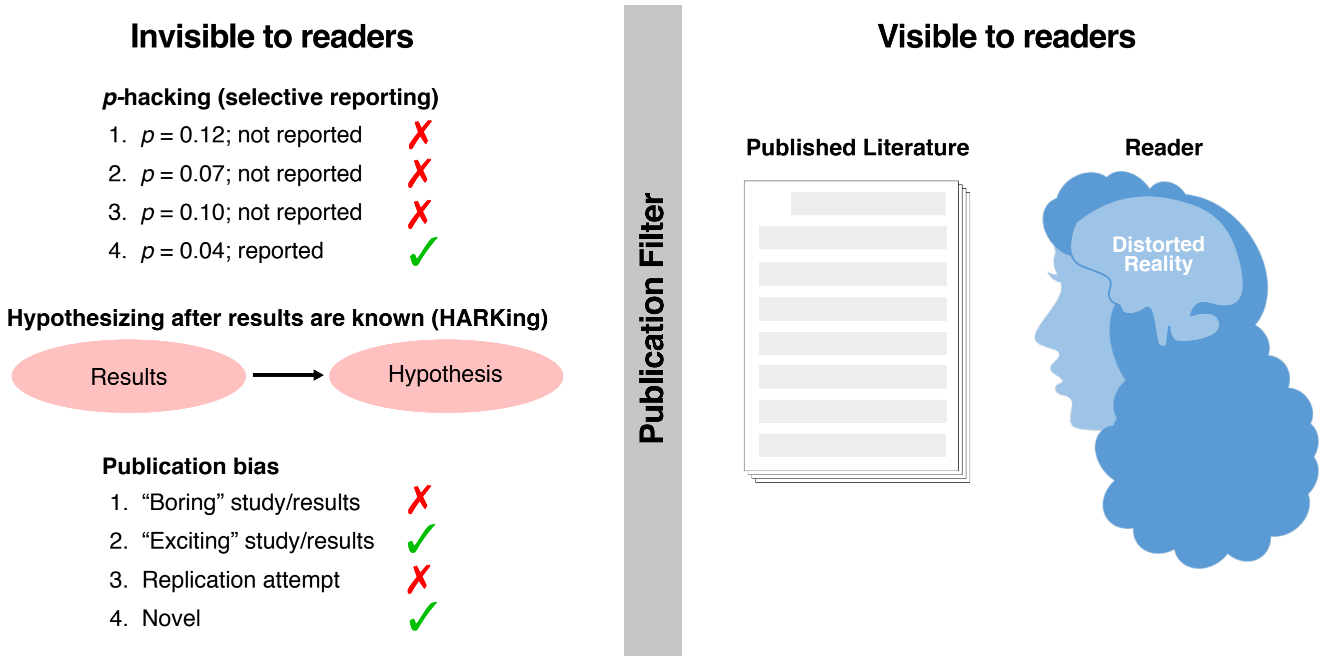

Figure 1: *Researchers' Distorted View of Reality.*
Researchers carry out numerous studies and perform many statistical tests, but not all of them are reported or published. Moreover, those results that are reported are not necessarily hypothesized *a priori*. These biases act as a filter, which distorts the findings present in the publish literature, providing readers (researchers) with a distorted view of reality.

negative results and help reduce the prevalence of publication bias (e.g., *Journal of Articles in the Support of the Null Hypothesis*, *Negative Results: Scientific Journal*), these journals are not popular among sport and exercise scientists. It is doubtful that sport and exercise science researchers will readily invest time to write manuscripts to submit to these less prestigious outlets. Such biases have likely contributed to the current replication crisis by inflating the rate of false positives in the scientific literature [20]. In addition to false positives, more extreme observations, or larger effect sizes, are more often published because small studies have to report a large effect size in order to reach statistical significance thresholds [36, 37]. Overall, the current publication system favors and incentivizes a number of practices that distort reality by preferentially selecting for likely false or misleading effects.

3 Solutions

3.1 Reclassifying the Types of Research

We support a general publishing framework which classifies all empirical research (including meta-analyses and systematic reviews) as either exploratory or confirmatory. Exploratory research is theoretically defined as research where the goal is to gain familiarity with a phenomenon and develop hypotheses [38]. Confirmatory analyses theoretically occur when a specific research question is being asked based on theory and a predefined statistical hypothesis is tested. In the practice of publishing, we propose that the practical difference between exploratory and confirmatory analyses is made transparent through study preregistration. Exploratory analyses are subject to greater researcher degrees of freedom [39] and, while there is a great potential for highly innovative findings, there is also a higher risk that the results will not be reproducible or will reproduce with a far smaller effect size [40]. Ideally, confirmatory research would have to be registered in advance of data collection on a publicly available medium. This approach would prevent changes to the original hypotheses and statistical plans after observing the data or, in the rare case that deviations to the analysis plan are necessary, the process ensures the deviations are transparently reported and justified. [19]. To date, there are a variety of ways to register the protocol of a study. First, researchers can utilize preregistration by posting falsifiable hypotheses and specific analysis plan commitments to independent registries; for example those operated by the National Institute of Health (ClinicalTrials.gov), private publishers such as BMC (ISRCTN registry), or by the nonprofit Center for Open Science (Open Science Framework). These registries can then independently preserve the committed analysis plan and archive these plans for use in the future. Second, a new format of publication has also been created in academic journals to allow researchers to register their study. While some journals support the publication of the protocol only as a complete paper, other journals also now offer a new format, called "Registered Reports," which includes the registration of the study protocol as a first step of the reviewing process before publishing the completed study with its results. After detailing these different options, we explain why we believe Registered Report is an appropriate solution to promote rigorous and less biased confirmatory research and elevate scientific standards in sport and exercise science.

3.2 Preregistration

Preregistration allows the reader to distinguish between which discoveries or findings were predicted or hypothesized (confirmatory), and which were made after the fact (exploratory). This will ensure that confirmatory findings were indeed hypothesis-driven from the outset of the experiment, and thus are more robust than the uncertainty of *post hoc* or exploratory analyses. Preregistration in no way precludes authors from performing and presenting exploratory analyses, but it does require authors to label them as such. Indeed, by making the distinction between confirmatory and exploratory work more clear, preregistration is likely to encourage unplanned discoveries, as was found when seven Registered Reports were conducted on a controversial finding in social psychology [41]. As Jonas et al. [41] stated in their review of power poses, "…a strong contribution of preregistration is evident in the exploratory analyses conducted across the different studies. Most of the studies did reveal some effects of power poses on [several psychological outcomes in] non-preregistered, exploratory analyses. The preregistration format, rather than inhibiting scientific discovery or exploration, actually then points researchers to the next direction for their research, while at the same time making it clear to the reader that such obtained effects were exploratory and not confirmatory." As an indicator of preregistration efficacy, compared to the original studies, preregistered replications often find smaller and non-statistically significant effects [42, 43, 44].

While preregistration can improve the quality and transparency with which science is conducted, it is not without its shortcomings. First, preregistration does not prevent researchers from making theoretically or biologically implausible hypotheses or predictions. For example, there is no mechanism in place to prevent an ardent astrologer from predicting that zodiac signs influence athletic performance [45]. No matter where they are hosted, preregistrations are not typically reviewed by peers prior to data collection and analysis, possibly harming the quality of the final publication [46]. Second, while the researcher declares their beliefs or hypotheses when using preregistration, there is no assurance that reviewers will agree with the preregistered approach. Peer reviewers are also likely to be influenced by their preexisting beliefs which can bias their review [47]; for example, the data itself may influence a reviewer's decision rather the quality of the methods. Therefore, a researcher may not feel motivated to do the additional work to preregister a study when there is no mechanism to prevent such hindsight bias in reviewers and editors [44].

3.3 Registered Reports

A new publication format, Registered Reports, address many of the shortcomings of the traditional publication process, in addition to preregistration alone. At the most basic level, Registered Reports function similarly to the traditional publishing process, except reviewers and editors of a journal approve a study prior to data collection, and the final results are then interpreted and again reviewed by the editors and reviewers. Registered Reports are submitted to a journal, instead of a repository, prior to data collection for an initial review by an editor and reviewers. The manuscript is then approved for data collection, and the final manuscript is reviewed again prior to publication to ensure the approved protocol and analyses were followed and interpreted correctly. In doing so, reviewers ensure that any deviations from the original protocol are detailed and justified.

Registered Reports are reviewed twice: once before data collection, and again after results are known and discussed. The initial submission includes an introduction and a methods section that reviewers can critique and provide suggestions for prior to the start of data collection. Following a successful "Stage 1" peer review,[1] the article is given an "in-principle acceptance" (IPA). The authors can then proceed to collect data that adhere to their IPA plan. When data collection and analyses are completed, and a discussion is written, the authors then submit a finalized manuscript, at which point "Stage 2" peer review occurs. In this stage, the reviewers and editors evaluate the entire manuscript. The primary aims of the Stage 2 review are to determine adherence to the IPA plan and evaluate the presentation and interpretation of the results. This review process ensures that the experimental design, methods, and statistical analysis are appropriate for the proposed study. Furthermore, publication occurs regardless of the results of the study (i.e., reduces publication bias). An outline of the Registered Reports process can be found in Figure 2.

3.4 How do Registered Reports differ from preregistration?

Registered Reports are more formal and undergo peer-review before the experiment is carried out. Furthermore, Registered Reports provide authors peace of mind that publication is not dependent on results, and the Registered Reports system cannot be "cheated." For example, it is possible to preregister multiple analytic plans for a single experiment under separate preregistrations, then only report the results from the most favorable preregistration.

Registered Reports are a natural and logical extension of the preregistration process. This process allows researchers to pursue questions and hypotheses regardless of the outcome, and publication in a relevant journal regardless of the novelty or "statistical significance" of the results. Reviewers and editors can have the peace of mind that the methods and rationale are sound before they see the data. In the domain of sport medicine, a study indicated that less than 60% of the registered clinical trials resulted in publication [48, 49], and many studies do not disclose changes to the data collection or analysis plans [49, 50]. Registered Reports avoid this problem; the Stage 1 review and IPA process lock authors into a set of hypotheses and procedures. Finally, if the authors were to withdraw their IPA, then the journal could publish a withdrawal notice, which in concept is similar to an article retraction notice [51].

Registered Reports help avoid some of the problems of the current published literature, including publication bias, hindsight bias, and undisclosed statistical analysis flexibility [19, 52, 53]. The current publication system often tempts au-

[1]The reviewers find that the research question makes some meaningful contribution to the field and that the proposed methods are sound.

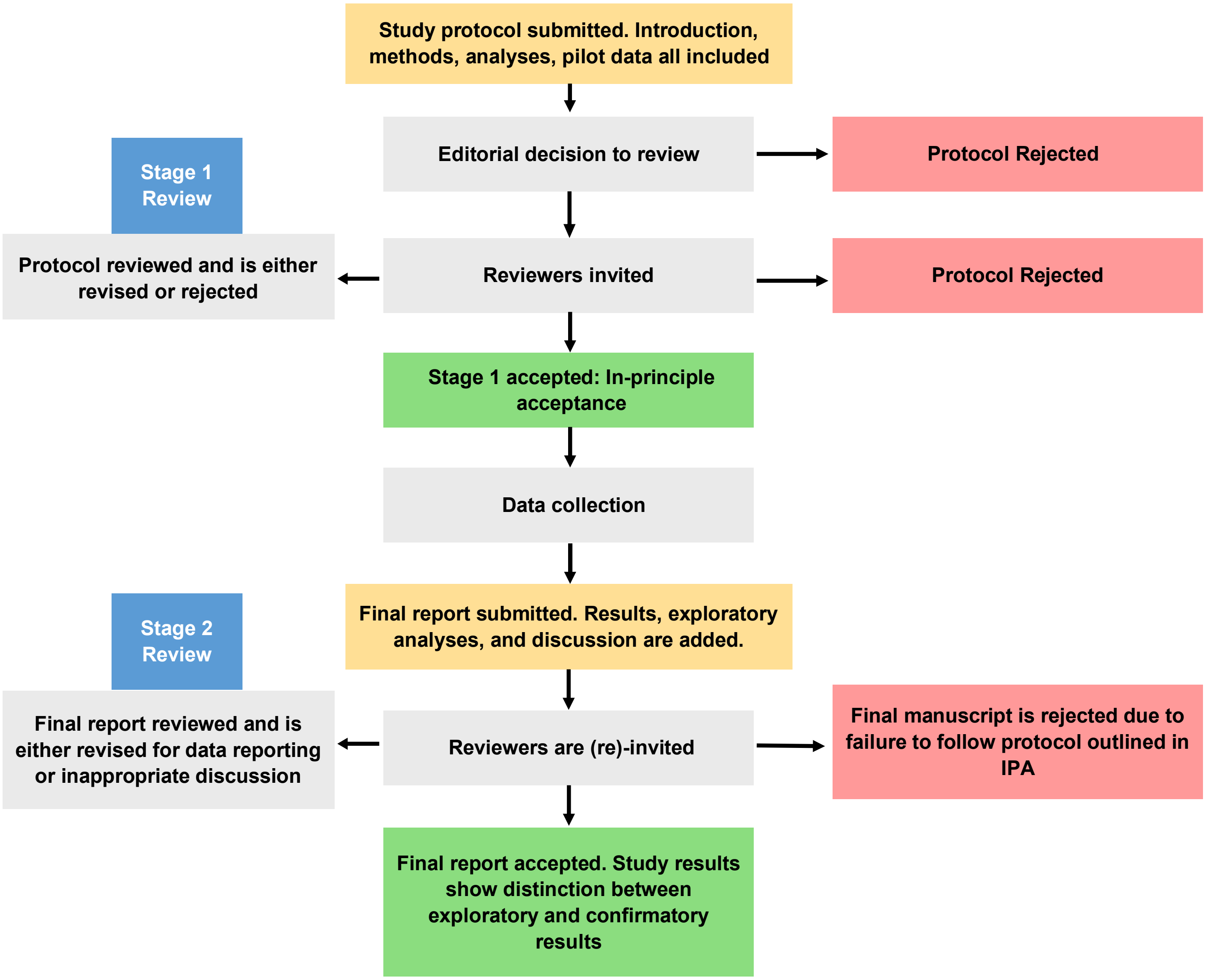

Figure 2: *The Registered Reports Process.*
Before starting data collection, the authors submit the study rationale and methods for peer-review (Stage 1). After the study is scrutinized by the editor and reviewers, it will either receive an in-principle acceptance (IPA) or is rejected. If the study receives an IPA, the authors may proceed to data collection. Once the authors complete the study, they are to analyze and interpret the data in accordance with the Registered Report that was accepted in Stage 1. The authors then re-submit the completed study for Stage 2 review, which is ideally quickly accepted under the condition that the results are interpreted reasonably, and the study was completed in accordance with the methods proposed in Stage 1. Yellow = submission by the authors; red = rejection; green = acceptance.

thors to perform questionable research practices for several reasons. There is strong empirical evidence from other fields (e.g., psychology) that, under the current publication system, authors will often pick analyses, and change hypotheses, to create a more publishable narrative [7]. Registered Reports can avoid this pitfall via the Stage 1 review process. Authors will have to adhere to sound methodological and analysis plans they agreed upon in Stage 1 which prevents hypotheses switching, hacking analyses for significance, and selective reporting of outcomes or analyses.

3.5 Possible Barriers, Gaps, or Problems

Registered Reports are a relatively new phenomena with the earliest journals adopting the practice in 2013 [54]. There is, however, emerging evidence regarding Registered Reports efficacy [51]. Numerous journals have adopted the practice (see cos.io/rr), with psychology and medical journals being the most prevalent adopters [51]. Unfortunately, sport and exercise science journals are still under-represented on this list, which presents a major difficulty for sport and exercise science researchers who would like to adopt this practice.

The primary cause for concern in Registered Reports is a lack of transparency [51]. In most cases, the IPA is publicly available following final publication of the Registered Report, so readers can view the original data collection, analysis plans, and potentially pilot data. It is also very encouraging to see that, at the time of publication of this manuscript, there have been no reports of author withdrawal following the IPA. Specifically, Hardwicke and Ioannidis [51] expressed concerns regarding (1) a lack of consistency in policies between journals and (2) a lack of transparency regarding the IPA. These problems should easily be solved with time, as journal editors determine the best policies for their respective fields and determine an appropriate way to catalog the initial IPA. Moreover, there are efforts to assist journals that allow Registered Reports by providing centralized quality control through the Open Science Framework [51].

Scientists may worry that this new publication format will raise the bar or move the goalposts for what is necessary to produce publishable science. However, neither Registered Reports nor preregistration are meant to replace current publishing practices. Instead, Registered Reports complement the current publishing system by providing a new path to publication. Further, Registered Reports do not diminish the importance of exploratory research, but rather, allow the reader to understand and separate what is exploratory versus what is confirmatory. In fact, it is entirely possible to include *post hoc* analyses in Registered Reports, but the authors will have to distinguish this from other results by creating an "Exploratory Analyses" section. In other words, Registered Reports encourage transparent science without affecting traditional publication routes or the ability to include exploratory analyses.

Opposition to Registered Reports may also come from both authors and editorial boards worried about the time commitment involved, considering there are two (rather than one) stages of peer review. Authors may be concerned about the increased time committed to amending ethics documents to appease reviewers suggested changes to the protocol. Further, editors and reviewers may require changes to the methods that conflict with those outlined in an already-awarded grant. In cases where an agreement between the authors and the reviewers cannot be reached, a Registered Report may not be possible. Finally, it is up to the editor to decide if the required revisions to the protocol are feasible.

The Stage 1 review process reportedly takes nine weeks on average to reach a first decision (cos.io/rr). However, the Stage 2 review process is undoubtedly considerably faster than the typical handling of a final manuscript. First, the reviewers are already identified and have agreed to review the Stage 2 submission. Second, the reviewers have already agreed upon the study rationale, methodology, and analysis plan. Traditionally, it is not uncommon for manuscripts to be submitted for review to multiple journals and reviewers prior to an eventual acceptance—a process which often takes months. Registered Reports can help alleviate two major publication problems that lead to systematic rejection and increased reviewer workload: (1) methodological shortcomings and (2) low perceived contribution and/or novelty of the study results. Indeed, the Stage 1 review helps prevent methodologically flawed research from being performed in the first place, by allowing reviewers to comment on the methods and design prior to data collection. The IPA policy reassures authors that they are evaluated based on the importance of their research questions and the quality of their study design; not on the perceived novelty or originality of the results.

Notwithstanding the inherent limitations of Registered Reports—or, indeed, any publishing format—we believe the benefits greatly outweigh the challenges. Science should prioritize quality over quantity. To this end, Registered Reports may be worth the extra time for increased transparency and, potentially, replicability [44].

4 Example Vignette for Comparing Publication Models

To help illustrate the benefits of Registered Reports, in addition to what it may look like in our field, we will draw a hypothetical scenario that researchers may find familiar. Let us assume a hypothetical research group is interested in the effects of a supplement on muscular strength based on previous research. To answer this question, the hypothetical research group decides to measure several variables (e.g., handgrip strength, isokinetic knee extension and flexion strength, leg press strength, and bench press strength) in an arbitrary sample of 20 "recreationally active young adults," randomly assigned to two groups. Researchers train both control and supplementation groups over a period of eight weeks. The pre- and post-intervention data are collected and analyzed; most of the results are negative, and the data are more variable than expected. Therefore, the Principal Investigator suggests log-transforming the data, dropping the handgrip strength and isokinetic data due to its low practical importance to weight lifters, and excluding 3 participants with less than two years of training prior to the start of the study. The final results indicate a statistically significantly greater improvement in the experimental group for bench press but not leg press. The research group then theorizes in the final manuscript that a) the study was underpowered to detect a difference in leg press given the variability of the effect, b) the results were "trending towards significance" [55, 56], and more time would be needed to detect a difference in leg press strength, assuming a positive effect of the supplement, or c) the supplement only has a positive effect on bench press strength in these participants. In reality, it is highly plausible that the observed effects of the supplement are spurious, and that the *post hoc* data analysis and accompanying narrative are dubious, speculative, and intellectually dishonest.

Instead, let us suppose the hypothetical research group decides to use the Registered Reports system. First, the Stage 1 review would identify the analyses as exploratory or confirmatory; in this case, the analyses are intended to be confirmatory. This Stage would also flag the problems regarding the measurement of numerous, likely correlated dependent variables collected in the study, assumptions regarding the practical importance of observed changes, sample size justification (e.g., *a priori* power analysis), and the participant inclusion/exclusion criteria. In particular, Stage 1 review would reveal the degrees of freedom in the data analysis plan. For example, reviewers would likely require the authors to detail the criteria for data analysis, including the application of specific statistical tests, thereby limiting the number of "forking paths" [26]. At the very least, the research group would have to report all of the results from the initial analyses. Reporting additional outcomes as exploratory analyses—involving exclusion of certain participants—or descriptive statistics could then be presented as additional information with sufficient justification. The final manuscript would be both more reliable and transparent to the reader due to the Stage 1 review, and the full representation of the results since the authors were required to report all the results and originally planned analyses. Registered Reports can improve the quality of sport and exercise science research by limiting analytic flexibility, improving methodological quality, and ensuring honest analyses and transparent reporting.

5 Conclusion

The categorization of analyses into exploratory and confirmatory facilitates the publication of all types of research while highlighting their respective strengths and weaknesses. Meanwhile, Registered Reports are a critical tool for moving sport and exercise science into more transparent scientific practices. This new publication format is not a catch-all solution to problematic scientific practices,[2] but, as highlighted above (see vignette), it does provide a new incentive structure that will help to minimize issues in this regard. For those who are unable or not interested in submitting a Registered Report, we highly recommend utilizing the existing resources for preregistration such as the Open Science Framework (osf.io) or AsPredicted (AsPredicted.org). Those interested in adopting Registered Reports are highly encouraged to read more at the Center for Open Science (cos.io/rr/), and contact the editors of journals in which they would like to publish Registered Reports. Editors may be resistant to adopting a new publication format, and it is unlikely that every journal will need to use or offer Registered Reports as an avenue to publication. However, a number of researchers, as evidenced by the author line, now endorse and will utilize the Registered Reports if some sport and exercise science journals were to adopt such a format.

6 Acknowledgements

ARC and ADV devised and lead the writing of this manuscript. Co-authors participated in the brainstorming, drafting and editing, or support the initiatives included within the manuscript. Author order—except for ARC and ADV—was determined via randomization, as per majority vote. Finally, we would like to thank Dr. Matthew N. Cramer for his feedback in the preparation of this manuscript. All readers are encouraged to engage the available authors on social media if they would like to discuss the ideas and issues highlighted within the manuscript. The following authors can be reached on Twitter: Aaron R. Caldwell (@ExPhysStudent), Andrew D. Vigotsky (@avigotsky), Matthieu P. Boisgontier (@MattBoisgontier), Jason L. Neva (@jasonlneva), Richard S. Severin (@PTReviewer), David Nunan (@dnunan79), Zad R. Chow (@dailyzad), Andreas Kreutzer (@akreutzer82), Julia Schmidt (@juliayve), Jakob Škarabot (@JSkarabot), Greg Nuckols (@GregNuckols), Ian D. Boardley (@MDSportex), Sue Peters (@smpeters9), Bret Contreras (@bretcontreras), Jean-Benoit Morin (@jb_morin), Matthew S. Tenan (@TenanATC), Rosie Twomey (@Rosie_Twomey), Cody T. Haun (@drcodyhaun), Brooke M. Bouza (@BrookeBouza), Megan E. Rosa-Caldwell (@Meg_R2014), Mitchell Naughton (@Mitch_Naughton_), and David Mellor (@EvoMellor).

7 Conflicts of Interest

DTM is an employee of the Center for Open Science, a nonprofit organization whose mission includes advocating for increased transparency in scientific research, which includes the Registered Reports format. JPM is the Founder of SportRχiv.

References

[1] Open Science Collaboration. Estimating the reproducibility of psychological science. *Science*, 349(6251):aac4716, 8 2015. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.aac4716. URL `http://science.sciencemag.org/content/349/6251/aac4716`. PMID: 26315443.

[2] K Bollen, J.T. Cacioppo, R.M. Kaplan, J.A. Krosnick, and J.L. Olds. Social, behavioral, and economic sciences perspectives on robust and reliable science: Report of the subcommittee on replicability in science advisory committee to the national science foundation directorate for social, behavioral, and economic sciences. 2015. URL `https://www.nsf.gov/sbe/AC_Materials/SBE_Robust_and_Reliable_Research_Report.pdf`.

[3] Brian A Nosek and Timothy M Errington. Making sense of replications. *eLife*, 6:e23383, 1 2017. ISSN 2050-084X. doi: 10.7554/eLife.23383. URL `https://doi.org/10.7554/eLife.23383`. [Online; accessed 2018-10-22].

[4] Colin F. Camerer, Anna Dreber, Felix Holzmeister, Teck-Hua Ho, Jrgen Huber, Magnus Johannesson, Michael Kirchler, Gideon Nave, Brian A. Nosek, Thomas Pfeiffer, Adam Altmejd, Nick Buttrick, Taizan Chan, Yiling Chen, Eskil Forsell, Anup Gampa, Emma Heikensten, Lily Hummer, Taisuke Imai, Siri Isaksson, Dylan Manfredi, Julia Rose, Eric-Jan Wagenmakers, and Hang Wu. Evaluating the replicability of social science experiments in nature and science between 2010 and 2015. *Nature Human Behaviour*, page 1, 8 2018. ISSN 2397-3374. doi: 10.1038/s41562-018-0399-z. URL `https://www.nature.com/articles/s41562-018-0399-z`. [Online; accessed 2018-08-27].

[5] John P. A. Ioannidis. Why most published research findings are false. *PLoS medicine*, 2(8):e124, 8 2005. ISSN 1549-1676. doi: 10.1371/journal.pmed.0020124. PMID: 16060722 PMCID: PMC1182327.

[6] Michael A. Clemens. The meaning of failed replications: A review and proposal. *Journal of Economic Surveys*, 31(1):326–342, 2 2017. ISSN 1467-6419. doi: 10.1111/joes.12139. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/joes.12139`. [Online; accessed 2018-10-22].

[2] Registered Reports are only one step in a long process for improving sport and exercise science research; therefore, groups such as the Society for Transparency, Openness, and Reproducibility in Kinesiology (STORK, http://storkinesiology.org/) have formed to help address these issues. Additionally, for researchers who want their work to be freely available, but cannot publish open access, most journals allow for pre- and post-print archiving; for instance, in SportRχiv (osf.io/preprints/sportrxiv).

[7] Leslie K. John, George Loewenstein, and Drazen Prelec. Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological science*, 23(5):524532, 2012.

[8] Hannah Fraser, Tim Parker, Shinichi Nakagawa, Ashley Barnett, and Fiona Fidler. Questionable research practices in ecology and evolution. *PLOS ONE*, 13(7):e0200303, 7 2018. ISSN 1932-6203. doi: 10.1371/journal.pone.0200303. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0200303. [Online; accessed 2018-10-22].

[9] Klaus Fiedler and Norbert Schwarz. Questionable research practices revisited:. *Social Psychological and Personality Science*, 10 2015. doi: 10.1177/1948550615612150. URL http://journals.sagepub.com/doi/pdf/10.1177/1948550615612150. [Online; accessed 2018-10-22].

[10] Daniele Fanelli. Opinion: Is science really facing a reproducibility crisis, and do we need it to? *Proceedings of the National Academy of Sciences*, page 201708272, 3 2018. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1708272114. URL http://www.pnas.org/content/early/2018/03/08/1708272114. PMID: 29531051.

[11] Daniele Fanelli. How many scientists fabricate and falsify research? a systematic review and meta-analysis of survey data. *PLOS ONE*, 4(5):e5738, 5 2009. ISSN 1932-6203. doi: 10.1371/journal.pone.0005738. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0005738. [Online; accessed 2018-10-22].

[12] Howard Bauchner. Notice of retraction: Wansink b, cheney mm. super bowls: Serving bowl size and food consumption. jama. 2005;293(14):1727-1728. *JAMA*, 2018. ISSN 1538-3598. doi: 10.1001/jama.2018.14249. PMID: 30265737.

[13] Taylor L. Buchanan and Keith R. Lohse. Researchers' perceptions of statistical significance contribute to bias in health and exercise science. *Measurement in Physical Education and Exercise Science*, 20 (3):131–139, 7 2016. ISSN 1091-367X. doi: 10.1080/1091367X.2016.1166112. URL https://doi.org/10.1080/1091367X.2016.1166112. [Online; accessed 2018-11-28].

[14] Israel Halperin, Andrew D. Vigotsky, Carl Foster, and David B. Pyne. Strengthening the practice of exercise and sport-science research. *International Journal of Sports Physiology and Performance*, 13(2):127–134, 2 2018. ISSN 1555-0273. doi: 10.1123/ijspp.2017-0322. PMID: 28787228.

[15] P. D. White, M. C. Sharpe, T. Chalder, J. C. DeCesare, R. Walwyn, and Pace trial group. Protocol for the pace trial: a randomised controlled trial of adaptive pacing, cognitive behaviour therapy, and graded exercise, as supplements to standardised specialist medical care versus standardised specialist medical care alone for patients with the chronic fatigue syndrome/myalgic encephalomyelitis or encephalopathy. *BMC Neurol*, 7:6, 2007. ISSN 1471-2377 (Electronic) 1471-2377 (Linking). doi: 10.1186/1471-2377-7-6. URL https://www.ncbi.nlm.nih.gov/pubmed/17397525.

[16] Keith Lohse, Taylor Buchanan, and Matthew Miller. Underpowered and overworked: Problems with data analysis in motor learning studies. *Journal of Motor Learning and Development*, 4(1):37–58, 6 2016. ISSN 2325-3193. doi: 10.1123/jmld.2015-0010. URL https://journals.humankinetics.com/doi/abs/10.1123/jmld.2015-0010. [Online; accessed 2018-10-27].

[17] N. L. Kerr. Harking: hypothesizing after the results are known. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc*, 2(3):196–217, 1998. ISSN 1088-8683. doi: 10.1207/s15327957pspr0203_4. PMID: 15647155.

[18] B. Fischhoff. Hindsight not equal to foresight: the effect of outcome knowledge on judgment under uncertainty. 1975. *Quality & Safety in Health Care*, 12(4):304–311; discussion 311–312, 8 2003. ISSN 1475-3898. PMID: 12897366 PMCID: PMC1743746.

[19] Brian A. Nosek, Charles R. Ebersole, Alexander DeHaven, and David Mellor. The preregistration revolution. *Proceedings of the National Academy of Sciences of the United States of America*, 6 2017. doi: 10.1073/pnas.1708274114. URL http://www.pnas.org/content/early/2018/03/08/1708274114#ref-3. [Online; accessed 2018-06-19].

[20] Joseph P. Simmons, Leif D. Nelson, and Uri Simonsohn. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11):1359–1366, 11 2011. ISSN 1467-9280. doi: 10.1177/0956797611417632. PMID: 22006061.

[21] A. D. de Groot. The meaning of "significance" for different types of research [translated and annotated by eric-jan wagenmakers, denny borsboom, josine verhagen, rogier kievit, marjan bakker, angelique cramer, dora matzke, don mellenbergh, and han l. j. van der maas]. 1969. *Acta Psychologica*, 148:188–194, 5 2014. ISSN 1873-6297. doi: 10.1016/j.actpsy.2014.02.001. PMID: 24589374.

[22] Marcus R. Munafó, Brian A. Nosek, Dorothy V. M. Bishop, Katherine S. Button, Christopher D. Chambers, du Nathalie Percie Sert, Uri Simonsohn, Eric-Jan Wagenmakers, Jennifer J. Ware, and John P. A. Ioannidis. A manifesto for reproducible science. *Nature Human Behaviour*, 1(1):0021, 1 2017. ISSN 2397-3374. doi: 10.1038/s41562-016-0021. URL https://www.nature.com/articles/s41562-016-0021. [Online; accessed 2018-07-29].

[23] Vera E. Heininga, Albertine J. Oldehinkel, Ren Veenstra, and Esther Nederhof. I just ran a thousand analyses: benefits of multiple testing in understanding equivocal evidence on gene-environment interactions. *PloS One*, 10(5):e0125383, 2015. ISSN 1932-6203. doi: 10.1371/journal.pone.0125383. PMID: 26016887 PMCID: PMC4446037.

[24] Chirag J. Patel, Belinda Burford, and John P. A. Ioannidis. Assessment of vibration of effects due to model specification can demonstrate the instability of observational associations. *Journal of Clinical Epidemiology*, 68(9):1046–1058, 9 2015. ISSN 1878-5921. doi: 10.1016/j.jclinepi.2015.05.029. PMID: 26279400 PMCID: PMC4555355.

[25] Joshua Carp. On the plurality of (methodological) worlds: estimating the analytic flexibility of fmri experiments. *Frontiers in Neuroscience*, 6:149, 2012. ISSN 1662-453X. doi: 10.3389/fnins.2012.00149. PMID: 23087605 PMCID: PMC3468892.

[26] Andrew Gelman and Eric Loken. The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time. 11 2013. URL https://stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf.

[27] Sara Steegen, Francis Tuerlinckx, Andrew Gelman, and Wolf Vanpaemel. Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*, 11(5):702–712, 9 2016. ISSN 1745-6916, 1745-6924. doi: 10.1177/1745691616658637. URL http://journals.sagepub.com/doi/10.1177/1745691616658637. [Online; accessed 2018-10-09].

[28] R. Silberzahn, E. L. Uhlmann, D. P. Martin, P. Anselmi, F. Aust, E. Awtrey, . Bahnk, F. Bai, C. Bannard, E. Bonnier, R. Carlsson, F. Cheung, G. Christensen, R. Clay, M. A. Craig, A. Dalla Rosa, L. Dam, M. H. Evans, I. Flores Cervantes, N. Fong, M. Gamez-Djokic, A. Glenz, S. Gordon-McKeon, T. J. Heaton, K. Hederos, M. Heene, A. J. Hofelich Mohr, F. Hgden, K. Hui, M. Johannesson, J. Kalodimos, E. Kaszubowski, D. M. Kennedy, R. Lei, T. A. Lindsay, S. Liverani, C. R. Madan, D. Molden, E. Molleman, R. D. Morey, L. B. Mulder, B. R. Nijstad, N. G. Pope, B. Pope, J. M. Prenoveau, F. Rink, E. Robusto, H. Roderique, A. Sandberg, E. Schlter, F. D. Schnbrodt, M. F. Sherman, S. A. Sommer, K. Sotak, S. Spain, C. Sprlein, T. Stafford, L. Stefanutti, S. Tauber, J. Ullrich, M. Vianello, E.-J. Wagenmakers, M. Witkowiak, S. Yoon, and B. A. Nosek. Many analysts, one data set: Making transparent how variations in analytic choices

affect results. *Advances in Methods and Practices in Psychological Science*, page 2515245917747646, 8 2018. ISSN 2515-2459. doi: 10.1177/2515245917747646. URL https://doi.org/10.1177/2515245917747646. [Online; accessed 2018-08-27].

[29] Tim van der Zee, Jordan Anaya, and Nicholas J. L. Brown. Statistical heartburn: an attempt to digest four pizza publications from the cornell food and brand lab. *BMC Nutrition*, 3(1):54, 12 2017. ISSN 2055-0928. doi: 10.1186/s40795-017-0167-x. URL https://link.springer.com/article/10.1186/s40795-017-0167-x. [Online; accessed 2018-08-03].

[30] Megan L. Head, Luke Holman, Rob Lanfear, Andrew T. Kahn, and Michael D. Jennions. The extent and consequences of p-hacking in science. *PLoS biology*, 13(3):e1002106, 3 2015. ISSN 1545-7885. doi: 10.1371/journal.pbio.1002106. PMID: 25768323 PMCID: PMC4359000.

[31] Uri Simonsohn, Leif D. Nelson, and Joseph P. Simmons. p-curve and effect size: Correcting for publication bias using only significant results. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, 9(6):666–681, 11 2014. ISSN 1745-6924. doi: 10.1177/1745691614553988. PMID: 26186117.

[32] Dorothy V. M. Bishop and Paul A. Thompson. Problems in using p-curve analysis and text-mining to detect rate of p-hacking and evidential value. *PeerJ*, 4:e1715, 2016. ISSN 2167-8359. doi: 10.7717/peerj.1715. PMID: 26925335 PMCID: PMC4768688.

[33] Adam N. Sanborn and Thomas T. Hills. The frequentist implications of optional stopping on bayesian hypothesis tests. *Psychonomic Bulletin & Review*, 21(2):283–300, 4 2014. ISSN 1531-5320. doi: 10.3758/s13423-013-0518-9. PMID: 24101570.

[34] Daniël Lakens. Performing high-powered studies efficiently with sequential analyses. *European Journal of Social Psychology*, 44 (7):701–710, 12 2014. ISSN 1099-0992. doi: 10.1002/ejsp.2023. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/ejsp.2023. [Online; accessed 2018-09-13].

[35] Annie Franco, Neil Malhotra, and Gabor Simonovits. Social science. publication bias in the social sciences: unlocking the file drawer. *Science (New York, N.Y.)*, 345(6203):1502–1505, 9 2014. ISSN 1095-9203. doi: 10.1126/science.1255484. PMID: 25170047.

[36] John P. A. Ioannidis. Why most discovered true associations are inflated. *Epidemiology (Cambridge, Mass.)*, 19(5):640–648, 9 2008. ISSN 1531-5487. doi: 10.1097/EDE.0b013e31818131e7. PMID: 18633328.

[37] Katherine S. Button, John P. A. Ioannidis, Claire Mokrysz, Brian A. Nosek, Jonathan Flint, Emma S. J. Robinson, and Marcus R. Munafò. Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14 (5):365–376, 5 2013. ISSN 1471-0048. doi: 10.1038/nrn3475. URL https://www.nature.com/articles/nrn3475. [Online; accessed 2018-07-29].

[38] C. R. Kothari. *Research Methodology: Methods and Techniques*. New Age International, 2004. ISBN 978-81-224-1522-3. Google-Books-ID: hZ9wSHysQDYC.

[39] Robert D. McIntosh. Exploratory reports: A new article type for cortex. *Cortex*, 96:A1–A4, 11 2017. ISSN 0010-9452. doi: 10.1016/j.cortex.2017.07.014. URL http://www.sciencedirect.com/science/article/pii/S0010945217302393. [Online; accessed 2018-07-29].

[40] Daniël Lakens and Ellen RK Evers. Sailing from the seas of chaos into the corridor of stability: Practical recommendations to increase the informational value of studies. *Perspectives on Psychological Science*, 9(3):278–292, 2014.

[41] Kai J. Jonas, Joseph Cesario, Madeliene Alger, April H. Bailey, Dario Bombari, Dana Carney, John F. Dovidio, Sean Duffy, Jenna A. Harder, van Dian Huistee, Benita Jackson, David J. Johnson, Victor N. Keller, Lukas Klaschinski, Onawa LaBelle, Marianne LaFrance, Ioana M. Latu, Margot Morssinkhoff, Kelly Nault, Vaani Pardal, Caroline Pulfrey, Nicolas Rohleder, Richard Ronay, Laura Smart Richman, Marianne Schmid Mast, Konrad Schnabel, Michaela Schrder-Ab, and Josh M. Tybur. Power poses – where do we stand? *Comprehensive Results in Social Psychology*, 2(1):139–141, 1 2017. ISSN 2374-3603. doi: 10.1080/23743603.2017.1342447. URL https://doi.org/10.1080/23743603.2017.1342447. [Online; accessed 2018-10-22].

[42] Rémi Radel, Gavin Tempest, Gauthier Denis, Pierre Besson, and Raphael Zory. Extending the limits of force endurance: Stimulation of the motor or the frontal cortex? *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 97:96–108, 2017. ISSN 1973-8102. doi: 10.1016/j.cortex.2017.09.026. PMID: 29101820.

[43] M. S. Hagger, N. L. D. Chatzisarantis, H. Alberts, C. O. Anggono, C. Batailler, A. R. Birt, R. Brand, M. J. Brandt, G. Brewer, S. Bruyneel, D. P. Calvillo, W. K. Campbell, P. R. Cannon, M. Carlucci, N. P. Carruth, T. Cheung, A. Crowell, D. T. D. De Ridder, S. Dewitte, M. Elson, J. R. Evans, B. A. Fay, B. M. Fennis, A. Finley, Z. Francis, E. Heise, H. Hoemann, M. Inzlicht, S. L. Koole, L. Koppel, F. Kroese, F. Lange, K. Lau, B. P. Lynch, C. Martijn, H. Merckelbach, N. V. Mills, A. Michirev, A. Miyake, A. E. Mosser, M. Muise, D. Muller, M. Muzi, D. Nalis, R. Nurwanti, H. Otgaar, M. C. Philipp, P. Primoceri, K. Rentzsch, L. Ringos, C. Schlinkert, B. J. Schmeichel, S. F. Schoch, M. Schrama, A. Schütz, A. Stamos, G. Tingög, J. Ullrich, M. vanDellen, S. Wimbarti, W. Wolff, C. Yusainy, O. Zerhouni, and M. Zwienenberg. A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science*, 11(4):546–573, 7 2016. ISSN 1745-6916. doi: 10.1177/1745691616652873. URL https://doi.org/10.1177/1745691616652873. [Online; accessed 2018-10-22].

[44] Christopher P G Allen and David Marc Anton Mehler. Open science challenges, benefits and tips in early career and beyond. *PsyArXiv*. doi: 10.31234/osf.io/3czyt. URL https://osf.io/3czyt. [Online; accessed 2018-11-27].

[45] Peter C. Austin, Muhammad M. Mamdani, David N. Juurlink, and Janet E. Hux. Testing multiple statistical hypotheses resulted in spurious associations: a study of astrological signs and health. *Journal of Clinical Epidemiology*, 59(9):964–969, 9 2006. ISSN 0895-4356. doi: 10.1016/j.jclinepi.2006.01.012. PMID: 16895820.

[46] Coosje Lisabet Sterre Veldkamp, Marjan Bakker, Marcel A. L. M. van Assen, Elise Anne Victoire Crompvoets, How Hwee Ong, Brian A. Nosek, Courtney K. Soderberg, David Thomas Mellor, and Jelte M. Wicherts. Ensuring the quality and specificity of preregistrations. *PsyArXiv*, 2018. doi: 10.31234/osf.io/cdgyh. URL https://osf.io/cdgyh. [Online; accessed 2018-11-28].

[47] Anthony Bastardi, Eric Luis Uhlmann, and Lee Ross. Wishful thinking: belief, desire, and the motivated evaluation of scientific evidence. *Psychological Science*, 22(6):731–732, 6 2011. ISSN 1467-9280. doi: 10.1177/0956797611406447. PMID: 21515736.

[48] Jaskarndip Chahal, S. Sebastian Tomescu, Bheeshma Ravi, Bernard R. Bach, Darrell Ogilvie-Harris, Nizar N. Mohamed, and Rajiv Gandhi. Publication of sports medicine-related randomized controlled trials registered in clinicaltrials.gov. *The American Journal of Sports Medicine*, 40(9):1970–1977, 9 2012. ISSN 1552-3365. doi: 10.1177/0363546512448363. PMID: 22679295.

[49] Holly N. Smith, Mohit Bhandari, Nizar N. Mahomed, Meryam Jan, and Rajiv Gandhi. Comparison of arthroplasty trial publications after registration in clinicaltrials.gov. *The Journal of Arthroplasty*, 27(7):1283–1288, 8 2012. ISSN 1532-8406. doi: 10.1016/j.arth.2011.11.005. PMID: 22226609.

[50] Padhraig S. Fleming, Despina Koletsi, Kerry Dwan, and Nikolaos Pandis. Outcome discrepancies and selective reporting: Impacting the leading journals? *PLOS ONE*, 10(5):e0127495,

5 2015. ISSN 1932-6203. doi: 10.1371/journal.pone.0127495. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0127495. [Online; accessed 2018-10-22].

[51] Tom E. Hardwicke and John P. A. Ioannidis. Mapping the universe of registered reports. *Nature Human Behaviour*, page 1, 10 2018. ISSN 2397-3374. doi: 10.1038/s41562-018-0444-y. URL https://www.nature.com/articles/s41562-018-0444-y. [Online; accessed 2018-10-22].

[52] Brian A. Nosek and Danil Lakens. Registered reports. *Social Psychology*, 45(3):137–141, 1 2014. ISSN 1864-9335. doi: 10.1027/1864-9335/a000192. URL https://econtent.hogrefe.com/doi/full/10.1027/1864-9335/a000192. [Online; accessed 2018-07-29].

[53] Jonathan Kimmelman, Jeffrey S. Mogil, and Ulrich Dirnagl. Distinguishing between exploratory and confirmatory preclinical research will improve translation. *PLoS biology*, 12(5):e1001863, 5 2014. ISSN 1545-7885. doi: 10.1371/journal.pbio.1001863. PMID: 24844265 PMCID: PMC4028181.

[54] Christopher D. Chambers. Registered reports: a new publishing initiative at cortex. *Cortex*, 49(3):609–610, 3 2013. ISSN 1973-8102. doi: 10.1016/j.cortex.2012.12.016. PMID: 23347556.

[55] John Wood, Nick Freemantle, Michael King, and Irwin Nazareth. Trap of trends to statistical significance: likelihood of near significant p value becoming more significant with extra data. *Bmj*, 348: g2215, 2014.

[56] Sander Greenland, Stephen J Senn, Kenneth J Rothman, John B Carlin, Charles Poole, Steven N Goodman, and Douglas G Altman. Statistical tests, p values, confidence intervals, and power: a guide to misinterpretations. *European journal of epidemiology*, 31 (4):337–350, 2016.