Communications in Kinesiology

Registered Report

Please cite as:

Name et al. (2021). The Nature of Our Literature: An Observational Study of the Positive Result Rate and Reporting Practices in Kinesiology. *Communications in Kinesiology*. doi: 10.51224/XXXXXXXXX

Subject Areas:

metascience

Keywords:

key1, key2, key3

Author for correspondence:

S. Name

e-mail:

s.name@fake.edu

STORK
SOCIETY FOR TRANSPARENCY OPENNESS AND REPLICATION IN KINESIOLOGY

The Nature of Our Literature: An Observational Study of the Positive Result Rate and Reporting Practices in Kinesiology

Some Name[1], Same Name[1], Another Nombre[2]

[1]Fake University

[2]Fake Place

Scientists rely upon an accurate scientific literature in order to build and test new theories about the natural world. In the past decade, observational studies of the scientific literature have indicated that numerous questionable research practices and poor reporting practices may be hindering scientific progress. In particular, 3 recent have indicated a implausibly high rate of studies with positive (i.e., hypothesis confirming) results. In Sports Medicine, a field closely related to Kinesiology, studies that tested a hypothesis indicated support for their primary hypothesis ~70% of the time. However, a study of journals that cover the entire field of Kinesiology has yet to be completed, and the quality of other reporting practices, such as clinical trial registration, has not been evaluated. Therefore, in this study we will retrospectively evaluate 300 original research articles from the flagship journals of America (Medicine and Science in Sport and Exercise), Europe (European Journal of Sport Science), and Australia (Journal of Science and Medicine in Sport).

1. Introduction

Scientists and knowledge-users who make decisions based on scientific evidence rely on the published literature to be reported transparently and to be an accurate representation of the research that scientists conduct. The ability to replicate scientific findings is vital to establish the credibility of scientific claims and to allow research to progress (Nosek & Errington, 2019). A large-scale collaborative effort estimated the replicability of findings in psychological science and found that most replication effects are smaller than originally reported Collaboration (2015). Whilst this is a complex issue, questionable research practices (QRPs) and publication bias explain at least some of the differences between the original and replication effect sizes (Head et al., 2015; John et al., 2012; Simmons et al., 2011). Alongside psychology Collaboration (2015), other fields have struggled to replicate or reproduce findings, including neuroscience (Boekel et al., 2015; Kharabian Masouleh et al., 2019; Turner et al., 2018), cancer biology (Nosek & Errington, 2017), human genetics ("Replicating Genotype–Phenotype Associations," 2007) and pharmacology (Prinz et al., 2011). This type of systematic replication and evaluation of previously published results has not yet been attempted in kinesiology (alternatively known as sport and exercise science). However, considering the similarities (e.g,. the study of human behaviour) and overlap (e.g. sport and exercise psychology) between psychology and kinesiology, we have reason to believe it may suffer from the same QRPs. Replication appears to be rare in kinesiology, which is perhaps surprising considering that kinesiology has been the focus of significant critique due to overly optimistic inferences (Sainani et al., 2019) and a history of underpowered studies (Abt et al., 2020). Furthermore, a lack of sample size estimation (Abt et al., 2020), misuse of p-values and statistical significance testing, limited collaboration with statisticians (Sainani et al., 2020), minimal or arbitrary use of effect sizes (Caldwell & Vigotsky, 2020), and other reporting issues (D. N. Borg, Lohse, et al., 2020) appear to be commonplace.

In the past few years, a community of researchers in kinesiology have been advocating for and adopting open and replicable research practices (D. N. Borg, Bon, et al., 2020; D. N. Borg, Lohse, et al., 2020; A. R. Caldwell et al., 2020; Sainani et al., 2020, p. @Caldwell_Vigotsky_2020; Vigotsky et al., 2020). Some journals in the field have started to adopt the Registered Report format for manuscripts which is commendable (see \url{www.cos.io/rr for a list of participating journals). However, such practices include openly sharing data and code, pre-registration, and using the registered reports format (for a primer, see A. R. Caldwell et al. (2020) for details). Some of the issues that motivated the open science movement in psychology and other fields Munafò et al. (2017) are comparatively unexplored in kinesiology, and in addition currently, the number of kinesiology researchers adopting open research practices is largely unknown. There is some indication that both pre-registration and sharing of data is uncommon (D. N. Borg, Lohse, et al., 2020; Tamminen & Poucher, 2018) and flagship journals of our field (e.g., Medicine & Science in Sport & Exercise, European Journal of Sport Science) do not include a statement encouraging data sharing in the author guidelines (Oct 2020). Evaluating a recent sample of the kinesiology literature for such practices may help draw attention to these potential issues.

Another issue that warrants consideration is the positive result rate (the rate at which a published study finds support for its hypothesis) of published kinesiology studies. Recently, Büttner et al. (2020) estimated the positive result rate in three high ranking sports journals and one high ranking sports physiotherapy journal. In line with previous research in other scientific disciplines (Fanelli, 2010; Scheel et al., 2020), the positive result rate exceeded 80%. What are the mechanisms for the suspiciously high positive result rates in the scientific literature? Given the assumption of a completely unbiased literature, such a high positive result rate could only occur if both statistical power and the proportion of true hypotheses that researchers have chosen to test is consistently high Scheel et al. (2020). The more plausible explanation perhaps, corroborated in previous work (John et al., 2012; Simmons et al., 2011), is that the literature is distorted by undisclosed flexibility in analysis and other QRPs, and the incentive to publish positive results. Registered reports are specifically designed to help mitigate these issues Chambers et al. (2015). Therefore, Scheel et al. (2020) assessed the positive result rate in research articles published using the traditional format in comparison to registered reports in a sample of the psychology literature. The positive result rate was an implausibly high 96% for traditional articles and a significantly lower 46% for registered reports. The increased methodological rigour inherent to the registered report format has clearly led to an increase in the reporting of null findings in the psychological literature.

The equivalent findings regarding standard and registered reports have not been reported for kinesiology, and simply would not be possible given the current literature; unlike psychological science Scheel et al. (2020), and to our knowledge, kinesiology has not accumulated more than 70 RRs to evaluate against traditional publication formats. Nevertheless, the adoption of registered reports in kinesiology is progressing slowly. One reason for this may be a lack of awareness regarding the replication crisis and movement towards more rigorous and transparent research practices. However, the slow adoption of registered reports could also be due to a lack of concern about the kinesiology literature given the limited evidence that a problem exists. Therefore, the primary aim of this study is to assess the positive result rate of reported hypotheses in the recent kinesiology literature, using society-affiliated flagship journals from the field. Considering the majority of scientific disciplines documented by Fanelli (2009) had a positive rate of 80%, we hypothesize that the $>$80% of the published studies in kinesiology will report positive results (i.e, support for the hypothesis) for their first stated hypothesis. Our secondary aims are to assess a number of related research practices, including whether the kinesiology literature includes replications of previous effects, the detail of statistical reporting and adoption of other transparent reporting practices.

2. Methods

(a) Sample

Research articles were sampled from three flagship kinesiology journals: Medicine and Science in Sport and Exercise (MSSE), the European Journal of Sport Science (EJSS) and the Journal of Science and Medicine in Sport (JSAMS), which represent three major kinesiology societies of North America (American College of Sports Medicine), Europe (European College of Sport Science) and Australia (Sports Medicine Australia), respectively. We selected three major societies and their official flagship journals because we believed they represent a diverse selection of research in kinesiology and provide insights into the practices of the field as a whole. In addition, we chose to focus on these three journals rather than a random sample of the entire literature because these journals should represent the best research in the field (compared to any published article which could be sampled from a possible predatory publisher). We selected 100 original research articles per journal, 300 in total, excluding study protocols, methodological tutorials/reports, opinions, commentaries, perspectives, conference proceedings, narrative reviews, systematic reviews and meta-analyses. We also excluded research articles if they have been retracted or contain insufficient information to reach coding decisions. To sample a recent selection of the literature, research articles were sampled consecutively backwards from December 31, 2019, by the data analyst (ARC).

(b) Data Extraction

We dentified nine coders who were responsible for data extraction. Coders underwent standardized training that has been designed based on the queries raised and clarification required during pilot testing (see later section). These nine coders will form three teams of three, and each team will be randomly allocated the research articles from one journal (MSSE, EJSS, or JSAMS). All coders extracted data independently and entered this directly into a Qualtrics survey. The Qualtrics survey was refined after pilot testing and a copy is available at https://osf.io/nwcx6/?view_only=a41116388e9244b7821bfb9fe5496bd2 . Each team was coordinated by a team leader trained at a doctoral level in a kinesiology discipline (RT, VY and JW). Once independent coding was complete, interrater reliability was assessed using Fleiss's Kappa. Team leaders were responsible for resolving all conflicts (any instance where there was not agreement between all group members) within their team through group review of the item and group discussion. Where conflicts could not resolved (and revised if necessary) using this process, the team leader consulted the other two team leaders. All data (original coder responses and summary decisions) is available on study's Open Science Framework repository.

(c) Measures and Coding Procedure

All articles were categorized as basic physiology (animal and cell physiology), applied exercise physiology (human), environmental physiology (heat, cold, and altitude), clinical research, biomechanics, motor learning/control/behaviour, epidemiology, sport/exercise psychology, sport performance, or other (the category that best describes the article). Research articles thtat did include explicit statements that a hypothesis was tested were not included in the analysis of the positive result rate. However, all articles (i.e., 300) were included in analysis related to replication status, statistical reporting and other reporting practices, as described in the following sections.

(d) Support for a Hypothesis in the Kinesiology Literature

From the 300 articles, we expected that approximately 60% would include explicit statements that a hypothesis was tested as part of the study (e.g., "We hypothesized that…") (Büttner et al., 2020). Therefore, we expected to extract data on the positive results rate from approximately 180 research articles. The main dependent variable was whether the *first* stated hypothesis was supported or not, as reported by the authors. We planned to closely follow the coding procedure used by Fanelli (2010) and Scheel et al. (2020), which is described as follows: By examining the abstract and/or full text, it will be determined whether the authors of each paper had concluded to have found a positive (full or partial) or negative (null or negative) support. If more than one hypothesis was being tested, only the first one to appear in the text was considered. The coding of support for the hypothesis was based on the author's description of their results. In line with previous work (Büttner et al., 2020; Scheel et al., 2020), we coded a hypothesis as having received "support," "partial support," "no support" or "unclear or not stated". We added this fourth option after pilot indicated that some authors failed to state whether or not the study's hypotheses were, or were not, supported in the discussion section of the manuscript. This was re-coded into a binary "support" (full or partial) vs. "no support" variable, with "unclear or not stated" removed, for the main analysis. The language used to state a hypothesis and support for the first tested hypothesis were included in the data extraction and are included in the final dataset.

(e) Replication Status

Coders assessed whether a study is a replication of a previously published one, as reported by the authors. Coders searched the full texts of all papers for the string 'replic*' and, for papers that contained it, indicated whether the coded hypothesis was a close replication with the goal to verify a previously published result (Scheel et al., 2020).

(f) Statistical Reporting

Coders assessed whether authors included language related to statistical significance and if p-values were included in the results (relating to all analyses and not only the first hypothesis). If yes, coders assessed if the p-value was interpreted as significant and if the exact or relative p-value was reported (i.e., $p = 0.049$ vs. $p < 0.05$). If a relative p-value was reported, the level of the reported p-value (e.g., $p < 0.05$, $p < 0.01$) were coded. Coders also extracted whether an effect size was reported, including, but not limited to: Cohen's d, correlation coefficients, mean differences, and measures of model fit (e.g., coefficient of determination: R^2). Finally, coders assessed whether the information on sample size was provided, and if provided, the total sample size (the number of participants included in the analyses, rather than the planned sample size) will be extracted. Finally, coders assessed whether any sample size justification (e.g. power analysis) were included in the manuscript.

(g) Other Reporting Practices

Coders assessed whether the study was a clinical trial, according to the ICJME definition (https://hub.ucsf.edu/clinicaltrialsgov-definition-clinical-trial). If yes, coders assessed if a clinical trial registration was reported in the manuscript. For all other types of studies, coders assessed whether studies were pre-registered (as reported within the manuscript). Coders assessed if a manuscript provides a statement on data availability, and if yes, whether there was open access to the original data and/or code via a link or supplementary file.

(h) Pilot Testing

To ensure that our questionnaire for our raters accurately and consistently reflects the above-detailed information from relevant articles, we conducted pilot testing before submission of the Stage 1 manuscript. Fifteen original research articles published in 2018, five from each of our three chosen journals, were selected to be used for pilot testing. One team of naive coders (i.e., were not trained prior to coding) extracted all data from these articles and entered this into Qualtrics. Independent coding was checked for disagreements, and this was used to inform training procedures. Pilot aggregated data were generated, and further adjustments were made to refine the planned extraction and analysis process. A summary report of the pilot work can be found at https://osf.io/nwcx6/?view_only=

a41116388e9244b7821bfb9fe5496bd2. Overall, our pilot work indicated minimally acceptable agreement among the raters on the questions essential to our study such as whether a hypothesis was tested (κ= 0.903; complete agreement = 14/15) and if the authors found support for this hypothesis (κ= 0.586; complete agreement = 6/9). For all items with rater disagreement, at least two coders were in agreement on the rating. After the conclusion of pilot testing, a forum among the team was completed in order to appropriately adjust the questionnaire and refine future instructions/training for the coding teams in the full study. Prior to coding, all coding team members underwent formal training and were presented with example articles (not from the study sample) in order to improve consistency in the coding process.

(i) Analysis Plan

First, we estimated the rate at which kinesiology research finds support for the first tested hypothesis (as reported by the authors). Further, we planned to compare this to the majority of disciplines surveyed in Fanelli (2010) which reported a positive result rate in excess of 80% (16 of 20 disciplines). We believed it unlikely that kinesiology would have a positive result rate lower than 80%, and believe that the actual rate is closer to the social sciences at approximately 85% (Fanelli, 2010). Considering we had a good prior information, and a belief we wanted to test, we opted to use a Bayesian analysis to test our hypothesis. Therefore, we planned to test our hypothesis that the positive result rate is greater than 80% using a generalized Bayesian regression model (Bürkner, 2017). We assumed a prior of $\beta(17, 3)$ on the intercept of the model (i.e., the rate of positive results). Evidence for our hypothesis is reported as the posterior probability, $pr(Intercept > .8|data)$, of our hypothesis and the Bayes Factor (BF), the ratio of evidence for our hypothesis versus the null. We performed a Monte Carlo simulation assuming we obtained 150 studies which contained hypotheses from a population where 85% will contain a positive result for the first stated hypothesis. This simulation indicated that our model would have an 87% chance of being able to obtain some evidence (BF in favor of our hypothesis > 3) for our hypothesis. All other data is summarized descriptively and as frequencies and proportions with chi-squared and binomial proportions tests where appropriate. A detailed summary of the planned hypothesis test and "power" analysis can be found at https://osf.io/nwcx6/?view_only=a41116388e9244b7821bfb9fe5496bd2.

3. Results

(a) Confirmatory Results

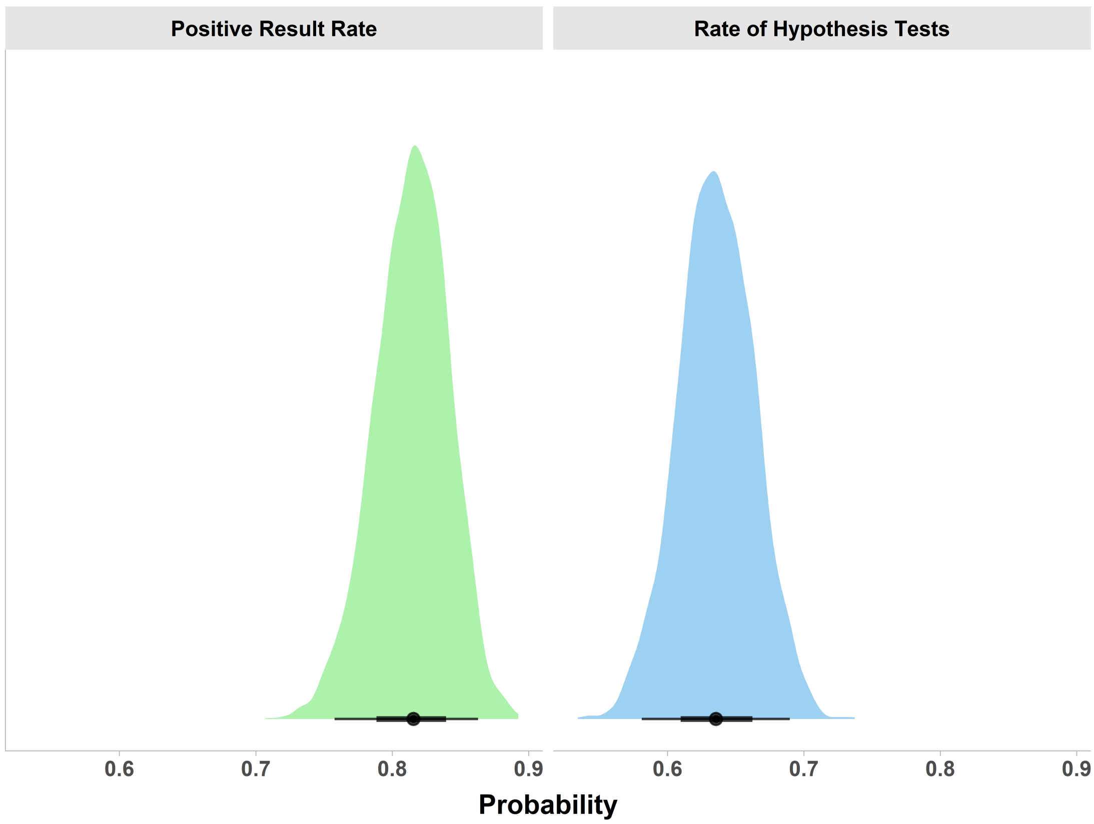

Figure 1. Posterior distributions from Bayesian model with the 50% and 95% compatibility intervals represented by the error bars at the bottom.

(b) Exploratory Results

(i) Statistics Reporting

(ii) Other Important Reporting Practices

(iii) Breakdown by Journal

(iv) Breakdown by Discipline

(v) Analysis of RCT and Clinical Trials

4. Discussion

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Aenean commodo ligula eget dolor. Aenean massa. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Donec quam felis, ultricies nec, pellentesque eu, pretium quis, sem. Nulla consequat massa quis enim. Donec pede justo, fringilla vel, aliquet nec, vulputate eget, arcu. In enim justo, rhoncus ut, imperdiet a, venenatis vitae, justo. Nullam dictum felis eu pede mollis pretium. Integer tincidunt. Cras dapibus. Vivamus elementum semper nisi. Aenean vulputate eleifend tellus.

Aenean leo ligula, porttitor eu, consequat vitae, eleifend ac, enim. Aliquam lorem ante, dapibus in, viverra quis, feugiat a, tellus. Phasellus viverra nulla ut metus varius laoreet. Quisque rutrum. Aenean imperdiet. Etiam ultricies nisi vel augue. Curabitur ullamcorper ultricies nisi. Nam eget dui. Etiam rhoncus. Maecenas tempus, tellus eget condimentum rhoncus, sem quam semper libero, sit amet adipiscing sem neque sed ipsum. Nam quam nunc, blandit vel, luctus pulvinar, hendrerit id, lorem. Maecenas nec odio et ante tincidunt tempus.

Donec vitae sapien ut libero venenatis faucibus. Nullam quis ante. Etiam sit amet orci eget eros faucibus tincidunt. Duis leo. Sed fringilla mauris sit amet nibh. Donec sodales sagittis magna. Sed consequat, leo eget bibendum sodales, augue velit cursus nunc, quis gravida magna mi a libero. Fusce vulputate eleifend sapien. Vestibulum purus quam, scelerisque ut, mollis sed, nonummy id, metus. Nullam accumsan lorem in dui. Cras ultricies mi eu turpis hendrerit fringilla. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; In ac dui quis mi consectetuer lacinia.

Nam pretium turpis et arcu. Duis arcu tortor, suscipit eget, imperdiet nec, imperdiet iaculis, ipsum. Sed aliquam ultrices mauris. Integer ante arcu, accumsan a, consectetuer eget, posuere ut, mauris. Praesent adipiscing. Phasellus ullamcorper ipsum rutrum nunc. Nunc nonummy metus. Vestibulum volutpat pretium libero. Cras id dui. Aenean ut eros et nisl sagittis vestibulum. Nullam nulla eros, ultricies sit amet, nonummy id, imperdiet feugiat, pede. Sed lectus. Donec mollis hendrerit risus. Phasellus nec sem in justo pellentesque facilisis. Etiam imperdiet imperdiet orci. Nunc nec neque. Phasellus leo dolor, tempus non, auctor et, hendrerit quis, nisi. Curabitur ligula sapien, tincidunt non, euismod vitae, posuere imperdiet, leo. Maecenas malesuada. Praesent congue erat at massa. Sed cursus turpis vitae tortor. Donec posuere vulputate arcu. Phasellus accumsan cursus velit. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Sed aliquam, nisi quis porttitor congue, elit erat euismod orci, ac

(a) Conclusion

orem ipsum dolor sit amet, consectetuer adipiscing elit. Aenean commodo ligula eget dolor. Aenean massa. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Donec quam felis, ultricies nec, pellentesque eu, pretium quis, sem. Nulla consequat massa quis enim. Donec pede justo, fringilla vel, aliquet nec, vulputate eget, arcu. In enim justo, rhoncus ut, imperdiet a, venenatis vitae, justo. Nullam dictum felis eu pede mollis pretium. Integer tincidunt. Cras dapibus. Vivamus elementum semper nisi. Aenean vulputate eleifend tellus. Aenean leo ligula, porttitor eu, consequat vitae, eleifend ac, enim. Aliquam lorem ante, dapibus in, viverra quis, feugiat a.

5. Additional Information

(a) Data Accessibility

The authors agree to share the raw data, digital study materials and analysis code. All study materials can be found on our OSF repository https://osf.io/nwcx6/?view_only=a41116388e9244b7821bfb9fe5496bd2.

(b) Author Contributions

- Contributed to conception and design: TBA
- Contributed to acquisition of data: TBA
- Contributed to analysis and interpretation of data: TBA
- Drafted and/or revised the article: TBA
- Approved the submitted version for publication: TBA

(c) Conflict of Interest

Authors have no conflicts of interest to declare.

(d) Funding

The proposed research is not funded. All necessary support is provided by the authors institutions. This study is an analysis of published research and does not require ethical approval.

(e) Acknowledgments

We thank everybody.

(f) Preregistration

Following Stage 1 in-principle acceptance, the authors agreed to pre-registration of the approved protocol on the Open Science Framework. The IPA registration can be found here: https://osf.io/3pqr7

(g) Conflicts of Interest

ARC, RT, and VRY currently serve as executive committee members for the Society of Transparency, Openness, and Replication in Kinesiology (STORK). VRY is a section editor and ARC is on the Steering Board for Registered Reports in Kinesiology. Neither will be involved in any aspect of handling this manuscript except as authors. The opinions or assertions contained herein are the private views of the author(s) and are not to be construed as official or reflecting the views of the Army or the Department of Defense. Any citations of commercial organizations and trade names in this report do not constitute an official Department of the Army endorsement of approval of the products or services of these organizations. No authors have any conflicts of interest to disclose. Approved for public release; distribution is unlimited.

6. References

List of Names

Abt, G., Boreham, C., Davison, G., Jackson, R., Nevill, A., Wallace, E., & Williams, M. (2020). Power, precision, and sample size estimation in sport and exercise science research. *Journal of Sports Sciences*, *38*(17), 1933–1935. https://doi.org/10.1080/02640414.2020.1776002

Boekel, W., Wagenmakers, E.-J., Belay, L., Verhagen, J., Brown, S., & Forstmann, B. U. (2015). A purely confirmatory replication study of structural brain-behavior correlations. *Cortex*, *66*, 115–133. https://doi.org/10.1016/j.cortex.2014.11.019

Borg, D. N., Bon, J. J., Sainani, K. L., Baguley, B. J., Tierney, N. J., & Drovandi, C. (2020). Comment on: "Moving sport and exercise science forward: A call for the adoption of more transparent research practices". *Sports Medicine*, *50*(8), 1551–1553. https://doi.org/10.1007/s40279-020-01298-5

Borg, D. N., Lohse, K. R., & Sainani, K. L. (2020). Ten common statistical errors from all phases of research, and their fixes. *PM&R*, *12*(6), 610–614. https://doi.org/10.1002/pmrj.12395

Bürkner, P.-C. (2017). Brms: An r package for bayesian multilevel models using stan. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01

Büttner, F., Toomey, E., McClean, S., Roe, M., & Delahunt, E. (2020). Are questionable research practices facilitating new discoveries in sport and exercise medicine? The proportion of supported hypotheses is implausibly high. *British Journal of Sports Medicine*. https://doi.org/10.1136/bjsports-2019-101863

Caldwell, A. R., & Vigotsky, A. D. (2020). *Does one effect size fit all? The case against default effect sizes for sport and exercise science*. SportRxiv. https://doi.org/10.31236/osf.io/tfx95

Caldwell, A. R., Vigotsky, A. D., Tenan, M. S., Radel, R., Mellor, D. T., Kreutzer, A., Lahart, I. M., Mills, J. P., Boisgontier, M. P., Boardley, I., Bouza, B., Cheval, B., Chow, Z. R., Contreras, B., Dieter, B., Halperin, I., Haun, C., Knudson, D., Lahti, J., … Consortium for Transparency in Exercise Science (COTES) Collaborators. (2020). Moving sport and exercise science forward: A call for the adoption of more transparent research practices. *Sports Medicine*, *50*(3), 449–459. https://doi.org/10.1007/s40279-019-01227-1

Chambers, C. D., Dienes, Z., McIntosh, R. D., Rotshtein, P., & Willmes, K. (2015). Registered reports: Realigning incentives in scientific publishing. *Cortex*, *66*, A1–2. https://doi.org/10.1016/j.cortex.2015.03.022

Collaboration, O. S. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251). https://doi.org/10.1126/science.aac4716

Fanelli, D. (2009). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PloS One*, *4*(5), e5738. https://doi.org/10.1371/journal.pone.0005738

Fanelli, D. (2010). "Positive" results increase down the hierarchy of the sciences. *PLOS ONE*, *5*(4), e10068. https://doi.org/10.1371/journal.pone.0010068

Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of p-hacking in science. *PLOS Biology*, *13*(3), e1002106. https://doi.org/10.1371/journal.pbio.1002106

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, *23*(5), 524–532. https://doi.org/10.1177/0956797611430953

Kharabian Masouleh, S., Eickhoff, S. B., Hoffstaedter, F., & Genon, S. (2019). Empirical examination of the replicability of associations between brain structure and psychological variables. *eLife*, *8*. https://doi.org/10.7554/elife.43464

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 1–9. https://doi.org/10.1038/s41562-016-0021

Nosek, B. A., & Errington, T. M. (2017). Making sense of replications. *eLife*, *6*. https://doi.org/10.7554/elife.23383

Nosek, B. A., & Errington, T. M. (2019). *What is replication?* Center for Open Science. https://doi.org/10.31222/osf.io/u4g6t

Prinz, F., Schlange, T., & Asadullah, K. (2011). Believe it or not: How much can we rely on published data on potential drug targets? *Nature Reviews Drug Discovery*, *10*(9), 712–712. https://doi.org/10.1038/nrd3439-c1

Replicating genotype–phenotype associations. (2007). *Nature*, *447*(7145), 655–660. https://doi.org/10.1038/447655a

Sainani, K. L., Borg, D. N., Caldwell, A. R., Butson, M. L., Tenan, M. S., Vickers, A. J., Vigotsky, A. D., Warmenhoven, J., Nguyen, R., Lohse, K. R., & al. (2020). Call to increase statistical collaboration in sports science, sport and exercise medicine and sports physiotherapy. *British Journal of Sports Medicine*, bjsports–2020–102607. https://doi.org/10.1136/bjsports-2020-102607

Sainani, K. L., Lohse, K. R., Jones, P. R., & Vickers, A. (2019). Magnitude-based inference is not bayesian and is not a valid method of inference. *Scandinavian Journal of Medicine & Science in Sports*, *29*(9), 1428–1436. https://doi.org/10.1111/sms.13491

Scheel, A. M., Schijen, M., & Lakens, D. (2020). An excess of positive results: Comparing the standard psychology literature with registered reports. *PsyArXiv*. https://doi.org/10.31234/osf.io/p6e9c

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Tamminen, K. A., & Poucher, Z. A. (2018). Open science in sport and exercise psychology: Review of current approaches and considerations for qualitative inquiry. *Psychology of Sport and Exercise*, *36*, 17–28. https://doi.org/10.1016/j.psychsport.2017.12.010

Turner, B. O., Paul, E. J., Miller, M. B., & Barbey, A. K. (2018). Small sample sizes reduce the replicability of task-based fMRI studies. *Communications Biology*, *1*(1). https://doi.org/10.1038/s42003-018-0073-z

Vigotsky, A. D., Nuckols, G. L., Heathers, J., Krieger, J., Schoenfeld, B. J., & Steele, J. (2020). *Improbable data patterns in the work of barbalho et al.* SportRxiv. https://doi.org/10.31236/osf.io/sg3wm