A REPRODUCIBILITY OR AN IDENTITY CRISIS IN SPORTS AND EXERCISE SCIENCES? A VIEWPOINT ON CURRENT CHALLENGES AND SOLUTIONS

[1] Guilherme Weiss Freccia*, PhD, 0000-0002-7393-9751

[1] Tales de Carvalho, MD, PhD, 0000-0002-8817-1255

1 Universidade do Estado de Santa Catarina (UDESC), Centro de Ciências da Saúde e do Esporte (CEFID), Florianópolis, Santa Catarina, Brazil

*Corresponding author and guarantor of the data

Prof. Guilherme Weiss Freccia, PhD

Universidade do Estado de Santa Catarina (UDESC)

Rua Pascoal Simone, 358 - Coqueiros - Florianópolis – SC. ZIP: 88080-350

Department of Physical Education, Universidade do Estado de Santa Catarina, Florianópolis, Brazil
guilherme.freccia@udesc.br, gwfreccia@gmail.com

This is a pre-print submitted to the BMJ Open Sport and Exercise Medicine.

How to cite (example):

Freccia GW, Carvalho T. A reproducibility or an identity crisis in sports and exercise sciences? A viewpoint on current challenges and solutions. SportRxiv 2022. DOI: 10.51224/SRXIV.171

ABSTRACT

Seeking to understand Sport and Exercise Sciences (SES) methodological scenarios, this viewpoint discusses (a) the methodological issues affecting reproducibility in SES, and (b) the initiatives intending to face these main problems, along with epistemological considerations. Reproducibility can be affected by inadequacies such as poor reporting of methods and outcomes. Results' reliability is an evidence-based cornerstone, and how these have been published in the SES field seem to be biased. Another contributing factor to a probable reproducibility crisis is the rising trend in the prevalence of positive results in comparison to negative results. Some suspicions boosted collaborations calling for the adoption of more transparent SES research, which currently lacks transparent research practices such as code and data sharing. To improve reproducibility, the reporting of methods, interventions and outcomes must be accurate and detailed. We may need to assess epistemology to better comprehend the identity of SES, and research methods the field has been relying on. This will require a collaborative effort and creativity, and more “theory-driven” research questions. It is time to slow down and rediscover the identity of the SES by establishing field-specific quality criteria and refining the study design under the scope of our own epistemological lens.

Key-words: sports and exercise, method, statistics, meta-research, reproducibility

Running title: Is there a reproducibility crisis in sports and exercise sciences?

Word-count: 198

INTRODUCTION

Recently, a research group raised a question for a Systematic Review with Meta-Analysis (SRMA), comparing the effects of different training interventions to fat loss in humans,(1)**.** As soon as the study was published in the *British Journal of Sports Medicine*, a letter in the journal's blog warned about the biological plausibility and discrepancies regarding accuracy of the statistical methods. Fortunately, some of the authors decided to give this study another narrative in light of good practices. After the retraction, the data was re-analyzed and the study pre-registered. The article was shared as pre-print, and the final version was published,(2) alongside with the proper sharing of data, codes, and materials (https://osf.io/6karz/). This study by Moher et al.,(3)**,** exemplifies how relevant it is to stimulate the adoption of accurate methodological and transparency practices in research. Despite these issues being relatively well-known in the biomedical field, this is not the case for the Sports and Exercise Sciences (SES),(4). In fields such as psychology, the topic has been extensively approached and the "replication crisis" was labelled as an overcome concern in the past decade. Yet, the crisis remains pervasive and the pattern may be followed in other areas of knowledge,(5,6). Nonetheless, addressing it in SES represents a challenge by itself, since the field has unique aspects in comparison to other health sciences and should be considered when tailoring solutions. Undoubtedly, it is high time for SES researchers to tackle such methodological issues under the field's epistemological lens.

The epistemological proximity with the biomedical sciences enables the SES scientist to transfer its methodological models, intending to answer research questions from performance to health,(7). Although the same inferential models can be used, such a strategy is not always appropriate. Many SES-related studies rely on surrogate outcomes even without proper prior validation of methods (e.g., performance technologies). Moreover, much of what is investigated in the biomedical field involves cutting-edge technology (e.g., surgical interventions), which does not always meet the SES' objectives: to understand the adaptations resulting from exercise, both for health outcomes and for performance (8).

In parallel, the quality of the methods and outcomes investigated in SES are influenced by common and unique features of the field. Firstly, methodological limitations are intrinsic to the research process and may occur from the conception of research design to the scientific dissemination of the final manuscript. Nonetheless, methodological errors and malpractices can arise either intentionally or not. Inadequate validation of methods and procedures, scarcity of longitudinal designs and replications, omission of null or negative results and unsatisfactory research transparency are some examples of how the process has been conducted without further solutions being provided,(7). Concomitantly, the SES researchers are frequently challenged by both low numbers of human resources (e.g., high-level athletes) and heterogeneous responses to the interventions (i.e., training adaptations). Yet, these issues should not be a justification to the omission of transparency or the underpowered statistical treatments, which hover over SES,(9,10).

Seeking to understand the problems mentioned above, this manuscript discusses (a) the methodological issues affecting reproducibility in SES, and (b) the initiatives intending to face these main problems. This viewpoint is not intended to explore the differences between the kinesiology and biomedical field, but to provide a rationale that explain why SES inherit a similar methodological crisis as in psychology and biomedical fields, and peculiarities that shall be considered when developing solutions.

THREATS TO REPRODUCIBILITY IN SPORTS AND EXERCISE SCIENCES

The reproducibility crisis in science is not recent. However, the issue received considerable attention after the publication of a remarkable piece by Douglas Altman, in which he states that *"We need less research, better research, and research done for the right reasons",*(11). Soon after, Ioannidis et al.,(12) warned about the high rate of false positives in biomedical research. As a consequence, the scientific community reinforced its efforts to understand, measure, and improve biomedical research quality, which led to the establishment of both collaborative projects,(13) and international organizations (e.g., EQUATOR -

Enhancing the QUAlity and Transparency Of health Research; ICMJE – International Committee of Medical Journal Editors).

Accordingly, underpowered studies have also been prevalent in SES for decades. One of the first studies to explore the application of the epidemiological methodology to SES,(14) discussed design issues and presented a revealing prevalence of underpowered studies in the field. The least popular study design was Randomized Controlled Trials (RCTs) (8%), and statistical power was not reported in most analyzed articles. Similarly, two studies examining the SES literature found a low proportion of RCTs ranging from 3%,(15) to 9.5%,(16). Interestingly, a relevant prevalence of both exercise and physical therapy interventions accounted for 39% among published trials,(17). More recently, three studies analyzed the SES literature and empirically confirmed the hypothesis of a general low statistical power in the field. Estimated sample sizes from various sub-disciplines were often below 40 subjects (10). Particularly, biomechanics articles had between 12-18 research participants (18); and the *Journal of Sports Science* presented a mean sample size of 19 (4). Furthermore, Twomey,(10) found that over 75% of analyzed studies did not include a justification for the sample size. Small sample sizes undermine the power to detect meaningful effects, but also skew the results and hamper the accuracy of evidence-based conclusions.

Reproducibility - also referred to as replicability – constitutes the researcher's ability to duplicate the results of previous studies. Nevertheless, this capability can be affected by inadequacies such as poor reporting of methods and outcomes,(19). A recent article in the SES field recalls a meta-research cornerstone: results' reliability,(20). The authors confirmed the hypothesis that studies comparing changes in $VO_{2máx}$ between Sprint Interval Training and Moderate Intensity Continuous Training have a high or unclear risk of bias and poor quality of reporting, which draws attention to the extent of biased research reporting in SES and that should be explored in other modes of intervention, such as resistance training.

Indeed, large, and questionable effect sizes are only a portion of the biases in resistance training studies. Possibly containing the above mentioned aspects, two studies brought suspicion of scientific misconduct, involving trained women,(21) and men,(22). They were retracted and received further attention in a whitepaper, which outlined important data abnormalities that suggested data fabrication,(23). Interestingly, these studies reported the same 6-week training regimen and instruments, resembling a form of "salami-science",(24). These suspicions boosted collaborations, such as the recent article calling for the “adoption of more transparent research",(9), which highlights the supporting role of registered reports to prevent malpractices and led to the proposal of the STORK (Society for Transparency, Openness, and Replication in Kinesiology). After that, Borg et al.,(25) sought to understand the lack of transparency in SES, and found a 4.3% prevalence of code and data-sharing among the 299 articles analyzed, addressing one pivotal transparency practice: the mandatory sharing of analytical material.

Another contributing factor to a probable reproducibility crisis is the rising trend in the prevalence of positive results in comparison to negative results,(26). When researchers aim to reach a p-value below 0.05 - which can be attained by defining obvious hypotheses and ensuring sufficient power - the search for a result seems to have more value than the answer to the research question itself. In the case of non-significance, researchers can still manipulate the database and statistical tests to reach the so-desirable value by "p-hacking" for statistical significance, (27). Buttner et al.,(28), assessed the proportion of hypothesis confirmation in highly ranked SES journals. The authors found 82% of positive results supporting the primary hypothesis, which shows a similar SES’ trend to the biomedical literature. In a recent study with 300 SES-related articles with different designs and sub-disciplines,(10), 64% of the sample reported results confirming its previous hypotheses. This is a considerably high rate of positive results, possibly driven by questionable research practices and publication bias. Moreover, one third of RCTs were not previously pre-registered, raising the issue of selective outcome reporting in SES clinical studies (29). Also, exercise scientists should take care when using statistical methods without a solid rationale and mathematical background, especially with small samples and intentions of doing

statistical inferences (i.e., clinical trials' effects). An example of this statistical misunderstanding emerge with the use of the magnitude-based inferences (MBI),(30), which may be increasing the false-positive rates,(31). Lack of transparency in the statistical reporting can make reproducibility difficult, as poor described statistical approaches cannot be replicated by independent researchers. Such context demonstrates the relevance of promoting collaborations between exercise scientists and statistical experts,(32).

IMPROVING SPORTS AND EXERCISE SCIENCE REPRODUCIBILITY: FINDING OUR IDENTITY

Registration is considered a helpful path in the case of a reproducibility crisis, and perhaps the first step towards a more transparent and reproducible SES. The reader should be able to understand if new findings had its hypotheses pre-specified, or if these hypotheses happened for convenience, after the results are known by the researchers leading the study,(33). Registered reports have also appeared as a promising publication model, where researchers register and submit the study design, hypotheses, and methods before its execution. In SES, we already have journal-led initiatives that adopted the Registered Report format (www.cos.io/rr), which facilitates the sharing of codes and data as well preregistration. In the field, these initiatives can re-establish of the importance of negative and/or trivial effects, an approach that can lead to an estimated drop of approximately 46% of positive results,(34).

As an example of efforts being carried out in low- and middle-income countries, a group of Brazilian meta-researchers has been evaluating the quality of clinical trials and systematic reviews in SES (https://sees-initiative.org) by assessing methodological rigor and transparency practices on publicly available databases,(35). The project has shown questionable practices in SES high-ranked journals, highlighting the omission of details provided in the standard recommendations for research methods and their publication (i.e., reporting guidelines), especially in RCTs. The initiative not only periodically assesses the methodological quality of publications in the exercise sciences, but also quantifies

the number of articles according to the study design, including RCTs, SRMAs, Guidelines, Letters to the Editor, and other categories.

To improve reproducibility, the reporting of methods, interventions and outcomes must be accurate and detailed. Despite well-established statements and guidelines such as CONSORT, SPIRIT and TIDieR exist to guide researchers on this concern,(36), SES clinical trials tend not to meet these recommendations and standards, particularly in regards to protocol accessibility, sample size calculations, number of analyzed subjects, or even specification of hypotheses,(10,37). Additionally, data availability statements are very uncommon and should be considered if we want to pursue a more reliable SES,(10,25).

Other practices that aim to improve the methodological quality should be considered by SES journals, such as optimizing submission rules for authors and in the peer review process. It is intriguing that some journals in the field did not have statements about transparent data availability until 2020. Along with this, the posture of journal editorial boards and reviewers should be more receptive to replication. Publishers can encourage their journals periodically promote special issues for articles centered on reproducibility or allocate space in specific sections for replication studies and manuscripts with negative results. The growth of SES as a field of scientific interest is undeniable. Under these circumstances, it is vital that we lay the foundations to accurately measure and monitor our methods and results,(16). More observatory studies of high scientific quality will elucidate the pathways for hypothesis-driven research questions that will support relevant RCTs and underpin the SRMAs we rely on in our clinical practice.

CONCLUSION

The task of addressing such relevant methodological issues that burden a reproducibility crisis will not be smooth. We may need to assess epistemology to better comprehend the identity of SES, and research methods the field has been relying on. This will require a collaborative effort and creativity, and more “theory-

driven” research questions. Even though innovations and inferences are pillars that ignite the scientific community, not all interventional effects are discoveries or have an intrinsic causality. We recommend that SES researchers dive into observational research more often, rather than relying most of its evidence on inferential study designs. It is time to slow down and rediscover the identity of the SES by establishing field-specific quality criteria and refining the study design under the scope of our epistemological lens.

ADDITIONAL INFORMATION

Acknowledgements:

None

Competing interests

None

Compliance with Reproducibility and Data Sharing Standards:

Not applicable

Funding Sources

This manuscript did not receive funding from any source to be conducted neither funding sources for their involved researchers.

Contributorship

Conception: GWF
Data Curation: not applicable
Formal Analysis: not applicable
Investigation: not applicable
Methodology: not applicable
Data treatment: not applicable
Writing: GWF, TC
Revision: TC
Final edition: GWF
Supervision: TC
Project administration: not applicable
Funding acquisition: not applicable

REFERENCES

1. Viana RB, Naves JPA, Coswig VS, de Lira CAB, Steele J, Fisher JP, et al. Is interval training the magic bullet for fat loss? A systematic review and meta-analysis comparing moderate-intensity continuous training with high-intensity interval training (HIIT). Br J Sports Med. 2019 May;53(10):655–64.

2. Steele J, Plotkin D, Van Every D, Rosa A, Zambrano H, Mendelovits B, et al. Slow and Steady, or Hard and Fast? A Systematic Review and Meta-Analysis of Studies Comparing Body Composition Changes between Interval Training and Moderate Intensity Continuous Training. Sports Basel Switz. 2021 Nov 18;9(11):155.

3. Moher D, Shamseer L, Cobey KD, Lalu MM, Galipeau J, Avey MT, et al. Stop this waste of people, animals and money. Nature. 2017 Sep;549(7670):23–5.

4. Abt G, Boreham C, Davison G, Jackson R, Nevill A, Wallace E, et al. Power, precision, and sample size estimation in sport and exercise science research. J Sports Sci. 2020 Sep;38(17):1933–5.

5. Iso-Ahola SE. Reproducibility in Psychological Science: When Do Psychological Phenomena Exist? Front Psychol. 2017 Jun 2;8:879.

6. Schweizer G, Furley P. Reproducible research in sport and exercise psychology: The role of sample sizes. Psychol Sport Exerc. 2016 Mar;23:114–22.

7. Halperin I, Vigotsky AD, Foster C, Pyne DB. Strengthening the Practice of Exercise and Sport-Science Research. Int J Sports Physiol Perform. 2018 Feb 1;13(2):127–34.

8. Gray. The Hazards of a Biomedical Exercise Paradigm: Exploring the Praxis of Exercise Professionals. Philosophies. 2019 Sep 12;4(3):54.

9. Consortium for Transparency in Exercise Science (COTES) Collaborators, Caldwell AR, Vigotsky AD, Tenan MS, Radel R, Mellor DT, et al. Moving Sport and Exercise Science Forward: A Call for the Adoption of More Transparent Research Practices. Sports Med. 2020 Mar;50(3):449–59.

10. Twomey R, Yingling V, Warne J, Schneider C, McCrum C, Atkins W, et al. Nature of Our Literature: A Registered Report on the Positive Result Rate and Reporting Practices in Kinesiology. Commun Kinesiol [Internet]. 2021 Dec 2 [cited 2022 Mar 8];1(3). Available from: https://storkjournals.org/index.php/cik/article/view/43

11. Altman DG. The scandal of poor medical research. BMJ. 1994 Jan 29;308(6924):283–4.

12. Ioannidis JPA. Why Most Published Research Findings Are False. PLoS Med. 2005 Aug 30;2(8):e124.

13. Open Science Collaboration. Estimating the reproducibility of psychological science. Science. 2015 Aug 28;349(6251):aac4716–aac4716.

14. Walter SD, Hart LE. Application of epidemiological methodology to sports and exercise science research. Exerc Sport Sci Rev. 1990;18:417–48.

15. Thompson B. A review of the British Journal of Sports Medicine 1991-5. Br J Sports Med. 1996 Dec 1;30(4):354–5.

16. Bleakley C. The quality of research in sports journals. Br J Sports Med. 2002 Apr 1;36(2):124–5.

17. Hopewell S, Dutton S, Yu L-M, Chan A-W, Altman DG. The quality of reports of randomised trials in 2000 and 2006: comparative study of articles indexed in PubMed. BMJ. 2010 Mar 23;340(mar23 1):c723–c723.

18. Knudson D. Confidence crisis of results in biomechanics research. Sports Biomech. 2017 Oct 2;16(4):425–33.

19. Goodman SN, Fanelli D, Ioannidis JPA. What does research reproducibility mean? Sci Transl Med. 2016 Jun 1;8(341):341ps12-341ps12.

20. Bonafiglia JT, Islam H, Preobrazenski N, Gurd BJ. Risk of bias and reporting practices in studies comparing VO2max responses to sprint interval vs. continuous training: A systematic review and meta-analysis. J Sport Health Sci. 2021 Mar;S2095254621000302.

21. Barbalho M, Coswig VS, Steele J, Fisher JP, Paoli A, Gentil P. Evidence for an Upper Threshold for Resistance Training Volume in Trained Women. Med Sci Sports Exerc. 2019 Mar;51(3):515–22.

22. Barbalho M, Coswig VS, Steele J, Fisher JP, Giessing J, Gentil P. Evidence of a Ceiling Effect for Training Volume in Muscle Hypertrophy and Strength in Trained Men - Less is More? Int J Sports Physiol Perform. 2020 Feb 1;15(2):268–77.

23. Vigotsky AD, Nuckols GL, Fisher J, Heathers J, Krieger J, Schoenfeld BJ, et al. Improbable data patterns in the work of Barbalho et al. [Internet]. SportRxiv; 2020 Jul [cited 2021 Jul 6]. Available from: https://osf.io/sg3wm

24. Norman G. Data dredging, salami-slicing, and other successful strategies to ensure rejection: twelve tips on how to not get your paper published. Adv Health Sci Educ. 2014 Mar;19(1):1–5.

25. Borg DN, Bon JJ, Sainani K, Baguley BJ, Tierney NJ, Drovandi C. Sharing Data and Code: A Comment on the Call for the Adoption of More Transparent Research Practices in Sport and Exercise Science [Internet]. SportRxiv; 2020

Mar [cited 2021 Sep 11]. Available from: https://osf.io/ftdgj

26. Fanelli D. Negative results are disappearing from most disciplines and countries. Scientometrics. 2012 Mar;90(3):891–904.

27. Marshall JR. Data Dredging and Noteworthiness: Epidemiology. 1990 Jan;1(1):5–7.

28. Büttner F, Toomey E, McClean S, Roe M, Delahunt E. Are questionable research practices facilitating new discoveries in sport and exercise medicine? The proportion of supported hypotheses is implausibly high. Br J Sports Med. 2020 Nov;54(22):1365–71.

29. Singh B, Fairman CM, Christensen JF, Bolam KA, Twomey R, Nunan D, et al. Outcome Reporting bias in Exercise Oncology trials (OREO): a cross-sectional study [Internet]. Sports Medicine; 2021 Mar [cited 2022 Mar 11]. Available from: http://medrxiv.org/lookup/doi/10.1101/2021.03.12.21253378

30. Batterham AM, Hopkins WG. Making meaningful inferences about magnitudes. Int J Sports Physiol Perform. 2006 Mar;1(1):50–7.

31. Lohse KR, Sainani KL, Taylor JA, Butson ML, Knight EJ, Vickers AJ. Systematic review of the use of “magnitude-based inference” in sports science and medicine. Bender R, editor. PLOS ONE. 2020 Jun 26;15(6):e0235318.

32. Sainani KL, Borg DN, Caldwell AR, Butson ML, Tenan MS, Vickers AJ, et al. Call to increase statistical collaboration in sports science, sport and exercise medicine and sports physiotherapy. Br J Sports Med. 2021 Jan;55(2):118–22.

33. Kerr NL. HARKing: Hypothesizing After the Results are Known. Personal Soc Psychol Rev. 1998 Aug;2(3):196–217.

34. Scheel AM, Schijen MRMJ, Lakens D. An Excess of Positive Results: Comparing the Standard Psychology Literature With Registered Reports. Adv Methods Pract Psychol Sci. 2021 Apr;4(2):251524592110074.

35. Umpierre D, De Nardi AT, Botton CE, Helal L, Pfeifer LO, Ricardo LI, et al. 52 Strengthening the evidence in exercise sciences initiative (SEES initiative): a prospective project based on openness, surveillance, and feedback. In: 3 Minute Quick Fire [Internet]. BMJ Publishing Group Ltd; 2019 [cited 2021 Jul 13]. p. A31.2-A32. Available from: https://ebm.bmj.com/lookup/doi/10.1136/bmjebm-2019-EBMLive.60

36. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ. 2014 Mar 7;348(mar07 3):g1687–g1687.

37. Bandholm T, Henriksen M, Thorborg K. Slow down to strengthen sport and exercise medicine research. Br J Sports Med. 2017 Oct;51(20):1453.