Communications in Kinesiology

Opinion

DOI:
10.51224/cik.2024.57

Subject Areas:
Meta-Science

Keywords:
Trainability, between-individual variability, adaptive variability, non-response

Editor:
David Borg

STORK SOCIETY FOR TRANSPARENCY OPENNESS AND REPLICATION IN KINESIOLOGY

Time to retire the raw analysis of individual responses

Arthur Henrique Bossi[1,2,*]

[1]School of Applied Sciences, Edinburgh Napier University, Edinburgh, United Kingdom
[2]Mountain Bike Centre of Scotland, Peel Tower, Glentress, Peebles, United Kingdom
[*]Correspondence: a.bossi@napier.ac.uk

This letter critiques the data analysis of a study investigating peak oxygen uptake responses to cycling and running sprint interval training (Digby et al., 2022 *J Strength Cond Res*, 37(4), e313-e316). While the study effectively demonstrates the specificity principle in the context of sprint interval training, concerns arise regarding the methodology used to categorise participants as responders or non-responders. The letter highlights the disregard for the recommendations of a number of academics advocating for specific experimental designs and statistical analyses to examine inter-individual variability. Furthermore, the reliability of within-individual adaptive responses to training and the potential impact of measurement errors and biological fluctuations are considered. It is suggested that the (non-)responder categorisation adds nothing to the main findings of the study and should be avoided. The importance of using appropriate experimental designs and statistical analyses when investigating inter-individual variability is emphasised. An open-source beta-version simulator is introduced as an educational resource to demonstrate the limitations inherent in the responder counting approach.

Letter

I am writing to share my thoughts regarding the data analysis employed by Digby et al. (2022). Using a between-subjects study design, Digby et al. (2022) investigated the response of peak oxygen uptake to cycling and running sprint interval training in 2 well-balanced groups of 9 young adults each, with sessions performed three times per week for three consecutive weeks. Regardless of the exercise mode, training sessions involved 30-s all-out sprints interspersed with 4-min resting periods, as well as a 5-min warm-up and cool-down at participants' preferred intensity. While 4 sprints per session were performed in the first week of training, a fifth and a sixth sprint were added to the sessions of weeks 2 and 3, respectively. A relevant aspect of the study was the measurement of adaptive responses using both a cycle ergometer and a treadmill, which allowed Digby et al. (2022) to examine crossover effects when tests are conducted with exercise modes that do not correspond with those of training. This study showed that cycling and running sprint interval training are equally effective in improving peak oxygen uptake, with mean increases of 5.2% and 5.6%, respectively. These improvements were nevertheless abolished when testing modes did not coincide with training modes, providing a compelling demonstration of the specificity principle (Reilly et al., 2009).

Alongside the usual cohort-based analysis, Digby et al. (2022) determined individual response rates (%) for peak oxygen uptake within each group. Essentially, participants were categorised as responders or non-responders depending on whether their change in peak oxygen uptake exceeded a threshold, calculated as $2 \times SD_{diff}/\sqrt{2}$. Here, SD_{diff} represents the standard deviation of the differences between the first and second measurements across participants in a reliability study. According to Hopkins (2000), cited by Digby et al. (2022), this procedure could be used for the monitoring of an individual (see page 7). However, Hopkins (2000) also dedicated a section to the estimation of the magnitude of individual differences (see page 9), where he presents an equation to gauge inter-individual variability in the experimental group(s) based on the variability associated with the control group or a reliability study. This approach is based on the premise that systematic effects can be attributed to the intervention only when outcomes in the absence of such intervention are accounted for—although relying on a shorter reliability study may inflate inter-individual variability estimates, which Hopkins (2000) failed to acknowledge. Despite Hopkins (2000) presenting a somewhat contradictory and incomplete perspective, Digby et al. (2022) unequivocally asserted that a control group was not required for their study. This wrong assumption, as explained below, indicates a potential oversight or deliberate disregard for the recommendations of many other academics advocating for specific experimental designs and statistical analyses to investigate inter-individual variability (Atkinson et al., 2019; Atkinson & Batterham, 2015; Chrzanowski-Smith et al., 2019; Egger et al., 2016; Hecksteden et al., 2015, 2018; Hrubeniuk et al., 2021; Ross et al., 2019; Senn, 2015; Voisin et al., 2019; Williamson et al., 2017).

Reliability research in exercise sciences has primarily concentrated on measurement errors and day-to-day biological fluctuations in response to a period without intervention (Hopkins, 2000). However, the sources of variability associated with an intervention are more difficult to quantify and control for (Atkinson et al., 2019; Atkinson & Batterham, 2015; Chrzanowski-Smith et al., 2019; Del Giudice et al., 2020a; Egger et al., 2016; Hecksteden et al., 2015, 2018; Hrubeniuk et al., 2021; Jamnick et al., 2020; Mann et al., 2013; Meyler et al., 2021; Ross et al., 2019; Senn, 2015; Voisin et al., 2019; Williamson et al., 2017). Indeed, a close examination of some old publications might suggest that within-individual adaptive responses are hardly replicable when endurance training programmes are repeated on distinct occasions (Cureton & Phillips, 1964; Örlander et al., 1977; Pedersen & Jørgensen, 1978; Simoneau et al., 1987); a hypothesis corroborated by more recent work (Del Giudice et al., 2020b). While these training studies may have failed to standardise the training load between and within individuals, given the limitations of relative intensity prescription methods (Egger et al., 2016; Jamnick et al., 2020; Mann et al., 2013; Meyler et al., 2021), their results reflect the difficulty in accurately estimating individual adaptive responses. As a minimum, it requires that training responses are measured at multiple points along the course of an intervention for modelling purposes, or the inclusion of a control group/period for comparative analysis based on standard deviations (Atkinson et al., 2019; Atkinson & Batterham, 2015; Chrzanowski-Smith et al., 2019; Hecksteden et al., 2015, 2018; Hrubeniuk et al., 2021; Ross et al., 2019; Senn, 2015; Voisin et al., 2019; Williamson et al., 2017). These methodological approaches account, at least partially, for the random measurement error and long-term biological fluctuations typical of training studies, in stark contrast with the simplistic post-hoc analysis conducted by Digby et al. (2022).

While I appreciate the hard work and dedication that Digby et al. (2022) have put into their study, effectively demonstrating the specificity principle in the context of sprint interval training, I believe that the (non-)responder categorisation adds nothing to their main findings. Objectively, my view is that the raw analysis of individual responses is fundamentally flawed. It overlooks the potential impact of compounded measurement errors and biological fluctuations associated with testing, both in the presence and absence of an intervention. It also neglects the possibility that within-individual adaptive responses are inherently variable from one training intervention to another. Put simply, there is no guarantee that the observed changes in peak oxygen uptake in the study of Digby et al. (2022) represent the true responses for each individual. In reality, the number of responders and non-responders within a sample is strongly influenced by the magnitude of the mean adaptive response, regardless of the response threshold chosen, as demonstrated nicely by simulations (https://ahbossi.shinyapps.io/Bossi_Simulator). Therefore, authors should avoid using this approach in an attempt to enhance the impact of their studies, as it could potentially detract readers from appreciating their primary findings. When the investigation of inter-individual variability in adaptive responses is the main research focus, an a priori decision should be made to use appropriate experimental designs and statistical analyses, for which a vast body of literature is available (Atkinson et al., 2019; Atkinson & Batterham, 2015; Chrzanowski-Smith et al., 2019; Hecksteden et al., 2015, 2018; Hrubeniuk et al., 2021; Ross et al., 2019; Senn, 2015; Voisin et al., 2019; Williamson et al., 2017).

Concluding Remarks and Future Directions

While this letter is deliberately succinct, readers are likely to benefit further from the literature previously referenced. My primary aim was to summarise the pitfalls associated with the raw analysis of individual responses, rather than providing a comprehensive review. Accompanying this letter is an interactive beta-version simulator, which I propose should undergo further optimisation to account for all sources of within-individual variability affecting the adaptive responses to training measured at the participant level. Future versions should also allow users to select from a variety of experimental designs to observe the impact on the (non-)responder categorisation, whilst presenting methodologically sound alternatives for the investigation of inter-individual variability. The source code for the simulator is publicly available (https://doi.org/10.5281/zenodo.8387674), and contributions to its continued development are highly encouraged. This is the first step in what must be a collective educational effort to enhance the rigour and validity of individual response research in exercise sciences.

Additional Information

Data Accessibility

The categorisation of research participants into responders and non-responders is critically examined through an interactive simulator, conceptualised to serve as a supplementary tool to this letter (https://ahbossi.shinyapps.io/Bossi_Simulator). The source code for the simulator has been archived and is publicly available (https://doi.org/10.5281/zenodo.8387674).

Author Contributions

A.H.B. conceived the idea, drafted, revised, and edited the manuscript. A.H.B. agrees to be accountable for all aspects of the work.

Conflict of Interest

A.H.B. has no conflicts of interest to declare.

Funding

None.

References

Atkinson, G., & Batterham, A. M. (2015). True and false interindividual differences in the physiological response to an intervention. *Experimental Physiology*, *100*(6), 577–588. https://doi.org/10.1113/ep085070

Atkinson, G., Williamson, P., & Batterham, A. M. (2019). Issues in the determination of 'responders' and 'non-responders' in physiological research. *Experimental Physiology*, *104*(8), 1215–1225. https://doi.org/10.1113/ep087712

Chrzanowski-Smith, O. J., Piatrikova, E., Betts, J. A., Williams, S., & Gonzalez, J. T. (2019). Variability in exercise physiology: Can capturing *intra*-individual variation help better understand true *inter*-individual responses? *European Journal of Sport Science*, *20*(4), 452–460. https://doi.org/10.1080/17461391.2019.1655100

Cureton, T., & Phillips, E. (1964). Physical fitness changes in middle-aged men attributable to equal eight-week periods of training, non-training and re-training. *The Journal of Sports Medicine and Physical Fitness*, *4*, 87–93.

Del Giudice, M., Bonafiglia, J. T., Islam, H., Preobrazenski, N., Amato, A., & Gurd, B. J. (2020a). Investigating the reproducibility of maximal oxygen uptake responses to high-intensity interval training. *Journal of Science and Medicine in Sport*, *23*(1), 94–99. https://doi.org/10.1016/j.jsams.2019.09.007

Del Giudice, M., Bonafiglia, J. T., Islam, H., Preobrazenski, N., Amato, A., & Gurd, B. J. (2020b). Investigating the reproducibility of maximal oxygen uptake responses to high-intensity interval training. *Journal of Science and Medicine in Sport*, *23*(1), 94–99. https://doi.org/10.1016/j.jsams.2019.09.007

Digby, L., McCarthy, S. F., Bornath, D. P. B., Copeland, J. L., & Hazell, T. J. (2022). Individual and Group Responses of Cardiorespiratory Fitness to Running and Cycling Sprint Interval Training. *Journal of Strength and Conditioning Research*, *37*(4), e313–e316. https://doi.org/10.1519/jsc.0000000000004340

Egger, F., Meyer, T., & Hecksteden, A. (2016). Interindividual Variation in the Relationship of Different Intensity Markers—A Challenge for Targeted Training Prescriptions. *PLOS ONE*, *11*(10), e0165010. https://doi.org/10.1371/journal.pone.0165010

Hecksteden, A., Kraushaar, J., Scharhag-Rosenberger, F., Theisen, D., Senn, S., & Meyer, T. (2015). Individual response to exercise training - a statistical perspective. *Journal of Applied Physiology*, *118*(12), 1450–1459. https://doi.org/10.1152/japplphysiol.00714.2014

Hecksteden, A., Pitsch, W., Rosenberger, F., & Meyer, T. (2018). Repeated testing for the assessment of individual response to exercise training. *Journal of Applied Physiology*, *124*(6), 1567–1579. https://doi.org/10.1152/japplphysiol.00896.2017

Hopkins, W. G. (2000). Measures of Reliability in Sports Medicine and Science. *Sports Medicine*, *30*(1), 1–15. https://doi.org/10.2165/00007256-200030010-00001

Hrubeniuk, T. J., Bonafiglia, J. T., Bouchard, D. R., Gurd, B. J., & Sénéchal, M. (2021). Directions for Exercise Treatment Response Heterogeneity and Individual Response Research. *International Journal of Sports Medicine*, *43*(01), 11–22. https://doi.org/10.1055/a-1548-7026

Jamnick, N. A., Pettitt, R. W., Granata, C., Pyne, D. B., & Bishop, D. J. (2020). An Examination and Critique of Current Methods to Determine Exercise Intensity. *Sports Medicine*, *50*(10), 1729–1756. https://doi.org/10.1007/s40279-020-01322-8

Mann, T., Lamberts, R. P., & Lambert, M. I. (2013). Methods of Prescribing Relative Exercise Intensity: Physiological and Practical Considerations. *Sports Medicine*, *43*(7), 613–625. https://doi.org/10.1007/s40279-013-0045-x

Meyler, S., Bottoms, L., & Muniz-Pumares, D. (2021). Biological and methodological factors affecting response variability to endurance training and the influence of exercise intensity prescription. *Experimental Physiology*, *106*(7), 1410–1424. https://doi.org/10.1113/ep089565

Örlander, J., Kiessling, K.-H., Karlsson, J., & Ekblom, B. (1977). Low Intensity Training, Inactivity and Resumed Training in Sedentary Men. *Acta Physiologica Scandinavica*, *101*(3), 351–362. https://doi.org/10.1111/j.1748-1716.1977.tb06017.x

Pedersen, P. K., & Jørgensen, K. (1978). Maximal oxygen uptake in young women with training, inactivity, and retraining. *Medicine and Science in Sports*, *10*(4), 233–237.

Reilly, T., Morris, T., & Whyte, G. (2009). The specificity of training prescription and physiological assessment: A review. *Journal of Sports Sciences*, *27*(6), 575–589. https://doi.org/10.1080/02640410902729741

Ross, R., Goodpaster, B. H., Koch, L. G., Sarzynski, M. A., Kohrt, W. M., Johannsen, N. M., Skinner, J. S., Castro, A., Irving, B. A., Noland, R. C., Sparks, L. M., Spielmann, G., Day, A. G., Pitsch, W., Hopkins, W. G., & Bouchard, C. (2019). Precision exercise medicine: understanding exercise response variability. *British Journal of Sports Medicine*, *53*(18), 1141–1153. https://doi.org/10.1136/bjsports-2018-100328

Senn, S. (2015). Mastering variation: variance components and personalised medicine. *Statistics in Medicine*, *35*(7), 966–977. https://doi.org/10.1002/sim.6739

Simoneau, J.-A., Lortie, G., Boulay, M. R., Marcotte, M., Thibault, M.-C., & Bouchard, C. (1987). Effects of two high-intensity intermittent training programs interspaced by detraining on human skeletal muscle and performance. *European Journal of Applied Physiology and Occupational Physiology*, *56*(5), 516–521. https://doi.org/10.1007/bf00635363

Voisin, S., Jacques, M., Lucia, A., Bishop, D. J., & Eynon, N. (2019). Statistical Considerations for Exercise Protocols Aimed at Measuring Trainability. *Exercise and Sport Sciences Reviews*, *47*(1), 37–45. https://doi.org/10.1249/jes.0000000000000176

Williamson, P. J., Atkinson, G., & Batterham, A. M. (2017). Inter-Individual Responses of Maximal Oxygen Uptake to Exercise Training: A Critical Review. *Sports Medicine*, *47*(8), 1501–1513. https://doi.org/10.1007/s40279-017-0680-8