



Published in final edited form as:

*J Physiol.* 2019 August ; 597(16): 4413–4414. doi:10.1113/JP278464.

## RESPONSE TO THE LETTER TO THE EDITOR

Jacob Graves McPherson, Albert Chen, Michael D. Ellis, Jun Yao, C. J. Heckman, Julius P. A. Dewald

In ‘Statistics and spin: it’s time to improve’, Gandevia and Heroux continue their pursuit of improved statistical reporting in biomedical research, and specifically within *The Journal of Physiology*. This is a worthy goal. Given their commitment to fairness and transparency, however, we were surprised by the amount of so-called ‘spin’ present in their Letter to the Editor. Below, we clarify a series of points raised by Gandevia and Heroux.

With regard to McPherson *et al.* (2018), first, we thank Gandevia and Heroux for noticing the accidental omission of the definition of error bars or asterisks in Fig. 2–4. The authors must have not noticed the error bar definitions in the legend of Fig. 6. The Statistics section of the manuscript was intended to have a statement reading, ‘Error bars are presented as standard error of the mean and asterisks are  $P < 0.05$  unless otherwise noted.’ It is unfortunate that a clerical error was missed during review and copy editing, resulting in the simple omission of this important information.

It is disappointing that Gandevia and Heroux suggest – without basis – that we purposefully omitted the definition of our error bars and did so because we are uncertain ‘which is better’. Such unqualified assertions are excellent examples of what they refer to as spin, and they do little to bring about meaningful change in the field.

The authors also raise a question about the demographics of the non-age-matched control participants. The cohort is stratified as two male/eight female, with a mean age of 24 years and a standard deviation of  $\pm 0.5$  years. Nine of ten participants reported right hand dominance and 1/10 did not report. We would have been happy to provide that information had it been requested by reviewers. However, to the best of our knowledge, no studies have reported a direct impact of biological sex, age or hand dominance on flexion synergy expression or reaching work area in this context.

Regarding statistical thresholding and our description of results, it appears that Gandevia and Heroux have highlighted portions of our manuscript and data presentation without full content or context. For example, our article actually reads, ‘a  $P$  value of 0.05 or less was considered *significant* in all cases.’ And indeed, readers of our manuscript will note that the statistical comparisons of all primary outcome measures fell below this threshold. That

---

### Author contributions

All authors have read and approved the final version of this manuscript and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All persons designated as authors qualify for authorship, and all those who qualify for authorship are listed.

### Competing interests

None declared.

being said, a full, fair and transparent accounting of our results necessitates reporting  $P$ -values both above and below this threshold. This is precisely why we have reported and commented on  $P$ -values between 0.05 and 0.1.

In another misleading characterization, Gandevia and Heroux extract the following portion of a statement from our manuscript:

Increased contralesional cortical activity in stroke was also strongly correlated with [...] and reduced angular velocity for shoulder flexion ( $r = 0.32$ ,  $P = 0.10$ ) and elbow extension ( $r = 0.37$ ,  $P = 0.06$ ).

They then question why ‘these  $P$ -values are implicitly considered statistically significant’. However, the full text of this section, which is omitted by Gandevia and Heroux, reads as follows (emphasis ours):

Increased contralesional cortical activity in stroke (i.e. a shift towards negative LI) was also strongly correlated with *poorer performance on kinematic movement parameters, including reduced angular excursion for shoulder flexion and elbow extension (Pearson’s  $r = 0.56$  and  $r = 0.59$ , respectively;  $P < 0.01$  for both)* and reduced angular velocity for shoulder flexion ( $r = 0.32$ ,  $P = 0.10$ ) and elbow extension ( $r = 0.37$ ,  $P = 0.06$ ). *These findings are consistent with the qualitative observation that ...*

It is curious that Gandevia and Heroux do not include mention of the statistically significant findings in the list, nor our assertion that amongst all of those data – statistically significant or not – we are making only a high-level, qualitative inference.

Similarly, Gandevia and Heroux point to two sentences that describe the same non-statistically significant difference between two groups – one sentence that occurs in the main body of the manuscript and one that appears in the legend of a figure. Despite the fact that the  $P$ -value of 0.08 was cited in both instances, the authors conclude that in the second case, our use of the word ‘trend’ constitutes a claim that the result was statistically significant. Furthermore, they omit the additional nearby statements that describe differences between the two groups that *were* statistically significant and that clearly provide a contrast to the result that did not meet the significance threshold. Despite the calculated excerpts selected by Gandevia and Heroux, we are confident that readers of our article are capable of interpreting these results using the language in the text.

Finally, it is unclear what point Gandevia and Heroux seek to make by noting the well-known mathematical relationship between  $P$ -values and study reproducibility. This appears to be another instance of spin in their Letter to the Editor. It leads the reader to the implied conclusion that our study in particular is somehow more unreliable than others. In actuality, all studies are bound by these same mathematical constraints. As Gandevia and Heroux well note, ‘these types of distortion are misleading and should not be present in the scientific literature (Wood *et al.* 2014)’.

If Gandevia and Heroux wish to have an unbiased, constructive discussion of statistical reporting and reproducibility, they will find us willing allies. These are core tenets of our

work. If, however, the authors wish to make unqualified accusations, to debate our diction, and to use us to advance their agenda, then we elect not to engage. If the authors' goal is alternatively to change the review process at *The Journal of Physiology*, then it strains credulity to think that singling out our manuscript in a Letter to the Editor is the most effective approach. It is, after all, likely to be read only by a subset of *The Journal's* subscribers. Instead, Gandevia and Heroux might be better served by recapitulating their 2018 review of statistical reporting (Diong *et al.* 2018) and submitting it directly to *The Journal of Physiology* as a full article.

Achieving optimal peer review and statistical reporting are worthy goals to which all scientists and publishers should aspire. Towards this end, we will continue our commitment to transparency by reporting all statistical outcomes and our interpretations thereof, allowing readers to evaluate the results for themselves. And, just as the field strives for precision and fairness in the reporting and review process, so too should it apply these criteria to editorial commentaries – avoiding spin and bias to maintain our academic discourse at the highest levels. In pursuit of their worthy agenda, we encourage Gandevia and Heroux to maintain sight of the forest as well as the trees.

## Acknowledgments

Funding

National Institutes of Health grants: R01NS054269, R01HD047569, R01 NS105759 and R01HD039343.

## References

- Diong J, Butler AA, Gandevia SC & Héroux ME (2018). Poor statistical reporting, inadequate data presentation and spin persist despite editorial advice. *PLoS One* 13, e0202121. [PubMed: 30110371]
- McPherson JG, Chen A, Ellis MD, Yao J, Heckman CJ & Dewald JPA (2018). Progressive recruitment of contralesional cortico-reticulospinal pathways drives motor impairment post stroke. *J Physiol* 596, 1211–1225. [PubMed: 29457651]
- Wood J, Freemantle N, King M, Nazareth I (2014). Trap of trends to statistical significance: likelihood of near significant P value becoming more significant with extra data. *BMJ* 348, g2215. [PubMed: 24687314]