Biological Psychology xxx (xxxx) 108763

Contents lists available at ScienceDirect

Biological Psychology

journal homepage: www.elsevier.com/locate/biopsycho

Commentary to: Standardization of facial electromyographic responses by van Boxtel and van der Graaff

Ursula Hess [a,*], Ottmar V. Lipp [b]

[a] Humboldt-University of Berlin, Germany
[b] Queensland University of Technology, Australia

ARTICLE INFO

Keywords:
Facial EMG

In their commentary *"Standardization of facial electromyographic responses"* van Boxtel and van der Graaff (2024) advocate to quantify facial EMG responses to stimuli as percentage of baseline amplitude. The rationale for this choice is that EMG responses have (theoretically) a resting value of zero. As such, this is good advice that can be followed in cases where indeed the theoretical baseline value of the expected EMG responses is zero.

However, the authors go further by making the claim that because of the theoretical zero level of resting EMG, the use of the percentage of baseline quantification is the only adequate quantification and that alternate approaches, specifically, standardized difference scores are misleading, because they are based on the assumption that EMG responses are interval scaled when in fact, they are ratio-scaled (i.e., have a true zero).

We argue that this claim is too strict. When considering scale levels, it is true that one cannot use statistics that are adequate for higher scale levels when working with lower scale levels but there is no such restriction in the other direction. As an example, "average" house prices are often expressed as medians because the existence of very few very expensive houses in a given area would distort the arithmetic mean normally used for averaging interval or ratio-scaled data. Thus, even though the arithmetic mean is an adequate statistic for prices, the median provides a more realistic estimate. In the same vein, the use of difference scores only requires minimally interval scaled data and as EMG data is ratio-scaled it fulfills that criterion. In short, the fact that EMG data is ratio-scaled does not by itself preclude the use of difference scores and in fact in some cases this metric provides a better estimate.

Specifically, standardized difference scores have a number of advantages. For example, contrary to the statement by the authors, they can be used to compare responses of different muscles. Standardization is achieved by using the z-transform. The result is that the data are now expressed in standard deviation units. As such, the transformed measures are now on the same scale and can hence be compared. That is, I can make for example the statement that whereas the activity of Orbicularis Oculi increased only one standard deviation unit, Zygomaticus Major activity increased by 1.5 standard deviation units.

Another reason that a given researcher may opt for standardized difference scores is that in reality facial muscles do not actually achieve a "true" zero baseline. Van Boxtel and van der Graaff advocate the use of relaxing videos as these tend to reduce EMG levels – but even this procedure may not achieve the desired goal. In fact, it may even be of interest to purposefully use a non-zero baseline. For example, if target emotional faces are placed into an emotion eliciting context, which itself also elicits facial expressions, one may use the context exposure as "baseline" to test for reactions to facial expressions over and above the reaction to the context. In this case the theoretical baseline is definitely different from zero and as such the advocated quantification may be quite misleading (as well as being statistically inappropriate).

Generally speaking, however, it is not the case that expressing trials as percent baseline results in a metric that is somehow more stable or reliable than z-transformed difference scores, because any artifact that affects either the baseline or the trial will affect both equally, since both metrics are computed based on the trial and the baseline.

In light of these considerations using within-subject z-transformed difference scores seems a conservative choice, because they do not rest on the (in practice problematic) assumption of EMG data measured at ratio-scale. However, van Boxtel and van der Graaff rightly note that this metric cannot be used when group comparisons are of interest. In this case, the choice of percentage baseline corrected scores presents a

* Correspondence to: Department of Psychology, Humboldt-University, Berlin, Germany.
E-mail address: Ursula.Hess@hu-berlin.de (U. Hess).

https://doi.org/10.1016/j.biopsycho.2024.108763
Received 7 January 2024; Received in revised form 8 February 2024; Accepted 8 February 2024
0301-0511/

solution – with its own problems. As such, neither approach is perfect, but perfection is often elusive.

In sum, van Boxtel and van der Graaff agree in principle with much of what we claim, but their conclusion is different. A more detailed appreciation of both approaches shows that both can be justified. Some experimental settings may in fact demand the use of one over the other. Researchers will have to make their own decision.

AI Statement

During the preparation of this work the authors used did not use any generative AI tool.

CRediT authorship contribution statement

Hess Ursula: Writing – original draft, Conceptualization. **Lipp Ottmar:** Writing – review & editing, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

Statement: The authors did not use generative AI technologies for preparation of this work.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Reference

van Boxtel, A. & van der Graaff, J. (in press) Standardization of facial electromyographic responses. Biological Psychology.