Biological Psychology 185 (2024) 108737

Contents lists available at ScienceDirect

Biological Psychology

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

# Standardization of facial electromyographic responses

Anton van Boxtel [a,*], Jolien van der Graaff [b]

[a] Department of Cognitive Neuropsychology, Tilburg University, Warandelaan 2, 5037 AB Tilburg, the Netherlands
[b] Department of Youth and Family, Faculty of Social and Behavioral Sciences, Utrecht University, Martinus J. Langeveldgebouw, Heidelberglaan 1, 3584 CS Utrecht, the Netherlands

ARTICLE INFO

*Keywords*
Facial EMG responses
Standardization
Percentage scores
Difference scores

ABSTRACT

In behavioral studies, facial electromyographic (EMG) responses to external stimuli or internal events are usually quantified relative to the resting state, presumed to represent a neutral baseline condition. In the large majority of recent studies, EMG responses were expressed as a difference score in terms of microvolts with the resting state. We argue that since EMG activity is measured on a ratio scale rather than on an interval scale, percentage scores should be used instead of difference scores. Reanalyzing results from an earlier study on the relationships between facial EMG responses and affective empathic responses to emotional video clips, we found that the two different types of EMG response quantification were differently related to affective empathy. Relationships between EMG responses and affective empathy were more consistent or stronger for percentage scores than for difference scores. In another study, facial EMG mimicry responses to pictures of emotional facial expressions were stronger for percentage scores than for difference scores. The adequacy of percentage scores relative to difference scores as indices of psychological variables may be simply checked by comparing both types of scores

In many studies, experimental effects of emotional conditions, sensory stimuli, cognitive tasks, or motor processes were investigated on electromyographic (EMG) activity of facial muscles. EMG responses were generally expressed relative to a resting baseline period immediately preceding the experimental condition. In the large majority of studies, EMG responses were quantified as the difference in microvolts between experimental condition and baseline amplitude. In a minority of studies, responses were quantified as a percentage of baseline amplitude. An overview of 100 recent (2000–2020) randomly selected EMG studies (see Supplementary Material) shows that difference scores were calculated in 68 studies and percentage scores in 12 studies. In 13 studies, raw EMG amplitudes or difference scores were intra-individually transformed into *z*-scores. Finally, 7 studies analyzed absolute EMG amplitudes during experimental conditions in terms of microvolts.

Contrary to the majority of commonly studied psychophysiological measures (e.g., brain potentials; cardiovascular, respiratory, or skin conductance measures) which are measured on an interval scale, muscular responses are quantified on a ratio scale with a zero origin. During complete relaxation, motoneurons are silent so that EMG activity is absent. Different from changes on an interval scale which are expressed as difference scores, changes on a ratio scale should be expressed as percentage scores. According to guidelines for reporting surface EMG responses of skeletal muscles being common in medical or physiological disciplines (Merletti, 1999), such responses are therefore expressed as percentage scores. Although facial muscles generally show some EMG activity during baseline conditions, their responses nevertheless should be quantified on a ratio scale since their activity in principle may decrease to zero as shown during REM sleep (Bliwise et al., 1974; Okura et al., 2006).

In their *Guidelines for human electromyographic research*, being based on a questionnaire among more than 100 investigators, Fridlund and Cacioppo (1986) recognize that a facial muscle at rest in principle shows electrical silence. However, they also recognize that within behavioral studies facial muscles in general are not completely inactive during baseline conditions, among others due to effects of previous experimental conditions or anticipation of forthcoming conditions. They propose that correcting for resting baseline levels could be performed by subtracting the resting EMG activity level from the experimentally induced level. To enable comparisons between different muscular sites within the same subject, or comparable sites in different subjects, they also propose other types of score corrections (e.g., range-corrected scores, standard scores). Nevertheless, they emphasize that there is insufficient evidence for the adequacy of a particular metric under all

* Corresponding author.
*E-mail address:* a.vanboxtel@tilburguniversity.edu (A. van Boxtel).

https://doi.org/10.1016/j.biopsycho.2023.108737
Received 11 June 2023; Received in revised form 16 November 2023; Accepted 18 December 2023
Available online 20 December 2023
0301-0511/

experimental circumstances. They notice that standardizing EMG scores has often been used to enable comparisons between different EMG sites within the same subjects, or comparable sites within different individuals, but that such scores may obscure or misrepresent treatment effects. They refer to a study by Fridlund et al. (1984) which would demonstrate that baseline-corrected (i.e., difference scores) or range-corrected EMG scores did not improve distinguishing between different emotions within the same subjects. These authors recommend performing between-subjects studies to evaluate whether baseline corrections or other types of correction are useful. They also suggest that researchers using corrected scores also furnish analyses based on uncorrected scores. Stating that it is complex to provide useful guidelines for standardizing facial EMG responses, Fridlund and Cacioppo (1986) do not recommend using a particular type of scores such as difference or percentages scores. The absence of such a recommendation may also be related to the fact that in this respect a predominant viewpoint of the investigators responding to their questionnaire was lacking.

As stated above, analogous to other popular psychophysiological measures, EMG difference scores are predominantly used in psychological studies. The major problem of such scores is that they are expressed in terms of microvolts and therefore considerably influenced by psychologically irrelevant physical factors determining the absolute amplitude of compound muscle action potentials like age or gender of participants (Jensen & Fuglsang-Frederiksen, 1994; Visser, 1974). In addition, facial muscles show considerable variability in functional and metabolic properties affecting absolute EMG amplitude like contraction speed, contraction strength, and fatigability (Freilinger et al., 1990; Schwarting et al., 1982). For example, frontalis and corrugator supercilii have a relatively high proportion of slow-twitch oxidative Type I muscles fibers with a relatively high resistance to fatigue, whereas orbicularis oculi and zygomaticus major have a high proportion of fast-twitch glycolytic Type II fibers with a lower resistance to fatigue. Such physical properties of muscle fibers are related to amplitude and duration of EMG action potentials (Lapatki et al., 2006) and may therefore affect EMG difference scores. In addition, when EMG measurements are repeated within the same person in different sessions, difference scores may vary between sessions. The reason is that minor changes in electrode location may have strong effects on absolute EMG amplitude in terms of microvolts (Lapatki et al., 2010). When measuring corrugator supercilii responses to angry expressions and zygomaticus major responses to happy expressions during two experimental sessions, test-retest correlations and intraclass correlations appeared to be considerably higher for percentage scores than for *z*-transformed difference scores (Hess et al., 2017).

EMG difference scores as an index of psychological processes may thus be strongly influenced by biophysical muscular properties and therefore have insufficient psychological validity. Such contaminating variables are better controlled for if, like in medical or physiological disciplines, EMG responses are expressed as a percentage of baseline activity rather than difference scores. Percentage scores enable intraindividual comparisons between responses of different muscles measured within a single experimental session, or responses of a specific muscle measured in different sessions (cf. Overbeek et al., 2012). They also enable interindividual comparisons (such as effects of age or gender) between responses of a particular muscle within or between sessions (cf. Overbeek et al., 2014).

Within-subjects *z*-score transformations applied on difference scores may be considered useful for intraindividual comparisons between responses of a particular muscle, but not for comparing responses of different muscles since EMG difference scores strongly depend on peculiar biophysical muscular properties. For the same reason, it is difficult to compare *z*-transformed difference scores between individuals.

Because during an experimental session resting baseline EMG levels may continuously vary under influence of experimental conditions, background stimulation, arousal, anxiety, expectancy of forthcoming events, habituation, mental fatigue, and other psychological factors, it should be recommended to measure a baseline value as shortly as possible before the onset of an experimental condition. In addition, random changes in resting baseline EMG activity may be reduced by presenting a relaxing video during the baseline period, such as a video of aquatic scenes (Piferi et al., 2000; see, for example, van der Graaff et al., 2016) or penguins (Vlemincx et al., 2011), instead of using a resting period without external stimulation. In accordance with a recommendation of Hastrup (1986) we also recommend starting a laboratory experiment after an introductory period which is long enough to dissipate the influence of initial physical or psychological activities on facial EMG activity. According to our experience, spontaneous EMG activity gradually declines during such an initial habituation period, like it also declines during the course of the entire experimental session. In an update of Fridlund and Cacioppo's guidelines paper (1986), Tassinary et al. (2017) mention several methods to reduce spontaneous baseline EMG activity like recording during pseudo trials or using a closed-loop baseline procedure during which the presentation of experimental conditions is programmed to be contingent on acceptably low levels of baseline EMG activity across the recording sites. However, this may be difficult to reconcile with the design of many studies.

Expressing EMG responses in terms of percentage of baseline level rather than as difference with baseline would enable a more reliable assessment of potential relationships between facial EMG responses and behavioral indices of emotional, perceptual, motor, or cognitive processes. Such relationships might be unreliable when using EMG difference scores. To illustrate differences between percentage scores and difference scores in this respect, we reanalyzed data which we collected in an earlier study (van der Graaff et al., 2016). In this study, zygomaticus and corrugator EMG responses were measured in 354 healthy adolescents (157 girls, mean age 16.9 years; 197 boys, mean age 17.0 years) as an index of empathic responses to happy or sad expressions exposed by peers in real-life video clips which were assembled from Dutch documentary films. Besides EMG responses, self-reported measures of affective state empathy (i.e., emotional valence) were collected. EMG responses during emotional film clips (two happy, two sad clips) were expressed as a percentage of EMG baseline level measured during the first 10 s of the emotionally neutral opening scene of each clip. As expected, significant positive Spearman's rank correlations were found between zygomaticus responses and affective state empathy responses to happy clips ($r_s = .174$, N = 354; one-tailed $p < .001$), as well as between corrugator responses and affective state empathy responses to sad clips ($r_s = .182$, N = 354; one-tailed $p < .001$). Similar analyses were performed for EMG responses expressed as a difference score (in microvolts) with baseline level. During happy clips, zygomaticus responses unexpectedly showed a negative correlation with affective state empathy ($r_s = -.192$, N = 354; one-tailed $p = 0.999$) rather than a positive correlation. During sad clips, corrugator responses as expected showed a significant positive relationship with affective state empathy ($r_s = .125$, N = 354; one-tailed $p = .009$). However, this relationship was less significant than was observed for percentage scores.

In another study comparing percentage scores with difference scores (Rutkowska et al., https://doi.org/10.31234/osf.io/t4dp6), binary relationships were investigated between presented pictures of posed happy or sad emotional expressions and associated EMG mimicry responses of zygomaticus and corrugator. Relationships were as expected but were stronger if EMG responses were expressed as a percentage of baseline activity rather than as difference scores. These studies illustrate that expressing EMG responses in terms of percentage scores or difference scores may result in varying relationships with psychological variables. Expecting that such relationships will be stronger for percentage than for difference scores, we suggest to follow Fridlund and Cacioppo's (1986) recommendation to compare distinct types of analyses.

## Declaration of Competing Interest

The authors declare that they have no competing interests to declare. The author(s) did not use generative AI technologies for preparation of this work.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.biopsycho.2023.108737.

## References

Bliwise, D., Coleman, R., Bergman, B., Wincor, M. Z., Pivik, R. T., & Rechtschaffen, A. (1974). Facial muscle tonus during REM and NREM sleep. *Psychophysiology, 11*, 497–508. https://doi.org/10.1111/j.1469-8986.1974.tb00578.x

Freilinger, G., Happak, W., Burggasser, G., & Gruber, H. (1990). Histochemical mapping and fiber size analysis of mimic muscles. *Plastic and Reconstructive Surgery, 86*, 422–428. https://doi.org/10.1097/00006534-199009000-00005

Fridlund, A. J., & Cacioppo, J. T. (1986). Guidelines for human electromyographic research. *Psychophysiology, 23*, 567–589. https://doi.org/10.1111/j.1469-8986.1986.tb00676.x

Fridlund, A. J., Schwartz, G. E., & Fowler, S. C. (1984). Pattern recognition of self-reported emotional state from multiple-site facial EMG activity during affective imagery. *Psychophysiology, 21*, 622–637. https://doi.org/10.1111/j.1469-8986.1984.tb00249.x

Hastrup, J. L. (1986). Duration of initial heart rate assessment in *Psychophysiology*: Current practices and implications. *Psychophysiology, 23*, 15–18. https://doi.org/10.1111/j.1469-8986.1986.tb00584.x

Hess, U., Arslan, R., Mauersberger, H., Blaison, C., Dufner, M., Denissen, J. J. A., & Ziegler, M. (2017). Reliability of surface electromyography. *Psychophysiology, 54*, 12–23. https://doi.org/10.1111/psyp.12676

Jensen, R., & Fuglsang-Frederiksen, A. (1994). Quantitative surface EMG of pericranial muscles. Relation to age and sex in a general population. *Electroencephalography and Clinical Neurophysiology, 93*, 175–183. https://doi.org/10.1016/0168-5597(94)90038-8

Lapatki, B. G., Oostenveld, R., van Dijk, J. P., Jonas, I. E., Zwarts, M. J., & Stegeman, D. F. (2006). Topographical characteristics of motor units of the lower facial musculature revealed by means of high-density surface EMG. *Journal of Neurophysiology, 95*, 342–354. https://doi.org/10.1152/jn.00265.2005

Lapatki, B. G., Oostenveld, R., van Dijk, J. P., Jonas, I. E., Zwarts, M. J., & Stegeman, D. F. (2010). Optimal placement of bipolar surface EMG electrodes in the face based on single motor unit analysis. *Psychophysiology, 47*, 299–314. https://doi.org/10.1111/j.1469-8986.2009.00935.x

Merletti, R. (1999). International Society of Electrophysiology and Kinesiology. Standards for reporting EMG data. *Journal of Electromyography and Kinesiology, 9*. III-IV.

Okura, K., Kato, T., Montplaisir, J. Y., Sessle, B. J., & Lavigne, G. J. (2006). Quantitative analysis of surface EMG activity of cranial and leg muscles across sleep stages in human. *Clinical Neurophysiology, 117*, 269–278. https://doi.org/10.1016/j.clinph.2005.10.003

Overbeek, T. J. M., van Boxtel, A., & Westerink, J. H. D. M. (2012). Respiratory sinus arrhythmia responses to induced emotional states: Effects of RSA indices, emotion induction method, age, and sex. *Biological Psychology, 91*, 128–141. https://doi.org/10.1016/j.biopsycho.2012.05.011

Overbeek, T. J. M., van Boxtel, A., & Westerink, J. H. D. M. (2014). Respiratory sinus arrhythmia responses to cognitive tasks: Effects of task factors and RSA indices. *Biological Psychology, 99*, 1–14. https://doi.org/10.1016/j.biopsycho.2014.02.006

Piferi, R. L., Kline, K. A., Younger, J., & Lawler, K. A. (2000). An alternative approach for achieving cardiovascular baseline: Viewing an aquatic video. *International Journal of Psychophysiology, 37*, 207–217. https://doi.org/10.1016/S0167-8760(00)00102-1

Rutkowska, J.M., Ghilardi, T., Vacaru, S.V., van Schaik, J.E., Meyer, M., Hunnius, S., Oostenveld, R. Optimal processing of surface facial EMG to identify emotional expressions: A data-driven approach. PsyArXiv Preprint. https://doi.org/10.31234/osf.io/t4dp6.

Schwarting, S., Schröder, M., Stennert, E., & Goebel, H. H. (1982). Enzyme histochemical and histographic data on normal human facial muscles. *Otorhinolaryngology, 44*, 51–59. https://doi.org/10.1159/000275567

Tassinary, L. G., Cacioppo, J. T., & Vanman, E. J. (2017). The somatic system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (4th ed.,, pp. 151–182). Cambridge: Cambridge University Press.

van der Graaff, J., Meeus, W., de Wied, M., van Boxtel, A., van Lier, P. A. C., Koot, H. M., & Branje, S. (2016). Motor, affective and cognitive empathy in adolescence: Interrelations between facial electromyography and self-reported trait and state measures. *Cognition and Emotion, 30*, 745–761. https://doi.org/10.1080/02699931.2015.1027665

Visser, S. L. (1974). Influence of sex and age on EMG contraction pattern. *European Neurology, 12*, 229–235. https://doi.org/10.1159/000114623

Vlemincx, E., Taelman, J., De Peuter, S., Van Diest, I., & Van den Bergh, O. (2011). Sigh rate and respiratory variability during mental load and sustained attention. *Psychophysiology, 48*, 117–120. https://doi.org/10.1111/j.1469-8986.2010.01043.x