Supplemental Information (SI) 1 for

Reading **your** emotions in **my** physiology? Reliable emotion interpretations in absence of a robust physiological resonance

Julia Folz[1,2], Donatella Fiacchino[1], Milica Nikolić[1,2,3], Henk van Steenbergen[1,2], Mariska E. Kret[1,2]

[1]Department of Cognitive Psychology, Institute of Psychology, Leiden University, 2333 AK Leiden, The Netherlands

[2]Leiden Institute for Brain and Cognition (LIBC), Leiden University, 2300 RCLeiden, The Netherlands

[3]Research Institute of Child Development and Education, University of Amsterdam, 1018 WS Amsterdam, The Netherlands

[*]Correspondence: m.e.kret@fsw.leidenuniv.nl

Outlier identification

Accuracy. In order to decrease the risk of including participants in the analyses who did not complete the task reliably (e.g., random response behavior), we searched for outliers in overall accuracy scores among our subjects. Hence, we calculated the median accuracy score ($Md_{Acc} = 0.85$) as well as the median absolute deviation ($MAD_{Acc} = 0.056$) of the subject-average accuracy scores across expressions. Values which were 3 *MAD*s smaller than the median were defined as outliers. According to this definition, one subject had to be excluded from the behavioral analyses ($M_{Acc} = 0.4$).

Pupil size. Due to the low sampling rate of the pupil size data (60Hz), time bins which had less than 100% signal coverage (i.e. less than 1 or 2 data points per bin) were not trusted and replaced by missing values. If there was no valid data point in the baseline window or more than 50% of the data points in the response window of interest (2s-4s after stimulus onset) were missing, the entire trial was excluded. Further, subjects with less than 50% valid trials in the pupil size data were excluded from later analyses. Overall, we excluded 760 trials out of the 7004 trials (incl. 5 entire datasets of subjects).

EMG. The outlier detection in the EMG data was performed on a trial-by-trial basis in Matlab, with a user-written automatic algorithm adapted from van Steenbergen (Dignath et al., 2019). If the absolute corrugator activity of a given time bin and/or its difference with the following time bin exceeded ± 3.5 *SD*s of the mean value/the mean difference value in the entire time window of interest (5 baseline time bins and 40 stimulus presentation time bins), the trial was said to contain an artifact. For the baseline window, it was additionally checked whether the absolute activity of a given time bin

exceeded the mean of a distribution of all time bins in the same position across trials by (±) 3.5 *SD*s. The same approach was used for the Zygomaticus, but the outlier criterion was elevated to ± 4 *SD*s as otherwise too many trials would have been excluded. Out of 7004 trials, 636 trials for the corrugator and 1007 trials for the Zygomaticus contained extreme values and were excluded. One subject was entirely excluded from EMG analyses due to excessive noise in both channels.

Skin conductance. Already during visual inspection, notes were taken if the signal for a specific trial seemed improbable (i.e. very noisy, unexpectedly large values/changes, or very little values/changes). In the next step, we applied theoretical criteria by searching for values outside of the expected range for skin conductance levels (2-20 µS) as well as changes which exceeded expected gradual changes in SCL (3 µS) within a trial (Dawson et al., 2016). Trials which were marked as improbable in both the visual inspection and the theoretical data exploration were immediately rejected. The remaining trials which exceeded the criteria were re-inspected and re-evaluated, given the overall impression of the participant's dataset. In total, we excluded 770 out of 7004 trials (incl. entire datasets of 7 participants).

Skin temperature. Similar to the skin conductance level, we noted conspicuous segments during visual inspections in the skin temperature recordings. Next, we checked for raw temperature values outside a range of 26 and 37 degrees as they seem to be rather unlikely on the cheek (Metzmacher et al., 2018; Voncken & Bögels, 2009). Changes within the 7.5 seconds of one trial exceeding 0.8°C were also marked (Ioannou et al., 2017). If visual inspection and value-based criteria overlapped, the data was excluded from the analysis. Since the exclusion criteria mostly affected participants' entire datasets, we eventually disregarded data from 7 participants.

References

Dawson, M. E., Schell, A. M., & Filion, D. L. (2016). The electrodermal system. *Handbook of Psychophysiology, Fourth Edition*, 217–243.

Dignath, D., Berger, A., Spruit, I. M., & van Steenbergen, H. (2019). Temporal dynamics of error-related corrugator supercilii and zygomaticus major activity: Evidence for implicit emotion regulation following errors. *International Journal of Psychophysiology*, *146*, 208–216.

Ioannou, S., Morris, P. H., Baker, M., Reddy, V., & Gallese, V. (2017). Seeing a blush on the visible and invisible spectrum: A functional thermal infrared imaging study. *Frontiers in Human Neuroscience*, *11*, 1–15.

Metzmacher, H., Wölki, D., Schmidt, C., Frisch, J., & van Treeck, C. (2018). Real-time human skin temperature analysis using thermal image recognition for thermal comfort assessment. *Energy and Buildings*, *158*, 1063–1078.

Voncken, M. J., & Bögels, S. M. (2009). Physiological blushing in social anxiety disorder patients

with and without blushing complaints: Two subtypes? *Biological Psychology*, *81*(2), 86–94.