UvA-DARE (Digital Academic Repository)

Emotional vocalizations are recognized across cultures regardless of the valence of distractors

Sauter, D.A.; Eisner, F.; Ekman, P.; Scott, S.K.

Link to publication

Emotional vocalisations are recognised across cultures regardless of distractor valence

Disa A. Sauter[1], Frank Eisner[2], Paul Ekman[3], Sophie K. Scott[4]

1. The Department of Social Psychology, University of Amsterdam, Weesperplein 4, 1018 XA Amsterdam, The Netherlands

2. Centre for Cognition, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Montessorilaan 3, 6525 HR Nijmegen, The Netherlands

3. The Paul Ekman Group, PO Box 26089, San Francisco, CA 94126, USA

4. Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London, WC1N 3AR, United Kingdom

Abstract

In a recent paper published in Psychological Science, Gendron et al. (2014) provided a replication of part of our study on cross-cultural recognition of emotional vocalisations (Sauter, Eisner, Ekman, & Scott, 2010b). Gendron et al. argued that listeners' recognition was better than chance only when distractors were of a different valence to the target. They concluded that affect, but not emotion, is communicated across cultures via vocalisations. We offer two alternative explanations for their results, and provide a re-analysis of our data. The results confirm our original conclusion that nonverbal vocalisations communicate specific emotions across cultures, extending our previous results by showing that this is the case also when distractor and target are of the same valence.

Main text

The question of whether emotional expressions have universal meanings has been at the centre of heated debates for many decades (e.g., Ekman, 1994; Russell, 1994). In a recent contribution, Gendron and colleagues (Gendron et al., 2014) took issue with our study of emotional vocalisations, in which we claimed to show that some nonverbal sounds reliably communicate emotional states across cultures (Sauter et al., 2010b). Gendron et al replicated part of our study (Himba listeners hearing British vocalisations), but analysed their data in terms of the valance and arousal of the distractors. They found that recognition was not better than chance for most types of distractors, and concluded that emotional vocalisations signal valence, but not specific emotions. We welcome their attempted partial replication of our work. Their paper has impelled us to re-analyse our original data in more depth. However, given the results of our re-analysis (described below), together with alternative explanations for Gendron et al.'s results, which we outline, we find that their conclusion does not hold.

In our study (Sauter et al., 2010b), participants heard a series of brief emotion stories. On each trial, they heard two vocalisations: one target (consistent with the story) and one distractor. Our original results, analysed per emotion but not distractor type, showed cross-cultural recognition of six emotions: Participants selected target vocalisations more often than would be expected by chance. However, three of the four positive emotions included in our study were not reliably recognised by Himba participants listening to British sounds. A fair test of whether recognition performance exceeds chance levels with different types of distractors should include only those emotions that are recognised across cultures at better-than-chance levels, since no claims have been made about the other emotions communicating specific emotional states across cultures. The fact that these emotions were included in Gendron et al.'s analysis will have resulted in a considerable reduction in overall accuracy (as shown by the particularly low accuracy for positive emotions found in their study). Our re-analysis excluded target vocalisations of the culture-specific signals of Triumph, Pleasure, and Relief.

In our original study, we varied distractors systematically, based on whether perceived valence was the same or different than the target (taken from Sauter et al., 2010a). For each emotion, every participant performed two trials with distractors of the same valence as the target, and two trials with distractors of opposite valence to the target[1]. We re-analysed the data from the 29 Himba participants in our original study who had heard British vocalisations (Sauter et al., 2010b), the part of our study that Gendron et al. replicated. Each trial was coded in terms of the relationship of the distractor to the target, Same Valence or Different Valence. Since surprise is valence neutral, trials in which surprise was either target or distractor were removed.

Following Gendron et al. (2014), scores were calculated as percentage correct classifications across emotions for trials with each type of distractor. One-sample t-tests comparing

performance to chance (50%) revealed that participants' performance was significantly above chance, both when distractors were of the opposite valence to the target ($t_{(28)}$ = 7.30, p < 0.001, Cohen's d:1.36; 95% CI [195.79, 348.55]), and when distractors were of the same valence as the target ($t_{(28)}$ = 4.05, p < 0.001, Cohen's d:0.75; 95% CI [66.38, 202.59]), see Figure 1.
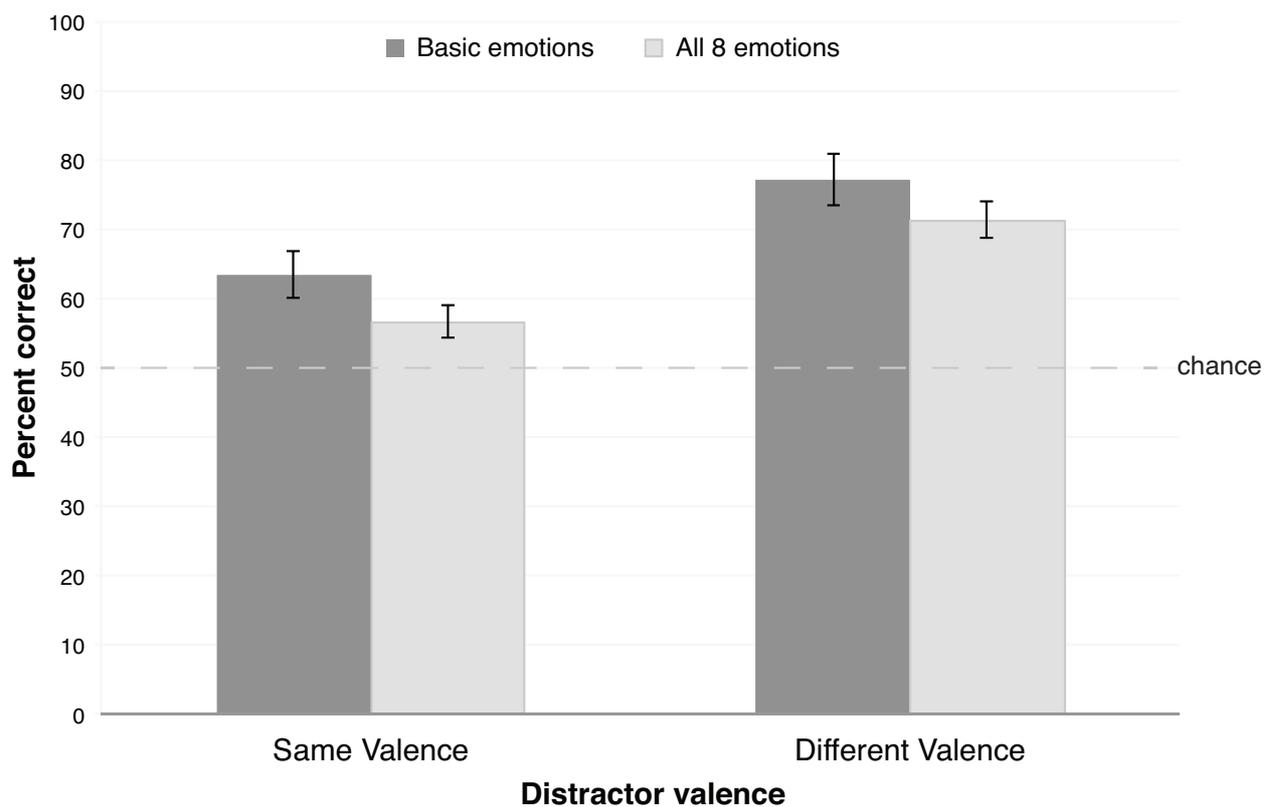

Figure 1. Himba participants' recognition performance, as percent correct, for basic emotions (dark bars) and all emotions (light bars). Distractors were of the same valence (left cluster) or different valence (right cluster), relative to the target. Error bars denote one standard error of the mean.

What might underlie the difference in results between our study and Gendron et al.'s (2014) study? As noted, the inclusion of the culture-specific signals of positive emotions in Gendron et al.'s study drove down recognition rates. In addition, the overall accuracy levels in their study are considerably lower than in our study (mean overall accuracy across all nine emotions in our

study: 64.46%; approximate mean based on Figure 4 in Gendron et al., 2014: 51%). We repeated our analysis including all eight valenced emotions, and found that participants' recognition was again significantly better than chance, both when distractors were of the opposite valence to the target, $t_{(28)} = 8.30$, $p < 0.001$ Cohen's d:1.54; 95% CI [160.91, 266.44], and when distractors were of the same valence as the target $t_{(28)} = 2.79$, $p < 0.01$, Cohen's d:0.52; 95% CI [17.89, 116.34], see Figure 1.

An apparent difference in procedure could account for the poor recognition rates in Gendron et al.'s (2014) study; In our study, each participant was asked, after each story, how the target person was feeling, in order to ensure that they had understood the story correctly. This is because in pilot testing, we found that participants would frequently say that they had understood a story, but when asked to explain it, they were unable to. In our study, therefore, participants could listen several times to the recorded story until they could explain the intended emotion in their own words. The inclusion of a rigorous manipulation check with experimenter verification rather than relying on participants' reports, was thus crucial. Gendron et al. do not report having included this check, only that participants who wished to, were allowed to hear the scenarios again. This raises the possibility that some participants in Gendron's study may not have correctly understood the intended emotional states.

This may also help to explain the perplexing pattern in Gendron et al.'s (2014) data, where recognition was better than chance only when the distractor matched the target in arousal and not in valence, but not when the distractor differed from the target in both arousal and valence. The authors acknowledge that this finding is unexpected; it does not fit with their account that vocalisations communicate valence-based information.

In conclusion, we have presented new analyses that show that nonverbal vocalisations communicate specific emotional states, regardless of the valence of the distractor. Furthermore, we have proposed alternative explanations for Gendron et al.'s (2014) failure to find this pattern in their data. We hope that this contribution will serve as a useful addition to the debate of what information can be inferred from emotional expressions across cultures.

Author contributions

D.A.S (re-)analysed the data and drafted the manuscript. All authors contributed to the original study design, and the writing of the manuscript.

Notes

1. In a small number of cases (<2%), participants did three trials with distractors of the same valence as the target for one emotion. Note that this, if anything, should make the task more difficult.

References

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's
mistaken critique. *Psychological Bulletin*, *115*(2), 268–287. doi: 10.1037/0033-
2909.115.

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Cultural Relativity in
Perceiving Emotion From Vocalizations. *Psychological Science*, *25*(4), 911–920.
doi:10.1177/0956797613517239

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review
of the cross-cultural studies. *Psychological Bulletin*, *115*(1), 102–141. doi:10.1037/0033-
2909.115.1.102

Sauter, D. A., Eisner, F., Calder, A. J., & Scott, S. K. (2010a). Perceptual cues in non-verbal vocal
expressions of emotion. *The Quarterly Journal of Experimental Psychology*, *63*(11),
2251–2272. doi:10.1080/17470211003721642

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010b). Cross-cultural recognition of basic
emotions through nonverbal emotional vocalizations. *Proceedings of the National
Academy of Sciences of the United States of America*, *107*(6), 2408–2412.
doi:10.1073/pnas.0908239106