



UvA-DARE (Digital Academic Repository)

Can perceivers recognise emotions from spontaneous expressions?

Sauter, D.A.; Fischer, A.H.

DOI

[10.1080/02699931.2017.1320978](https://doi.org/10.1080/02699931.2017.1320978)

Publication date

2018

Document Version

Final published version

Published in

Cognition & Emotion

License

CC BY-NC-ND

[Link to publication](#)

Citation for published version (APA):

Sauter, D. A., & Fischer, A. H. (2018). Can perceivers recognise emotions from spontaneous expressions? *Cognition & Emotion*, 32(3), 504-515.
<https://doi.org/10.1080/02699931.2017.1320978>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Can perceivers recognise emotions from spontaneous expressions?

Disa A. Sauter and Agneta H. Fischer

Department of Social Psychology, University of Amsterdam, Amsterdam, Netherlands

ABSTRACT

Posed stimuli dominate the study of nonverbal communication of emotion, but concerns have been raised that the use of posed stimuli may inflate recognition accuracy relative to spontaneous expressions. Here, we compare recognition of emotions from spontaneous expressions with that of matched posed stimuli. Participants made forced-choice judgments about the expressed emotion and whether the expression was spontaneous, and rated expressions on intensity (Experiments 1 and 2) and prototypicality (Experiment 2). Listeners were able to accurately infer emotions from both posed and spontaneous expressions, from auditory, visual, and audiovisual cues. Furthermore, perceived intensity and prototypicality were found to play a role in the accurate recognition of emotion, particularly from spontaneous expressions. Our findings demonstrate that perceivers can reliably recognise emotions from spontaneous expressions, and that depending on the comparison set, recognition levels can even be equivalent to that of posed stimulus sets.

ARTICLE HISTORY

Received 31 August 2016

Revised 6 April 2017

Accepted 6 April 2017

KEYWORDS

Nonverbal communication;
genuine; vocal expressions;
vocalisations; facial
expressions

The vast majority of research into nonverbal communication of emotions uses posed stimuli, because of the high degree of experimental control that they afford researchers. However, critics have argued that the use of posed expressions inflates recognition accuracy relative to spontaneous expressions (e.g. Nelson & Russell, 2013), and concerns have been raised over whether observers can in fact reliably recognise emotions from spontaneous expressions at all (Russell, 1994). Posed stimuli have also been criticised for being artificial and consequently not representative of expressions that occur in real life (see Scherer, Clark-Polner, & Mortillaro, 2011 for a discussion). But although some studies have examined the recognition of individual emotions from spontaneous expressions (e.g. Fernandez-Dols, Carrera, & Crivelli, 2011; Tracy & Matsumoto, 2008; Wagner, 1990), surprisingly few studies have directly compared recognition of emotions from spontaneous and posed stimuli within a single paradigm. But given the wealth of research into nonverbal emotional

expressions that uses posed expressions, it is important to establish whether it is scientifically sound to generalise from findings using posed expressions to real-life situations involving spontaneous emotional expressions. The current study aimed to contribute to addressing the question of how spontaneous emotional expressions are perceived compared to the typical stimuli used in the field of emotion research, that is, posed expressions.

Studies comparing recognition of posed and spontaneous expressions

As noted, only a handful of studies have directly compared the perception of spontaneous and posed facial expressions, and they have generally lent support to the proposal that recognition is more accurate for posed than for spontaneous expressions (Russell, 1994). In an early study, Zuckerman and colleagues examined whether viewers could judge valence and intensity from spontaneous facial expressions of positive

CONTACT Disa A. Sauter  d.a.sauter@uva.nl

 Supplemental data for this article can be accessed here. <https://doi.org/10.1080/02699931.2017.1320978>

© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

and negative emotion. Spontaneous expressions were elicited through viewing film clips and were compared to posed expressions from the same individuals (Zuckerman, DeFrank, Hall, & Rosenthal, 1976). Perceivers' ratings of valence were more often correct for posed than for spontaneous expressions, although whether viewers could judge specific emotional states from the expressions was not examined (see Wilting, Krahmer, & Swerts, 2006, for a similar approach using speech). A later study using a forced-choice emotion classification task also found that spontaneous emotional facial expressions were poorly recognised compared to posed expressions (Motley & Camden, 1988). Thus, the few studies conducted to date on facial expressions have supported the idea that observers are better at recognising emotions from posed, as compared to spontaneous, expressions.

In the auditory channel of nonverbal communication of emotion, two studies comparing emotion recognition for posed and spontaneous expressions have been conducted to date. Jürgens and colleagues have examined spontaneous emotional speech prosody from radio sequences containing fear, anger, joy, and sadness, and compared them to posed portrayals (Jürgens, Drolet, Pirow, Scheiner, & Fischer, 2013; Jürgens, Grass, Drolet, & Fischer, 2015). Their findings revealed a modest but statistically significant recognition advantage for posed as compared to spontaneous emotional speech in the 2013 study. In the 2015 study, however, recognition accuracy was highest in absolute terms for the spontaneous expressions overall, though the pattern differed across emotions. Thus, very little work has directly compared emotion recognition from spontaneous and posed stimuli and findings so far are mixed. No study to date has compared recognition of posed and spontaneous nonverbal vocalisations, nor for multimodal expressions.

The current study

The main goal of the current study was to examine the recognition of emotions from posed and spontaneous nonverbal vocalisations perceived from auditory and/or visual cues. Nonverbal vocalisations are brief vocal expressions that do not contain speech. They include screams, sighs, and laughs, but exclude lexicalised exclamations, such as *ouch*, *yuck*, or *yikes*. Based on theoretical arguments (e.g. Russell, 1994), as well as previous findings (e.g. Motley & Camden, 1988), we hypothesised that posed expressions would be better

recognised than spontaneous expressions. Based on research on speech prosody (Jürgens et al., 2013), we also predicted that recognition from spontaneous expressions would reach better-than-chance levels.

The present studies were designed to also test two supplementary hypotheses. We sought to establish whether two candidate features would contribute to recognition accuracy: perceived intensity (Experiments 1 and 2) and perceived prototypicality (Experiment 2). Several studies have found a link between perceived intensity and recognition rates. Hess and colleagues, for example, examined the recognition of posed facial expressions of varying levels of intensity. They found that recognition levels varied linearly with the intensity of expressions, that is, observers recognised more intense expressions more accurately (Hess, Blairy, & Kleck, 1997). There is also evidence for a role of perceived intensity in the recognition of vocal expressions of emotions from work on speech intonation. Juslin and Laukka (2001) tested recognition of happy, sad, angry, scared, and disgusted speech segments produced with either weak or strong emotion intensity. They found higher decoding accuracy for portrayals with strong emotion intensity. However, no study to date has examined the role of intensity in the recognition of emotion from nonverbal vocalisations.

In an examination of prototypicality, Laukka and colleagues tested the perception of segments of speech inflected with anger, fear, or joy (Laukka, Audibert, & Aubergé, 2012). Their results lend some support to the notion that expressions that are more prototypical are better recognised, but whether this relationship applies to nonverbal vocalisations is not yet clear. In sum, following these earlier findings, we hypothesised that expressions with higher perceived intensity and prototypicality would be recognised more accurately. Of particular interest was whether this pattern of results would be found for both spontaneous and posed expressions.

Finally, we aimed to test the generalisability of previous findings from the perception of speech prosody, which have shown that listeners are able to judge whether an emotional expression is genuine or posed (Audibert, Aubergé, & Rilliard, 2008; Jürgens et al., 2013; see also Jürgens et al., 2015). We hypothesised that perceivers would be better than chance in judging whether a stimulus was posed or spontaneous.

The hypotheses were tested using unimodal (Experiments 1 and 2) and multimodal (Experiment

2) expressions. Both experiments were approved by the University of Amsterdam Psychology ethics committee. The sample sizes of both experiments were pre-determined based on feasibility, and all measures taken are included in this report. No participants were excluded in Experiment 1; 10 participants were excluded in Experiment 2. This was due to a programming error yielding incomplete data (8 participants) or them completing the task unconsciously (pressing only one response button: 2 participants).

Experiment 1

Experiment 1 tested whether listeners could recognise emotions from spontaneously produced nonverbal vocalisations of emotion, and whether recognition accuracy would be lower for spontaneous than for posed expressions. Furthermore, we collected judgements of the perceived intensity of the emotional expressions, and finally, judgements of whether expressions were spontaneous or posed.

Method

Stimuli

Spontaneous stimuli were taken from online sources (e.g. youtube.com, soundcloud.com). Segments were extracted from shows including Expedition Robinson, Holland's Next Top Model, Try Before You Die, and Secret Story (see Supplementary Table 4). Inclusion was determined on the basis of two criteria: (1) whether the situational context allowed clear inference of a target emotion based on the core relational theme of each emotion (see Supplementary Table 1), and (2) the presence of a clearly audible nonverbal vocalisation, defined as any human vocalisation other than speech (e.g. screams, sighs, grunts, laughs). Thus, selection was not based on whether vocalisations were deemed to be emotional, but exclusively on the eliciting context and the mere presence of a nonverbal vocalisation. No clips were discarded for any reason other than failing to fulfil these two criteria. The collection of spontaneous expressions was done by research assistants who had not heard the posed expressions, and they were free to use any search terms relating to the emotions and core relational themes. A stopping criterion was applied such that searches were conducted until four stimuli (two male, two female) were found for each emotion category.

The posed vocalisations were taken from a validated set of nonverbal emotional vocalisations (Sauter, 2013). The posed stimuli were produced by lay people via enactment of felt or recalled emotions. The producers were completely unconstrained in terms of the form of their expressions, except that they must not contain speech. The subset of stimuli used in the current study was a random selection (constrained by emotion category and gender) from a set of well-recognised expressions by Dutch producers (Sauter, 2013).

For spontaneous and posed sounds, respectively, four stimuli (2 female) for each of nine emotions (anger, disgust, fear, sadness, surprise, triumph, amusement, sensual pleasure and relief) were included, yielding a total of 72 stimuli. Average duration was 1.15 seconds for the posed stimuli and 1.28 seconds for the spontaneous stimuli. For the spontaneous stimuli, each item was produced by a different individual; the posed stimuli were produced by seven different speakers (three female, four male). The full set of stimuli is available from the first author on request.

Participants

Thirty-three Dutch participants (25 female; average age 21 years) took part for course credits or payment.

Procedure

Participants were tested individually and provided informed consent. Sounds were delivered in a random order via headphones using the Psychophysics toolbox (Brainard, 1997) for MATLAB (Mathworks Inc., Natick, MA) running on a MacBook laptop. After every clip, participants identified the expressed emotion in a 9-way forced choice task, indicated the intensity of the emotion on a 7-point Likert scale, and judged whether the vocalisation was spontaneous or posed in a two-way forced choice. On each trial, participants thus made three judgements in a fixed order. All responses were given using the number keys.

Results and discussion

Recognition accuracy

Recognition rates were analysed using Hu scores (unbiased hit rates, Wagner, 1993), computed for each participant for spontaneous and posed responses, and arcsine transformed prior to statistical analyses. A score of 0 represents no correct responses, and a score of 1 is perfect performance. Note that an

observer's judgment that a stimulus expresses a given emotion does not necessarily entail them attributing an emotional state to the expresser; an observer's judgment of an expression as being of, for example, fear, was defined as being correct if the expression was produced in a situation that involved the core relational theme of fear.

To test whether performance was significantly better than chance, separate *t*-tests were performed for recognition of posed and spontaneous expressions. Chance was set to a stringent level of 1/4 responses correct, as there were four options of each valence. We set chance to 0.25 rather than 0.11 (i.e. 1/9) as this is considered a more conservative test (see e.g. Cordaro, Keltner, Tshering, Wangchuk, & Flynn, 2016): Arguably, most classification errors are made within valence categories, such that positive expressions are not likely to be (mis)taken to express negative emotions, and setting chance level to 0.11 may therefore be overly likely to yield significant results. In line with the primary hypothesis, performance was significantly better than chance for both posed ($t(32) = 17.57$, $p < .001$, Cohen's d : 3.06, 95% CI [0.32, 0.40]) and spontaneous ($t(32) = 5.96$, $p < .001$, Cohen's d : 1.04, 95% CI [0.06, 0.13]) expressions (see Figure 1; see Table 1 for breakdown per emotion and Supplementary Table 2 for arcsine transformed Hu scores per emotion). The results also confirmed the prediction that posed expressions would be recognised better than spontaneous ones. Performance was superior for posed (mean 0.88) as compared to spontaneous (mean 0.62) expressions, $t(32) = 9.86$, $p < .001$, Cohen's d : 1.72, 95% CI [0.21, 0.32]. These results show that participants

could reliably recognise emotions from both posed and spontaneous vocalisations, but that they were more accurate for posed vocal expressions.

Intensity

Our main interest in collecting ratings of intensity was to examine whether this feature would be related to how well perceivers could identify the expressed emotion. However, these data also allowed for a test of whether posed and spontaneous expressions differed in terms of how intense they were perceived to be. Indeed, a *t*-test showed a difference between the conditions in perceived intensity, with posed expressions being judged as more intense (mean 4.68; standard deviation: 0.55) than spontaneous expressions (mean 4.31; standard deviation: 0.55), $t(32) = 5.55$, $p < .001$, Cohen's d : 0.97, 95% CI [0.23, 0.51].

To examine the relationship between perceived intensity and recognition accuracy, the data were re-coded by items rather than by participants. As perceived intensity differed across conditions (see above), separate linear regressions were performed for posed and spontaneous expressions. Perceived intensity predicted recognition accuracy (Hu scores) for spontaneous ($\beta = 0.14$, $t(34) = 3.06$, $p < .004$, $r^2 = 0.22$), but not for posed expressions ($p > .1$). This demonstrates that participants were more accurate in recognising emotions from vocalisations that they perceived as expressing more intense states, but only for spontaneous vocalisations.

Differentiation between posed and spontaneous expressions

Listeners were on average 59% correct in the differentiation between posed and spontaneous expressions (see Table 2). In order to test listeners' accuracy on

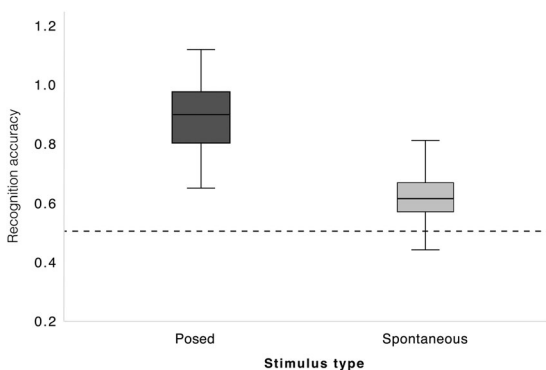


Figure 1. Performance (arcsine Hu scores) on the emotion recognition task in Experiment 1. Data are plotted by stimulus type. Lines through the boxes are the medians, box edges are the 25th and 75th percentiles, and the whiskers extend to the most extreme data points excluding outliers. The dashed line represents chance (calculated as 1/4 correct, as there were four options of each valence).

Table 1. Table showing recognition rates (raw Hu scores) in Experiment 1 ($n = 33$) for spontaneous (left) and posed (right) expressions (standard deviations in brackets). Means as arcsine transformed Hu scores (used in the statistical analyses) can be found in Supplementary Table 2.

Emotion	Spontaneous vocalisations	Posed vocalisations
Triumph	.30 (.20)	.37 (.22)
Amusement	.46 (.23)	.84 (.17)
Anger	.24 (.18)	.38 (.20)
Disgust	.11 (.08)	.70 (.17)
Fear	.34 (.17)	.48 (.24)
Relief	.25 (.19)	.66 (.20)
Sadness	.53 (.17)	.57 (.17)
Pleasure	.65 (.25)	.53 (.21)
Surprise	.32 (.18)	.62 (.20)
Total	.36 (.19)	.57 (.20)

Table 2. Correct judgments (%) of whether expressions were spontaneous or posed for Experiment 1 (above) and Experiment 2 (below) for each modality separately.

Modality	Experiment 1	
	Spontaneous	Posed
Auditory	58.33	59.09
	Experiment 2	
	Spontaneous	Posed
Auditory	48.08	54.96
Visual	74.17	75.28
AudioVisual	74.44	85.21

this measure, *d* prime scores were calculated and tested against chance. Participants were able to discriminate between posed and spontaneous vocalisations at significantly better than chance levels, $t(32) = 4.94$, $p < .001$, Cohen's *d*: 0.86, 95% CI [0.29, 0.69].

A previous study of spontaneous and posed emotional speech found evidence of a "truth bias" (Jürgens et al., 2013), such that listeners were more likely to judge stimuli as being genuine rather than posed. In order to test whether this effect would replicate with our data, we examined *c* scores, which reflect response biases in guessing tendencies. A *t*-test was conducted to compare *c* scores (actual guessing bias) computed from the posed/spontaneous judgement task, to zero (the absence of a guessing bias). No evidence of a guessing bias was found (mean: 0.01, $p > .8$).

Experiment 2

Experiment 2 tested whether the pattern of results found in Experiment 1 would hold using a different posed stimulus set, and also examined whether the results would generalise across modalities. Specifically, Experiment 2 tested whether recognition accuracy would be poorer for spontaneous than for posed stimuli across auditory, visual, and audiovisual presentation. In addition to intensity, judgments of prototypicality were recorded in order to test whether perceived prototypicality would contribute to improved recognition of emotional expressions (Laukka et al., 2012; see also Scherer et al., 2011). Finally, perceivers' ability to distinguish between spontaneous and posed expressions from uni- and multimodal cues was examined.

Method

Stimuli

For the posed stimulus set used in Experiment 1, no multimodal expressions are available. Therefore, in

Experiment 2, the posed expressions were taken from the Geneva Multimodal Expression Corpus (GEMEP, Bänziger, Mortillaro, & Scherer, 2012). The posed stimuli in the GEMEP set were produced by French-Swiss actors who were guided by a director in the enactment of felt or recalled emotions. All vocalisations had to consist of producers saying "aa", but were otherwise unconstrained in terms of form. The items used in the current study were a random set (constrained by emotion category and gender) from the full GEMEP set (see Bänziger et al., 2012).

The spontaneous stimulus set from Experiment 1 was modified by replacing tokens in which the face of the person vocalising could not be seen, or where background sound was present. Sources included TV series such as the Great British Bake Off, Fear Factor, and The Complete Sex Guide, as well as videos uploaded by the general public (see Supplementary Table 4). A moving oval mask was applied to the videos using Adobe After Effects (Adobe Systems, San Jose, CA) in order to remove all visual information except the face (and movement) of the target individual. A total of 72 stimuli (half spontaneous, half posed; half male, half female) were included, balanced across 9 emotions, with an Auditory, Visual, and AudioVisual version of each. Average duration was 1.96 seconds for the posed stimuli and 1.31 seconds for the spontaneous stimuli. As in Experiment 1, the spontaneous stimuli were all produced by different individuals. The posed stimuli were produced by 11 different speakers (6 female). The full set of stimuli is available from the first author on request.

Participants

A total of 122 Dutch participants took part for course credits or payment, with modality of the stimulus being a between-subjects factor; 42 participants (36 female; average age 22 years) were in the Audio condition, 40 participants (31 female; average age 23 years) were in the Visual condition, and 40 participants (29 female; average age 22 years) were in the AudioVisual condition.

Procedure

The procedure was identical to that of Experiment 1, except that ratings of prototypicality were added, with judgements made on a 7-point Likert scale. Specifically, participants were asked to what extent they found the expression prototypical of the emotion category that they had selected for that expression. Note that participants were not provided

with prototypical exemplars, but merely evaluated whether they found each stimulus prototypical according to their own judgment. The experiment was run using Presentation (Neurobehavioral Systems Inc., Berkeley, CA).

Results and discussion

Recognition accuracy

As in Experiment 1, arcsine transformed Hu scores were used to examine recognition accuracy. A series of *t*-tests compared performance against chance for each stimulus type (posed and spontaneous stimuli separately for Audio, Visual, and AudioVisual presentation, Bonferroni corrected for multiple tests). All stimulus types were recognised better than chance (set to 1/4 as in Experiment 1): (Audio: posed: $t(41) = 8.50$, $p < .001$, Cohen's $d: 1.31$, 95% CI [0.10, 0.16]; spontaneous: $t(41) = 5.08$, $p < .001$, Cohen's $d: 0.78$, 95% CI [0.06, 0.13]); (Visual: posed: $t(39) = 4.95$, $p < .001$, Cohen's $d: 0.78$, 95% CI [0.06, 0.13]; spontaneous: $t(39) = 3.47$, $p < .001$, Cohen's $d: 0.55$, 95% CI [0.03, 0.11]); (AudioVisual: posed: $t(39) = 9.53$, $p < .001$, Cohen's $d: 1.51$, 95% CI [0.19, 0.30]; spontaneous: $t(39) = 9.57$, $p < .001$, Cohen's $d: 1.51$, 95% CI [0.20, 0.31]). Thus, naïve observers reliably recognised emotional states from both posed and spontaneous emotional expressions. The results are displayed in Figure 2 (see Table 3 for a breakdown of results per

emotion and Supplementary Table 3 for arcsine transformed Hu scores per emotion).

To test if posed stimuli were recognised better than spontaneous ones, and whether this effect would interact with presentation Modality, an ANOVA was run with Modality as a between-subjects factor (3 levels: Audio, Visual, AudioVisual) and Spontaneity as a within-subject factor (2 levels: posed, spontaneous). Modality had a significant effect on recognition accuracy ($F(2,119) = 25.59$, $p < .001$, $\eta^2 = 0.30$), with recognition levels higher for multimodal presentation (mean: 0.77) as compared to the unimodal presentation conditions, which did not differ from one another (Audio mean: 0.63; Visual mean: 0.60). No main effect or interaction with spontaneity was found (both $p > .1$). In order to conduct a maximally tough test of the hypothesis that posed stimuli are recognised better than spontaneous expressions, follow-up simple effects tests were run within each modality, comparing recognition accuracy for posed and spontaneous expressions. No significant differences were found, that is, recognition accuracy for spontaneous expressions did not significantly differ from that of posed stimuli in any modality condition (Audio mean difference 0.035, $p > .1$; Visual mean difference 0.026, $p > .2$; AudioVisual mean difference 0.09, $p > .6$). These results fail to support the

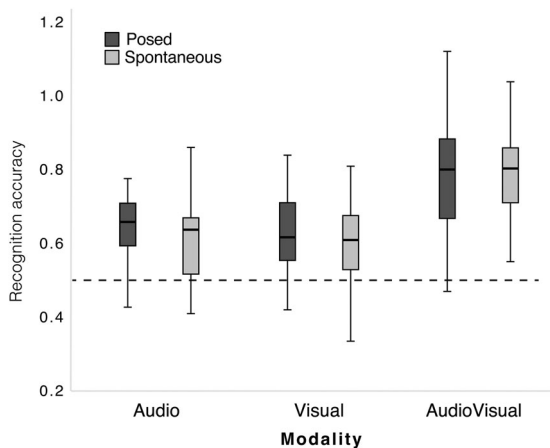


Figure 2. Emotion recognition in Experiment 2 (arcsine Hu scores) for posed (dark boxes) and spontaneous (light boxes) emotional expressions. Lines through the boxes are the medians, box edges are the 25th and 75th percentiles, and the whiskers extend to the most extreme data points excluding outliers. The dashed line represents chance (calculated as 1/4 correct, as there were four options of each valence).

Table 3. Table showing mean recognition rates (raw Hu scores) in Experiment 2 in each modality, separately for spontaneous (top) and posed (bottom) expressions. Means as arcsine transformed Hu scores (used in the statistical analyses) can be found in the Supplementary Materials.

Emotion	Audio ($n = 42$)	Visual ($n = 40$)	AudioVisual ($n = 42$)
	Spontaneous		
Triumph	.16 (.17)	.12 (.13)	.24 (.20)
Amusement	.54 (.16)	.35 (.16)	.47 (.20)
Anger	.47 (.27)	.21 (.15)	.40 (.23)
Disgust	.34 (.25)	.45 (.23)	.58 (.27)
Fear	.25 (.13)	.25 (.21)	.42 (.20)
Relief	.26 (.20)	.14 (.13)	.37 (.22)
Sadness	.59 (.28)	.63 (.18)	.89 (.19)
Pleasure	.58 (.20)	.63 (.32)	.72 (.25)
Surprise	.09 (.11)	.26 (.15)	.21 (.15)
Total	.36 (.20)	.34 (.18)	.48 (.21)
Emotion	Posed		
	Audio ($n = 42$)	Visual ($n = 40$)	AudioVisual ($n = 42$)
Triumph	.21 (.21)	.10 (.16)	.18 (.18)
Amusement	.55 (.12)	.36 (.13)	.52 (.20)
Anger	.74 (.19)	.76 (.25)	.84 (.22)
Disgust	.28 (.21)	.22 (.16)	.39 (.14)
Fear	.51 (.20)	.67 (.19)	.71 (.25)
Relief	.45 (.22)	.30 (.20)	.44 (.21)
Sadness	.07 (.11)	.37 (.17)	.22 (.21)
Pleasure	.39 (.19)	.12 (.13)	.46 (.31)
Surprise	.35 (.18)	.35 (.17)	.55 (.23)
Total	.40 (.18)	.36 (.18)	.48 (.22)

hypothesis that posed stimuli necessarily yield inflated recognition rates compared to spontaneous expressions of emotion.

Intensity. An ANOVA was used to test whether ratings of intensity differed depending on modality of presentation and whether stimuli were posed or spontaneous. No main effect of Modality was found, but there was a main effect of Spontaneity ($F(1,119) = 24.27$, $p < .001$, $\eta^2 = 0.17$), qualified by an interaction ($F(2,119) = 23.33$, $p < .001$, $\eta^2 = 0.28$). Follow-up simple effects analyses revealed that intensity ratings of spontaneous and posed stimuli differed significantly, but in different directions depending on the modality of presentation. Consistent with Experiment 1, in Audio presentation, posed expressions (mean: 5.11; standard deviation: 0.58) were rated as more intense than spontaneous expressions (mean: 4.93; standard deviation: 0.65; $p = .008$). In contrast, for Visual presentation, spontaneous expressions (mean: 4.97; standard deviation: 0.54) were perceived as more intense than posed expressions (mean: 4.64; standard deviation: 0.58; $p < .001$), and this pattern was also found for AudioVisual presentation (spontaneous mean: 5.09; standard deviation: 0.59; posed mean: 4.66; standard deviation 0.75; $p < .001$). This suggests that there is not a straight-forward relationship between perceived intensity and the spontaneity of emotional expressions, but rather that this relationship depends on the modality of presentation.

Prototypicality. Ratings of prototypicality were analysed using the same approach as the intensity ratings, with an ANOVA. No main effects were found, but there was an interaction between Modality and Spontaneity $F(2,119) = 17.33$, $p < .001$, $\eta^2 = 0.23$. Follow-up simple effects analyses showed that there were no differences in perceived prototypicality between spontaneous and posed stimuli for Visual presentation (spontaneous mean: 4.70; standard deviation: 0.60; posed mean: 4.75; standard deviation 0.57), but differences were found in opposite directions for Audio and AudioVisual presentation. In Audio presentation, posed expressions (mean: 4.83; standard deviation: 0.73) were rated as more prototypical than spontaneous expressions (mean: 4.50; standard deviation: 0.71; $p < .001$). In contrast, for AudioVisual presentation, spontaneous expressions (mean: 4.93; standard deviation: 0.78) were perceived as more prototypical than posed expressions (mean: 4.68; standard deviation: 0.99; $p < .001$). This pattern

of results is consistent with the ratings of intensity, and indeed ratings of intensity and prototypicality were found to correlate significantly for judgments of both posed ($r = 0.60$, $p < .001$) and spontaneous ($r = 0.59$, $p < .001$) expressions.

Intensity and prototypicality in relation to recognition accuracy

To examine whether recognition accuracy could be predicted from perceived intensity and/or prototypicality, the data were re-coded by item. As intensity and prototypicality ratings differed across conditions (see above), separate linear regressions were performed for each condition. Intensity did not predict recognition in any model. When prototypicality was added, however, it predicted recognition for posed stimuli only in the AudioVisual condition ($\beta = 0.32$, $t(33) = 3.75$, $p < .001$, r^2 change = 0.36). For spontaneous expressions, prototypicality predicted recognition in all modalities (Auditory: $\beta = 0.21$, $t(33) = 3.03$, $p < .005$, r^2 change = 0.22; Visual: $\beta = 0.26$, $t(33) = 2.93$, $p < .006$, r^2 change = 0.25; AudioVisual: ($\beta = 0.28$, $t(33) = 4.70$, $p < .001$, r^2 change = 0.43). Prototypicality thus relates to recognition accuracy, particularly for spontaneous expressions.

Differentiation between posed and spontaneous expressions

Participants were, in terms of per cent correct, accurate on 52% of trials in the Auditory condition, 75% of trials in the Visual only condition, and 80% of trials for AudioVisual stimuli (See Table 2) As in Experiment 1, d prime scores were tested against chance to examine discrimination between posed and spontaneous expressions. Participants were able to discriminate posed from spontaneous expressions for Visual ($t(39) = 10.39$, $p < .001$, Cohen's d : 1.64, 95% CI [1.26, 1.88]) and AudioVisual stimuli ($t(39) = 14.23$, $p < .001$, Cohen's d : 2.25, 95% CI [1.70, 2.27]), but not for Audio stimuli.

Similarly to Experiment 1, we tested for the existence of a "truth bias" (Jürgens et al., 2013), that is, listeners being more likely to think that stimuli are genuine as compared to posed. Specifically, we conducted a t -test comparing c scores (actual guessing bias) computed from the posed/spontaneous judgment task, to zero (the absence of a guessing bias), separately for each modality. No difference was found for Audio presentation (mean: 0.12, $p > .1$) or Visual presentation (mean 0.08, $p > .2$). However, in the AudioVisual condition, a truth bias was found,

with participants being significantly more likely to guess that stimuli were spontaneous than that they were posed (mean: 0.26, $t(39)=4.16$, $p<.001$). In sum, these results provide limited support for the truth bias hypothesis.

General discussion

Emotion recognition from posed and spontaneous expressions

Across two experiments, perceivers could reliably recognise emotions from spontaneous expressions. This finding was consistent across modalities (Auditory, Visual, and AudioVisual), across four participant samples, and using stringently set chance levels.

Whether posed stimuli were recognised more accurately than spontaneous expressions differed depending on the posed stimulus set used: the stimuli from Sauter (2013) used in Experiment 1 were better recognised than the spontaneous expressions, while this was not the case for the GEMEP (Bänziger et al., 2012) in Experiment 2. These findings highlight the importance of the stimulus set used when comparing recognition from spontaneous and posed expressions, as recognition levels for the posed stimuli differed considerably (raw Hu scores mean 0.57 in Experiment 1; 0.40 in Experiment 2 Audio condition; note that recognition levels for the spontaneous auditory expressions were the same across the experiments with a mean of 0.36). The difference in recognition rates for the posed stimuli is noteworthy, particularly because the ways in which the two sets of stimuli were produced were largely consistent (and both sets of posed expressions were produced in contexts that matched the core relational theme of each emotion, as were the spontaneous expressions, see Supplementary Materials). Both posed stimulus sets were produced via enactment of felt or recalled emotions, with producers relatively unconstrained in terms of the form of their expressions, that is, how their voice should sound or how they should move their faces.

Multiple factors may nevertheless help explain the difference in recognition levels. Firstly, there was a difference in the elicitation procedures used for the two posed stimulus sets: The posed stimuli employed in Experiment 1 were produced by lay individuals, while the GEMEP stimuli were produced by trained actors who were guided by a director. However, recent work on speech prosody found no difference

in recognition accuracy when comparing posed stimuli produced by individuals with and without acting experience (Jürgens et al., 2015), and so this factor is unlikely to have had a large influence. Secondly, the posed stimulus sets differed in pre-testing: the stimulus set used in Experiment 1 was a selection of well-recognised items from a larger set. In Experiment 2, a subset of the GEMEP was used, but it was not possible to use the core set (i.e. the best recognised items) because this set does not include nonverbal vocalisations. A third possible source of the difference in recognition levels between posed sets may be the cultural origins of the stimuli. The posed expressions used in Experiment 1 were produced by Dutch people who were from the same cultural group as the listeners. The posed stimuli employed in Experiment 2, in contrast, were produced by French-Swiss expressers. Though the cultural differences between these two groups are unlikely to be dramatic, the current pattern of results for the posed expressions is in line with previous findings showing a cultural in-group advantage for posed nonverbal vocalisations of emotions (e.g. Sauter & Scott, 2007). Finally, the two posed stimulus sets also differ in terms of phonetic properties: the GEMEP affect bursts all consist of speakers saying “aa”, whereas the stimuli used in Experiment 1 were completely unconstrained (other than that they could not contain speech). Further work will be needed to establish the influence of each of these factors on the recognition of posed expressions.

The high recognition accuracy for the spontaneous expressions is remarkable given that the spontaneous stimuli were selected based on situational rather than expressional features, and had not been pre-selected based on pilot testing. In addition, the quality of the spontaneous videos was inferior to that of the posed stimuli, which had been produced in controlled laboratory conditions. Nevertheless, the spontaneous expressions were recognised at similar levels to posed stimuli in Experiment 2. It should be noted, however, that although the research assistants who selected the spontaneous clips were specifically instructed to include all clips that met the criteria (a match with core relational theme and the presence of a nonverbal vocalisation), it cannot be comprehensively ruled out that their selection could nevertheless have been influenced by some degree of implicit bias. For example, they may have inadvertently selected exemplars containing particularly prototypical nonverbal expressions. It is also possible that clips readily

available on sources like youtube may differ from expressions occurring during most events in daily life, in terms of, for example, prototypicality. This may especially be the case for nonverbal expressions produced as part of reality TV shows. Hopefully, it will be possible to rule out these possibilities with future studies. Replications will also be important in order to establish the generalisability of the current findings to other spontaneous expressions of emotions. Future work may also want to include other designs and response formats, such as employing intensity ratings or including a “none-of-the-above” alternative (see Beaupre & Hess, 2005; Frank & Stennett, 2001; Yik, Widen, & Russell, 2013), as well as directly comparing expressions produced in contexts in which the expresser was aware of being observed or not.

Given the small number of stimuli per emotion in the current study, we do not wish to draw conclusions regarding the recognition of individual emotions. However, some preliminary comparisons with previous work may nevertheless be informative, as the recognition of spontaneous expression has been reported to vary dramatically across emotion categories: Jürgens et al. (2013) reported that posed expressions of anger were recognised more accurately than their spontaneous counterparts, while spontaneous expressions of sadness were recognised better than posed ones (see also Jürgens et al., 2015), and no differences were found for expressions of fear and joy. This pattern of results partially fits with those of the current study: Inspection of Tables 1 and 3 suggests that anger expressions were consistently better recognised from posed as compared to spontaneous expressions. However, this appears to also be the case for fear expressions, which differs from Jürgens and colleagues’ findings. We found partial support for Jürgens and colleagues’ result that sad expressions are better recognised from spontaneous expressions, with this pattern borne out in Experiment 2 (across modality conditions), but not in Experiment 1. Results for happiness cannot be compared across studies, as the current study did not use the label “happiness” but rather included multiple categories of specific positive emotions.

Finally, the current study made use of posed and spontaneous expressions, but it is worth noting that this distinction has been questioned. Arguably all emotional expressions that occur in public are subject to some social constraints, including cultural norms about socially appropriate behaviour.

Therefore, it has been suggested that the distinction between posed and spontaneous may be more gradual than categorical (Scherer et al., 2011).

Expressions, feelings, and intensity

The studies presented in this paper concern the perception of emotional expressions, and thus do not allow for firm conclusion regarding whether the producers of the spontaneous expressions necessarily felt the emotions they expressed (see Lench, Flores, & Bench, 2011 for a meta-analysis finding support for the link between emotional expressions and self-reported emotion, but also Fernández-Dols & Crivelli (2013) for a critique of this notion). Future studies should ideally measure the felt emotional experience of the producer, though emotion induction methods are typically limited by low emotion intensity. Our results also provide some evidence suggesting that high intensity and prototypicality can facilitate recognition from spontaneous expressions. Note however that the expressions in the current study were overall of high intensity. Most nonverbal vocalisations are probably unlikely to occur in low-intensity situations: for example, one would be likely to emit a scream of fear if one is suddenly petrified, but not if one is only a little bit scared. It is worth noting though that episodes in which no nonverbal vocalisations were produced were not examined in the current study. Thus, the current dataset cannot be used to establish in what kinds of situational contexts nonverbal vocalisations, or particular kinds of nonverbal vocalisations, occur. There is, however, empirical support for the association between emotional nonverbal behaviours (including both nonverbal vocalisations and facial expressions) and intensity of subjective experience (e.g. Mauss, Levenson, McCarter, Wilhelm, & Gross, 2005; but see Reisenzein, Studtmann, & Horstmann, 2013 for a critical perspective). Though this association between felt emotional intensity and emotional expressions concerns the production of expressions, it does seem at odds with recent research on emotion perception that has found that perceivers cannot differentiate expressions of extremely intense positive and negative emotions (Aviezer, Trope, & Todorov, 2012). This raises the possibility of an inverted U-shaped curve, with signal clarity being maximal at high, but not extreme, levels of intensity. However, it will need to be established whether these results replicate with expressions other than static facial cues, and it will also be

important to use measures other than valence judgments: Studies of facial expressions associated with sexual pleasure have found great overlap in terms of muscular movements with facial configurations seen during pain (Fernandez-Dols et al., 2011), but naïve viewers are able to differentiate between them in a two-way forced-choice task with the response alternatives “pain” and “sexual pleasure” (Hughes & Nicholson, 2008).

Differentiating between posed and spontaneous expressions

Consistent with findings from emotional speech prosody (Jürgens et al., 2013; but see Jürgens et al., 2015), perceivers in our study were sometimes able to differentiate between posed and spontaneous expressions. However, performance was far from ceiling, especially from only auditory cues, suggesting that posed vocalisations do not sound dramatically different to spontaneous ones. In contrast, differentiation between posed and spontaneous expressions was accurate for visual stimuli (the Visual and AudioVisual conditions in Experiment 2). Might this be explained by the fact that the recordings of the posed stimuli were technically of a better quality than the spontaneous stimuli? We consider this unlikely since substantial differences in recording quality would likely yield perceptible auditory differences as well. A candidate explanation for the high discrimination of visual stimuli may be physical cues that differ between authentic and inauthentic facial expressions. For example, movement-onset asymmetry differs between spontaneous and posed facial expressions (e.g. Ross & Pulusu, 2013). In the domain of facial expressions (e.g. Hess & Kleck, 1990) and particularly on smiling (e.g. Hess, Kappas, McHugo, Kleck, & Lanzetta, 1989; Krumhuber & Kappas, 2005; Krumhuber & Manstead, 2009), substantial advances have been made towards establishing the physical cues that differentiate between posed and spontaneous expressions. Recent work has examined the acoustic cues associated with posed and spontaneous speech segments (Jürgens, Hammerschmidt & Fischer, 2011; Jürgens et al., 2015), but for nonverbal vocalisations these features have not yet been mapped out, though recent work has described the acoustic cues differentiating posed from spontaneous laughter (Bryant & Aktipis, 2014; Lavan, Scott, & McGettigan, 2016). It will also be important to establish which cues perceivers rely on when making judgments

about authenticity. It will thus be a worthwhile task to extend such studies further from facial configurations to the vocal domain.

Spontaneous expressions across cultures

Previous studies have found that emotions can be recognised from spontaneous expressions across cultural boundaries (e.g. Matsumoto, Olide, Schug, Willingham, & Callan, 2009a). Hopefully, future comparisons of posed and spontaneous expressions can incorporate cross-cultural comparisons to test whether cross-cultural differences in recognition occurs for posed but not spontaneous expressions (see Matsumoto, Olide, & Willingham, 2009b). Relating to the point of cross-cultural consistency, it is worth noting that the current set of results suggests that the relationship between prototypicality and spontaneity may be rather complex, questioning the notion that posed stimuli are overall disproportionately prototypical (e.g. Scherer, 2003). In the current study, in Experiment 2, no differences in perceived prototypicality were found between spontaneous and posed stimuli for Visual presentation, and differences were found in opposite directions for Audio and AudioVisual presentations (see also Jürgens et al., 2015). This suggests that the use of posed stimuli in cross-cultural studies does not necessarily introduce a confound in terms of prototypicality (see Russell, 1994). However, more cross-cultural studies of spontaneous expressions are needed to establish whether the nonverbal communication of a wide range of emotional states via spontaneous facial, vocal, and multimodal expressions is universal. Our results point to considerable differences between emotions for recognition accuracy of posed and spontaneous expressions, which may be interesting given the considerable consistency in the specificity of cross-cultural recognition of different emotion categories from posed expressions (see e.g. Cordaro et al., 2016). Further studies of spontaneous expressions that allow for clear inferences regarding individual emotions will thus be particularly informative.

Conclusions

In sum, this study demonstrates that emotions can be recognised from spontaneous expressions from both auditory, visual, and audiovisual cues. Whether recognition accuracy for spontaneous expressions was inferior to that of posed expressions depended on

the posed stimulus set, and so our results provide limited support for the proposal that the use of posed expressions necessarily inflates recognition accuracy relative to spontaneous expressions. Finally, in line with theoretical predictions, our findings suggest a role for intensity and prototypicality in the recognition of spontaneous emotional expressions.

Acknowledgements

The authors would like to thank Paul Barker, Bob Bramson, Laura Poell, Narda Schenk, and Friederike Windel for assistance with finding spontaneous expressions and collecting data, and Marc Heerdink for statistical advice.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The funding is an Innovational Research Incentives Scheme Veni grant to Disa Sauter from The Netherlands Organisation for Scientific Research/the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), [grant number (dossier-nummer) 275-70-033].

References

- Audibert, N., Aubergé, V., & Rilliard, A. (2008). How we are not equally competent for discriminating acted from spontaneous expressive speech. In P. A. Barbosa, S. Madureira, & C. Reis (Eds.), *Proceedings of the Speech Prosody 2008 Conference* (pp. 693–696). Campinas: Editora RG/CNPq.
- Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science*, 338(6111), 1225–1229. doi:10.1126/science.1224313
- Beaupre, M. G., & Hess, U. (2005). Cross-cultural emotion recognition among Canadian ethnic groups. *Journal of Cross-Cultural Psychology*, 35, 355–370. doi:10.1177/0022022104273656
- Bänziger, T., Mortillaro, M., & Scherer, K. R. (2012). Introducing the Geneva multimodal expression corpus for experimental research on emotion perception. *Emotion*, 12(5), 1161–79. doi:10.1037/a0025827
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436. doi:10.1163/156856897X00357
- Bryant, G. A., & Aktipis, C. A. (2014). The animal nature of spontaneous human laughter. *Evolution and Human Behavior*, 35(4), 327–335. doi:10.1016/j.evolhumbehav.2014.03.003
- Cordaro, D. T., Keltner, D., Tshering, S., Wangchuk, D., & Flynn, L. M. (2016). The voice conveys emotion in ten globalized cultures and one remote village in Bhutan. *Emotion*, 16(1), 117–128. doi:10.1037/emo0000100
- Fernandez-Dols, J. M., Carrera, P., & Crivelli, C. (2011). Facial behavior while experiencing sexual excitement. *Journal of Nonverbal Behavior*, 35(1), 63–71. doi:10.1007/s10919-010-0097-7
- Fernández-Dols, J. M., & Crivelli, C. (2013). Emotion and expression: Naturalistic studies. *Emotion Review*, 5(1), 24–29.
- Frank, M. G., & Stennett, J. (2001). The forced-choice paradigm and the perception of facial expressions of emotion. *Journal of Personality and Social Psychology*, 80(1), 75–85. doi:10.1037/0022-3514.80.1.75
- Hess, U., Blairy, S., & Kleck, R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behavior*, 21(4), 241–257. doi:10.1023/A:1024952730333
- Hess, U., Kappas, A., McHugo, G. J., Kleck, R. E., & Lanzetta, J. T. (1989). An analysis of the encoding and decoding of spontaneous and posed smiles: The use of facial electromyography. *Journal of Nonverbal Behavior*, 13(2), 121–137. doi:10.1007/BF00990794
- Hess, U., & Kleck, R. E. (1990). Differentiating emotion elicited and deliberate emotional facial expressions. *European Journal of Social Psychology*, 20(5), 369–385. doi:10.1002/ejsp.2420200502
- Hughes, S. M., & Nicholson, S. E. (2008). Sex differences in the assessment of pain versus sexual pleasure facial expressions. *Journal of Social, Evolutionary, and Cultural Psychology*, 2(4), 289.
- Jürgens, R., Drolet, M., Pirow, R., Scheiner, E., & Fischer, J. (2013). Encoding conditions affect recognition of vocally expressed emotions across cultures. *Frontiers in Psychology*, 4, 111. doi:10.3389/fpsyg.2013.00111
- Jürgens, R., Grass, A., Drolet, M., & Fischer, J. (2015). Effect of acting experience on emotion expression and recognition in voice: Non-actors provide better stimuli than expected. *Journal of Nonverbal Behavior*, 39, 195–214. doi:10.1007/s10919-015-0209-5
- Jürgens, R., Hammerschmidt, K., & Fischer, J. (2011). Authentic and play-acted vocal emotion expressions reveal acoustic differences. *Frontiers in Psychology*, 2, 180. doi:10.3389/fpsyg.2011.00180
- Justlin, P. N., & Laukka, P. (2001). Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion*, 1(4), 381–412. doi:10.1037/1528-3542.1.4.381
- Krumhuber, E., & Kappas, A. (2005). Moving smiles: The role of dynamic components for the perception of the genuineness of smiles. *Journal of Nonverbal Behavior*, 29(1), 3–24. doi:10.919-004-0887
- Krumhuber, E. G., & Manstead, A. S. (2009). Can Duchenne smiles be feigned? New evidence on felt and false smiles. *Emotion*, 9(6), 807–820. doi:10.1037/a0017844
- Laukka, P., Audibert, N., & Aubergé, V. (2012). Exploring the determinants of the graded structure of vocal emotion expressions. *Cognition and Emotion*, 26(4), 710–719. doi:10.1080/02699931.2011.602047
- Lavan, N., Scott, S. K., & McGettigan, C. (2016). Laugh like you mean it: Authenticity modulates acoustic, physiological and perceptual properties of laughter. *Journal of Nonverbal Behavior*, 40(2), 133–149. doi:10.1007/s10919-015-0222-8
- Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: A meta-analysis of experimental emotion elicitations. *Psychological Bulletin*, 137(5), 834–855. doi:10.1037/a0024244
- Matsumoto, D., Olide, A., Schug, J., Willingham, B., & Callan, M. (2009a). Cross-cultural judgments of spontaneous facial

- expressions of emotion. *Journal of Nonverbal Behavior*, 33(4), 213–238. doi:10.1007/s10919-009-0071-4
- Matsumoto, D., Ollide, A., & Willingham, B. (2009b). Is there an ingroup advantage in recognizing spontaneously expressed emotions? *Journal of Nonverbal Behavior*, 33(3), 181–191. doi:10.1007/s10919-009-0068-z
- Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., & Gross, J. J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion*, 5(2), 175–190. doi:10.1037/1528-3542.5.2.175
- Motley, M. T., & Camden, C. T. (1988). Facial expression of emotion: A comparison of posed expressions versus spontaneous expressions in an interpersonal communication setting. *Western Journal of Speech Communication*, 52(1), 1–22. doi:10.1080/10570318809389622
- Nelson, N. L., & Russell, J. A. (2013). Universality revisited. *Emotion Review*, 5(1), 8–15. doi:10.1177/1754073912457227
- Reisenzein, R., Studtmann, M., & Horstmann, G. (2013). Coherence between emotion and facial expression: Evidence from laboratory experiments. *Emotion Review*, 5(1), 16–23. doi:10.1177/1754073912457228
- Ross, E. D., & Pulusu, V. K. (2013). Posed versus spontaneous facial expressions are modulated by opposite cerebral hemispheres. *Cortex*, 49(5), 1280–1291. doi:10.1016/j.cortex.2012.05.002
- 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. (2013). The role of motivation and cultural dialects in the in-group advantage for emotional vocalizations. *Frontiers in Emotion Science*, 4, 814. doi:10.3389/fpsyg.2013.00814/abstract
- Sauter, D. A., & Scott, S. K. (2007). More than one kind of happiness: Can we recognize vocal expressions of different positive states? *Motivation and Emotion*, 31(3), 192–199. doi:10.1007/s11031-007-9065-x
- Scherer, K. R., Clark-Polner, E., & Mortillaro, M. (2011). In the eye of the beholder? Universality and cultural specificity in the expression and perception of emotion. *International Journal of Psychology*, 46(6), 401–435. doi:10.1080/00207594.2011.626049
- Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, 1-2(40), 227–256.
- Tracy, J. L., & Matsumoto, D. (2008). The spontaneous expression of pride and shame: Evidence for biologically innate nonverbal displays. *Proceedings of the National Academy of Sciences of the United States of America*, 105(33), 11655–11660. doi:10.1073/pnas.0802686105
- Wagner, H. L. (1990). The spontaneous facial expression of differential positive and negative emotions. *Motivation and Emotion*, 14(1), 27–43. doi:10.1007/BF00995547
- Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior*, 17(1), 3–28. doi:10.1007/BF00987006
- Witling, J., Krahmer, E., & Swerts, M. (2006). *Real vs. acted emotional speech*. Proceedings of the Annual Conference of the International Speech Communication Association (Interspeech) 2006 (pp. 805–808), Pittsburgh, PA.
- Yik, M., Widen, S. C., & Russell, J. A. (2013). The within-subjects design in the study of facial expressions. *Cognition and Emotion*, 27(6), 1062–1072. doi:10.1080/02699931.2013.763769
- Zuckerman, M., DeFrank, R. S., Hall, J. A., & Rosenthal, R. (1976). Encoding and decoding of spontaneous and posed facial expressions. *Journal of Personality and Social Psychology*, 34(5), 966–977. doi:10.1037/0022-3514.34.5.966