I thought that I heard you laughing: Contextual facial expressions modulate the perception of authentic laughter and crying

Nadine Lavan[1*], César F. Lima[2,3*], Hannah Harvey[1], Sophie K. Scott[3], and Carolyn McGettigan[1,3]

[1]Department of Psychology, Royal Holloway, University of London, U.K.

[2]Centre for Psychology, University of Porto, Portugal

[3]Institute of Cognitive Neuroscience, University College London, U.K.

[*] These authors have made equal contributions to the work

Author Note

César Lima is supported by the Portuguese Foundation for Science and Technology (SFRH/BPD/77189/2011). This work was also sponsored by a Research Strategy Fund award from Royal Holloway, University of London. The auditory stimulus preparation was funded by a Wellcome Trust Senior Research Fellowship (WT090961MA) awarded to Sophie K. Scott. Correspondence to: Carolyn McGettigan, Department of Psychology, Royal Holloway, University of London, Egham Hill, Egham TW20 0EX, UK. Email: Carolyn.McGettigan@rhul.ac.uk

Abstract

It is well established that categorizing the emotional content of facial expressions may differ depending on contextual information. Whether this malleability is observed in the auditory domain and in genuine emotion expressions is poorly explored. We examined the perception of authentic laughter and crying in the context of happy, neutral and sad facial expressions. Participants rated the vocalizations on separate unipolar scales of happiness and sadness, and on arousal. Although they were instructed to focus exclusively on the vocalizations, consistent context effects were found: For both laughter and crying, emotion judgments were shifted towards the information expressed by the face. These modulations were independent of response latencies and were larger for more emotionally ambiguous vocalizations. No effects of context were found for arousal ratings. These findings suggest that the automatic encoding of contextual information during emotion perception generalizes across modalities, to purely nonverbal vocalizations, and is not confined to acted expressions.

Keywords: context; cross-modal biases; emotion perception; nonverbal vocalizations

Introduction

In the emotion literature, emotional expressions are usually presented in isolation, and discrete emotional states can be reliably perceived. In social interactions, however, humans encounter a wide range of signals from different modalities at the same time, such as facial, body, vocal, and situational cues. How different sources of information are combined for social perception has long been investigated (e.g., Noller, 1985), but recently a renewed interest in this topic has been motivated by the debates on the neural bases of multisensory integration and on roles of context and ambiguity during the processing of basic emotions (e.g., Barrett, Mesquita, & Gendron, 2011; de Gelder & Van den Stock, 2012; Hassin, Aviezer, & Beint, 2013). A body of research shows that emotion categorizations of facial expressions is altered by information from the context in which they are embedded, including body postures (Aviezer et al., 2008), speech prosody (i.e., emotionally inflected speech; de Gelder & Vroomen, 2000; Paulmann & Pell, 2010), and nonverbal vocalizations (Sherman, Sweeny, Grabowecky, & Suzuki, 2012). These effects can be observed when participants are instructed to ignore the contextual cues and under increased cognitive load, suggesting that face-context relationships are automatically perceived (Aviezer, Bentin, Dudarev, & Hassin, 2011). Central to the understanding of these findings are the questions of whether they (a) reflect multisensory integration at an early perceptual stage and (b) generalize to genuine vocal expressions, i.e., can be observed for emotional signals beyond acted facial expressions, which informed the majority of previous work.

Context-related modulations of emotion perception have been investigated behaviorally by comparing categorization accuracy and reaction times for emotionally congruent versus incongruent stimulus pairs. Manipulated expressions are frequently used to increase the ambiguity of the stimuli (e.g., through the creation of morphed continua; de Gelder & Vroomen, 2000) while participants typically perform forced-choice categorization tasks, deciding whether a face is e.g., angry or fearful. The finding that the proportion of responses for a given category is biased towards the context in which the expression occurs has been taken as evidence for perceptual integration, i.e., sensory input from different modalities is combined into a unified percept (Barrett et al., 2011; de Gelder & Van den Stock, 2012; Wieser & Brosch, 2012). However, further evidence is needed to exclude the alternative hypothesis that these biases partly originate at a later response-selection stage: In forced-choice tasks, the emotions conveyed by the context and the response options overlap, such that incongruent pairings result in competition/conflict between the *correct* and *context* response categories. Participants are required to choose between mutually exclusive categories and may therefore tend to use the context response because they are unsure about the target, and not necessarily because the context modulated perceptual processes (for discussions on this issue, see Hietanen, Leppaenen, Illi, & Surakka, 2004; Vroomen, Driver & de Gelder, 2001).

Our knowledge on whether and how contextual information affects the processing of basic emotional expressions beyond the visual domain is fragmentary. There is evidence that categorizations of semantically neutral sentences produced in a fearful or happy tone of voice vary depending on concurrent facial (de Gelder & Vroomen, 2000) and body expressions (Van de Stock, Righart, & de Gelder, 2007). Yet, in addition to speech prosody, auditory emotions include nonverbal vocalizations, such as laughter and crying, which are far less well understood. Nonverbal vocalizations are the most phylogenetically continuous of all expressive channels, and differ from speech prosody in other important ways: they are not constrained by articulatory movements related to speech production, and involve distinct production mechanisms, resulting in distinct acoustic properties (Scott, Sauter, & McGettigan, 2010). Laughter and crying are deeply grounded in human biology – they emerge early in development, do not require auditory experience (Scheiner, Hammerschmidt, Juergens, & Zwirner, 2006), and are cross-culturally recognized (Sauter, Eisner, Ekman, & Scott, 2010). Importantly, apart from rare exceptions (e.g., Aviezer, Trope, & Todorov, 2012), the literature on context effects is based on acted displays, for which actors pose expressions without feeling the corresponding emotions. While this type of expressions is informative, a growing number of studies highlights the importance of investigating more naturalistic, genuine emotional expressions. For instance, posed and genuine laughter are perceptually and acoustically distinct (Bryant & Aktipis, 2014), and they engage different neurocognitive systems (McGettigan et al., 2013). Whether previously described context effects generalize to genuine nonverbal vocalizations remains an open question. Showing that contextual information is integrated into the processing of such basic expressions would provide unique evidence for the argument that context encoding may be an in-built feature of emotion perception (Aviezer et al., 2012; Barrett et al., 2011).

In the current study, we examined for the first time whether the perception of genuine laughter and crying is modulated by simultaneously presented facial expressions. To tease apart integration from response selection processes, we used two strategies. First, the competition/conflict between response options was minimized by using separate unipolar rating scales for each emotion category. Participants were thereby given the option to freely provide lower ratings (including the option no emotion at all; Frank & Stennett, 2001) in cases when they perceived the target emotion as being less clear/intense, instead of being forced to choose between a correct and a context response option. Crucially, neither the auditory nor the facial stimuli were altered to artificially inflate their ambiguity (i.e., no morphing techniques were used, e.g., de Gelder & Vroomen, 2000; and no visual or auditory noise was added, e.g., Collignon et al. 2008). Second, and in addition to the emotion judgments, participants judged vocalizations for a general affective feature, arousal, where different levels of potential competition/conflict are also present (high arousal auditory targets paired with high arousal facial expressions versus auditory targets paired with low arousal neutral faces) but cross-modal biases may be less pronounced. It was hypothesized that context may be more important for emotion perception than for arousal inferences, on the basis of evidence from facial expression perception showing that context is more frequently remembered when participants categorize specific emotions than when they just judge the affective significance of the stimuli (Barrett & Kensinger, 2010).

Laughter and crying were paired with happy, neutral or sad facial expressions, resulting in three audio-visual conditions that formed a gradient of congruency. The vocalizations were judged regarding how much they expressed two emotions, happiness and sadness, and in terms of the more general affective feature of arousal. A dissociation between emotion and arousal judgments was predicted: For emotion judgments, we predicted that ratings on the target emotion would be highest for congruent pairings (laughter paired with a happy face, crying with a sad face), intermediate for neutral pairings (laughter/crying with a neutral face), and lowest for incongruent pairings (laughter paired with a sad face, crying with a happy face). For arousal judgments, no context modulations were expected between congruent (i.e., high arousal vocalization and face) and neutral pairings (i.e., high arousal vocalization and low arousal face) despite the difference in the degree of congruence. If previous evidence on morphed facial expressions (de Gelder & Vroomen, 2000) extends to a more naturalistic variability in ambiguity, the magnitude of context-related modulations should be larger for more ambiguous vocalizations.

Methods

Participants

Thirty-eight healthy participants aged between 18 and 45 years (28 female; M_{age} = 22.15 years, SD = 5.49) were recruited via the Royal Holloway Experiment Management System. They had normal or corrected-to-normal vision, normal hearing, and were tested in individual sessions lasting around 1 hour. Ethical approval was obtained from the Departmental Ethics Committee at the Department of Psychology, Royal Holloway, University of London.

Materials and Methods

Genuine amusement laughter and crying stimuli, produced by 5 adults (3 female) were recorded in a sound-proof anechoic chamber at University College London (Lima et al., in preparation). Laughter was elicited as described in McGettigan et al. (2013) – speakers were shown video clips, which they identified beforehand as being amusing. For crying, they recalled upsetting events and/or initially posed crying to encourage a transition into genuine crying associated with genuine felt sadness. The speakers reported genuine feelings of amusement and sadness during and after recording the corresponding expressions. In a series of behavioral experiments, it was shown that participants reliably rate genuine laughter and crying as more authentic than their posed counterparts (Lima et al., in preparation; McGettigan et al., 2013). All vocalizations were extracted from the recordings and normalized for RMS amplitude using Praat (http:// www.praat.org/).

We ran two pilot studies on a larger set of auditory stimuli (56 crying and 60 laughter tokens) prior to the main experiment. In one of these studies ($N = 13$) we collected arousal and valence ratings (7-point scale), and in the other one ($N = 8$) emotional categorization ratings (7-point scale, *Is this crying or laughter?* with 1 = *crying* and 7 = *laughter*). None of the participants of the pilot studies took part in the main experiment. The final set of auditory stimuli included 30 laughs and 30 crying sounds. The two emotions were matched for duration (laughter, $M = 2{,}520$ ms, $SD = 0.24$; crying, $M = 2{,}640$ ms, $SD = 0.35$; $F[1,29] = 1.871$, $p = .182$) and for the number of stimuli from each speaker (6 per speaker per vocalization). The best possible match for arousal was ensured, although laughter was still rated significantly higher on this feature than crying (laughter, $M = 4.84$, $SD = 0.77$; crying: $M = 3.68$, $SD = 0.48$; $F[1,29] = 49.782$, $p < .001$). A range of emotional ambiguity was included based on the 7-point categorization ratings (laughter, $M = 5.85$, $SD = 0.86$, range = 4.12 – 6.88; crying, $M = 2.2$, $SD = 0.69$, range = 1.31 – 3.94). The visual stimuli were static images of happy, sad and neutral facial expressions (30 per emotion), selected from the Karolinska Directed Emotional Faces (Lundqvist, Flykt, & Öhman, 1998). Happy and sad faces were matched for category recognisability (happy faces, M = 93.53% correct, SD = 5.39; sad faces, M = 90.77%, SD = 5.28; $F[1,29] = 1.99$, $p = .169$) and arousal (happy faces, M = 3.53, SD = 0.44 on a scale from 1-9; sad faces, M = 3.49, SD = 0.43; $F[1,29] = 0.111$, $p = .741$). Neutral faces had lower overall recognisability (M = 82.68%, SD = 5.14; compared to happy faces, $F[1,29] = 35.249$, $p < .001$; compared to sad faces, $F[1,29] = 19.262$, $p < .001$) and arousal (M = 1.55, SD = 0.27; compared to happy faces, $F[1,29] = 306.14$, $p < .001$; compared to sad faces, $F[1,29] = 393.25$, $p < .001$). In a follow-up study (N = 10), we observed that laughter and crying are rated similarly when paired with neutral faces and with a checkerboard, suggesting that neutral faces are indeed perceived as a neutral pairing (happy scale: laughter, $F[1,29] = 0.223$, $p = .64$; crying, $F[1,29]< 0.001$, $p = .986$; sad scale: laughter, $F[1,29] = 0.366$, $p = .55$; crying, $F[1,29] = 0.77$, $p = .783$; arousal scale: laughter, $F[1,29] = 0.438$, $p = .518$; crying, $F[1,29] = 1.732$, $p = .199$).

The auditory stimuli were presented via headphones, in MATLAB (Mathworks, Inc., Natick, MA) using the Psychophysics Toolbox extension (http://psychtoolbox.org/). For all trials included in the main experiment, a vocalization was paired with a picture of a face of the same sex in three conditions: same emotion in both domains (congruent), neutral face, different emotions in both domains (incongruent). Audio-visual combinations were randomized and repeated once in each of the three blocks of the experiment (order of blocks randomized across participants): (1) *How happy is this person?*; (2) *How sad is this person?*; and (3) and *How aroused is the person?*. Participants used 7-point Likert scales to respond, from 1 (*not at all*) to 7 (*extremely*), using a key press. Note that happiness and sadness were rated separately on unipolar scales, in which lower ratings indicated absence of those emotions. We thus avoided the limitations of forced-choice formats, in which participants are always forced to select one of the manipulated categories (Frank & Stennett, 2001). For arousal, 1 was defined as *The person is drowsy or sleepy* and 7 as *The person is feeling very alert and energetic*. The face was visible for the duration of the sound. All the analyses reported and discussed here are based on instances in which participants were instructed to focus their attention on the vocalizations only; however, to ensure that they perceived the faces as well, “distractor” trials were included, in which participants were asked to evaluate the face and not the vocalization. A coloured frame appeared around the face 1.5 seconds after the trial onset, so participants were required to attend to both the visual and auditory modality from the outset of each trial: With the cue of the frame, participants were instructed to exclusively judge the vocalization when the frame was red (experimental trials), and the face when the frame was blue (distractor trials). Distractor trials were not analysed and were independent from the experimental trials. Additionally, response latencies were recorded, measuring the delay from the appearance of the rating scale that coincided with the offset of the sounds until the button press. The experiment included 540 trials in total, comprising 60 laughter/crying stimuli x 3 context conditions x 3 judgments, plus 30 distractor trials per block. Participants were instructed to respond as quickly as possible, and completed a short practice session.

Results

The average ratings provided for laughter and crying are depicted in Figure 1, as a function of Judgement Type (happiness, sadness, arousal) and context condition (same emotion, neutral, different emotion). As expected, ratings were higher for laughter on the happy scale, and for crying on the sad scale. A 3 (Context: same emotion, neutral, different emotion) x 3 (Judgement Type: happiness, sadness, arousal) x 2 (Vocalization: laughter, crying) repeated-measures ANOVA was conducted on the raw ratings. For ease of interpretation, the directionality of context effects was made consistent across happiness and sadness judgments by inverting sadness ratings for the laughter stimuli (8 minus each rating), and inverting happiness ratings for the crying stimuli. That is, the expected pattern of ratings for laughter was *same emotion* > *neutral* > *different emotion* on the happiness scale, and *same emotion* < *neutral* < *different emotion* on the sadness scale; inverting sadness ratings allowed us to make the direction of the expected effects similar across scales as *same emotion* > *neutral* > *different emotion*). Context had a large effect on the ratings (F[2,74] = 14.15, p < .001, η_p^2 = .28), which was similarly significant for laughter and crying (interaction Context x Vocalization ns, F[2,74] = 0.32, p = .72, η_p^2 = .01), but depended on the Judgement Type, as predicted (interaction Context x Judgement Type, F[4,148] = 4.714, p = .001, η_p^2 = .11; main effect of Judgement Type, F[2,74] = 35.397, p < .001, η_p^2 = .49; interaction Judgement Type x Vocalization, F(2,74) = 36.62, p < .001, η_p^2 = .5; interaction Context x Judgement Type x Vocalization ns, F[4,148] = 1.157, p = .33, η_p^2 = .03). Planned contrasts were computed using a Bonferroni-corrected significance level of p = .008 (6 comparisons). For the emotion-specific judgments, happy and sad, a linear trend contrast confirmed the predicted pattern of same emotion > neutral > different emotion, F[1,37] = 21.435, p < .001 (same emotion > neutral, F[1,37] = 15.185, p < .001; neutral > different emotion, F[1,37] = 7.108, p = .011). For the arousal judgments, the ratings did not significantly differ across context conditions: same emotion (i.e., high arousal) > neutral (i.e., low arousal), F[1,37] = 0.748, p = .39 (different emotion > neutral, F[1,37] = 1.267, p = .27; same vs. different emotions, F[1,37] = 0.031, p = .86). Additionally, the fact that the three-way interaction was not significant indicates that cross-modal biases had a similar magnitude when participants were judging the vocalizations directly on the relevant target emotion (i.e., happy scale for laughter, sad scale for crying) and when they were judging the vocalizations on a different target emotion (i.e., sad scale for laughter, happy scale for crying).

To directly compare the magnitude of the context effects for emotion-specific and arousal judgements, an additional repeated-measures ANOVA was conducted on the difference between same emotion and neutral pairings for emotion-specific (averaged across laughter and crying for the two emotion scales) and arousal judgements (averaged across laughter and crying)[1]. As expected, context had a significantly larger effect for emotion-specific judgements (*M* = 0.107*;* 95% CI [0.051, 0.163]) as compared to the general affective judgment of arousal (*M* = -0.028*;* 95% CI [-0.09, 0.035]), $F[1,37] = 10.703$, $p = .002$, $\eta_p^2 = .22$). Thus, contextual facial expressions biased ratings more when the goal was to perceive emotions than when the goal was to make a dimensional arousal judgment.

—- Figure 1 here —-

To examine possible links between ratings and differences in response latencies, a 3 (Context: same emotion, neutral, different emotion) x 3 (Judgement Type: happiness, sadness, arousal) x 2 (Vocalization: laughter, crying) repeated-measures ANOVA was conducted. On average, participants took 807 ms to respond (95% CI [731, 882]) after the offset of the sound. Responses were quicker for laughter (759 ms; 95% CI [688, 829]) than crying (855 ms, 95% CI [772, 937]; main effect of Vocalization, $F[1,37] = 52.062$, $p < .001$, $\eta_p^2 = .61$), but did not differ across Judgement Type and context conditions, indicating that context-related modulations of the ratings cannot be explained by differences in latencies (main effect of Judgement Type, $F(2,74) = 2.705$, $p = .07$, $\eta_p^2 = .07$; main effect of context, $F(2,74) = 1.181$, $p = .31$, $\eta_p^2 = .04$; interaction Judgement Type x Context, $F(4,148) = 1.398$, $p = .238$, $\eta_p^2 = .04$; interaction Judgement Type x Vocalization, $F(2,74) = 0.022$, $p = .98$, $\eta_p^2 = .00$; interaction Context x Vocalization, $F(2,74) = 0.134$, $p = .875$, $\eta_p^2 = .00$; interaction Judgement Type x Context x Vocalization, $F(4,148) = 1.134$, $p = .343$, $\eta_p^2 = .03$).

[1] Pairings of different emotions were not included in this analysis because they may not be directly comparable across emotion and arousal judgments – even though the expressed emotion differs between target and context (e.g., laughter paired with sad face), this may not cause incongruence in the case of arousal, as they both have high arousal.

We also predicted that context effects on happiness and sadness judgments would be larger for ambiguous vocalizations, i.e., vocalizations rated farther from the extremes of the scale *Is this crying or laughter?* used in a pilot study. A regression analysis was computed, with perceived ambiguity as a predictor and the magnitude of the context effect as the dependent variable (i.e., ratings for emotionally congruent pairs – ratings for incongruent pairs). The quadratic association was significant, $R^2 = .18$, $F(2, 57) = 6.07$, $p = .004$, showing that biases were larger for vocalizations rated closer to the scale mid-point (more ambiguous), and smaller for the sounds rated closer to the scale extremes (less ambiguous; see Figure 2).

—- Figure 2 here —-

Discussion

This study forms a novel demonstration of the automatic encoding of context during the perception of genuine emotional expressions in the auditory domain. Our results show that facial expressions significantly shift emotion judgments of authentic laughter and crying, even when participants are instructed to focus exclusively on the vocalizations. Previous research, predominantly using forced-choice tasks and acted stimuli, has reported consistent context effects for (mostly negative) facial emotions, but knowledge on whether these effects extend to different types of auditory expressions was relatively limited (but see, e.g., Collignon et al., 2008; de Gelder & Vroomen, 2000). We expand on the existing literature by showing that contextual information is integrated during the perception of evolutionarily ancient and naturalistic auditory expressions. Additionally, while forced-choice tasks may promote response biases due to competition/conflict between the correct and context response categories, this was minimized in the current study by using a rating task, in which participants could freely indicate the degree to which emotions were expressed in the stimuli (including *not at all*). Context effects in the ratings were independent of response latencies, which further suggests that competition/conflict at a response selection stage was indeed minimized, and that our task was probably tapping mostly into integration processes: Studies using forced-choice tasks typically find slower responses for incongruent pairings (e.g. de Gelder & Vroomen, 2000), which is likely to reflect increased task demands due to conflict.

The context effects were selective to the perception of emotion categories (happiness and sadness), and did not extend to the general affect judgment of arousal. If context were encoded during arousal judgments, ratings would have been higher for pairings that were congruent in arousal (e.g., laughter paired with happy face) than for neutral pairings (e.g., high-arousal laughter paired with a neutral face), as observed in the case of emotion-specific judgements. According to pilot data, crying stimuli were lower in arousal than laughter stimuli. However, this difference, if anything, would predict larger context effects for laughter compared to crying (because the distance in arousal between laughter and neutral faces would be larger), and we found no context-related modulations of arousal ratings for either vocalization. Directly comparing the size of the context effect for emotion-specific judgements and arousal judgements provided further evidence for the selective encoding of context. This result supports the argument that context may be particularly important for making inferences about specific emotional states. While isolated expressions may provide sufficient information to infer a general arousal state, the integration of contextual information may be routinely required to optimize the understanding of the specific emotional states (Barrett & Kensinger, 2010). Previous evidence for this comes from a memory task in the visual domain (ibd.), which we extend here to the auditory domain and to an emotion perception task.

In line with the emotion-specific findings, context effects were larger for more ambiguous vocalizations, indicating that participants integrated more contextual information when the cues to the target emotion were less clear. De Gelder and colleagues (de Gelder &

Vroomen, 2000; Van de Stock et al., 2007) showed analogous effects for stimuli morphed from fear to happiness, and Collignon and colleagues (2008) found that adding white noise to their stimuli increased the impact of context. We demonstrate that such modulations occur in unmodified nonverbal vocalizations. The similarities in the production mechanisms and acoustic profile may explain why even naturalistic laughter and crying afford a range of ambiguity: In contrast to other emotional vocalizations, which are based on a small number of bursts, both laughter and crying are characterized by a succession of high-pitched vocalic bursts that can last for several seconds as a result of spasms of the abdominal and intercostal muscles (Lloyd, 1938).

The findings of the current study are based on dynamic auditory stimuli paired with static facial expressions taken from an independent stimulus set. While we acknowledge that this aspect of the design constrains the ecological validity of the pairs (e.g., Noller, 1985), it results from the technical challenge of synchronizing the auditory bursts (characteristic of laughter and crying) with the corresponding dynamic visual information so that incongruent pairs do not appear less realistic than congruent ones. Using dynamic visual information would additionally be difficult the neutral condition, which is an informative baseline measure for the interpretation of the direction of (in)congruency effects. Nevertheless, future studies using more ecological multimodal materials will certainly contribute to a better understanding of cross-modal interactions. Other interesting areas of further investigation include examining whether context effects in vocalizations show other features of automaticity, such as exerting an influence even under condition of cognitive load (Aviezer et al., 2011); and whether the pattern of results is modulated as a function of the stimulus onset asynchrony, i.e., whether presenting the auditory and visual stimuli at different onset times (e.g., presenting the faces before the vocalizations) would lead to context effects distinct from the ones obtained here with simultaneous presentation. This matters, considering for instance the findings of prosody-face interactions showing differentiated electrophysiological responses for facial expressions depending on the duration of previous exposure to emotional prosody, 200 ms or 400 ms (Paulmann & Pell, 2010).

In sum, we showed that contextual information (facial expressions) is automatically encoded and integrated into the perception of basic emotions in genuine auditory expressions: laughter and crying. These findings suggest that even primitive and authentic auditory emotional expressions may be inherently ambiguous to a certain extent, and that the available contextual cues are routinely used for inferences about emotional states.

References

Aviezer, H., Bentin, S., Dudarev, V., & Hassin, R. R. (2011). The automaticity of emotional face-context integration. *Emotion, 11,* 1406-1414. doi:10.1037/a0023578

Aviezer, H., Hassin R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., …, & Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science, 19,* 724-732. doi:10.1111/j.1467-9280.2008.02148.x

Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science, 338,* 1225-1229. doi:10.1126/science.1224313

Barrett, L. F., & Kensinger, E. A. (2010). Context is routinely encoded during emotion perception. *Psychological Science, 21,* 595-599. doi:10.1177/0956797610363547

Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science, 20,* 286-290.

Bryant, G. A., & Aktipis, C. A. (2014). The animal nature of spontaneous human laughter. *Evolution and Human Behavior, 35*, 327-335. doi:10.1016/j.evolhumbehav.2014.03.003

Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Research, 1242,* 126-135. doi:10.1016/j.brainres.2008.04.023

de Gelder, B., & Van den Stock, J. (2012). Real faces, real emotions: Perceiving facial expressions in naturalistic contexts of voices, bodies, and scenes. In A. J. Calder, G. Rhodes, Mark H. Johnson & J. V. Haxby (Eds.), *The Oxford Handbook of Face Perception* (pp. 535-550). New York: Oxford University Press.

de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition & Emotion, 14,* 289-311. doi:10.1080/026999300378824.

Frank, M. F., & Stennett, J. (2001). The forced-choice paradigm and the perception of facial expressions of emotion. *Journal of Personality and Social Psychology, 80,* 75-85. doi:10.1037/0022-3514.80.1.75

Hassin, R. R., Aviezer, H., & Bentin, S. (2013). Inherently ambiguous: Facial expressions of emotions, in context. *Emotion Review, 5,* 60-65. doi:10.1177/1754073912451331

Hietanen, J. K., Leppeanen, J. M., Illi, M., & Surakka, V. (2004). Evidence for the integration of audiovisual emotional information at the perceptual level of processing. *European Journal of Cognitive Psychology, 16,* 769-790. doi:10.1080/09541440340000330.

Lima, C. F., Lavan, N., Evans, S., Chen, S., Boebinger, D., & Scott, S.K. (in preparation). Acoustic correlates of authenticity in laughter and weeping.

Lloyd, E. L. (1938). The respiratory mechanism in laughter. *Journal of General Psychology, 18*, 179-189. Retrieved from http://search.proquest.com/docview/1290511391?accountid=11455.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces – KDEF [CD]. Department of Clinical Neuroscience, Psychology section, Karolinska Institutet. ISBN 91-630-7164-9.

McGettigan, C., Walsh, E., Jessop, R., Agnew, Z. K., Sauter, D. A., Warren, J. E., & Scott, S. K. (2013). Individual differences in laughter perception reveal roles for mentalizing and sensorimotor systems in the evaluation of emotional authenticity. *Cerebral Cortex.* Advance online publication. doi:10.1093/cercor/bht227

Noller, P. (1985). Video primacy – a further look. *Journal of Nonverbal Behavior, 9,* 28-47. doi:10.1007/BF00987557

Paulmann, S., & Pell, M. D. (2010). Contextual influences of emotional speech prosody on face processing: How much is enough? *Cognitive, Affective, & Behavioral Neuroscience, 10,* 230-242. doi:10.3758/CABN.10.2.230

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

Scheiner, E., Hammerschmidt, K., Juergens, U., & Zwirner, P. (2006). Vocal expression of emotions in normally hearing and hearing-impaired infants. *Journal of Voice, 20,* 585-604. doi:10.1016/j.jvoice.2005.09.001

Scott, S. K., Sauter, D., & McGettigan, C. (2010). Brain mechanisms for processing perceived emotional vocalizations in humans. In M. B., Stefan (Ed.), *Handbook of Behavioral Neuroscience* (pp. 187-197). London, UK: Academic Press.

Sherman, A., Sweeny, T. D., Grabowecky, M., & Suzuki, S. (2012). Laughter exaggerates happy and sad faced depending on visual context. *Psychonomic Bulletin & Review, 19,* 163-169. doi:10.3758/s13423-011-0198-2

Van de Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion, 7,* 487-494. doi:10.1037/1528-3542.7.3.487

Vroomen, J., Driver, J., & de Gelder, B. (2001). Is cross-modal integration of emotional expressions independent of attentional resources? *Cognitive, Affective, & Behavioral Neuroscience, 1,* 382-387. doi:10.3758/CABN.1.4.382

Wieser, M., & Brosch, T. (2012). Faces in context: A review and systematization of contextual influences on affective face processing. *Frontiers in Psychology, 3,* article 471. doi:10.3389/fpsyg.2012.00471

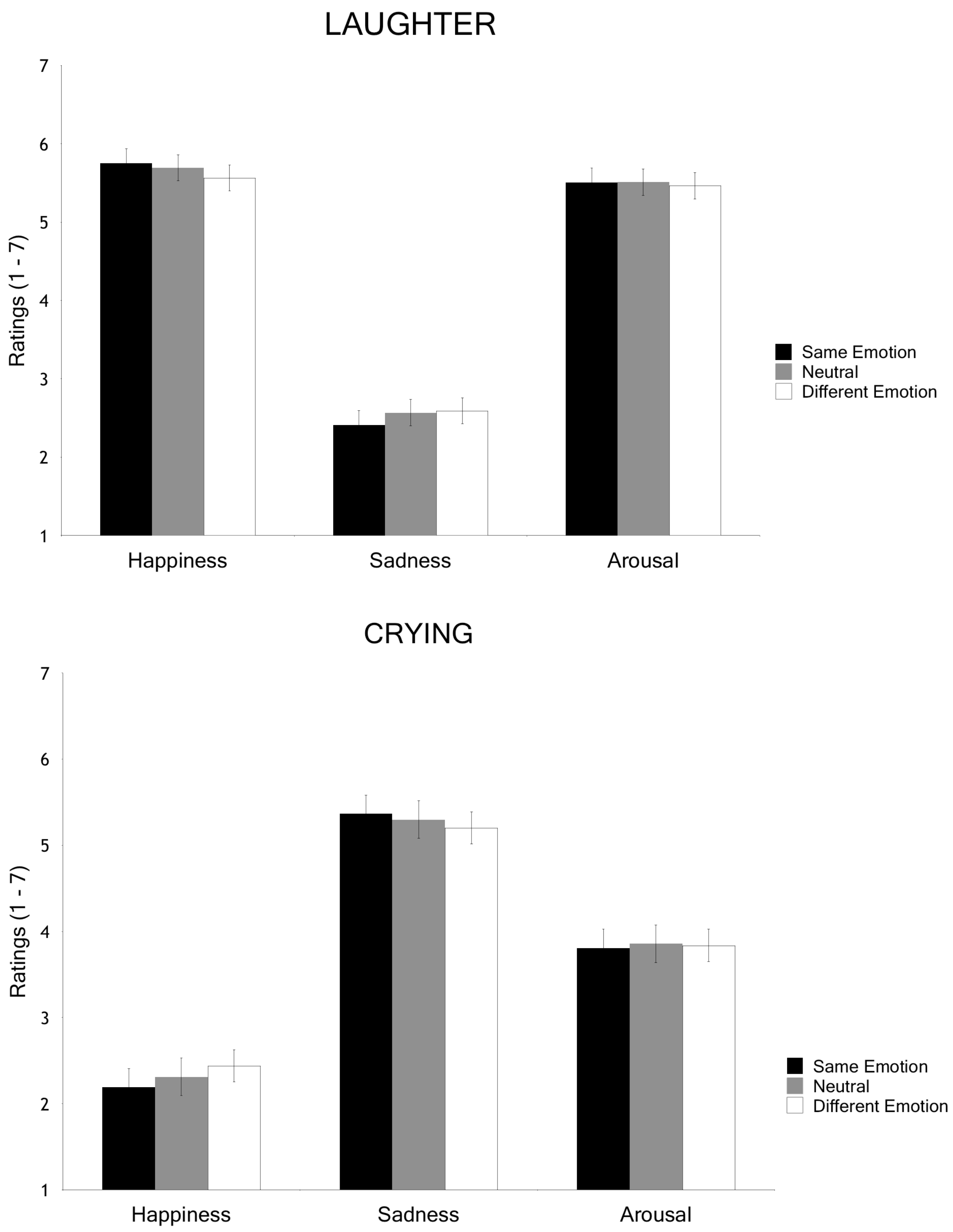

Figure 1. Average ratings (1 = *not at all; 7 = extremely*) for laughter and crying as a function of Judgement Type and context condition. Error bars indicate 95% confidence intervals.

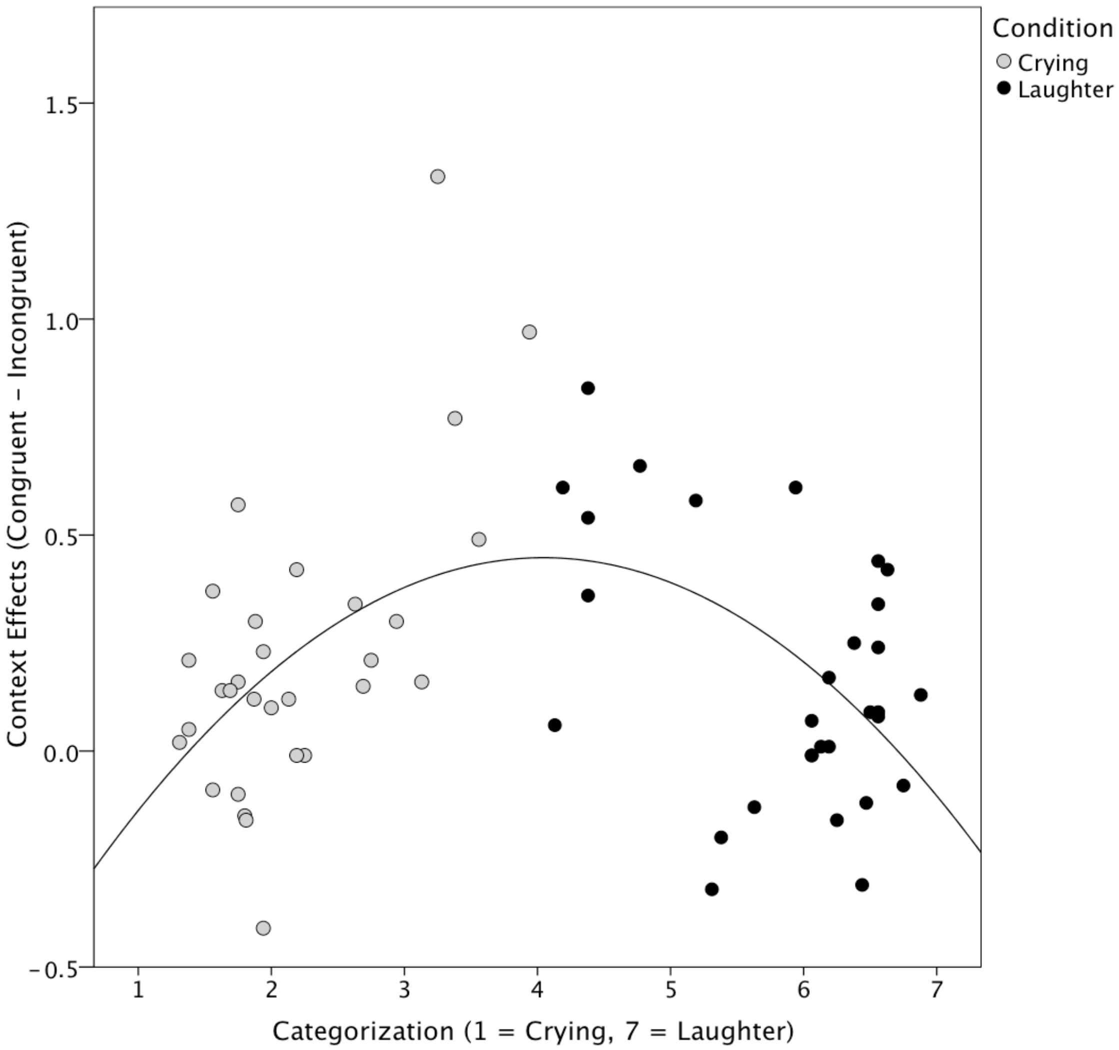

Figure 2. Magnitude of context effects (ratings for emotionally congruent pairs – ratings for incongruent pairs) as a function of laughter and crying ambiguity.