Journal of Experimental Psychology: General

0096-3445/19/$12.00

2019, Vol. 148, No. 10, 1842–1848
http://dx.doi.org/10.1037/xge0000535

BRIEF REPORT

Loud and Unclear: Intense Real-Life Vocalizations During Affective Situations Are Perceptually Ambiguous and Contextually Malleable

Doron Atias
Hebrew University of Jerusalem

Alexander Todorov
Princeton University

Shai Liraz, Arella Eidinger, Ilan Dror, and Yaacov Maymon
The Israel State Lottery—Mifal Hapais, Tel Aviv, Israel

Hillel Aviezer
Hebrew University of Jerusalem

A basic premise of emotion theories is that experienced feelings (whether specific emotions or broad valence) are expressed via vocalizations in a veridical and clear manner. By contrast, functional–contextual frameworks, rooted in animal communication research, view vocalizations as contextually flexible tools for social influence, not as expressions of emotion. Testing these theories has proved difficult because past research relied heavily on posed sounds which may lack ecological validity. Here, we test these theories by examining the perception of human affective vocalizations evoked during highly intense, real-life emotional situations. In Experiment 1a, we show that highly intense vocalizations of opposite valence (e.g., joyous reunions, fearful encounters) are perceptually confusable and their ambiguity increases with higher intensity. In Experiment 1b, we use authentic lottery winning reactions and show that increased hedonic intensity leads to lower, not higher valence. In Experiment 2, we demonstrate that visual context operates as a powerful mechanism for disambiguating real-life vocalizations, shifting perceived valence categorically. These results suggest affective vocalizations may be inherently ambiguous, demonstrate the role of intensity in driving affective ambiguity, and suggest a critical role for context in vocalization perception. Together, these findings challenge both basic emotion and dimensional theories of emotion expression and are better in line with a functional–contextual account which is externalist and by definition, context dependent.

Keywords: emotional vocalizations, context effects, valence perception, intense emotions, real-life

Supplemental materials: http://dx.doi.org/10.1037/xge0000535.supp

From sobs of grief to cheers of joy, emotional vocalizations are central in human experience, but researchers debate their nature. According to the basic emotion account, emotional vocalizations express specific emotions that are recognized universally (Sauter, Eisner, Ekman, & Scott, 2010, 2015), rapidly (Sauter & Eimer, 2010), and automatically (Lima, Anikin, Monteiro, Scott, & Castro, 2018), especially when they are intense (Tracy, 2014). By contrast, dimensional accounts argue that vocalizations accurately express core affective properties (e.g., valence), but not specific emotions[1] (Gendron, Roberson, & Barrett, 2015; Gendron, Roberson, van der Vyver, & Barrett, 2014; Russell, Bachorowski, & Fernández-Dols, 2003). Finally, functional–contextual accounts argue that vocalizations do not express inner emotion or affect, rather they are generated as a tool for influencing others in a flexible context-dependent manner (Bachorowski & Owren, 2003; Cheney & Seyfarth, 2018; Crivelli & Fridlund, 2018; Fridlund, 1994).

Despite their disagreements, basic-emotion and dimensional frameworks are both intrapsychic, assuming that affect-related meaning (e.g., specific emotions, valence) is encoded by vocalizers and then diagnostically decoded by listeners. Consequently, these theories make strong assumptions on the causal and veridical

This article was published Online First December 27, 2018.
Doron Atias, Department of Psychology, Hebrew University of Jerusalem; Alexander Todorov, Department of Psychology, Princeton University; Shai Liraz, Arella Eidinger, Ilan Dror, and Yaacov Maymon, The Israel State Lottery—Mifal Hapais, Tel Aviv, Israel; Hillel Aviezer, Department of Psychology, Hebrew University of Jerusalem.

Portions of these data were presented at the 2016 Consortium of European Research on Emotion in Leiden, the Netherlands. This work was supported by an Israel Science Foundation grant (259/18) to Hillel Aviezer.

Correspondence concerning this article should be addressed to Hillel Aviezer, Department of Psychology, Hebrew University of Jerusalem, Jerusalem 91905, Israel. E-mail: haviezer@huji.ac.il

[1] Similar to dimensional accounts, a component process model of appraisal (Banse & Scherer, 1996) would argue that the appraisal check of "intrinsic pleasantness" would activate different components of expression that should be highly distinct in positive and negative situations.

link between inner feelings and produced acoustics. By contrast, functional–contextual frameworks are externalist and thus not committed to a clear link between felt emotion and specific reactions (Cheney & Seyfarth, 2018; Fridlund, 2017; Seyfarth & Cheney, 2018; Waller, Whitehouse, & Micheletta, 2017).

Critically, past research relied largely on posed (e.g., Belin, Fillion-Bilodeau, & Gosselin, 2008; Gendron et al., 2014; Lima et al., 2018; Sauter et al., 2010) or lab-induced vocalizations evoked by acting or emotional reexperiencing (Lavan, Lima, Harvey, Scott, & McGettigan, 2015), which likely differ from naturalistic expression (Fernández-Dols & Crivelli, 2013) and inflate recognizability. Although recent work has started to move beyond posed emotional vocalizations (e.g., Anikin & Lima, 2018; Sauter & Fischer, 2018), some limitations remain. For example, recognition rates may be inflated due to reliance on forced choice paradigms (discussed in Gendron et al., 2015; Nelson & Russell, 2013), the extensive focus on select "basic emotions" (which may lack natural confusability), and the inclusion of TV show stimuli which may be staged and nonintense.

Indeed, recent evidence from animal and human expressions hints that the distinction between positive and negative vocalizations may not be clear. For example, the "high hoot," one of the most common Bonobo vocalizations is ambiguous, signaling different meaning depending on context (Cheney & Seyfarth, 2018). Similarly, research with posed human vocalizations has shown that a small subset (13%) of posed fear vocalizations is confused with amusement (Sauter et al., 2010). Yet, while suggestive, these promising hints are inconclusive: Animal and human vocalizations may differ vastly, and caution must be taken when interpreting a minority of posed vocalizations which display confusability.

One particularly interesting case is that of expressions occurring in highly intense situations. In humans, the valence of real-life facial expressions may turn ambiguous, especially when affect is intense (e.g., Aviezer, Trope, & Todorov, 2012; Israelashvili, Hassin, & Aviezer, 2018). However, research on human vocalizations suggests the opposite is true. Sauter and Fischer (2018) examined the relationship between perceived intensity and recognition accuracy for posed and spontaneous expressions. Their results demonstrate that participants are more accurate in recognizing emotions from spontaneous (but not posed) vocalizations that are perceived as expressing more intense states. However, posed vocalizations in that study had higher intensity than spontaneous vocalizations, suggesting that spontaneous stimuli originating from TV shows may be more controlled and rehearsed. Thus, using real-life intense vocalizations may yield different results.

Considering the literature, we hypothesized that highly intense, real-life human vocalizations may be more ambiguous and contextually malleable than previously assumed. To this end, in Experiment 1a, we tested the perception of valence and intensity of real-life vocalizations occurring in intense positive (e.g., reacting to reunions with loved ones) versus negative (e.g., reacting to an attacker invading one's home) situations. In Experiment 1b, we utilized real-life vocalizations of lottery winners in which the hedonic situational value increased parametrically in an objective manner. In both cases, we predicted that participants would display poor valence perception, with higher confusability associated with increased intensity.

Finally, in Experiment 2, we examined the hypothesis that contextual information may serve an efficient mechanism to strongly impact and shape the perceived valence of real-life vocalizations. Although a recent attempt to contextualize vocalizations with images produced only minimal Stroop-like effects and failed to produce cross-valence shifts (Lavan et al., 2015), the stimuli used in that study were lab-produced (e.g., posing sadness until a feeling pursued) and were arguably not ecological or highly intense. Thus, contextualizing real-life intense stimuli may show different patterns.

In line with current theories of constructed emotions (Barrett, Mesquita, & Gendron, 2011), and in accordance with the pivotal role of context in functional–contextual frameworks (e.g., Fridlund, 2017; Seyfarth & Cheney, 2018), we hypothesized that affective contextual visual information would strongly impact the perceived valence of both positive and negative vocalizations but that real-life vocalizations would be more susceptible to context than posed vocalizations.

In contrast to lab produced stimuli, we took a "field study" approach using reactions occurring in uncontrolled environments and thus self-report data on accompanying feelings were unavailable. Thus, when we refer to "positive" and "negative" reactions, they should be viewed as "vocalizations produced in stereotypically positive or negative situations." Furthermore, while we classify vocalizations as positive or negative, the world of emotions is likely mixed even for stereotypically hedonic events such as winning the lottery. Nevertheless, we refer to situations in which the presumed overarching cumulative experience is likely positive or negative. Despite these inherent limitations, real-life stimuli may shed new light on affective behavior.

Experiment 1A: Real-Life Vocalizations in Highly Intense Opposing-Valence Situations

Method

Stimuli. Forty spontaneous vocalizations (20 positive, 20 negative) were obtained from online home videos (www.youtube.com) documenting real-life situations of extreme joy (e.g., when announcing a new baby to grandparents) and extreme fear (e.g., when pranked to believe an attacker is in one's home). Stimuli included brief vocalization "bursts" which portrayed the first reaction of the vocalizer (Scherer, 1994; see Section S1 in the online supplemental materials for a detailed description of the stimuli selection, exclusion, and technical editing procedure). The full set of vocalizations is available upon request from the authors.

Participants. Vocalizations were rated by 39 students (31 female) from the Hebrew University of Jerusalem, $M_{age} = 24$, participating for credit or payment. Sample size (~35–40) was determined based on effect sizes in previous studies with emotional stimuli (e.g., Aviezer et al., 2012; Sauter & Fischer, 2018), and was estimated to be of sufficient power for detecting a medium to large effect size. All experiments were approved by the IRB ethics committees at the Hebrew University (Experiments 1a and 2) and Princeton University (Experiment 1b).

Procedure and design. In each trial, participants heard a vocalization delivered through headphones with fixed amplitude and were asked to rate the valence and intensity of the stimuli. No information was given about the origin of stimuli. Valence was rated on a Likert scale ranging from 1 (*extremely negative*), through 5 (*neutral*), to 9 (*extremely positive*). Intensity[2] was rated

[2] An online replication with Arousal ratings is reported in the supplemental information.

using the Self-Assessment Manikin combined with a 9-point Emotional Intensity scale (Bradley & Lang, 1994). All stimuli were presented in a single block in randomized order using E-prime 3.0 software. The experiment followed a one-factor (Vocalization Category: positive vs. negative) within-subjects design, with mean valence and intensity as dependent variables.

Results

For convenience, in this and the following experiments, valence ratings were transformed (by X – 5) such that positive values represent increasingly positive valence judgments (1 to 4) and negative values represent increasingly negative valence judgments (−1 to −4).

Mean valence ratings of vocalizations from both positive ($M = -2.06$, $SE = .15$) and negative ($M = -1.8$, $SE = .15$) situations were in the negative valence range indicating that participants failed to recognize their objective valence (see Figure 1A). Positive vocalizations were judged as slightly more negative than negative vocalizations, $t(38) = 3.27$, $p < .001$, $d = 0.52$, but this difference did not hold when data were recoded by items, $t(19) = .9$, $p = 36$.

An analysis of the intensity ratings demonstrated that vocalizations occurring in positive situations ($M = 6.58$, $SE = .19$) were perceived as more intense than those occurring in negative situations ($M = 6.08$, $SE = .2$), $t(38) = 6.2$, $p < .0001$, $d = 1.00$. Most importantly, Pearson correlations between participants' mean intensity and valence judgments revealed a strong negative correlation between vocalization intensity and perceived valence, for both positive, $r(37) = -.61$, $p < .0001$, and negative vocalizations, $r(37) = -.58$, $p < .0001$, see Figure 1B and 1C. Similarly, when data were recoded by items, a strong negative correlation between vocalization intensity and perceived valence was found for both positive, $r(18) = -.68$, $p < .0001$, and negative vocalizations, $r(18) = -.82$, $p < .0001$.

Furthermore, we ran an online replication of this experiment with "arousal" instead of "intensity" ratings, which yielded even

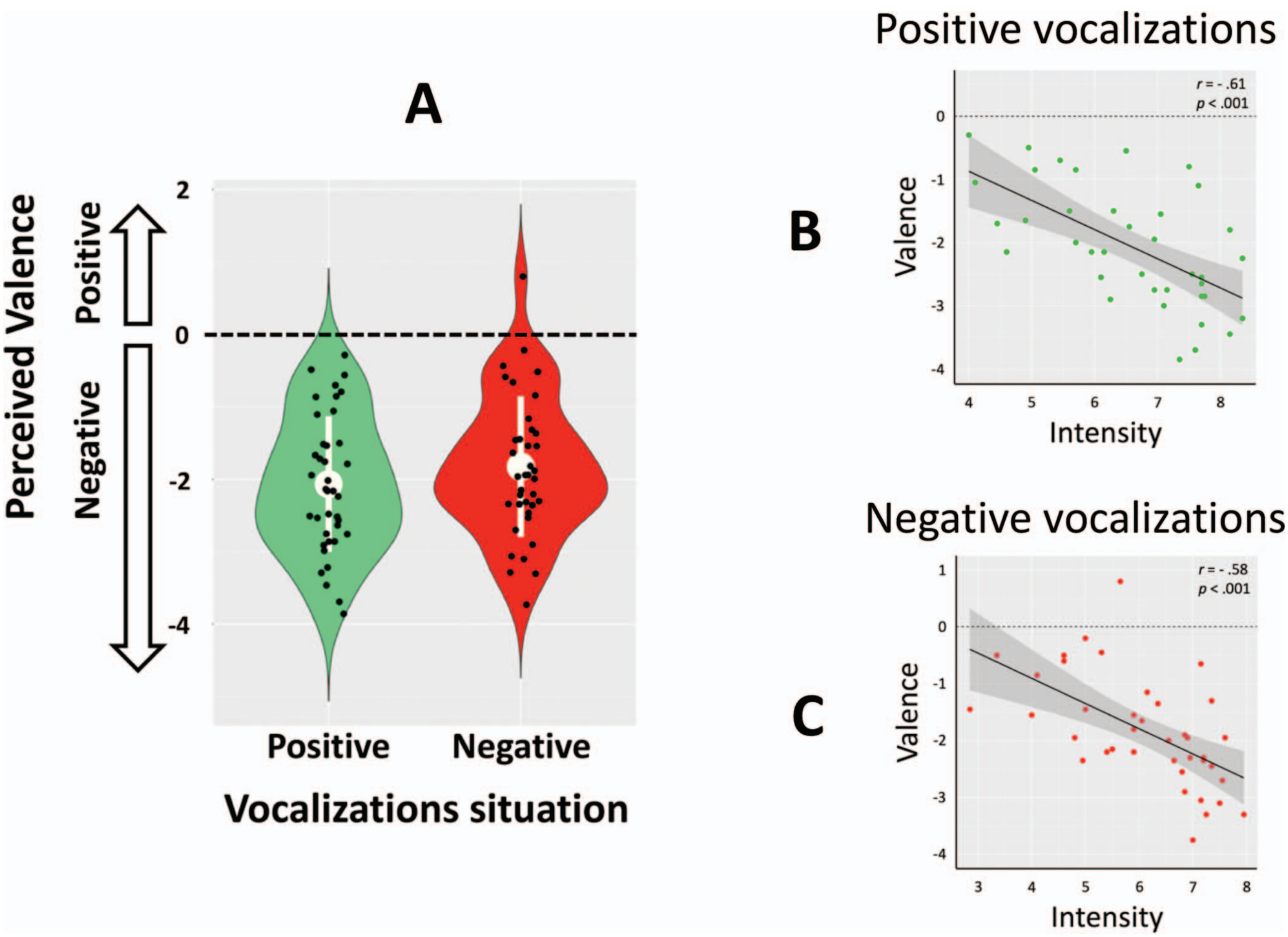

Figure 1. (A) Violin plots of perceived valence ratings of real-life vocalizations evoked in highly positive and negative situations. Despite the situations being highly positive or negative, perceivers failed to correctly classify the valence of the accompanying vocalizations (central white circles and vertical white bars reflect mean and *SD*). Correlation scatterplots between participants' mean intensity and valence ratings of (B) positive and (C) negative vocalizations. See the online article for the color version of this figure.

stronger correlations in the same direction, positive vocalizations, $r(40) = -.85$, $p < .001$, negative vocalizations $r(40) = -.83$, $p < .001$ (see Section S2 in the online supplemental materials). Thus, higher perceived intensity/arousal was negatively correlated with perceived valence and was associated with perceptual ambiguity between positive and negative vocalizations.

Experiment 1B: Vocalizations of Real-Life Lottery Winners With Increasing Prize Sums

Method

Stimuli. Vocalizations of 153 lottery winners were obtained for analysis. Winners were enrolled to the subscription program of the Israeli National Lottery at an annual cost of \$200. In this program, subscribers sign up for a weekly lottery and the winners are called (and recorded) by an official lottery representative (coauthor Arella Eidinger) who notifies them of their win and prize sum. Because the distribution of prize categories was expectedly uneven, we grouped the winners into four groups: Group 1 (~\$15,000, $N = 77$), Group 2 (~\$30,000, $N = 30$), Group 3 (~\$60,000, $N = 25$), and Group 4 (≥~\$125,000, $N = 21$). Prizes in the latter group were comprised of a majority of ~\$125,000 winners ($N = 15$) and a very small number of ~\$435,000 winners ($N = 6$). Excerpts of the winner reactions were later included in a radio advertising campaign of the Israeli National Lottery. All winners used in this study consented to have their recordings posted on the radio website and used in the public radio campaign. For privacy reasons, no identifying personal or demographic information was supplied about the winners.

Stimuli editing. Stimuli were edited to include only the initial spontaneous "bursting" reactions (Scherer, 1994). Vocalizations included both nonverbal bursts as well as brief Hebrew exclamations. Resulting clips ranged from 1 to 4 s long ($M_{\text{duration}} = 1.5$ s; $\text{mode}_{\text{duration}} = 1$ s).

Participants. Because some of the vocalizations included Hebrew exclamations, and because the lottery advertising campaign was well known in Israel, vocalization valence ratings were conducted by 21 non-Hebrew-speaking students from Princeton University. Sample size was estimated based on studies in our lab using posed emotional expressions, indicating that an N of 20–25 was sufficient for detecting medium to large effect sizes, and confirmed by a post hoc power analysis (see Section S3 in the online supplemental materials). Due to a technical computer error, gender was only documented for 18 (13 female) participants and age ($M = 23.3$) was only documented for 13 participants. However, samples in Princeton University are highly homogenous with respect to age.

Procedure, design, and analysis. Participants rated the valence of the stimuli as in Experiment 1a, with each participant rating the entire stimuli set. The experiment followed a one-factor (lottery win sum: \$15,000; \$30,000; \$60,000; ≥\$125,000) within-subjects design with mean valence as the dependent variable.

Results

A one-way repeated-measures analysis of variance revealed a highly significant effect of prize category on vocalization valence, $F(3, 60) = 55.38$, $p < .0001$, $\eta_p^2 = .735$. As demonstrated in Figure 2, the mean perceived valence of the vocalizations was in the positive range for the lower prize categories, but it shifted to the negative range in the higher prize categories.

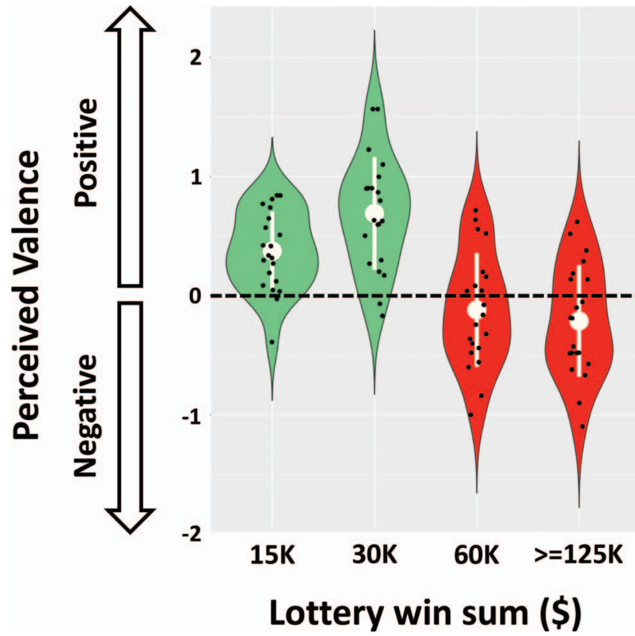

Figure 2. Violin plots of perceived valence ratings of vocalizations produced in response to winning the lottery. While low prizes up to \$30,000 were perceived as conveying positive valence (green [light grey] colored plots), higher prizes (\$60,000 and up), were perceived as conveying negative valence (red [dark gray] colored plots). Central white circles and vertical white bars reflect mean and *SD*. See the online article for the color version of this figure.

Pairwise comparisons revealed that vocal reactions to ~\$30,000 wins ($M = 0.69$, $SE = 0.1$) were rated as more positive than reactions to ~\$15,000 wins ($M = 0.37$, $SE = 0.07$), $t(20) = 3.8$, $p < .001$. However, vocal reactions to ~\$60,000 wins ($M = -0.12$, $SE = 0.1$) were rated as more negative than reactions to both ~\$30,000 wins $t(20) = 9.4$, $p < .001$ and ~\$15,000 wins, $t(20) = 7.57$, $p < .001$. Vocal reactions to ≥~\$125,000 wins ($M = -0.21$, $SE = 0.1$) were rated as more negative than both ~\$15,000, $t(20) = 9.37$, $p < .001$, and ~\$30,000, $t(20) = 8.9$, $p < .001$. The ~\$60,000 and ≥ ~\$125,000 winning reactions did not differ significantly, $p > .1$.

Experiment 2: Real-Life and Posed Vocalizations in Dynamic Context

Method

Stimuli. Fifty-eight vocalizations (29 positive, 29 negative) were embedded in 116 short home-video clips documenting real-life situations of extreme joy and extreme fear, as in Experiment 1a. The vocalizations set consisted of the 40 spontaneous vocalizations used in Experiment 1a (20 positive, 20 negative) and 18 additional posed vocalizations (nine positive, nine negative) obtained from standardized sets of emotional vocalizations (Belin et al., 2008; Cordaro, Keltner, Tshering, Wangchuk, & Flynn, 2016;

Lima, Castro, & Scott, 2013). To maximize similarity of the posed stimuli to the real-life vocalizations, we chose for the posed stimuli highly intense positive joyous reactions of triumph and achievement and highly intense negative vocalizations of fear (see Section S4 in the online supplemental materials). The full set of audiovisual combinations is available upon request from the authors.

Stimuli editing. Each vocalization was embedded in two unique home-video clips (one positive and one negative), resulting a total set of 116 short video clips (58 positive, 58 negative), each embedded with either positive or negative vocalizations. The video clips were obtained from online home-videos (www.youtube.com) and were selected according to the documented scenario in the exact same fashion as in Experiment 1a. All the video clips were edited using Sony Vegas Pro 13 video editing software (MAGIX Software GmbH, Berlin, Germany) and converted to .MP4 format. The original video clips were muted and edited with the vocalizations such that the onset of the vocalizations was synced with the fitting frame in the video in which the reaction would naturally be expected. The duration of the video clips ranged from 1 s to 36 s ($M_{duration}$ = 10.47 s, SD = 6.76; $Mdn_{duration}$ = 9 s).

Participants. Vocalizations were rated online by 42 participants (19 female, mean age = 30) that were recruited from an Israeli participant pool (https://www.panel4all.co.il/). Because no prior estimates of effect size existed for the contextual phenomenon, we determined sample size on the *N* used in Experiment 1a which demonstrated large behavioral effects with similar stimuli. All participants watched audio-visual clips, rated the vocalizations online and received payment for their participation in the study.

Procedure. Participants viewed the videos and were asked to rate the valence of each vocalization using a computerized bipolar scale ranging from 1 (*most negative*) to 8 (*most positive*). No information was given about the origin of vocalizations. All stimuli were presented in a single block in randomized order.

Results

As seen in Figure 3, the mean valence of the vocalizations shifted categorically from positive to negative, and vice versa, as a function of their context. All the main effects were significant, indicating that perceived valence was higher (i.e., more positive) with positive vocalizations, $F(1, 41) = 42.5$, $p < .0001$, $\eta_p^2 = .509$; with posed stimuli, $F(1, 41) = 61.7$, $p < .0001$, $\eta_p^2 = .6$; and with positive context $F(1, 41) = 143.9$, $p < .0001$, $\eta_p^2 = .77$. Importantly, two interactions were revealed. First, conceptually replicating the results of Experiment 1a, we found a Stimuli Source × Vocalization Valence interaction, indicating that in real life, but not in the posed stimuli, the valence of positive and negative vocalizations was highly similar, $F(1, 41) = 48.1$, $p < .0001$, $\eta_p^2 = .54$. Critically for our main prediction, a stimuli source × context interaction was found, indicating that the real-life vocalizations were more susceptible to the effect of context than were the posed vocalizations, $F(1, 41) = 34.3$, $p < .0001$, $\eta_p^2 = .45$.

Thus, participants did not merely ignore the vocalization and answer based on the visual context, nor did they merely respond due to demand characteristics. Rather, the perceptual valence of the real-life vocalizations shifted due to the accompanying context. No other interaction effects were significant (all ps > .15).

General Discussion

Across these studies we demonstrate that highly intense real-life vocalizations are perceptually ambiguous and contextually malleable. Together, these findings challenge basic emotion theories (Ekman & Cordaro, 2011; Sauter et al., 2010) and dimensional theories (Russell, 1980), as both accounts would predict a robust perceptual distinction between expressions during positive and negative situations. By contrast, our results fit well with functional–contextual frameworks (e.g., behavioral ecology theory), which argue that reactions during affective situations are flexible tools for social influence, and do not reflect

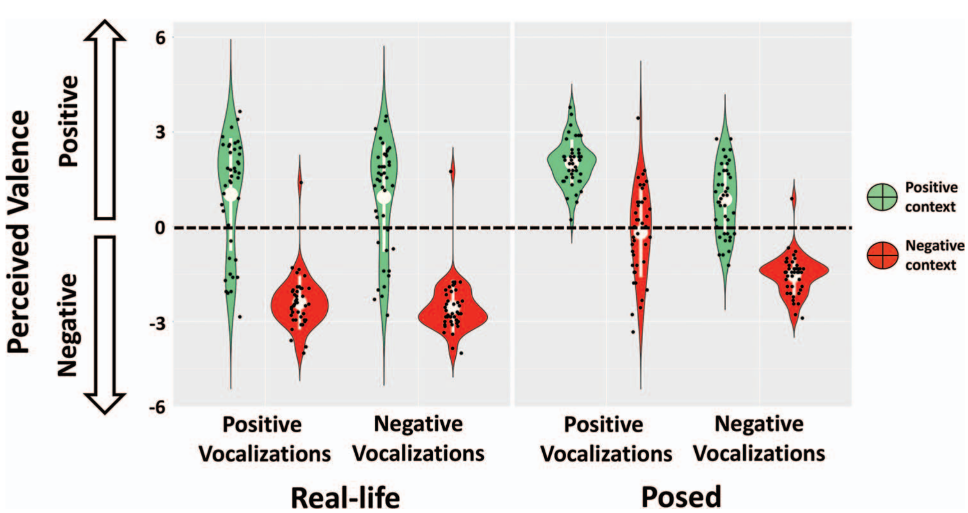

Figure 3. Violin plots of perceived valence of real-life and posed, positive and negative vocalizations as a function of positive and negative context (central white circles and vertical white bars reflect mean and *SD*). Context dramatically influenced the perceived valence of vocalizations such that the same vocalizations sounded positive or negative when paired with differently valenced visual context, an effect more robust for the real-life stimuli than for the posed stereotypical vocalizations. See the online article for the color version of this figure.

a veridical signal of felt emotion (Crivelli & Fridlund, 2018). Indeed, the data support a functional and externalist account which stems from animal communication and is, by definition, context-dependent (Bachorowski & Owren, 2003; Seyfarth & Cheney, 2018; Waller et al., 2017).

Highly intense vocalizations could operate by strongly and reliably influencing others by calling the attention of conspecifics to the high significance of an ongoing event. Depending on the situation in which they are heard (e.g., while at a party or while walking down a dark alley), they could raise predictions about the next move of the vocalizer, trigger approach or avoidance behavior in receivers, and once contextualized they could be perceived as conveying specific affective valence and emotions (Barrett & Kensinger, 2010; Barrett et al., 2011).

The current results suggest that increased intensity plays an important role in vocal ambiguity, as it does for facial ambiguity (Aviezer et al., 2012; Israelashvili et al., 2018). Whether similar patterns will emerge for real-life vocalizations displaying moderate and subtle levels of intensity is still unknown, as research with such stimuli is in its infancy. Nevertheless, we suspect such vocalizations will display more perceptual ambiguity and contextual sensitivity than previously assumed.

References

Anikin, A., & Lima, C. F. (2018). Perceptual and acoustic differences between authentic and acted nonverbal emotional vocalizations. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 71,* 622–641.

Anikin, A., & Persson, T. (2017). Nonlinguistic vocalizations from online amateur videos for emotion research: A validated corpus. *Behavior Research Methods, 49,* 758–771. http://dx.doi.org/10.3758/s13428-016-0736-y

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

Bachorowski, J.-A., & Owren, M. J. (2003). Sounds of emotion: Production and perception of affect-related vocal acoustics. *Annals of the New York Academy of Sciences, 1000,* 244–265. http://dx.doi.org/10.1196/annals.1280.012

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70,* 614–636. http://dx.doi.org/10.1037/0022-3514.70.3.614

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

Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science, 20,* 286–290. http://dx.doi.org/10.1177/0963721411422522

Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal affective voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods, 40,* 531–539. http://dx.doi.org/10.3758/BRM.40.2.531

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25,* 49–59. http://dx.doi.org/10.1016/0005-7916(94)90063-9

Cheney, D. L., & Seyfarth, R. M. (2018). Flexible usage and social function in primate vocalizations. *Proceedings of the National Academy of Sciences, 115,* 1974–1979.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

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,* 117–128. http://dx.doi.org/10.1037/emo0000100

Crivelli, C., & Fridlund, A. J. (2018). Facial displays are tools for social influence. *Trends in Cognitive Sciences, 22,* 388–399. http://dx.doi.org/10.1016/j.tics.2018.02.006

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review, 3,* 364–370. http://dx.doi.org/10.1177/1754073911410740

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191. http://dx.doi.org/10.3758/BF03193146

Fernández-Dols, J.-M., & Crivelli, C. (2013). Emotion and expression: Naturalistic studies. *Emotion Review, 5,* 24–29. http://dx.doi.org/10.1177/1754073912457229

Fridlund, A. J. (1994). *Human facial expression: An evolutionary view*. San Diego, CA: Academic Press.

Fridlund, A. J. (2017). The behavioral ecology view of facial displays, 25 years later. In J.-M. Fernandez-Dols & J. A. Russell (Eds.), *The science of facial expression* (pp. 77–92). New York, NY: Oxford University Press.

Gendron, M., Roberson, D., & Barrett, L. F. (2015). Cultural variation in emotion perception is real: A response to Sauter, Eisner, Ekman, and Scott (2015). *Psychological Science, 26,* 357–359. http://dx.doi.org/10.1177/0956797614566659

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Cultural relativity in perceiving emotion from vocalizations. *Psychological Science, 25,* 911–920. http://dx.doi.org/10.1177/0956797613517239

Israelashvili, J., Hassin, R. R., & Aviezer, H. (2018). When emotions run high: A critical role for context in the unfolding of dynamic, real-life facial affect. *Emotion*. Advance online publication. http://dx.doi.org/10.1037/emo0000441

Lavan, N., Lima, C. F., Harvey, H., Scott, S. K., & McGettigan, C. (2015). I thought that I heard you laughing: Contextual facial expressions modulate the perception of authentic laughter and crying. *Cognition and Emotion, 29,* 935–944.

Lima, C. F., Anikin, A., Monteiro, A. C., Scott, S. K., & Castro, S. L. (2018). Automaticity in the recognition of nonverbal emotional vocalizations. *Emotion*. Advance online publication. http://dx.doi.org/10.1037/emo0000429

Lima, C. F., Castro, S. L., & Scott, S. K. (2013). When voices get emotional: A corpus of nonverbal vocalizations for research on emotion processing. *Behavior Research Methods, 45,* 1234–1245. http://dx.doi.org/10.3758/s13428-013-0324-3

Nelson, N. L., & Russell, J. A. (2013). Universality revisited. *Emotion Review, 5,* 8–15. http://dx.doi.org/10.1177/1754073912457227

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39,* 1161–1178. http://dx.doi.org/10.1037/h0077714

Russell, J. A., Bachorowski, J.-A., & Fernández-Dols, J.-M. (2003). Facial and vocal expressions of emotion. *Annual Review of Psychology, 54,* 329–349. http://dx.doi.org/10.1146/annurev.psych.54.101601.145102

Sauter, D. A., & Eimer, M. (2010). Rapid detection of emotion from human vocalizations. *Journal of Cognitive Neuroscience, 22,* 474–481. http://dx.doi.org/10.1162/jocn.2009.21215

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 of the United States of America, 107,* 2408–2412. http://dx.doi.org/10.1073/pnas.0908239106

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2015). Emotional vocalizations are recognized across cultures regardless of the valence of distractors. *Psychological Science, 26,* 354–356. http://dx.doi.org/10.1177/0956797614560771

Sauter, D. A., & Fischer, A. H. (2018). Can perceivers recognise emotions from spontaneous expressions? *Cognition and Emotion, 32,* 504–515.

Scherer, K. R. (1994). Affect bursts. In S. H. M. van Goozen, N. E. Van de Poll, & J. A. Sergeant (Eds.), *Emotions: Essays on emotion theory* (pp. 161–196). Hillsdale, NJ: Erlbaum.

Seyfarth, R., & Cheney, D. (2018). Pragmatic flexibility in primate vocal production. *Current Opinion in Behavioral Sciences, 21,* 56–61. http://dx.doi.org/10.1016/j.cobeha.2018.02.005

Tracy, J. L. (2014). An evolutionary approach to understanding distinct emotions. *Emotion Review, 6,* 308–312. http://dx.doi.org/10.1177/1754073914534478

Waller, B. M., Whitehouse, J., & Micheletta, J. (2017). Rethinking primate facial expression: A predictive framework. *Neuroscience and Biobehavioral Reviews, 82,* 13–21. http://dx.doi.org/10.1016/j.neubiorev.2016.09.005

Received July 29, 2018
Revision received October 1, 2018
Accepted October 4, 2018 ■