The Role of Context and Development in Rating the Intensity of Facial Emotion

By

Brian T. Leitzke

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Psychology)

at the

UNIVERSITY OF WISCONSIN-MADISON

2019

Date of final oral examination: 5/30/2019

The dissertation is approved by the following members of the Final Oral Committee:

Seth Pollak, Professor, Psychology, Anthropology, Pediatrics, and Public Affairs
C. Shawn Green, Associate Professor, Psychology
Michael Koenigs, Associate Professor, Psychiatry
Kristin Shutts, Associate Professor, Psychology
Linda Camras, Professor Emerita, Psychology, DePaul University

Table of Contents

I.	Acknowledgements	ii
II.	Abstract	iii
III.	Overview	1-2
IV.	Introduction	3-11
V.	Method	11-18
VI.	Results	18-22
VII.	Discussion	23-31
VIII.	References	32-43
IX.	Tables	44-46
X.	Figures	47-49

Acknowledgments

This research would not be possible without the participation of the many children and their families, as well as undergraduates, for whose collaboration we are extremely appreciative. I also thank the undergraduate research assistants who facilitated data collection throughout the course of this study, including Olivia Bartell, Gwyneth DeLap, Selin Gok, Rachel Posener, and Abbey Schlicksup, as well as Yue Sun for his support with data processing. I would like to thank Anna Bechner for her support throughout this project and Aaron Cochrane for his expertise and assistance with data analysis. This research project was supported by a grant to Seth Pollak from the National Institute of Mental Health (R01-MH61285), a core grant to the Waisman Center from the National Institute of Child Health and Human Development (U54-HD090256), and a Ruth L. Kirschstein National Research Service Award to Brian Leitzke from the National Institute of Mental Health (F31-MH106179).

I also would like to thank Rachel Hamilton and Rista Plate for their incredible support and feedback throughout this project and beyond. I especially would like to thank Seth Pollak for his support and guidance throughout my graduate career. His mentorship has allowed me to accomplish much of what I have and has the stage for my future endeavors. Finally, I am forever thankful for the encouragement, wisdom, and unwavering support from Anne Loevinger Leitzke, without whom none of this would be possible.

Abstract

Background: While it is widely recognized that the ability to recognize and interpret others' emotional expressions is important for successful social communication, how facial emotion is perceived remains controversial. This dissertation focused on three key aspects of emotion perception often omitted in common emotion recognition procedures: (1) the evaluation of the intensity of portrayed emotion, rather than categorical labeling, (2) the role of contextual information in the perception of facial emotion, and (3) developmental changes in how children perceive co-occurring facial and contextual information. **Method:** This experiment examined developmental effects on the influence of congruent, incongruent, and neutral situational contexts on ratings of facial emotion intensity. **Results:** These data demonstrated that participants, particularly children, rated emotional intensity higher when contextual information was emotionally congruent with facial emotion. This effect was strongest for images depicting the lowest emotional intensity. Additionally, adults and adolescents were better able to discriminate between varying levels of face intensity than children. **Conclusions and Implications:** Contextual information is not only influential in determining what emotion someone may be feeling, but also how much of that emotion they are feeling. However, through experience and maturation, the ability to identify subtle changes in facial emotion improves and emotion judgments become less amenable to contextual influence. Understanding developmental patterns in processing complex emotional scenes may help identify problematic attentional patterns toward emotional signals that may be related to deficits in social and emotional functioning. Such deficits may place individuals at risk for the development of psychopathology.

Overview

This study focused on the role of context and development in how people evaluate facial expressions of emotion. Accurately gauging others' emotions guides social interactions and behavioral responses to social partners. However, there are multiple cues that perceivers could use to infer another's emotional state, including cues within the emoters face and within the larger context. We are just beginning to understand how individuals integrate multiple information sources to make emotion judgments. To date, there is a paucity of understanding how such cues interact when the perceiver needs to determine *how intensely* a social partner might be feeling a particular emotion. Yet, accurately assessing whether someone is, for example, mildly annoyed versus very angry might change how one modulates their own behavior. The present study sought to characterize developmental effects of contextual information on ratings of emotional intensity. We first discuss the role of facial expressions in the process of emotion perception and discuss the importance of recognizing subtle changes in expressions to guide adaptive responses. We highlight the influence of contextual information that often co-occurs with facial expressions in real-world interactions, and then discuss differences in emotion perception across development.

These issues informed the current experiment that examined developmental differences in rating facial emotion intensity when faces are presented in differing situational contexts. Employing a generalized, multilevel Bayesian regression modeling approach, we provide novel empirical evidence demonstrating an interactive effect of facial emotional intensity and contextual congruity on emotion ratings, which differs across development. Overall, findings from this study indicate age-related differences in the amenability of contextual influence in rating facial intensity as well the ability to detect differences in facial intensity. We conclude by

discussing the impact of this research, emphasizing the role of context and experience in the development of both adaptive and maladaptive emotion perception abilities. We also discuss limitations of this experiment, and implications for future research.

Introduction

Attending to and correctly interpreting the emotional states of others allows for appropriate behavioral responses to facilitate successful social interactions (Adolphs, 2002; Brosch, Pourtois, & Sander, 2010). To accurately identify another's emotion, one must decode meaning from a multitude of sources. These sources include facial musculature (Ekman, 1992), bodily movements (Aviezer, Ensenberg, & Hassin, 2017), situational contexts (Van den Stock & de Gelder, 2012), vocal cues (Sauter, Eisner, Ekman, & Scott, 2010), and past experiences (Barrett, 2017; Feldman-Barrett, Mesquita, & Gendron, 2011). Research examining the ways in which facial emotion is perceived focused primarily the broad categorization of emotional faces (e.g., angry, disgusted, sad, afraid), with disregard for emotional intensity. Yet full-blown expressions of emotion are not often encountered in the real world (Calvo & Nummenmaa, 2016). While one might see someone expressing extreme fear or anger in the movies, they are rarely seen on the playground or in the classroom. Rather, more nuanced evaluations, such as the degree to which an emotion is expressed, may more closely map on to real-world emotion perception (Calder, Young, Keane, & Dean, 2000). For example, varied expressions of less intense emotional expressions (annoyance vs. rage, contentment vs. ecstasy) are more frequently encountered (Calvo & Nummenmaa, 2016). Additionally, faces rarely appear in isolation. Facial expressions are typically embedded in a situation or context, which are influential in the perception and understanding of emotion (Aviezer et al., 2017; Barrett, Mesquita, & Gendron, 2011; Martinez, 2019). Beyond the contribution of specific visual components that influence emotion perception, experience and maturation also play a role in the development of emotion recognition skills. The ability to accurately identify different emotions improves with age (Rodger, Vizioli, Ouyang, & Caldera, 2015) though the sensitivity in recognizing some emotions

is dependent on experience and exposure to certain emotional expressions (Pollak, Messner, Kistler, & Cohn, 2009). Together, consideration of contextual cues and development may shed light on the process of emotion perception.

Perception of Facial Emotion

Research examining the perception of emotion has focused on information that can be gleaned from facial expressions. Early emotion theories stated that facial expressions of emotion provide a direct readout of internal emotional states (Buck, 1994), are biologically hardwired to facilitate communication and social functioning (Izard, 1992; Calder, Burton, Miller, Young, & Akamatsu, 2001), and are universally recognized (Ekman & Rosenberg, 2005). These categorical accounts of facial emotion perception proposed that the emotions perceived and recognized in facial expressions fell within specific categories of emotions, termed 'basic emotions' (Calvo & Nummenmaa, 2016; Ekman, 1992; Gendron, Roberson, van der Vyver, & Barrett, 2014). In contrast, dimensional views stipulate that emotional expressions can be perceived on the dimensions of valence and arousal (Russell, 1980; 2003). Rather than viewing faces as discrete emotion categories, these accounts propose that varying levels of valence and arousal are perceived in facial displays and integrated to determine another's emotional state (Carroll & Russell, 1996; Russell, 1997). A dimensional account describes emotions on dimensions, which is intended to capture more nuanced displays of emotion to allow for fine-grained emotion gradients that more accurately reflect real-world emotion processing (Russell, 2003). While early theories of facial emotion processing focused on discrete emotional states, recent research indicates a more nuanced approach to understanding the perception of emotional states.

To date, there is little research on how affective intensity informs emotion perception from either a categorical or dimensional perspective. Emerging literature suggests that a more nuanced approach to emotion perception that focuses on affective intensity may better capture real-world emotion processing. In many situations, to accurately identify what another is feeling, one must not only recognize the presence or absence of some discrete emotion (e.g., happy, angry), but also gauge the intensity of the emotion someone is expressing. Whether a child perceives their teacher to be mildly annoyed or outright furious will have considerable influence on subsequent behaviors. A recent nuanced approach places facial musculature on continua of muscular activation that allows for subtle distinctions of emotional expression, regardless of categorical label (Martinez, 2017). For example, furrowing of the brow and tightening of the lips constitute specific muscular configurations, known as facial action units (Ekman & Friesen, 1978), that depict the emotion of anger. Such changes in facial musculature allow for not only identification of discrete emotions, but also discrimination of subtle changes in affect and arousal that may help facilitate recognition of an expresser's emotional state. Further, the degree to which an action unit is activated is presumed to represent the intensity of the underlying emotional state (Martinez, 2017). In this way, recognizing subtle distinctions in facial musculature is likely beneficial when assessing facial emotional intensity. This process of recognizing emotional states from a facial display, however, assumes that faces are the sole source of information at one's disposal. Yet faces rarely occur in isolation.

Role of Context in Emotion Perception

There is a growing body of literature demonstrating the effect of context on emotion perception (Barrett, Mesquita, & Gendron, 2011; Chen & Whitney, 2019; Martinez, 2019).

Contextual information has the capacity to modify the emotion perceived in various facial expressions (Aviezer et al., 2017; Righart & de Gelder, 2008; Xu, Yang, Tan, & Zhang, 2017). For example, when two facial expressions of emotion are perceptually similar (e.g., anger and disgust), bodily context (i.e., posture) becomes the primary source of information used to make emotion judgments (Aviezer et al., 2008).

Aviezer and colleagues (2008) proposed that facial emotions are characterized by "emotion seeds," whereby certain facial displays of emotion share facial features, described as seeds. For example, the emotions of anger and disgust both involve a furrowing of the brow region, while the expressions of sadness and fear are highly similar in terms of their physical characteristics (Mondloch, 2012; Susskind, Littlewort, Bartlett, Movellan, & Anderson, 2007). When faces depicting a particular emotion are paired with body postures that are associated with a facial emotion that shares physical characteristics, the context emotion is more often selected as consistent with the expresser's emotion than the facial emotion (Aviezer et al., 2008; Mondloch, 2012). For instance, when viewing an expression of an individual who just won or lost a competitive sporting match (e.g., tennis, soccer), individuals are unable to accurately identify the expresser's emotion without any accompanying bodily context (e.g., fist pumping above head to indicate a win or arms at side to indicate loss; Aviezer et al., 2015). However, when situational context is emotionally congruent with the emotion depicted in facial expressions, emotion recognition is better and faster than when the context is emotionally incongruent (Lee, Choi, & Cho, 2012; Righart & de Gelder, 2008; Van den Stock & de Gelder, 2012; Xu et al., 2017). Contextual influence on face perception has been found in both laboratory and observational studies and has been reported across development (Aviezer et al., 2017;

Leitzke & Pollak, 2016; Noh & Isaacowitz, 2013; Rajhans, Jessen, Missana, & Grossman, 2016). These findings highlight the importance of context in emotion perception.

Context and Ratings of Emotional Intensity

There is some research on the effects of context that has similarly used paradigms that consist of forced choice and categorical labeling of emotions. Moreover, the extent to which context influences ratings of emotional intensity has yet to be examined. With regard to categorizing facial expressions of varied degrees of intensity, high-intensity faces are more accurately identified than low-intensity faces (Calder et al., 2000; Gao & Maurer, 2009; Hess, Blairy, & Kleck, 1997). Low intensity faces tend to be more difficult to identify as they are relatively ambiguous (Motley & Camden, 1988); thus, context is particularly helpful for informing emotion judgments in these cases (Aviezer, Trope, & Todorov, 2012).

Yet, contextual information may also influence ratings of how intensely an emotion is being conveyed. A fearful facial expression of low or intermediate intensity may be rated as more or less fearful depending on the accompanying context. For example, we might wonder how perceivers would assess the same facial cues within the context of a snake pit versus a dark room. Perhaps, perceivers use the context to gauge how strongly they think an emoter should feel given the circumstances and then integrate that judgment with their assessment of the emoter's facial cues. Such decisions are critical in social situations, as emotions of different intensities signify the necessity of different behavioral approaches. Exploration of the impact of context on ratings of facial emotional intensity, and how this may differ at varied levels of facial intensity, may provide a better parallel to the requirements of real-world emotion perception.

Context, Intensity, and Development

Attending to and recognizing facial emotion changes with age. Throughout development, individuals are exposed to a wide array of emotional cues including facial expressions, gestures, sounds, and smells. Children learn to sift through and hone their attention toward particular information and early on, prioritize affective over non-affective stimuli (Todd, Cunningham, Anderson & Thompson, 2012). Facial expressions provide a particularly important source of information. They can reference safety or danger (Feinman, 1982) and communicate feelings and intentions (Adolphs, 2002). Preferential attention toward faces emerges early in life (Izard, 1992). Indeed, infants younger than six months in age give preferential attention toward faces (Morton & Johnson, 1991) and are sensitive to categorical boundaries between some, but not all, facial expressions of emotion (White et al., 2019). By the age of 12 months, infants demonstrate the ability to reference others' emotions to guide behaviors (Feinman, 1982; Sorce, Emde, Campos, Klinnert, 1985). While some may view these early face-oriented abilities as the emergence of an innate, hard-wired ability (e.g., Izard, 1992), there is considerable evidence that the emergence of emotional attention is a developmental process, subject to learning and experience.

Developmental and cross-cultural research demonstrate that experience and learning inform emotion perception. The approach individuals use to process facial expressions changes from an analytical strategy during childhood, to a more holistic approach in adulthood (Schwarzer, 2000). While adult-like proficiency in recognizing some emotions, particularly positive emotions, may emerge early in childhood (Camras & Allison, 1985; Widen & Russell, 2003), there is also evidence that emotion perception develops throughout childhood and adolescence (Lawrence, Campbell, & Skuse, 2015; Rodger, Vizioli, Ouyang, & Caldera, 2015;

Thomas, De Bellis, Graham, & LaBar, 2007; Widen & Russell, 2008). Indeed, early event-related potential components that are sensitive to face processing in adults do not emerge until late adolescence (Batty & Taylor, 2006). Cross-cultural research has also yielded evidence for the role of learning and experience in the development of emotion perception. For example, people are able to discern cultural membership, such as whether another is Japanese versus Japanese-American, by merely viewing facial expressions of emotion (Marsh, Elfenbein, & Ambady, 2003). There are also differences in the way Eastern and Western cultures display and visually process facial emotion. Western cultures tend to prioritize information from the mouth to interpret facial emotion, while individuals from Eastern cultures are more likely to emphasize information extracted from eye regions (Eisenbarth & Alpers, 2011; Yuki, Maddux, & Masuda, 2007). Together, these data provide support for a developmental process of emotion perception that is sensitive to learning and experience.

The ability to integrate facial expressions and contextual information also changes throughout development. The development of perceptual systems that guide emotion recognition allow for individuals to identify consistencies in their environment and form associations between certain stimuli and outcomes (Pollak et al., 2009). Children as young as four years of age demonstrate a general ability to integrate competing emotional cues when making broad, categorical emotion judgments in a manner similar to adults (Mondloch, Horner, & Mian, 2013). However, sensitivity to changes in facial musculature, as well as other internal characteristics such as distances between certain facial features (e.g., eyes, nose, mouth), appears to develop relatively slowly throughout childhood (Mondloch, Geldart, Maurer, & Le Grand, 2003). Moreover, relative to adults, younger children tend to pay greater attention to external features of faces (e.g., hair; Bruce et al., 2000).

Leitzke & Pollak (2016) sought to characterize developmental changes in how individuals view contextual and facial information when making emotion judgments. In their study, individuals from four developmental epochs (early childhood, middle childhood, early adolescence, and early adulthood) were asked to apply emotion labels to pictures of people posing various facial expressions of emotion paired with congruent (e.g., angry face paired with body depicting context) and incongruent contexts (e.g., angry face paired with body depicting digust). Participants completed this task on an eye tracker to allow for quantification of eye movements to various components of the presented images. While older children and adults tended to allocate the majority of their attention toward faces relative to contextual information, younger children divided their attention to both sources of information equally. Moreover, contextual congruence influenced gaze patterns toward certain features of the face. These results indicate a developmental trend of increasing prioritzation of facial over contextual information with age.

Current Study

Despite emerging literature demonstrating the influence of development on the ability to process and integrate competing emotional cues, there is little research examining age-related changes in the effect of context on intensity ratings of facial emotion. Examining contextual effects across development will help inform our understanding of how nuanced emotion processing changes with age and experience. This information would aid the identification of normative and maladaptive emotion processing skills and inform the design of targeted, attention-based interventions. In this experiment, we tested children, adolescents, and young adults as they viewed images of facial expressions of emotion of different intensities presented

with emotion-congruent, incongruent, and neutral contexts. Participated were asked to rate the level of emotional intensity they believed each person in each image was feeling. Based on previous findings demonstrating the effect of context on judgments of facial emotion (Aviezer et al., 2017), developmental differences in attention to contextual cues (Leitzke & Pollak, 2016), and differences in emotion differentiation by age and experience (Pollak et al., 2009; Thomas et al., 2007), we made three hypotheses. First, we predicted that ratings of intensity would be greater when faces were presented in a congruent, relative to an incongruent or neutral context. Second, we predicted that differences in intensity ratings between congruent and incongruent displays would be greater for children than adolescents and adults. Finally, we predicted that adolescents and adults would be better able to discriminate between faces of different levels of intensity than children.

Method

Participants

One hundred sixty-two individuals from three age groups were recruited for this study. Children (N = 56; M_{age} = 7 years, 11 months, SD = 7 months; 48% Female; 7% Asian or Asian-American, 6% Black or African-American, 7% Hispanic, 73% White, and 7% Other Racial/Ethnic identification) and pre-adolescents (N = 54; M_{age} = 13 years, 1 months, SD = 7 months; 44% Female; 8% Asian or Asian-American, 6% Black or African-American, 4% Hispanic, 81% White, and 1% Other Racial/Ethnic identification) were recruited from the local community via television commercials, radio ads, and posted flyers as well as through a local school district registry where parents volunteered their contact information for research study recruitment. Young adults (N = 52; M_{age} = 19 years, 7 months, SD = 11 months; 37% Female;

26% Asian or Asian-American, 6% Black or African-American, 8% Hispanic, 58% White, and 2% Other Racial/Ethnic identification) were recruited from an introductory to psychology course or from the campus community via posted flyers.

Stimuli

Facial stimuli. Facial stimuli were borrowed from the Interdisciplinary Affective Science Laboratory (IASLab) Facial Stimuli Set.[1] This stimulus set consists of photographs of individuals posing various emotional expressions with both direct and averted gaze. We included models with averted gaze to direct participants' attention toward the contextual information displayed in each image. We selected facial displays of anger, disgust, fear, happiness, neutrality, and sadness from four models (two females: models F17 and F19, and two males: models M01 and M07). This stimulus set did not require posers to self-identify their race, though a pilot study demonstrated that all selected models were unanimously identified as White in a free-response format. We selected all White models to reduce biases in emotion recognition that may be due to race (Elfenbien & Ambady, 2002; Hu, Wang, Han, Weare, & Fu, 2017). While emotion recognition biases by gender also exist (Adams, Hess, & Kleck, 2015; Adams, Nelson, Soto, Hess, & Kleck, 2012), we included both male and female models to ensure some extent of natural variability in the presented stimulus set.

To create variation in emotion intensity, we morphed the facial expressions of each model between their display of each emotion and that same model's neutral expression. We used j.psychomorph (see Tiddeman, Stirrat, & Perrett, 2005) to morph each image to create 10% increments in intensity and selected the 20%, 50%, and 80% images to represent, low-, intermediate-, and high-intensity expressions respectively (Figure 1).

[1] Development of the Interdisciplinary Affective Science Laboratory (IASLab) Face Set was supported by the National Institutes of Health Director's Pioneer Award (DP1OD003312) to Lisa Feldman Barrett. More information is available online at www.affective-science.org

Contextual stimuli. One hundred contextual images were downloaded from the Internet and rated by 301 participants via Mechanical Turk (see Buhrmester, Kwang, & Gosling, 2011). In this rating task, participants viewed each image and assigned emotion labels to each one. They chose from 27 emotion options which were selected based on a recent study which identified highly used emotional terms via a free-labeling task from different cultures (Srinivasan & Martinez, 2018). To avoid a forced-choice bias, participants were able to assign one or as many options as they felt were appropriate in response to how they believed someone might feel if they viewed each scenario in real life. The top-rated images for each emotion category were selected (all images achieved 50% endorsements for each emotion category; chance performance would be 3.7%). We selected the top rated 24 images with four images for each emotion (anger, disgust, fear, happiness, sadness, neutral). For each of these emotions, two contextual images included people and two did not include people. Faces of all people in the situational context were blurred to ensure participants rated only the target face in each image.

Composites. Facial and contextual images were combined such that faces were superimposed on top of each context to create composites that displayed each face appearing to be responding to a contextual situation. The four facial expressions of anger, disgust, fear, and sadness were fully crossed with the three contextual conditions (congruent, incongruent, neutral) for 12 general composite images. The congruent condition paired each facial emotion with a context of the same emotion (e.g., anger face within anger context). Incongruent pairings were chosen based on the confusability of anger and disgust (Aviezer et al., 2008) and sadness and fear (Mondloch, 2012; Susskind et al., 2007); anger faces were paired with disgust contexts, and vice versa, and sad faces were paired with fear faces, and vice versa. All 12 face

emotion/congruence pairings were fully crossed with the three intensities (low, intermediate, high), creating 36 face emotion/congruence/intensity composites (see Figure 2).

Each face emotion/congruence/intensity composite appeared evenly with the four different models (two males, two females) and context exemplars (two with people, two without). The task consisted of 144 trials comprised of four presentations of each of the 36 face emotion/congruence/intensity composites, each with a different model and context exemplar. Sixteen trials (10%) consisted of happy faces and contexts to provide variation in stimuli valence. Specifically, happy faces were morphed to create low (20%) and high (80%) intensity happy faces in congruent and neutral contexts paired with each of the four facial emotion posers. In total, this experiment consisted of 160 total trials, with each of the four models and context exemplars appearing together ten times across the experiment with different composites.

Measures

Emotion recognition task. Participants viewed each image displayed at full screen resolution for 1000 msecs before being asked to rate the image. They then responded to the question “How is this person feeling?” by using a computer mouse to move a cursor along a visual analog scale that ranged from “Not at all [displayed facial emotion]?” to “Extremely [displayed facial emotion]?” to indicate the intensity of emotion they believed each person was expressing. Participants were randomly assigned to rate increasing intensity from left to right or right to left to reduce any directional bias. Resulting scores ranged from 0 (not at all displaying that emotion) to 100 (displaying that emotion very intensely). Participants were instructed to focus explicitly on the face in each image when making their ratings of each specific emotion to ensure ratings reflected the intensity of facial expressions in different contexts.

The experiment was divided into four 40-trial blocks, which participants viewed in a random order to ensure that all participants saw all stimuli and to eliminate order effects. Participants also completed an additional block that consisted of 48 trials of all face emotion/intensity/model combinations presented in the absence of any contextual information. The Emotion Recognition Task was created and presented with E-Prime 3.0 software (Psychology Software Tools, Pittsburgh, PA).

Cognitive ability and socioeconomic status. To address and control for possible differences in our university- versus community-recruited samples, participants completed both the two-scale version of the Weschler Abbreviated Scale of Intelligence, second edition (WASI-II; Weschler, 2011) and the Hollingshead 2-factor measure of social status (Hollingshead, 2011). The WASI-II is normed for individuals ages 6 through 90 and is designed to provide an estimate of general cognitive ability or intelligence (IQ) through administration of two subtests. The *Vocabulary* subtest involves participants defining words that are presented visually and orally and is designed to measure word knowledge and verbal concept formation. The second subtest, *Matrix Reasoning*, involves participants viewing an incomplete matrix and selecting a response option that completes the matrix. *Matrix Reasoning* measures fluid intelligence, broad visual intelligence, classification and spatial ability, knowledge of part–whole relationships, simultaneous processing, and perceptual organization. The WASI-II has evidenced high reliability, test-retest stability, and concurrent validity with commonly used measures of intelligence for both child and adult samples (McCrimmon & Smith, 2013).

We collected information about all participants' parental education and occupations to obtain a general measure of socioeconomic status (SES; Hollingshead, 2011). This two-factor measure of SES derives a score ranging from 8 to 66, based on the average scores from both

parents' education and occupation, or from just one parent if only one parent is present in a household. Education and occupation scores were based on guidelines reported in Hollingshead (2011) and updated to reflect modern professions. Young adult participants reported their parents' education and occupation information; children's parents self-reported their education and occupation. See Table 1 for IQ and SES scores.

Procedure

The University of Wisconsin – Madison Institutional Review Board approved all procedures and informed written consent/assent was obtained from all participants. Participants completed consent and assent procedures in a waiting room before moving to an individual testing room where they completed the study tasks. Participants completed all four blocks of the Emotion Recognition test, with the *Vocabulary* task of the WASI-II occurring between the second and third blocks, and the *Matrix Reasoning* task occurring after the fourth block. Stimuli were presented on a 21-inch Dell computer monitor at a resolution of 1920 by 1080 pixels and displayed at 75% of the width and height of the screen. A research assistant remained with each participant throughout the study to encourage and remind them to pay attention to the change in emotion label and direction of the scale and to use the face of each individual to make their determination of emotion intensity.

Adult participants were compensated with either course credit, if recruited from introduction to psychology course, or $10 cash payment if recruited from the campus community. Children received a prize for completing this experiment and their parents were compensated with $20 cash payment.

Data Analytic Plan

For the purpose of analyses, we excluded all trials where response times were under 200ms to account for responses made prior to the time required to initiate perceptual and motor processes to respond following stimulus presentation (Whelan, 2008; 0.1% of all responses). The data in this study were analyzed in R (R Core Team, 2018) using the brms package (Bürkner, 2017) to implement a Bayesian fitting approach. As this study had both within- (Congruence, Face Intensity) and between- (Age Group) subjects variables, we fit a generalized multilevel Bayesian regression examining the relationship between ratings of emotional intensity, face/context congruence, face intensity, and age. We added random intercepts for participant and context emotion and allowed ratings of face intensity to vary by participant, as well as stimulus model and facial emotion.

Because intensity ratings could take values between 0 and 100 (rescaled to 0 and 1), we modeled these responses as arising from a beta distribution. We modeled the expected value of the ratings while including a subject-level intercept term for the precision around the predicted expected value. To assist with interpretation, we centered the variables for face/context congruence and face intensity around zero. We added SES, IQ, and gender as covariates. All parameters were estimated as log odds. All *brms* models used Hamiltonian Monte Carlo sampling implemented in Stan (Stan Development Team, 2018). Each model included a warm up of 5,000 steps and four chains with 3,750 iterations each for a total of 15,000 used samples. With regard to model convergence, all Rhat scores were below 1.02 (score of 1 indicating convergence), all estimated samples were greater than 675, and visual inspection of the trace plots indicated adequate convergence of the models.

We first entered all theoretically meaningful variables in the full model and dropped variables that accounted for minimal variance, resulting in a model with effects and interactions that contributed most to the overall variance. We next examined the effectiveness of our manipulations of face intensity and congruence to determine whether all levels of each variable should remain in the final model. To interpret the primary results of the models, we examined the highest density interval (HDI) of the posterior samples to determine the location of the estimated posterior distribution relative to zero. Variables with an HDI closely centered around zero indicate a non-reliable predictor of the dependent variable, while an HDI that does not include zero indicates a reliable predictor.

We calculated expected values for specific groups and trial types for each sample of parameters, corresponding to 15,000 separate estimates of these values. We then tested for differences in ratings by examining the overlaps in the distributions of parameter samples (Kruschke, 2013). By subtracting values in paired samples, we calculated a distribution corresponding to predicted differences between two groups (e.g., children vs. adults) or trial types (e.g., low vs. high intensity). If the distribution of differences was largely positive or negative, we considered the difference reliable (using conventional 95% quantiles of parameter differences). We reported these comparisons in terms of the median and 95% credible intervals of the differences, as well as the predicted proportion of posterior samples that were on the same side of zero as the median value, which we reported as Congruent Density (cD). We reported the median as opposed to mean values as mean values are subject to bias with potentially skewed distributions. Calculated values were transformed back from log-odds ratios to raw values (0-100) to aid in interpretation.

Results

Preliminary Analyses

We first examined the influence of the covariates. The three covariates of SES, IQ, and Gender all had clearly zero-centered and non-reliable slopes when predicting variation in intensity ratings (SES: HDI [-.001, .013]; IQ: HDI [-.009, .005]; Gender: HDI [-.244, .101]). As all three covariates had no effect on our outcome variable, we excluded them in our final analyses.

An examination of the random effects of face emotion, context emotion, and stimulus model indicated marginal differences in the relationship between these variables and intensity ratings (see Table 2). While disgust faces were generally rated at a higher intensity than faces demonstrating anger, fear, and sadness, the confidence intervals for all levels overlapped considerably. Thus, there did not appear to be any systematic impact of any emotion or model on intensity ratings. Due to the observed differences, however, we retained model, face emotion, and context emotion as random effects in our final analyses.

Our preliminary analysis revealed that the manipulation of intensity was mostly successful. Collapsed across age groups, low-intensity faces were rated reliably less intense than intermediate intensity faces (Low: *Mdn* = 42.4; Intermediate: *Mdn* = 65.9; Difference: *Mdn* = -22.2, 95% CI [-13.4, -32.4], *cD* = 100%) and high-intensity faces (Low: *Mdn* = 42.4; High: *Mdn* = 70.3; Difference: *Mdn* = -26.8, 95% CI [-36.2, -16.2], *cD* = 100%). Intermediate-intensity faces were only marginally less intense than high-intensity faces (Intermediate: *Mdn* = 65.5; High: *Mdn* = 69.7; Difference: *Mdn* = 4.2, 95% CI [-2.6, 10.9], *cD* = 88.4%). This pattern suggests that the relationship between facial intensity and rated intensity was non-linear.

Regardless, we opted to retain all three intensity levels in our final model as the data did not appear to support the combination of intermediate- and high- intensity faces into one category.

Hypothesis #1: Does Context Alter Ratings of Emotional Intensity?

For our first hypothesis, we tested whether faces paired with an emotionally congruent context received higher ratings of intensity than faces paired with emotionally incongruent contexts. Across all age groups, faces in congruent contexts were rated as more intense than faces in neutral contexts (Congruent: *Mdn* = 63.2; Neutral: *Mdn* = 56.7; Difference: *Mdn* = 6.2, 95% CI [-2.1, 15.0], *cD* = 95.4%) and incongruent contexts (Congruent: *Mdn* = 63.2; Incongruent: *Mdn* = 58.3; Difference: *Mdn* = 5.3, 95% CI [0.4, 8.1], *cD* = 98.6%). However, faces rated in neutral contexts were generally indistinguishable from faces in incongruent contexts in both adults (Neutral: *Mdn* = 58.0; Incongruent: *Mdn* = 59.1; Difference: *Mdn* = 0.3, 95% CI [-8.3, 8.8], *cD* = 55.6%) and adolescents (Neutral: *Mdn* = 60.2; Incongruent: *Mdn* = 58.8; Difference: *Mdn* = 1.2, 95% CI [-5.3, 8.3], *cD* = 69.2%) and were rated as moderately less intense than in incongruent contexts among children (Neutral: *Mdn* = 59.0; Incongruent: *Mdn* = 65.1; Difference: *Mdn* = -5.5, 95% CI [-13.5, 2.6], *cD* = 93.8%).

Given the lack of differentiation in intensity ratings of faces in neutral and incongruent contexts, we combined neutral and incongruent contexts into one category (non-congruent) in all subsequent analyses. Here, we found participants rated faces in congruent trials as more intense than those in non-congruent trials (Congruent: *Mdn* = 68.5; Non-congruent: *Mdn* = 65.1; Difference: *Mdn* = 4.0, 95% CI [2.2, 11.0], *cD* = 100%; see Table 3 for summary statistics). Generally, the congruency-related patterns of results support our hypotheses, indicating that congruent contexts yield increased ratings of emotional intensity relative to any non-congruent context.

Examining intensity ratings by congruency across levels of facial intensity, results indicated that the difference in intensity ratings of faces between congruent and non-congruent contexts was greater for low-intensity faces than either intermediate- (Low: *Mdn* = 6.6; Intermediate: *Mdn* = 2.2; Difference: *Mdn* = 4.2, 95% CI [0.0, 8.7], *cD* = 97.9%) or high-intensity faces (Low: *Mdn* = 6.6; High: *Mdn* = 2.0; Difference: *Mdn* = 4.4, 95% CI [0.1, 9.2], *cD* = 99.5). There was no difference in intensity ratings by congruence between intermediate- and high-intensity faces (Intermediate: *Mdn* = 2.2; High: *Mdn* = 2.0; Difference: *Mdn* = 0.2, 95% CI [-0.7, 1.4], *cD* = 80.9%). This suggests that faces of lower intensity, those that appear more ambiguous due to the lesser strength of the emotional signal, are more subjective to congruency effects than higher intensity faces (see Figure 3a).

Hypothesis #2: Does the Effect of Congruence on Intensity Ratings Differ by Age?

To test our second hypothesis, we examined age-related differences in the impact of contextual congruence on intensity ratings. All three age groups rated faces in congruent contexts as greater than non-congruent contexts (all *Mdns* > 1.9, all *cDs* = 100%). However, the difference in intensity ratings between faces in congruent and non-congruent contexts was greater for children than both adolescents (Children: *Mdn* = 7.0; Adolescents: *Mdn* = 1.9; Difference: *Mdn* = 5.2, 95% CI [2.5, 13.5], *cD* = 100%) and adults (Children: *Mdn* = 7.0; Adults: *Mdn* = 2.1; Difference: *Mdn* = 4.7, 95% CI [2.1, 9.0], *cD* = 99.9%). Adolescents and adults, on the other hand, showed comparable congruency related changes in intensity ratings (Adolescents: *Mdn* = 1.9; Adults: *Mdn* = 2.1; Difference: *Mdn* = -0.1, 95% CI [-6.8, 0.1], *cD* = 77.9%). These findings indicate that the influence of context on ratings of emotional intensity decreases with age (see Figure 3b).

Hypothesis #3: Are there Age-related Differences in Discriminating Between Levels of Facial Intensity?

To examine our third hypothesis, we analyzed age-related differences in intensity ratings for faces of low, intermediate, and high intensity. Generally, both intermediate-intensity faces and high-intensity faces were rated as more intense than low-intensity faces (Intermediate: *Mdn* = 68.5; Low: *Mdn* = 48.1; Intermediate vs. Low: *Mdn* = 19.5, 95% CI [11.0, 26.7], *cD* = 100%; High: *Mdn* = 70.3; Low: *Mdn* = 48.1; High vs. Low: *Mdn* = 24.4, 95% CI [11.3, 35.2], *cD* = 100%).

However, high-intensity faces were rated as similarly intense compared to intermediate-intensity faces (High: *Mdn* = 73.6; Intermediate: *Mdn* = 68.5; Difference: *Mdn* = 4.8, 95% CI [-0.7, 16.1], *cD* = 80.9%). When comparing ratings of intensity between age groups, adolescents and adults showed a greater ability to distinguish between faces of varying intensity than children. Relative to children, adolescents and adults were better able to distinguish between intermediate- and high-intensity faces (Adolescents: *Mdn* = 6.9; Children: *Mdn* = 2.3; Adolescents vs. Children: *Mdn* = 4.5, 95% CI [1.0, 8.9], *cD* = 99.0%; Adults: *Mdn* = 5.6; Children: *Mdn* = 2.3; Adults vs. Children: *Mdn* = 3.4, 95% CI [-0.4, 7.7], *cD* = 95.6%) as well as between low- and high-intensity faces (Adolescents: *Mdn* = 28.1; Children: *Mdn* = 20.0; Adolescents vs. Children: *Mdn* = 8.1, 95% CI [2.7, 14.4], *cD* = 99.1%; Adults: *Mdn* = 24.6; Children: *Mdn* = 20.0; Adults vs. Children: *Mdn* = 4.4, 95% CI [1.2, 9.4], *cD* = 97.3%). Additionally, the difference in intensity ratings between low- and intermediate-intensity faces was less for children than adolescents (Children: *Mdn* = 17.9; Adolescents: *Mdn* = 21.2; Children vs. Adolescents: *Mdn* = -3.5, 95% CI [-9.4, -0.1], *cD* = 98.6%) and moderately less than adults (Children: *Mdn* = 17.9; Adults: *Mdn* = 19.1; Children vs. Adults: *Mdn* = -1.7, 95% CI [-4.9, 0.6],

cD = 93.9%). Adolescents' and adults' ratings of facial intensity, however, was comparable between all levels of intensity (Low vs. Intermediate: *Mdn* = -1.0, 95% CI [-7.0, 0.5], *cD* = 75.7%; Low vs. High: *Mdn* = -3.8, 95% CI [-8.5, 2.2], *cD* = 86.5%; Intermediate vs. High: *Mdn* = -1.3, 95% CI [-5.2, 2.6], *cD* = 73.8%; see Figure 3c).

Discussion

This experiment examined developmental differences on the impact of situational context on continuous ratings of facial emotion intensity. Specifically, this research demonstrates that contextual influence extends beyond facial emotion categorization to influence ratings of emotion intensity. Moreover, this effect was moderated by age, such that adolescents and adults were less influenced by context when rating emotional intensity. Adolescents and adults were also better able to distinguish between varying levels of facial intensity. Together, these findings indicate a developmental trend toward prioritizing facial emotion in making judgments of emotional intensity. Such findings can help inform techniques to identify maladaptive emotion processing abilities related to deficits in social functioning and guide the formation of novel, attention-based interventions for those at-risk for psychopathology.

The present findings indicate that context not only influences judgments of what another person feels, but also perceptions of how much they are feeling. The fact that faces in the current study were rated as more intense in congruent contexts suggests that the consistency of co-occurring emotional signals moderates the perception of emotional signal strength. Specifically, congruency between facial emotion and situational context resulted in higher intensity ratings than incongruency. Although speculative, one reason for this finding may be that incongruence decreased subjective levels of confidence in facial emotion intensity ratings. Notably, the

integration of context in making emotion judgments has been posited to be "automatic and effortless" (Aviezer, Dudarev, Bentin, & Hassin, 2011; Martinez, Falvello, Aviezer, & Todorov, 2015), with context influencing the perception of facial expressions at early stages of processing (Aviezer at al., 2008; Barrett, Lindquist, & Gendron, 2007). In the current paradigm, participants were allowed unlimited time to rate the emotional intensity of the presented stimuli; this extra time may have allowed for an elaborated integration of contextual information with the emotional faces. If contextual information is inconsistent with the co-occurring facial emotion, it may lead to confusion, which may in turn prompt decreased ratings of intensity. While observers' perception of facial emotion may be influenced early during viewing, with extended time to make emotion ratings, the inconsistency between facial and contextual information can be identified, and intensity ratings can be subsequently adjusted. In this study, the automaticity of face/context integration may have led to initial predictions that conflicted with the presented facial emotion. This interference then resulted in a reduction of perceived intensity.

Our initial statistical model found no difference in intensity ratings between incongruent and neutral contexts. One possibility for this lack of difference is that participants differentiated between congruent and *non*-congruent contexts. In other words, any situational context that did not match the emotion predicted from the displayed facial expression was considered inconsistent. If an emoter is expressing anger, an observer would expect the context to elicit anger. However, if an emoter is expressing anger and the context is disgust, or devoid of any emotion-eliciting situation, the observer is tasked with making their own attribution as to why the individual is experiencing anger. Such interpretations have been proposed to explain congruency effects for categorical emotion judgments (Carroll & Russell, 1996). In accordance

with this literature, the current finding suggests that any context that was incongruent with that perceived in the face altered emotion judgments.

There are a multitude of sources in the environment that contribute to one's ability to determine the emotional state of another. Such sources include a social partner's face, voice, or body posture, the surrounding context, prior knowledge of the individual, past experiences in similar contexts, etc. The face and context provide two sources of visual information, though the type of information they provide differs. Contexts may not elicit the same degree of emotion, or even the same emotion, among all people. If one dropped and broke their phone, how angry they may feel, or whether they are angry at all, would likely depend on certain factors, such as the availability of financial resources to replace it or the need to use it at that moment. Similarly, judging how another might respond in a particular situation will likely depend on an observer's experience with that situation. In this way, how contextual information is incorporated into emotion judgments might be influenced by the observer's focus on how they themselves would feel in that situation. So, while context is likely helpful in determining how an emoter might feel, further individualized information, such as facial expressions, is needed to accurately determine what *this* emoter is feeling. Context allows an individual to make a judgment of how they believe another could or should feel in a certain situation, whereas facial emotion provides individualized information about the emotional state of the emoter. Integrating facial and contextual information is likely required to make accurate emotion judgments.

This study also revealed that the influence of context on intensity ratings was more pronounced among younger emotion learners. Developmental patterns are evident in the emergence of emotion recognition (Camras & Allison, 1985; Leppanen & Nelson, 2006; Thomas et al., 2007; Widen & Russell, 2003) as well as different aspects of the perception of facial

emotion (Pollak et al., 2009). Moreover, the influence of context on labeling facial emotion appears to change from childhood into late adulthood (Leitzke & Pollak, 2016; Mondloch, 2012; Noh & Isaacowitz, 2012). Early emotion learners have access to a small set of emoters and limited contexts in their environment (Jayaraman, Fausey, & Smith, 2015). As they age, they gain experience with a greater number of emoters and are exposed to more diverse situations. Children, then, need to learn to sift through this increased variability of co-occurring emoters and contexts to make predictions about how people might feel in different situations. The current results indicate that this learning process may lead to a prioritization of facial information when making complex emotion judgments. Adults appear better able to focus on facial information despite competing contextual information. This is consistent with Leitzke & Pollak's (2016) findings that attention toward the face relative to context increases with age. This is not to say adults do not attend to context. Rather, they are likely able to quickly glean information from contextual information to shift attention toward information deemed more informative (Bar, 2004; de Gelder et al., 2006; Schyns & Oliva, 2010). The current study cannot determine the degree to which participants attended to different components of the stimuli, however, the reduced effect of congruency indicates that the information extracted and used from the context was less influential among older emotion learners.

The current study further demonstrated that the ability to discriminate between faces of varying levels of intensity improved with age. This finding is consistent with research that links enhanced facial emotion processing to maturation of certain brain regions (e.g., amygdala and prefrontal cortex; see Thomas et al., 2007). A biological account may also inform the understanding of the influence of context. As younger emotion learners are less adept at identifying subtle differences in facial emotion, they may draw upon contextual information to

aid in their emotion judgments. This would be particularly adaptive if certain facial expressions are difficult to assess. As individuals age and are better able to discern subtle changes in facial emotion, their increasing confidence in judging facial emotion may limit the necessity of complementary information from contextual cues. It is important to note that it can be difficult to differentiate between experiential growth and maturation as they inherently co-occur with aging. It is likely the two processes work in parallel, with increased brain maturation and experience informing one another and guiding enhancements in making efficient emotion judgments.

Abnormalities in either maturation or experience may lead to different developmental trajectories or different interpretations or prioritizations of emotional cues. Aberrant learning environments are associated with differential emotion perception. Children who were raised in neglectful or abusive environments are exposed to more anger (Plate et al., 2019) and exhibit atypical attentional patterns toward emotional stimuli. For example, abused children recognize anger with fewer perceptual cues than non-abused children (Pollak et al., 2009). Additionally, a range of psychological disorders are associated with difficulties in emotion recognition, including alcohol use disorder (Philippot et al., 1999), anxiety (Button, Lewis, Penton-Voak, & Munafo, 2013), autism (Celani, Battacchi, & Arcidiacono, 1999), bipolar disorder (Derntl, Seidel, Kryspin-Exner, Hasmann, & Dobmeier, 2009), depression (Rubinow & Post, 1992), and schizophrenia (Addington, Saeedi, & Addington, 2006). It is possible that aberrant emotional attention underlies the social and emotional deficits characteristic of these disorders (Morales, Fu, & Pérez-Edgar, 2016). Given the putative relationship between attentional biases towards emotional cues and psychopathological symptoms, there have been numerous efforts at targeting attentional processes toward emotional cues as a means of intervention (Mogoaşe, David, & Koster, 2014).

Overall, the current data indicate a need for a more nuanced approach to emotion recognition research that considers the role of context in discerning emotional intensity. This approach may be particularly important in experimental research designed to identify maladaptive attentional patterns among those at risk for the development of deficits in social functioning or psychopathology. Identification of differential attention to context among individuals with psychopathology can also inform the development of attention-based interventions; these interventions can thus focus on the recognition and perception of isolated facial expressions of emotion in a more targeted manner.

Limitations and Future Directions

This study also has important limitations to consider when making conclusions about the results. First, the facial stimuli in this study were not naturally-occurring expressions, rather, they were artificially created to represent low, intermediate, and high intensity facial emotion. As such, these facial expressions likely do not elicit naturalistic emotion processing. While it is possible to use pictures of individuals expressing emotion in a natural setting, it is not clear how one would be able to capture degrees of intensity in any systematic way. As the hypotheses and design of this study required a level of control over the expression of facial intensity, we included conventional static images with artificially-generated expressions created via morphing. The use of artificial facial expressions of emotion is an on-going concern in emotion recognition procedures. While this is the conventional method in the field, the use of stimuli should be dependent on the aims of a particular study. Further efforts to create naturalistic facial expressions will be helpful in improving emotion recognition procedures.

Second, the context stimuli were not unanimously determined to elicit the particular emotion they were assigned in this study. While each context was rated by a majority of

participants in the piloting phase, most stimuli were seen as eliciting other emotions as well. The images in this study were chosen to provide an unambiguous situation to contrast or support the facial emotions with which they were paired. This was designed to provide clear congruent and incongruent pairings. However, it is still possible that some situational contexts were not reliably interpreted as having the same meaning as intended. As such, the contextual influence hypothesized based on previous research demonstrating confusability between specific emotions (Aviezer et al., 2008; Mondloch, 2012) may not be supported. To account for this, we provided four different situational stimuli for each context emotion to allow for variability among stimuli for all participants. Further piloting to ensure validity of context emotions would be helpful in future research.

Additionally, the present study focused on congruence effects between anger and disgust as well as fear and sadness given their associated facial expressions overlap in physical characteristics (Aviezer et al., 2008; Mondloch, 2012). Whether similar congruency effects would emerge with different emotions is not yet known. Prior studies examining these congruency effects used body postures (Aviezer et al., 2008; Leitzke & Pollak, 2016), not situational context. Whether this study's congruency effects were driven by the same mechanism that yielded changes in emotion recognition based on differing body postures is not known from these data. Further exploration of differential congruency effects by emotion and emotion combinations would be fruitful in expanding our understanding of the influence of context.

While the facial stimuli were designed to depict low, intermediate, and high intensity, participants did not distinguish between intermediate and high intensity. In other words, intensity ratings for all ages in the current study appeared to have plateaued at intermediate and high intensity. This is consistent with previous research showing stable performance in identifying

emotional expressions at higher levels of intensity (Pollak & Kistler, 2002). However, while intermediate intensity faces did not provide a distinct level of intensity, both intermediate and high intensity faces were still rated as higher than low intensity faces. This issue likely had a role in reducing variability in participants' responses as two groups of facial intensity intended to be different levels of intensity were not perceived to be distinct. Future research intending to capture linearity in intensity ratings would benefit from ensuring facial stimuli of varying levels are discriminated as intended.

The results from this study are informative for the identification of attention processes that may be implicated in certain social deficits and psychiatric disorders. However, this study did not test clinical populations, nor did it directly associate individual differences in performance with clinical outcomes. It will be informative to explore differences in functioning in specific populations of interest and determine whether the ability to rate facial intensity in context is related to social or emotional functioning. For example, it might be that overreliance on the face or context may be associated with interpersonal difficulties, problems at school, or psychopathology. Moving forward, it will be important to clarify these associations.

Conclusion

The present experiment examined the role of context and development on ratings of emotional intensity. The ability to discriminate between levels of emotion intensity and incorporate contextual information appears to be a developmental process whereby individuals learn to track competing signals and prioritize information processing based on facial and contextual cues. With age and experience, people learn which situations may result in more reliable emotional responses and which are likely to result in highly variable responses, requiring more information to make more accurate judgments. While context provides important

information about the likelihood of an individual experiencing a particular emotional state, ambiguous social situations require further information. In this way, information extracted from facial expressions may provide the necessary complementary information to hone emotion judgments.

This study provides a framework for characterizing abilities in discerning degrees of facial intensity when in competition with contextual information. Such findings may provide the impetus for expanding research on identifying maladaptive attentional processes and designing attention-based interventions. These types of procedures may benefit from the inclusion of context as well as using a more nuanced approach to measuring emotion ratings. The current research aims to further our understanding of the role of development in the relationship between learning, emotion processing, and social functioning.

References

Adams, R. B., Jr., Hess, U., & Kleck, R. E. (2015). The intersection of gender-related facial appearance and facial displays of emotion. *Emotion Review, 7,* 5-13. https://doi.org/10.1177/1754073914544407

Adams, R. B., Jr., Nelson, A. J., Soto, J. A., Hess, U., & Kleck, R. E. (2012). Emotion in the neutral face: A mechanism for impression formation? *Cognition and Emotion, 26,* 131-141. https://doi.org/10.1080/02699931.2012.666502

Addington, J., Saeedi, H., & Addington, D. (2006). Facial affect recognition: A mediator between cognitive and social functioning in psychosis? *Schizophrenia Research, 85*, 142-150. https://doi.org/10.1016/j.schres.2006.03.028

Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion in Neurobiology, 12(2)*, 169-177. https://doi.org/10.1016/S0959-4388(02)00301-X

Aviezer, H., Dudarev, V., Bentin, S., & Hassin, R. R. (2011). The automaticity of emotional face-context integration. *Emotion, 11*(6), 1406-1414. https:/doi.org/10.1037/a0023578

Aviezer, H., Ensenberg, N., & Hassin, R. R. (2017). The inherently contextualized nature of facial emotion perception. *Current Opinion in Psychology, 17,* 47-54. https://doi.org/10.1016/j.copsyc.2017.06.006

Aviezer, H., Hassin, R., Bentin, S., & Trope, Y. (2008). Putting facial expressions back in context. In N. Ambady & J. J. Skowronski (Eds.), *First Impressions* (pp. 255-286). New York, NY: The Guilford Press.

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(7), 724-732. https://doi.org/10.1111/j.1467-9280.2008.02148.x

Aviezer, H., Messinger, D. S., Zangvil, S., Mattson, W. I., Gangi, D. N., & Todorov, A. (2015). Thrill of victory or agony of defeat? Perceivers fail to utilize information in facial movements. *Emotion*, *15*(6), 791. https://doi.org/10.1037/emo0000073

Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science, 338*(6111), 1225-1229. https://doi.org/10.1126/science.1224313

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5,* 617-629. https://doi.org/10.1038/nrn1476

Barrett, L. F. (2017). *How emotions are made: The secret life of the brain.* New York, NY: Houghton Mifflin Harcourt.

Barrett, L. F., Lindquist, K. A., & Gendron, M. (2007). Language as context for the perception of emotion. *Trends in Cognitive Sciences, 11,* 327-332. https://doi.org/10.1016/j.tics.2007.06.003

Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science, 20*(5), 286-290. https://doi.org/10.1177/0963721411422522

Batty, M., & Taylor, M. J. (2006). The development of emotional face processing during childhood. *Developmental Science, 9*(2), 207-220. https://doi.org/10.1111/j.1467-7687.2006.00480.x

Brosch, T., Pourtois, G., & Sander, D. (2010). The perception and categorisation of emotional stimuli: A review. *Cognition and Emotion, 24*(3), 377-400. https://doi.org/10.1080/02699930902975754

Bruce, V., Campbell, R. N., Doherty-Sneddon, G., Langton, S., McAuley, S., & Wright, R. (2000). Testing face processing skills in children. *British Journal of Developmental Psychology, 18*(3), 319-333. https://doi.org/10.1348/026151000165715

Buck, R. (1994). Social and emotional functions in facial expression and communication: The readout hypothesis. *Biological Psychology, 38*, 95-115. https://doi.org/10.1016/0301-0511(94)90032-9

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*(1), 1-28. https://doi.org/10.18637/jss.v080.i01

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data?. *Perspectives on Psychological Science, 6*(1), 3-5. https://doi.org/10.1177/1745691610393980

Button, K., Lewis, G., Penton-Voak, I., Munafo, M. (2013). Social anxiety is associated with general but not specific biases in emotion recognition. *Psychiatry Research, 210,* 199-207. https://doi.org/10.1016/j.psychres.2013.06.005

Celani, G., Battacchi, M. W., Arcidiacono, L. (1999). The understanding of the emotional meaning of facial expressions in people with autism. *Journal of Autism and Developmental Disorders, 29,* 57-66. https://doi.org/10.1023/A:1025970600181

Calder, A. J., Burton, A. M., Miller, P., Young, A. W., & Akamatsu, S. (2001). A principal component analysis of facial expressions. *Vision Research, 41*(9), 1179-1208. https://doi.org/10.1016/S0042-6989(01)00002-5

Calder, A. J., Rowland, D., Young, A. W., Nimmo-Smith, I., Keane, J., & Perrett, D. I. (2000). Caricaturing facial expressions. *Cognition, 76*(2), 105-146. https://doi.org/10.1016/S0010-0277(00)00074-3

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*(2), 527. https://doi.org/10.1037/0096-1523.26.2.527

Calvo, M. G., & Nummenmaa, L. (2016). Perceptual and affective mechanisms in facial expression recognition: An integrative review. *Cognition and Emotion, 30*(6), 1081-1106. https://doi.org/10.1080/02699931.2015.1049124

Camras, L. A., & Allison, K. (1985). Children's understanding of emotional facial expressions and verbal labels. *Journal of Nonverbal Behavior, 9*(2), 84-94. https://doi.org/10.1007/BF00987140

Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology, 70*(2), 205. https://doi.org/10.1037/0022-3514.70.2.205

Chen, Z., & Whitney, D. (2019). Tracking the affective state of unseen persons. *Proceedings of the National Academy of Sciences, 201812250*. https://doi.org/10.1073/pnas.1812250116

de Gelder, B., Meeren, H. K., Righart, R., Van den Stock, J., Van de Riet, W. A., & Tamietto, M. (2006). Beyond the face: Exploring rapid influences of context on face processing. *Progress in Brain Research, 155*, 37-48. https://doi.org/10.1016/S0079-6123(06)55003-4

Derntl, B., Seidel, E-M., Kryspin-Exner, I., Hasmann, A., Dobmeier, M. (2009). Facial emotion recognition in patients with bipolar I and bipolar II disorder. *British Journal of Clinical Psychology, 48*, 363-375. https://doi.org/10.1348/014466509X404845

Eisenbarth, H., & Alpers, G. W. (2011). Happy mouth and sad eyes: Scanning emotional facial expressions. *Emotion, 11*(4), 860. https://doi.org/10.1037/a0022758

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin, 128*(2), 203-235. https://doi.org/10.1037/0033-2909.128.2.203

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion, 6*(3-4), 169-200. https://doi.org/10.1080/02699939208411068

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system.* Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., & Rosenberg, E. (2005). *What the face reveals* (2nd ed.). New York, NY: Oxford University Press.

Feinman, S. (1982). Social referencing in infancy. *Merrill-Palmer Quarterly, 28*(4), 445-470.

Gao, X. Q. & Maurer, D. (2009). Influence of intensity on children's sensitivity to happy, sad, and fearful facial expressions. *Journal of Experimental Child Psychology, 102*(4), 503-521. https://doi.org/j.jecp.2008.11.002

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Perceptions of emotion from facial expressions are not culturally universal: evidence from a remote culture. *Emotion, 14(2),* 251. https://doi.org/10.1037/a0036052

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. https://doi.org/10.1023/A:1024952730333

Hollingshead, A. B. (2011). Four factor index of social status. In P. Smith (Ed.), *Yale Journal of Sociology* (pp. 21-52). New Haven, CT: Department of Sociology at Yale University.

Hu, C. S., Wang, Q., Han, T., Weare, E., & Fu, G. (2017). Differential emotion attribution to neutral faces of own and other races. *Cognition & Emotion, 31*(2), 360-368. https://doi.org/10.1080/02699931.2015.1092419

Izard, C. E. (1992). Basic emotions, relations among emotions, and emotion-cognition relations. *Psychological Review, 99*(3), 561-565. http://doi.org/10.1037/0033-295X.99.3.561

Jayaraman, S., Fausey, C. M., & Smith, L. B. (2015). The faces in infant-perspective scenes change over the first year of life. *PLOS ONE*, *10*(5), e0123780. https://doi.org/10.1371/journal.pone.0123780

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General, 142*(2), 573. https://doi.org/10.1037/a0029146

Lawrence, K., Campbell, R., & Skuse, D. (2015). Age, gender, and puberty influence the development of facial emotion recognition. *Frontiers in Psychology, 6,* 761. https://doi.org/10.3389/fpsyg.2015.00761

Lee, T. H., Choi, J. S., & Cho, Y. S. (2012). Context modulation of facial emotion perception differed by individual difference. *PLOS ONE, 7*. https://doi.org/10.1371/journal.pone.0032987

Leitzke, B. T., & Pollak, S. D. (2016). Developmental changes in the primacy of facial cues for emotion recognition. *Developmental Psychology, 52*(4), 572. https://doi.org/10.1037/a0040067

Leppanen, J. M. & Nelson, C. A. (2006). The development and neural bases of facial emotion recognition. *Advances in Child Development and Behavior, 34,* 207-246. https://doi.org/10.1016.S0065-2047(06)80008-X

Marsh, A. A., Elfenbein, H. A., & Ambady, N. (2003). Nonverbal "accents" cultural differences in facial expressions of emotion. *Psychological Science, 14*(4), 373-376. https://doi.org/10.1111/1467-9280.24461

Martinez, A. M. (2017). Computational models of face perception. *Current Directions in Psychological Science, 26*(3), 263-269. https://doi.org/10.1177/0963721417698535

Martinez, A. M. (2019). Context may reveal how you feel. *Proceedings of the National Academy of Sciences, 201902661.* https://doi.org/10.1073/pnas.1902661116

Martinez, L., Falvello, V. B., Aviezer, H., & Todorov, A. (2016). Contributions of facial expressions and body language to the rapid perception of dynamic emotions. *Cognition & Emotion, 30*(5), 939-952. https://doi.org/10.1080/02699931.2015.1035229

McCrimmon, A. W., & Smith, A. D. (2013). Review of the Wechsler Abbreviated Scale of Intelligence, (WASI-II). *Journal of Psychoeducational Assessment, 3,* 337-341. https://doi.org/10.1177/0734282912467756

Mogoaşe, C., David, D., & Koster, E. H. (2014). Clinical efficacy of attentional bias modification procedures: An updated meta-analysis. *Journal of Clinical Psychology, 70*(12), 1133-1157. https://doi.org/10.1002/jclp.22081

Mondloch, C. J. (2012). Sad or fearful? The influence of body posture on adults' and children's perception of facial displays of emotion. *Journal of Experimental Child Psychology, 111*(2), 180-196. https://doi.org/10.1016/j.jecp.2011.08.003

Mondloch, C. J., Geldart, S., Maurer, D., & Le Grand, R. (2003). Developmental changes in face processing skills. *Journal of Experimental Child Psychology, 86*(1), 67-84. https://doi.org/10.1016/S0022-0965(03)00102-4

Mondloch, C. J., Horner, M., & Mian, J. (2013). Wide eyes and drooping arms: Adult-like congruency effects emerge early in the development of sensitivity to emotional faces and body postures. *Journal of Experimental Child Psychology, 114*(2), 203-216. https://doi.org/10.1016/j.jecp.2012.06.003

Morales, S., Fu, X., & Pérez-Edgar, K. E. (2016). A developmental neuroscience perspective on affect-biased attention. *Developmental Cognitive Neuroscience, 21*, 26–41. https://doi.org/10.1016/j.dcn.2016.08.001

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review, 98*(2), 164. https://doi.org/10.1037/0033-295X.98.2.164

Noh, S. R., & Isaacowitz, D. M. (2013). Emotional faces in context: Age differences in recognition accuracy and scanning patterns. *Emotion, 13*(2), 238. https://doi.org/10.1037/a0030234

Philippot, P., Kornreich, C., Blairy, S., Baert, I., Dulk, A. D., Bon, O. L., Streel, E., Hess, U., Pelc, I., Verbanck, P. (1999). Alcoholics' deficits in the decoding of emotional facial expression. *Alcoholism: Clinical and Experimental Research, 23*, 1031-1038. https://doi.org/10.1097/00000374-199906000-00010

Plate, R. C., Bloomberg, Z., Bolt, D. M., Bechner, A. M., Roeber, B. J., & Pollak, S. D. (2019). Abused children experience high anger exposure. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00440

Pollak, S. D., Messner, M., Kistler, D. J., & Cohn, J. F. (2009). Development of perceptual expertise in emotion recognition. *Cognition, 110*(2), 242-247. https://doi.org/10.1016/j.cognition.2008.10.010

Psychology Software Tools, Inc. [E-Prime 3.0]. (2016). Retrieved from http://www.pstnet.com.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rajhans, P., Jessen, S., Missana, M., & Grossmann, T. (2016). Putting the face in context: Body expressions impact facial emotion processing in human infants. *Developmental Cognitive Neuroscience, 19*, 115-121. https://doi.org/10.1016/j.dcn.2016.01.004

Righart, R., & De Gelder, B. (2008). Rapid influence of emotional scenes on encoding of facial expressions: An ERP study. *Social Cognitive and Affective Neuroscience, 3*(3), 270-278. https://doi.org/10.1093/scan/nsn021

Rodger, H., Vizioli, L., Ouyang, X., & Caldara, R. (2015). Mapping the development of facial expression recognition. *Developmental Science, 18*(6), 926-939. https://doi.org/10.1111/desc.12281

Rubinow, D. R., Post, R.M. (1992). Impaired recognition of affect in facial expression in depressed patients. *Biological Psychiatry, 31*, 947-953. https://doi.org/10.1016/0006-3223(92)90120-O

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

Russell, J. A. (1997). Reading emotions from and into faces: Resurrecting a dimensional contextual perspective. In J. A. Russell & J. M. Fernandez-Dos (Eds.), *The Psychology of Facial Expressions* (pp. 295-320). New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9780511659911.015

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145. https://doi.org/10.1037/0033-295X.110.1.145

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*(6), 2408-2412. https://doi.org/10.1073/pnas.0908239106

Schwarzer, G. (2000). Development of face processing: The effect of face inversion. *Child Development, 71*(2), 391-401. https://doi.org/10.1111/1467-8624.00152

Schyns, P. G. & Oliva, A. (2010). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science, 5,* 195-200. https://doi.org/10.1111/j.1467-9280.1994.tb00500.x

Sorce, J. F., Emde, R. N., Campos, J. J., & Klinnert, M. D. (1985). Maternal emotional signaling: its effect on the visual cliff behavior of 1-year-olds. *Developmental Psychology, 21*(1), 195. https://doi.org/10.1037/0012-1649.21.1.195

Srinivasan, R., & Martinez, A. M. (2018). Cross-cultural and cultural-specific production and perception of facial Expressions of emotion in the wild. *IEEE Transactions on Affective Computing, 14*, 1-12. https://doi.org/10.1109/TAFFC.2018.2887267

Stan Development Team (2018). RStan: The R interface to Stan. R package version 2.17.3. http://mc-stan.org/.

Susskind, J. M., Littlewort, G., Bartlett, M. S., Movellan, J., & Anderson, A. K. (2007). Human and computer recognition of facial expressions of emotion. *Neuropsychologia, 45*(1), 152-162. https://doi.org/10.1016/j.neuropsychologia.2006.05.001

Thomas, L. A., De Bellis, M. D., Graham, R., & LaBar, K. S. (2007). Development of emotional facial recognition in late childhood and adolescence. *Developmental Science, 10*(5), 547-558. https://doi.org/10.1111/j.1467-7687.2007.00614.x

Tiddeman, B. P., Stirrat, M. R., & Perrett, D. I. (2005). Towards realism in facial image transformation: Results of a wavelet MRF method. *Computer Graphics Forum, 24*(3), 449-456. https://doi.org/10.1111/j.1467-8659.2005.00870.x

Todd, R. M., Cunningham, W. A., Anderson, A. K., & Thompson, E. (2012). Affect-biased attention as emotion regulation. *Trends in Cognitive Sciences, 16*(7), 365-372. https://doi.org/10.1016/j.tics.2012.06.003

Van den Stock, J., & de Gelder, B. (2012). Emotional information in body and background hampers recognition memory for faces. *Neurobiology of Learning and Memory, 97*(3), 321-325. https://doi.org/10.1016/j.nlm.2012.01.007

Weschler, D. (2011). *Weschler abbreviated scale of intelligence, Second edition (WASI-II)*. San Antonio, TX: NCS Pearson.

Whelan, R. (2008). Effective analysis of reaction time data. *Psychological Record. 58*(3), 475-482. https://doi.org/10.1007/BF03395630

White, H., Chroust, A., Heck, A., Jubran, R., Galati, A., & Bhatt, R. S. (2019). Categorical perception of facial emotions in infancy. *Infancy, 24*(2), 139-161. https://doi.org/10.1111/infa.12275

Widen, S. C., & Russell, J. A. (2003). A closer look at preschoolers' freely produced labels for facial expressions. *Developmental Psychology, 39*(1), 114. https://doi.org/10.1037/0012-1649.39.1.114

Widen, S. C., & Russell, J. A. (2008). Children acquire emotion categories gradually. *Cognitive Development, 23*(2), 291-312. https://doi.org/10.1016/j.cogdev.2008.01.002

Xu, Q., Yang, Y., Tan, W., & Zhang, L. (2017). Facial expressions in context: Electrophysiological correlates of the emotional congruency of facial expressions and

background scenes. *Frontiers in Psychology, 8,* 2175. https://doi.org/10.3389/fpsyg.2017.02176

Yuki, M., Maddux, W. W., & Masuda, T. (2007). Are the windows to the soul the same in the East and West? Cultural differences in using the eyes and mouth as cues to recognize emotions in Japan and the United States. *Journal of Experimental Social Psychology, 43*(2), 303-311. https://doi.org/10.1016/j.jesp.2006.02.004

Table 1

Intelligence (IQ) and Socioeconomic Status (SES) by Age Group

Age Group	IQ	SES
Children	112.4 (12.8)	50.9 (11.8)
Adolescents	113.6 (13.8)	53.2 (10.0)
Adults	106.1 (12.0)	50.2 (13.4)

Note. Values represent means (standard deviation). IQ values from Weschler Abbreviated Scale for Intelligence, 2nd edition (WASI-II; Weschler, 2011). SES (Socioeconomic Status) values from Hollingshead two factor index of social status (Hollingshead, 2011) which ranges from 8-66.

Table 2

Estimates and 95% CI of Random Effects

Stimulus Model	Face Emotion	*b*	95% CI Lower	95% CI Upper
	Anger	0.432	-0.443	1.378
	Disgust	0.973	0.112	1.909
	Fear	0.713	-0.149	1.659
	Sad	-0.025	-0.894	0.912
	Anger	-0.123	-0.880	0.932
	Disgust	0.849	-0.026	1.788
	Fear	0.339	-0.523	1.281
	Sad	0.571	-0.293	1.510
	Anger	0.577	-0.292	1.509
	Disgust	1.093	0.225	2.034
	Fear	0.610	-0.263	1.551
	Sad	0.367	-0.504	1.305
	Anger	-0.050	-0.909	0.880
	Disgust	1.325	0.465	2.259
	Fear	0.078	-0.793	1.012
	Sad	0.705	-0.156	1.647
Context Emotion				
Anger		0.500	-0.401	1.458
Disgust		0.618	-0.283	1.575
Fear		0.534	-0.369	1.494
Sad		0.480	-0.423	1.434

Note. All model images depict anger at low (20%) intensity. Models are, from top to bottom, F17, F19, M01, and M07 from the International Affective Science Laboratory (IASLab) Face Set.

Table 3

Summary Statistics for Intensity Ratings in Congruent and Non-congruent contexts by Face Intensity and Age Group

				95% CI	
Congruence	Intensity	Age	*Mdn*	Lower	Upper
Congruent	Low	Children	59.3	50.7	67.6
		Adolescents	47.3	38.6	56.1
		Adults	51.1	42.2	60.1
	Intermediate	Children	74.5	65.0	81.9
		Adolescents	68.6	58.1	77.3
		Adults	68.0	57.3	77.0
	High	Children	76.6	65.1	85.1
		Adolescents	75.3	63.4	84.5
		Adults	73.6	61.0	83.0
Non-congruent	Low	Children	46.5	37.4	56.3
		Adolescents	45.1	36.6	53.9
		Adults	44.3	35.3	53.8
	Intermediate	Children	68.1	57.5	76.8
		Adolescents	66.7	55.8	75.6
		Adults	66.1	55.0	75.4
	High	Children	70.5	57.5	80.6
		Adolescents	73.6	61.4	83.3
		Adults	71.8	58.8	81.6

Note. Values represent raw values converted from log odds. Intensity levels represent 20% (low), 50% (intermediate), and 80% (high) morphs.

Figure 1. Example of low, intermediate, and high intensity for fearful facial expression. Model depicted is model F17 from the International Affective Science Laboratory (IASLab) Face Set.

Figure 2. Exemplars of Face Emotion/Context composites. All four emotions (anger, disgust, sad, fear) depicted in congruent, incongruent, and neutral contexts. Model depicted is model F17 from the International Affective Science Laboratory (IASLab) Face Set expressing high (80%) intensity for each face emotion.

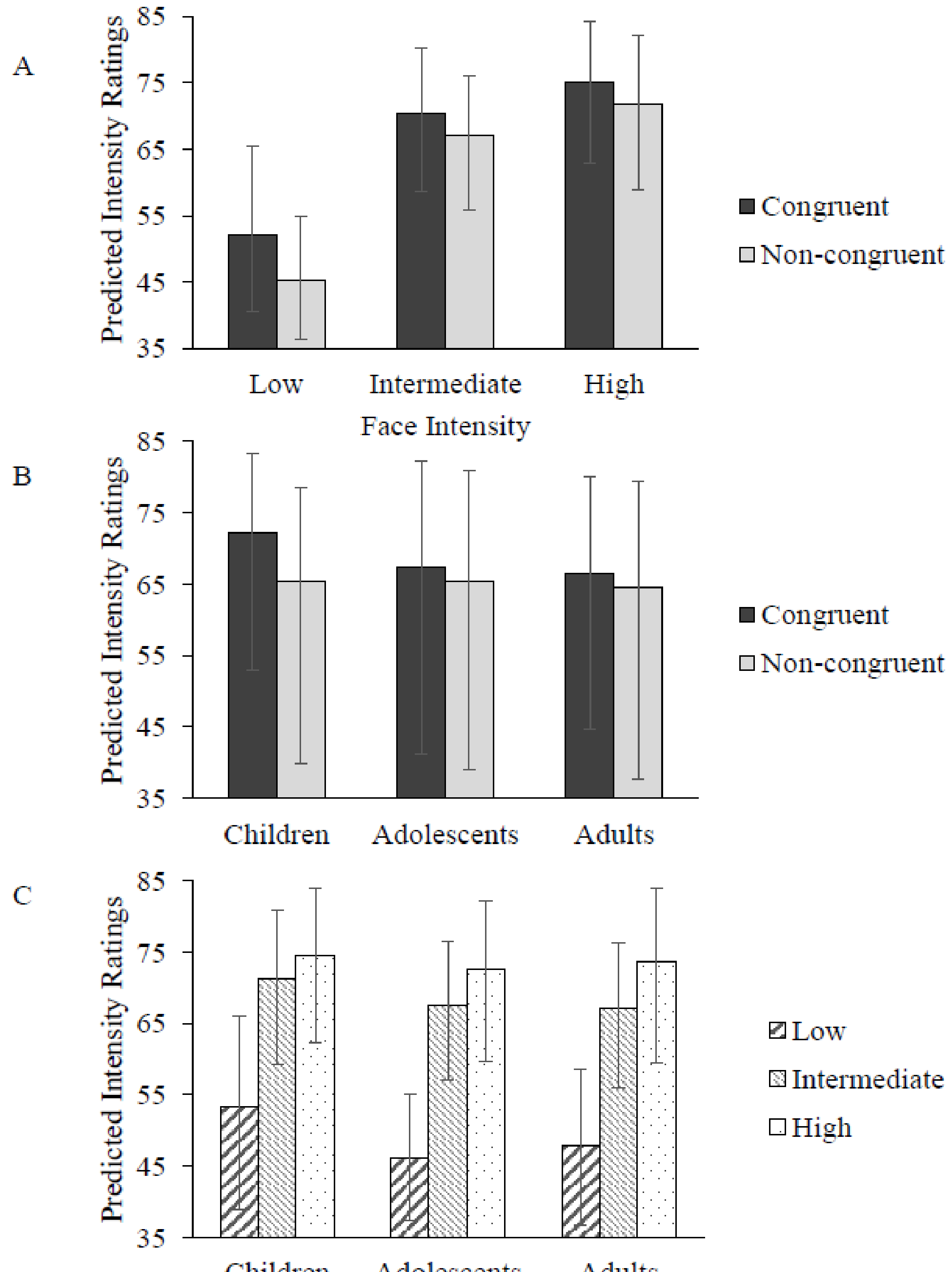

Figure 3. Predicted intensity ratings for congruent and non-congruent contexts by face intensity (A), for congruent and non-congruent contexts between children, adolescents, and adults (B), and for low, intermediate, and high face intensity between children, adolescents, and adults (C). Predicted ratings reflect posterior samples calculated from generalized, multi-level Bayesian regression. Error bars represented 95% credible intervals.