AMERICAN PSYCHOLOGICAL ASSOCIATION

Emotion

ISSN: 1528-3542

2025, Vol. 25, No. 3, 588–600
https://doi.org/10.1037/emo0001441

Perceptual Sensitivity to Labeling Stereotyped Emotion Expressions: Associations With Age and Subclinical Psychopathology Symptoms From Childhood Through Early Adulthood

David G. Weissman[1, 2], Henna I. Vartiainen[3], Erik C. Nook[3], Hilary K. Lambert[4, 5], Stephanie F. Sasse[1], Leah H. Somerville[1], and Katie A. McLaughlin[1, 6]

[1] Department of Psychology, Harvard University
[2] Department of Psychology, California State University, Dominguez Hills
[3] Department of Psychology, Princeton University
[4] Behavioral Health Partial Hospital Program, McLean Hospital, Belmont, Massachusetts, United States
[5] Department of Psychiatry, Harvard Medical School, Harvard University
[6] The Ballmer Institute for Children's Behavioral Health, University of Oregon

This study investigates (a) age-related differences in how the intensity of stereotyped facial expressions influence the emotion label children, adolescents, and adults assign to that face and (b) how this perceptual sensitivity relates to subclinical symptoms of psychopathology. In 2015–2016, 184 participants aged 4–25 years viewed posed stereotypes of angry, fearful, sad, and happy expressions morphed with neutral expressions at 10%–90% intensity. Thin plate regression smoothing splines were used to chart nonlinear associations between age and the perceptual threshold participants needed to assign the emotion label expected based on cultural consensus. Results suggest that sensitivity to labeling stereotypical happy faces as "happy" peaked by age 4. Sensitivity to perceiving stereotypical angry faces as "angry" increased from ages 4 to 7 and then plateaued. In contrast, sensitivity to perceiving stereotypical fearful and sad faces demonstrated protracted development, not reaching a plateau until ages 15 and 16, respectively. Reduction in selecting the "I don't know" response was the primary driver of these age-related changes. Stereotyped fear expressions required the highest intensity to be labeled as such and showed the most marked change in perceptual threshold across development. Interestingly, lower intensity morphs of stereotypical fear faces were frequently labeled "sad." Furthermore, perceiving lower intensity fear morphs was associated with fewer internalizing and externalizing symptoms in participants aged 7–19. This study describes the development of perceptual sensitivity to labeling stereotypical expressions of emotion according to cultural consensus and shows that how people perceive and categorize ambiguous facial expressions is associated with vulnerability to psychopathology.

Keywords: emotion expression labelling, fear, development, spline

Humans commonly interpret emotional meaning in the facial configurations of others. Though these expressions can be highly stereotyped and not necessarily faithful to an individual's internal affective state, understanding how individuals interpret facial expressions and what categories they use to label these expressions is useful as a lens into individuals' conceptualization of what different emotions "look like." Indeed, emerging theories posit that people deploy *concepts* to construct an understanding of what other

This article was published Online First October 7, 2024.
Amanda Guyer served as action editor.
David G. Weissman https://orcid.org/0000-0001-9966-3338
The authors report how we determined sample size, all data exclusions, all manipulations, and all measures in the study, and we follow the journal article reporting standards (Kazak, 2018). This study was not preregistered. All data and code for these analyses are publicly available at https://osf.io/6jms7/.

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or nonfinancial interest in the subject matter or materials discussed in this article. This research was funded by the National Institute of Mental Health (R37-MH119194 to Katie A. McLaughlin and K99-MH127248 to David G. Weissman), an Early Career Research Fellowship from the Jacobs Foundation (Katie A. McLaughlin), a OneMind Institute Rising Star Award (Katie A. McLaughlin), and a National Science Foundation Graduate Research Fellowship (DGE1144152 to Erik C. Nook).

David G. Weissman played a lead role in conceptualization, formal analysis, visualization, and writing–original draft. Henna I. Vartiainen played a supporting role in conceptualization, formal analysis, visualization, and writing–review and editing. Erik C. Nook played a lead role in data curation, investigation, and methodology and a supporting role in conceptualization, formal analysis, supervision, and writing–review and editing. Hilary K. Lambert played a supporting role in conceptualization, data curation, investigation, and project administration. Stephanie F. Sasse played a supporting role in data curation and investigation. Leah H. Somerville played a supporting role in conceptualization, funding acquisition, methodology, resources, supervision, and writing–review and editing. Katie A. McLaughlin played a lead role in conceptualization, funding acquisition, supervision, and writing–review and editing and a supporting role in investigation, methodology, and resources.

Correspondence concerning this article should be addressed to David G. Weissman, Department of Psychology, California State University, Dominguez Hills, 1000 East Victoria Street, Carson, CA 90747-0001, United States. Email: dweissman@csudh.edu

people feel based on ambiguous information (Barrett et al., 2019; Gendron & Barrett, 2018; Hoemann et al., 2019; Nook et al., 2015), meaning that how people perceive others' emotions is a reflection of their emotion concept knowledge. The present study examines age-related changes in this aspect of emotion concept knowledge, specifically characterizing how the strength of visual cues affects emotion conceptualization at different ages.

Labeling of facial configurations is often referred to according to its "accuracy," presuming a one-to-one mapping between a suite of facial configurations (i.e., an "expression") and an internal emotional state held by the expressor. Indeed, some traditions in affective science have argued these mappings are veridical expressions of internal emotion, and although they can be posed, they are largely faithful to the internal state of the expresser (e.g., Ekman et al., 1969, 1987; critiqued by Russell, 1994). An alternative viewpoint presumes that this one-to-one link emerges because cultures build scripts that bind stereotyped expressions with emotion concepts (e.g., scowls and anger, smiles and joy; Barrett et al., 2019; Gendron & Barrett, 2018; Hoemann et al., 2019; Nook et al., 2015). Under this view, these expressions are not assumed to "accurately" reflect an individual's internal state, as such accuracy requires nuanced synchrony between expresser in how they understand the expresser's internal states (Barrett et al., 2019; Gendron & Barrett, 2018; Hoemann et al., 2019; Nook et al., 2015). Nonetheless, how people match stereotyped expressions to culturally agreed-upon labels gives insight into their internal conceptual system, even if this has limited connection to accuracy of emotion perception "in the wild." Thus, the emotion label that a cultural script typically assigns to a stereotyped expression can be called a *consensus judgment*.

Facility at labeling facial configurations in alignment with consensus judgments is important for the development of social competency (Nowicki & Duke, 1992; Nowicki & Mitchell, 1998). Conversely, using labels that are aberrant relative to consensus judgments has been documented in children and adolescents with several forms of psychopathology (Boakes et al., 2007; Collin et al., 2013; Ellis et al., 1997; Martin-Key et al., 2018; Simonian et al., 2001; Singh et al., 1998). This study investigates age-related differences in the amount of perceptual information needed to categorize stereotypical facial expressions in line with the consensus judgments, which we call *perceptual sensitivity*. Further, we examine the associations of this perceptual sensitivity with subclinical psychopathology symptoms in a community sample.

Studies show that judging faces in alignment with cultural consensus improves across development (Carey et al., 1980; Durand et al., 2007; Mondloch et al., 2003). However, stimuli used in studies on the development of emotion expression perception often depict highly intense facial configurations unlike those usually observed in daily life and that are not being experienced emotionally by the actor posing them (Barrett et al., 2019; Somerville & Whalen, 2006). Participants' facility at applying the consensus labels to these expressions is likely supported by the use of forced choice paradigms, where limited options are available and participants are forced to choose between them, which artificially inflate consensus, compared to open ended responses (DiGirolamo & Russell, 2017; Gendron et al., 2018; Hoemann et al., 2019). These tasks may therefore be limited in what they can tell us about how people perceive emotions in facial expressions in daily life. At the same time, literacy in matching stereotypical posed expressions with the consensus label does appear to increase with age (Gao & Maurer, 2009, 2010; Rodger et al., 2018; Thomas et al., 2007) and relate to lower mental health vulnerability (Collin et al., 2013; Ellis et al., 1997) and may therefore nonetheless serve as a proxy for affective skills that are important for healthy emotional functioning.

Prior research has demonstrated that older children exhibit increasingly robust perceptual sensitivity. This increased sensitivity in linking emotion words with more subtle facial cues explains a large portion of age-related gains in consensus matching of facial configurations across childhood (Barisnikov et al., 2021). Several studies have examined age-related differences in perceptual sensitivity to facial configurations by manipulating the intensity of the intended emotion being displayed, with varying results. One study found no age-related differences among 4–15 year olds in sensitivity to categorizing facial configurations at different intensities (Herba et al., 2006). Another study with an age range of 7–57 years found that sensitivity to labeling scowling faces as "angry" at lower intensity increased from adolescence to adulthood, whereas the sensitivity to label wide eyed, grimacing faces with the label "fear" showed a more gradual increase across from childhood to early adulthood (Thomas et al., 2007). Two other studies found that sensitivity to labeling smiling faces as "happy" at lower intensity reached consensus with adults by the age of 5, demonstrating no age-related change after that. Between 5 and 10 years, sensitivity to applying surprise, disgust, and fear labels in line with cultural consensus to expressions with lower intensity rose, and sensitivity to applying sadness and anger labels according to cultural consensus continued to rise into adulthood (Gao & Maurer, 2009, 2010). The largest and most recent study on this topic found that across ages, participants were least likely to apply the label "scared" and most likely to apply the label "happy," but reported no age-related changes in sensitivity to the intensity of those stereotyped facial configuration from age 5 to 25 years; whereas sensitivity to applying sad, angry, disgust, and surprise labels to lower intensities of the corresponding posed expressions increased with age (Rodger et al., 2018). Together these studies consistently suggest that sensitivity to labeling happiness based on less intensely smiling faces reaches adultlike levels by early childhood, but age-related patterns of sensitivity to assigning consensus labels to less intense negative facial expressions is less consistent across studies. This may be due to having used small samples, a focus only on linear associations between age and perceptual sensitivity or designs that involve binning participants into age groups.

Besides these relations with age, deviations in consistently applying the consensus labels to facial expressions have been documented in children and adolescents with numerous forms of psychopathology (Collin et al., 2013; Ellis et al., 1997), including bipolar disorder (Brotman et al., 2008), social phobia (Simonian et al., 2001), conduct disorder (Martin-Key et al., 2018), and attention deficit hyperactivity disorder (Boakes et al., 2007; Singh et al., 1998). However, studies have often been small and tended to focus on individual disorders. Difficulties applying consensus labels to posed facial configurations intended to depict emotional expressions, which are associated with most forms of psychopathology in children and adolescents, do not tend to be specific to any emotion category, or when they are, are specific to the configurations that are least consistently labeled, on average (e.g., fear and disgust). However, there is consistent evidence in adults that antisocial behaviors tend to be specifically associated with a decreased tendency to label facial configurations as fearful, even accounting for the lesser overall tendency of individuals to perceive fear in posed expressions (Marsh & Blair, 2008).

There are many reasons why less sensitivity to perceiving emotion in facial configurations may be related to psychopathology. Difficulty interpreting social–emotional cues communicated by facial expressions could contribute to greater distress and worry and impair social functioning. Alternatively, or in addition, greater sensitivity to perceiving facial expressions, even posed expressions, may be reflective of a broader capacity in categorizing, conceptualizing, and labeling emotions that contributes to both greater sensitivity perceiving emotion in facial expressions and lower psychopathology. Having well differentiated emotion concepts, or high emotional granularity, is a skill that has consistently been found to support coping with psychosocial stress in the short term and contribute to well-being and resilience over time (Barrett et al., 2001; Lennarz et al., 2018; Nook et al., 2021; Starr et al., 2020). Conversely, low emotional awareness, or alexithymia, is consistently associated with both difficulty labeling one's own and others' emotions and with psychopathology across development (Aaron et al., 2018; Grynberg et al., 2012; Hemming et al., 2019; Weissman et al., 2020).

Sensitivity to matching facial expressions with emotion words increase at different rates and reach adult levels at different ages, depending on the emotion in question (Gao & Maurer, 2009; Rodger et al., 2018). The present study builds on and extends past work by examining nonlinear age trajectories of sensitivity to labeling posed expressions that are interpreted as conveying fear, anger, sadness, and happiness in a sample spanning ages 4–25 years, as well as the association between sensitivity to perceiving each facial expression and transdiagnostic (internalizing and externalizing) psychopathology. We hypothesized that we would replicate the findings of Rodger et al. (2018) demonstrating age-related increases in perceptual sensitivity for negatively valenced facial expressions, but not happiness. Further, we aimed to extend these findings by (a) characterizing the nonlinear age trajectories for each emotion, (b) testing for systematic patterns in youths' responses that differed from the cultural consensus emotion of the target expression, and (c) assessing whether sensitivity to perceiving facial expressions associated with negative emotion is linked with subclinical psychopathology symptoms even after accounting for age-related increases in psychopathology symptoms.

Method

Participants

Data were drawn from a cross-sectional study of emotional development conducted in a community sample of 203 participants spanning ages 4–25 years. The age distribution was uniform, with nine to 11 participants in each 1 year age bin between age 4 and 21 and 21 participants between the ages of 22 and 25. Of the 203 participants, 190 completed the expression labeling task, six of whom were excluded from analyses because their responses were unreliable (see details in Measures section). Thus, the present study includes data from 184 participants (97 female). This exceeds the sample size of any previous study on the development of perceptual sensitivity to our knowledge and is sufficient to detect effect sizes of β = .21 at 80% power. Participants' races were Asian or Pacific Islander (n = 18, 9%), Black (n = 22, 12%), Middle Eastern descent (n = 3, 2%), multiracial (n = 22, 12%), Native American (n = 1, 1%), and White (n = 116, 61%). Five participants (3%) declined to self-report their race. Sixteen participants (8%) were Hispanic or Latinx. Income-to-needs ratios ranged from 0.07 to 13.32, (M = 4.95, SD = 3.12, 9.23% below poverty line). All participants were native English speakers who were compensated for their time and recruited from communities surrounding Harvard University and the University of Washington in 2015–2016. The sample recruited from the Harvard site was considerably older and more female on average than the University of Washington site. Because of this, sensitivity analyses were conducted controlling for site. Results of these analyses can be found with the reproducible analysis code (RMarkdown) https://osf.io/6jms7/ (Weissman, 2023). As expected, given the collinearity between site and age, the estimates are changed somewhat in these analyses, but the overall results are very similar. Participants provided informed written consent/assent, and minor participants received written permission for their participation from a parent or legal guardian. The study was performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki. The Committee on the Use of Human Subjects at Harvard University (Institutional Review Board #15–2214: "Development of fundamental emotion processes") and the University of Washington Institutional Review Board (#50239: "Typical emotional development") approved all research procedures.

Measures

Facial Emotion Expression Perception

Stimuli included morphs of faces drawn from the NimStim stimulus set (Tottenham et al., 2009). Images used were faces intended to depict prototypical (closed mouth) expressions of four different emotions (angry, fear, happy, and sad). We started by selecting two expressions of eight different actors (four male and four female; five White, two Black, and one Asian). In the initial validation study (Tottenham et al., 2009), adults matched the angry, happy, and sad stereotypical expressions used in this study with their consensus labels 89%–98% of the time. Because stereotyped fear faces were labeled as such less frequently in this initial validation study (30%–80% of the time), we chose stimuli to include in this task that were labeled as fear more consistently (64%–74%). Thus, the sources for all stereotyped facial expressions used in this study are normatively matched with heir consensus label more often than not.

These stereotypical facial expressions were each morphed with "calm" expressions for each actor to produce a set of stimuli with fine-grained variations in emotion intensity. Specifically, faces labeled sad, angry, fearful, and happy were each morphed with calm faces at 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, and 90% intensity thresholds. This resulted in nine intensities of the 16 faces (i.e., 144 stimuli total), with each expression at each morph presented four times by four different actors. We evenly split trials into two runs. Each trial was preceded by a 500 ms blank intertrial interval followed by a 500 ms fixation cross. Each morph was presented for 1 s followed by a self-paced response window, during which participants were given six response options: "nothing," "angry," "scared," "happy," "sad," and "I don't know." Trials in which response times were less than 200 ms were excluded from analyses (3.4% of all trials), as it is implausible that participants could perceive and evaluate facial expressions that quickly (Ratcliff, 1993; Whelan, 2008). Six participants were excluded from analyses because more than 20% of their trials had to be excluded.

In addition, participants were excluded from analysis of perceptual sensitivity for a given emotion category if they were unable to define that emotion on an emotion vocabulary task. This task, as described in detail elsewhere (Nook et al., 2020), was created by adapting the format of the vocabulary test of the Wechsler Abbreviated Scale of

Intelligence (Wechsler, 2011). If participants were unable to provide a definition of "angry," "scared," "happy," or "sad" that was either a general definition of the word, a synonym of the word, or an example of a situation that would give rise to that emotion and not others, a perceptual sensitivity threshold was not computed for that participant, and they were excluded from further analyses using that threshold. This resulted in exclusion of 10 participants for "angry," seven participants for "scared," 12 participants for "happy," and 14 participants for "sad." Participants who were excluded from analyses were significantly younger than participants who were included (Cohen's $d = -1.11$, $p < .001$) but did not differ significantly in sex ($\chi^2 = 0.53$, $p = .466$).

Symptoms of Psychopathology

Psychopathology measures were validated in prior research for participants aged 7–19 years. Thus, our analytic sample for analyses of associations between emotion expression perception and psychopathology were restricted to the 120 participants (62 female) in this age range. Depression symptoms were assessed via child self-report with the Children's Depression Inventory–2, a recently revised version of the widely used self-report measure of depressive symptoms in children and adolescents (Kovacs, 2011). The Children's Depression Inventory–2 demonstrated excellent internal consistency in our sample ($\alpha = .90$). Anxiety symptoms were assessed via child self-report with the Screen for Child Anxiety Related Emotional Disorders, which measures anxiety disorder symptoms across five domains: panic/somatic, generalized anxiety, separation anxiety, social phobia, and school phobia (Birmaher et al., 1997). The Screen for Child Anxiety Related Emotional Disorders has excellent internal consistency in our sample ($\alpha = .93$). The Children's Depression Inventory–2 and Screen for Child Anxiety Related Emotional Disorders were both child self-report, as it is well established that children provide more valid reports of internalizing problems than parents (Aebi et al., 2017; Angold et al., 1987; Bird et al., 1992; Cantwell et al., 1997; Grills & Ollendick, 2003; Moretti et al., 1985).

Externalizing symptoms were assessed using both child and caregiver reports on the Youth Self-Report (YSR) and Child Behavior Checklist (CBCL; Achenbach, 1991). The YSR/CBCL scales are among the most widely used measures of youth emotional and behavioral problems. Following prior work (Kessler et al., 2012; Merikangas et al., 2010), the higher raw score between the CBCL and YSR was used from the Attention Problems, Rule Breaking Behaviors, and Aggressive Behavior subscales as our measures of these constructs. The use of the higher of parent or child report on the CBCL and YSR is an implementation of the standard "or" rule used in combining parent and child reports of psychopathology. In this approach, if either a parent or child endorses a particular symptom, it is counted, and the reporter endorsing the higher level of symptoms or impairment is used. This is a standard approach in the literature on child psychopathology—for example, it is how mental disorders are diagnosed in population-based studies of psychopathology in children and adolescents (e.g., Kessler et al., 2012; Merikangas et al., 2010). However, in interpreting our findings, it is important to acknowledge that the measure of externalizing problems is based on multiple reporters and is therefore less dependent on adolescents' own perceptions of their symptoms compared to the measure of internalizing problems. Descriptive statistics and intercorrelations among all psychopathology measures are provided in Table 1.

As reported previously (Weissman et al., 2020), we performed a confirmatory factor analysis (CFA) to fit a correlated-factors model specifying Internalizing and Externalizing latent factors. In order to ensure that our latent factors were not being driven by one or more indicators simply because of measurement differences across psychopathology instruments (i.e., different number of items, scoring, etc.), we binned scores on each indicator into deciles prior to CFA analyses. We used this method to transform skewed and zero-inflated data and place it on the same scale for CFA. All CFA analyses were performed in MPlus Version 8.1. Given that our observed indicator variables were slightly skewed and kurtotic, we used the robust maximum likelihood estimator (MLR), which employs a sandwich estimator to arrive at standard errors robust to nonnormality of observations. Latent factor scores for internalizing and externalizing psychopathology were extracted from the model for further analyses.

Analysis

Age-Related Patterns in Perceptual Sensitivity Thresholds

A Weibull function was fit to the data with the quickpsy package (Linares & López-Moliner, 2016) in R Version 4.0.0 for each facial stimulus viewed by each participant to estimate the perceptual sensitivity threshold, defined as the percentage morph from 0 to 100 at which the probability was 0.5 that the consensus category was

Table 1
Descriptive Statistics and Correlations of Psychopathology Measures

				Correlation							
Psychopathology measure	*N*	*M*	*SD*	1	2	3	4	5	6	7	8
1. CDI-2	103	6.29	6.56	—							
2. SCARED	101	13.8	12.0	.624*	—						
3. YSR attention problems	120	5.66	3.21	.578*	.360*	—					
4. CBCL attention problems	118	4.11	3.63	.513*	.084	.490*	—				
5. YSR rule breaking	120	3.28	3.11	.450*	.233*	.374*	.343*	—			
6. CBCL rule breaking	118	1.88	3.10	.512*	.131	.471*	.557*	.513*	—		
7. YSR aggressive behavior	120	6.00	4.34	.486*	.329*	.645*	.341*	.515*	.452*	—	
8. CBCL aggressive behavior	118	4.54	4.29	.455*	.052	.307*	.653*	.408*	.711*	.449*	—

Note. CDI-2 = Children's Depression Inventory–2; SCARED = Screen for Child Anxiety Related Emotional Disorders; YSR = Youth Self-Report; CBCL = Child Behavior Checklist.
* $p < .05$.

selected. While chance level performance on the task is technically 20%, we reasoned that the presence of the "nothing" option encouraged participants to choose that option instead of guessing randomly when they could not discern an emotion. Therefore, the discrimination threshold is estimated by computing the percentage morph that corresponds to a probability of .5. Nonetheless, we also conducted analyses based on a .2 threshold. As would be expected, perceptual sensitivity thresholds were lower based on this quantification of above chance, but results were similar overall. Results of these analyses can be found with the reproducible analysis code (RMarkdown) at https://osf.io/6jms7/ (Weissman, 2023).

Participant age was mean centered for all analyses. Four different models were used to examine relations between age, emotion category, and the Age × Emotion Category interaction in relation to the perceptual sensitivity threshold: (a) linear, (b) quadratic, (c) cubic, (d) thin plate regression smoothing spline. Orthogonalized polynomials for the quadratic and cubic models were created using the "poly" function, specifying second and third degree polynomials, respectively. Generalized additive mixed models were used in the spline model. All models were fit with the "gam" function in the mgcv package in R (Wood, 2022) to enable model fit comparison. Model fit was compared with a chi-squared deviance test.

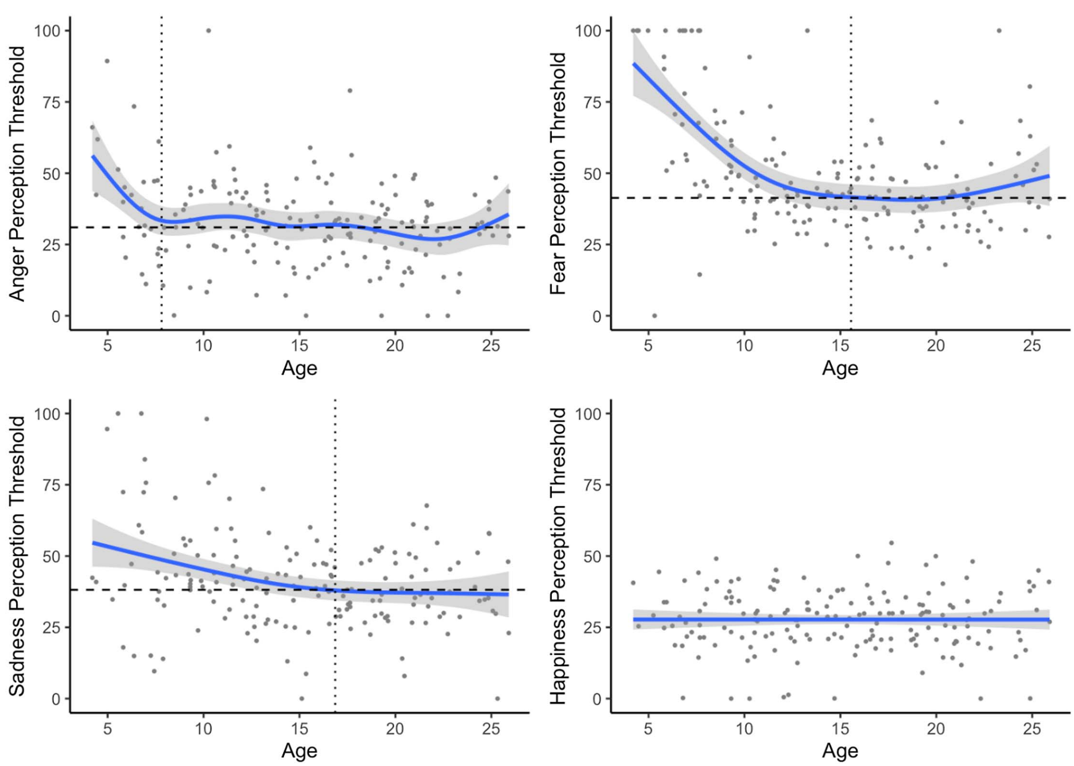
Visual inspection of the data and the spline model (Figure 1) suggested that perceptual sensitivity thresholds for the different emotion categories reached plateaus at different ages across development. To identify the age at which each of the perceptual sensitivity thresholds reached their plateau, we used a data-driven approach previously used to define the plateau points for the development of emotion vocabulary in this same sample (Nook et al., 2020). Two criteria identified plateaus: (a) reduced rate of change in the dependent variable and (b) the dependent variable reaching the maximum/minimum of its development. First, we extracted the best-fit line summarizing age-related change for each variable and computed its first derivative (Simpson, 2014). Because the first derivative quantifies the slope of a curve, the plateau of an emergent curve can be defined as the earliest age at which the first derivative approaches 0 (i.e., the curve flattens). Hence, we set the thresholds for a plateau as occurring when the change in perceptual sensitivity threshold slowed to less than 0.5%/year.

Sources of Age-Related Change in Expression Labeling

We conducted a series of analyses to understand whether the consensus-deviating labels selected for a given face (those that may be considered "incorrect" relative to the intention of the face stimuli developers) demonstrated systematic confusions. First, we computed the proportion of responses participants chose for each target stimulus and created a confusion matrix (i.e., a table indicating the proportion of each label assigned to each target). We then used mediation models to test whether systematic confusions explained

Figure 1

Results of Plateau Analyses of Thin Plate Regression Spline Model



Note. Dotted vertical lines correspond to the age at which the association between age and the decrease in perceptual sensitivity threshold was less than .5% per year. Dashed horizontal lines correspond to the perceptual threshold at which that plateau occurred. No dashed or dotted lines are provided for the happiness category, as no age-related changes were observed for labeling of happy faces. Error bars represent 95% confidence intervals. See the online article for the color version of this figure.

age-related differences in perceptual sensitivity thresholds. Specifically, we created four "mediation" path models using the lavaan package in R (Rosseel, 2012), one for each of the four target emotion categories (i.e., happy, angry, fear, and sad). In each model, the "independent variables" were age and sex, the "dependent variable" was the target's perceptual sensitivity threshold, and the "mediators" were the proportion of responses that were "nothing," "I don't know," and each of the other three nonconsensus emotion labels (e.g., happy, angry, and sad if the consensus emotion was scared). The indirect effects in these models should not be interpreted as causal explanations but rather statistical explanations: They provide insight into how age-related differences in perceptual sensitivity thresholds varying by emotion category may be explained by age-related differences in response patterns.

Associations Between Perceptual Sensitivity Thresholds and Psychopathology

Finally, linear regression was used to test associations between perceptual sensitivity thresholds and externalizing and internalizing psychopathology. The dependent variables were internalizing and externalizing symptom factor scores, and the regressors were, age, sex, and the perceptual sensitivity thresholds for anger, fear, sadness, and happiness consensus categories. Because of the study-specific nature of our latent measures of internalizing and externalizing psychopathology based on multiple different measures from both parent and child, in the interest of reproducibility and amenability to meta-analysis, we conducted additional analyses using the higher of the parent or child-reported Externalizing Problems total score on the CBCL/YSR and using the child-reported Internalizing Problems total score on the YSR as measures of psychopathology.

Transparency and Openness

We report how we determined sample size, all data exclusions, all manipulations, and all measures in the study, and we follow the journal article reporting standards (Kazak, 2018). This study was not preregistered. Data and code are available at https://osf.io/6jms7/ (Weissman, 2023).

Results

Sample characteristics, descriptive statistics, and intercorrelations among study variables are provided in Table 2. Across ages 4–25 years, only the perceptual sensitivity for the fear category was significantly correlated with age, such that older participants labeled a face as fear—in line with the cultural consensus—at lower intensities. The spline model fit significantly better than the linear (df = 14.7, deviance = 18,248, $p < .001$), quadratic (df = 10.7, deviance = 5,082, $p = .027$), and cubic models ($df = 6.7$, deviance = 5,047, $p = .015$). Results therefore focus on describing the age-related patterns reflected in those spline models (Figure 1).

Age-Related Patterns in Perceptual Sensitivity Thresholds

Happiness

Faces with the consensus judgment of happy were matched with the term "happy" 70.4% of the time across all morphs and 95.0% of the time for the 90% morph. Age was not associated with the perceptual threshold for faces with the consensus judgment of happiness, which was labeled as such more than 50% of the time at morphs above 27.7% on average across the whole sample. This was the lowest sensitivity threshold of any of the emotion categories.

Anger

Faces with the consensus judgment of angry were matched with the term "angry" 64.5% of the time across all morphs and 90.4% of the time for the 90% morph and were labeled as "angry" more than 50% of the time at morphs above 32.7% intensity on average across the whole sample. The plateau analyses suggested that the perceptual threshold for labeling anger decreased from ages 4.2 to 7.8, where it reached a plateau at an intensity threshold of 31.0%.

Sadness

Faces with the consensus judgment of sad were matched with the term "sad" 55.2% of the time across all morphs and 87.4% of the time for the 90% morph and were labeled as "sad" more than 50% of the time at morphs above 41.3% intensity on average across the whole sample. The plateau analyses suggested that the perceptual threshold for labeling sadness decreased from ages 4.2 to 16.9, where it reached a plateau at an intensity threshold of 38.2%.

Fear

Faces with the consensus judgment of fear were matched with the term "scared" 48.9% of the time across all morphs and 75.1% of the time for the 90% morph and were labeled as "scared" more than 50%

Table 2
Descriptive Statistics and Correlations of Study Variables

				Correlation						
Variable	N	M	SD	1	2	3	4	5	6	7
1. Age	203	14.2	5.94	—						
2. Anger perception threshold	175	32.7	15.5	−.04	—					
3. Fear perception threshold	177	49.2	20.2	−.40*	.21*	—				
4. Happiness perception threshold	173	27.7	10.7	−.01	.05	.04	—			
5. Sadness perception threshold	171	41.3	17.1	−.17	.33*	.10*	.08	—		
6. Internalizing symptoms	120	0.001	0.941	.20*	.16	.19	.18	−.03	—	
7. Externalizing symptoms	120	0.001	0.905	.01	.03	.21*	.11	−.13	.81*	—

* $p < .05$.

of the time at morphs above 49.2% intensity on average across the whole sample. The plateau analyses suggested the perceptual threshold for labeling fear decreased from ages 4.2 to 15.6, where it reached a plateau at an intensity threshold of 41.3%.

Sources of Age-Related Changes in Emotion Perception

When the consensus emotion was not perceived, the most common alternative response given for all categories was "nothing" (Figure 2). Selection of "I don't know" was more common among younger than older participants (β = −.402, CI [−.529, −.275], Figure 3). The confusion matrix indicated that faces with a consensus judgment of fear were reported to be "sad" 15% of the time across all participants and were more likely to be seen as sad among younger participants than older participants (β = −.278, CI [−.412, −.143]; Figure 3). Younger participants were also more likely than older participants to label faces intended to convey anger (β = −.212, CI [−.337, −.087]), sadness (β = −.328, CI [−.440, −.217]), and fear (β = −.236, CI [−.375, −.098]) as "happy" and to label faces intended to convey anger as "scared" (β = −.130, CI [−.252, −.007]; Figure 3).

We first tested what response patterns could explain age-related reductions in the perceptual sensitivity for faces with a consensus judgment of anger. The strongest indirect effect of older age on a lower anger perceptual sensitivity threshold emerged through a lower proportion of "I don't know" responses (β = −.164, CI [−.254, −.073]), followed by a lower proportion of "happy" labels (β = −.090, CI [−.162, −.019]). The indirect effects of older age on a lower perceptual threshold for the anger category via the proportion of "sad," "scared," or "nothing" responses were not significant (Figure 4a).

For explaining age-related reductions in the perceptual sensitivity for faces with a consensus judgment of fear, we again found the strongest indirect effect of older age on a lower fear perceptual threshold via a lower proportion of "I don't know" responses (β = −.292, CI [−.390, −.195]), followed by a lower proportion of "sad" responses (β = −.140, CI [−.220, −.060]) and a lower proportion of "happy" responses (β = −.083, CI [−.144, −.021]). Indirect effects of older age on a lower perceptual threshold for the fear category via the proportion of "angry" or "nothing" responses were not significant (Figure 4b).

Finally, the strongest indirect effect of older age on the perceptual sensitivity for faces with a consensus judgment of sadness was via a lower proportion of "I don't know" responses (β = −.271, CI [−.377, −.166]), followed by a lower proportion of "happy" responses (β = −.143, CI [−.214, −.073]). There was also a significant indirect effect of older age on a lower perceptual sensitivity threshold via a greater proportion of "nothing" responses (β = .159, CI [.018, .301]). The indirect effects of older age on a lower sadness perceptual sensitivity threshold via the proportion of "scared" or "angry" responses were not significant (Figure 4c).

Associations Between Perceptual Sensitivity Thresholds and Psychopathology

When accounting for age and sex, there was a positive association between the perceptual sensitivity for faces intended to convey fear and both internalizing (B = .0173, SE = .00611, β = .316, p = .006) and externalizing (B = .0133, SE = .00827, β = .254, p = .035) symptom factor scores (Figure 5). Higher perceptual sensitivity thresholds for faces intended to convey fear were associated with higher levels of psychopathology symptoms. Perceptual sensitivity thresholds for faces intended to convey anger, sadness, and happiness were not significantly associated with internalizing or externalizing factor scores (Table 3).

To test whether our measurement approach may have affected our results, we repeated these analyses using the total scores for Internalizing Problems and Externalizing Problems on the YSR. There were no significant associations between perceptual sensitivity thresholds and internalizing scores. The largest, nonsignificant association was

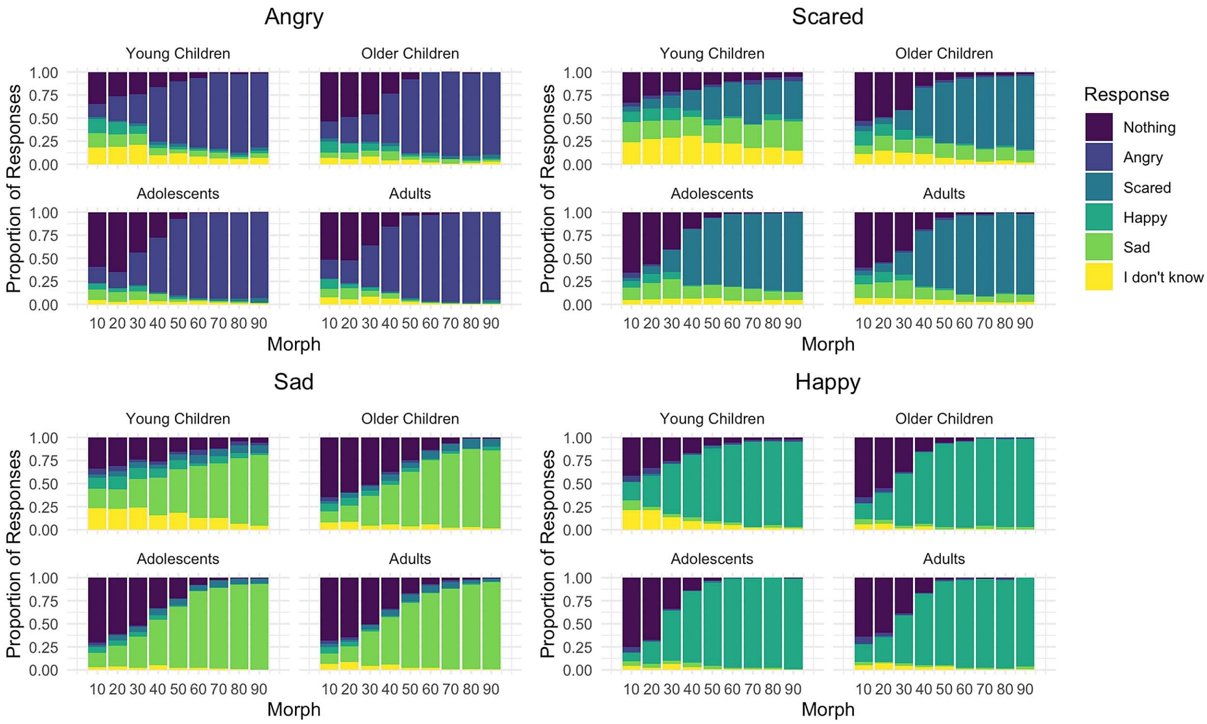
Figure 2

The Distribution of Participants' Responses Across All Ages

Note. The rows correspond to consensus emotion labels for a given target, whereas the columns represent the proportion of responses. Each cell in the confusion matrix represents the proportion of responses for a given emotion category. See the online article for the color version of this figure.

Figure 3

Proportions of Emotion Labels Within Four Different Age Groups



Note. Proportions of labels assigned to each stereotypical expression category within four different age groups: young children (4–7 years), older children (8–12), adolescents (13–17 years), and adults (18+ years). Additionally, columns represent different levels of morph (ranging from 10 to 90) of the target. See the online article for the color version of this figure.

for the fear category's perceptual sensitivity threshold (B = .0774, SE = 0.0618, β = .154, p = .213). However, higher fear perceptual thresholds were significantly associated with higher levels of externalizing problems (B = .112, SE = 0.0465, β = .281, p = .018). Expression perceptual sensitivity thresholds for anger, sadness, and happiness categories were not significantly associated with externalizing problems. Thus, the relationship between fear thresholds and internalizing scores requires the factor approach, whereas the relationship with externalizing scores is present only when analyzing the YSR. This helps protect against the concern that the externalizing result depends on parent-report.

Discussion

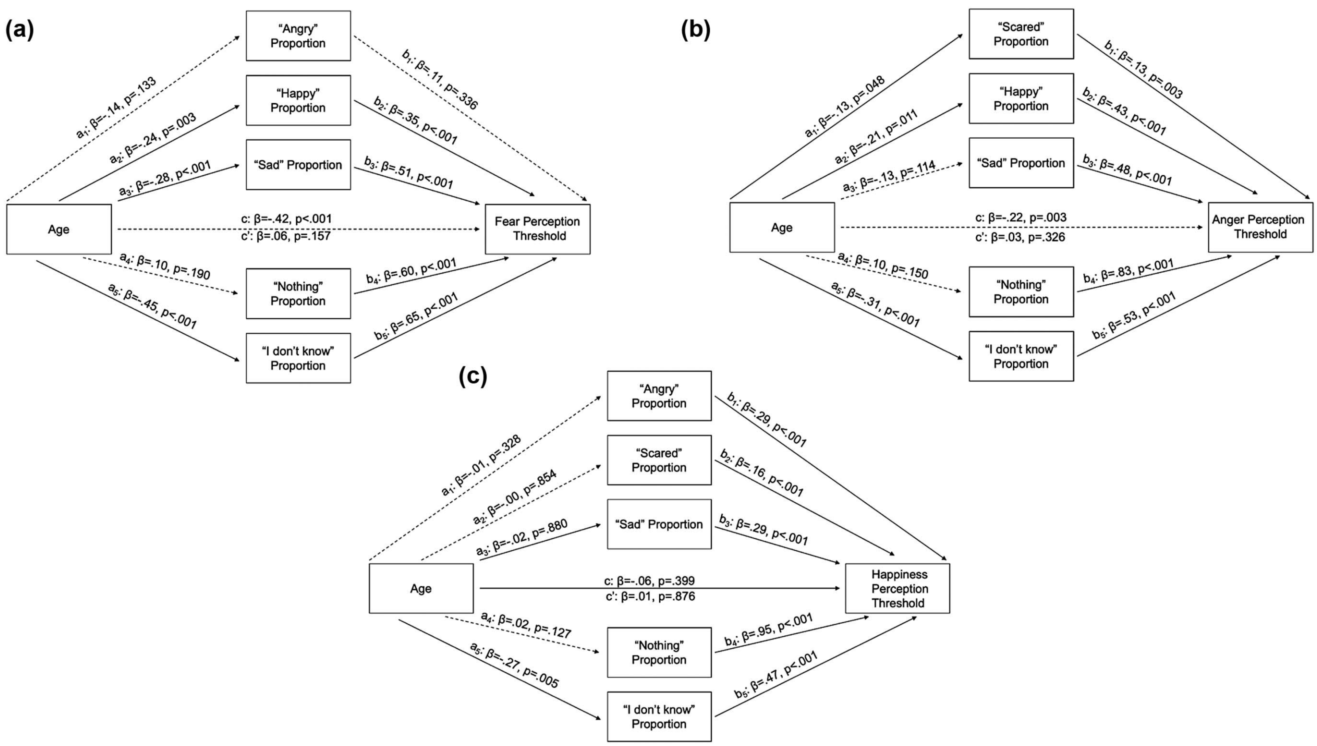
General Discussion

In this study, we examined age-related differences in perceptual sensitivity to applying consensus labels to posed facial expressions intended to depict fear, anger, sadness, and happiness in a sample spanning ages 4–25 years. The perceptual sensitivity thresholds for aligning labels to these emotion consensus categories differed significantly on average and demonstrated distinct nonlinear associations with age. The most common response when the consensus emotion was not chosen was "nothing." Younger participants were more prone than older participants to respond "I don't know," to use the "happy" label for all three canonical negative emotion displays, and to label poses with a consensus label of "scared" as "sad." In addition, being less sensitive to labeling fear (i.e., lower tendency to label fear when it was depicted at mild intensity) was associated with higher internalizing and externalizing psychopathology symptoms.

Applying emotion labels to facial expressions in a way that aligns with cultural consensus is a crucial affective skill, and results from this study add to our knowledge of how this skill develops. First, we found that sensitivity to assigning happiness labels to smiling faces has already peaked by age 4 and does not improve with increasing age. Notably, happiness was the only positively valenced emotion category included in this task. By preschool age, children already categorize facial expressions along dimensions of valence and arousal in a manner similar to adults (Bullock & Russell, 1984; Russell & Bullock, 1985), and the differentiation of emotion concepts based on valence develops very early, while differentiation across other dimensions occurs more slowly as vocabulary develops (Nook et al., 2017). Our results therefore likely reflect that the ability to map the term "happy" to its prototypical expression emerges early in development and that perceptual cues of emotional valence are easier to distinguish early in development than the characteristics of emotion expressions that differentiate among negative emotions (Becker & Srinivasan, 2014; Widen & Russell, 2008). Younger participants were also more prone to labeling stereotyped poses of negative emotional expressions as "happy," suggesting the presence of a positivity bias in children's emotion judgments that may decrease with age. This type of positivity bias has been observed frequently in studies examining the perception of emotional facial expressions (Kauschke et al., 2019), especially in children (Vesker et al., 2018).

Figure 4

Path Models Examining the Sources of Age-Related Changes in Perceptual Thresholds for Stereotypical Expressions

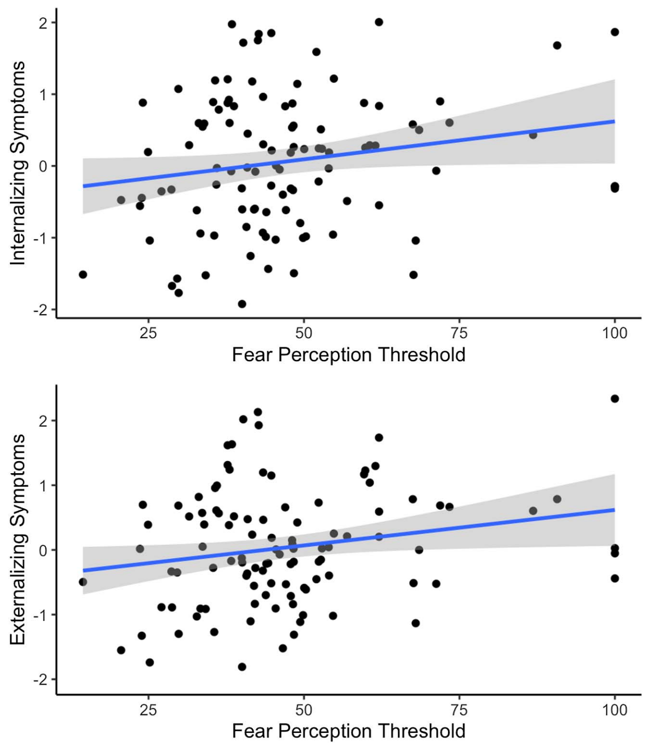


Note. (a) There were significant indirect effects of older age on lower anger perceptual thresholds through fewer "I don't know" responses (β = −.164, CI [−.254, −.073]) and "happy" responses (β = −.090, CI [−.162, −.019]). (b) There were significant indirect effects of older age on lower fear perceptual thresholds through fewer "I don't know" responses (β = −.292, CI [−.390, −.195]), "sad" responses (β = −.140, CI [−.220, −.060]), and "happy" responses (β = −.083, CI [−.144, −.021]). (c) There were significant indirect effects of older age on lower sadness perceptual thresholds through fewer "I don't know" responses (β = −.271, CI [−.377, −.166]) and "happy" responses (β = −.143, CI [−.214, −.073]) and a greater proportion of "nothing" responses (β = .159, CI [.018, .301]). CI = confidence interval.

Labeling canonically angry faces according to the cultural consensus becomes increasingly sensitive from ages 4 to 8 years, where it reached a plateau. Labeling canonical expressions of fear and sadness according to the cultural consensus demonstrated more protracted development, not reaching a plateau until age 17 years. Younger participants were more sensitive to judging canonical expressions of sadness, so the improvements with age were much more modest than the improvement in sensitive to judging canonical expressions of fear. One potential explanation of these findings is that they may reflect the frequency with which children encounter and have the opportunity to label these expressions in daily life. Among the emotions included in this study, adults report experiencing happiness most frequently in their daily lives, followed by sadness, anger, then fear (Trampe et al., 2015). Similarly, trained observers report that happiness is the most commonly encountered facial expression, followed by sadness, anger, then fear (Calvo et al., 2014; Somerville & Whalen, 2006). It is therefore plausible that children tend to encounter expressions of happiness, anger, and sadness from peers and caregivers with greater frequency than expressions of fear and therefore learn to label these emotion expressions with greater sensitivity at an earlier age. Indeed, prior work has found that children exposed to physical abuse, and who therefore may have seen more expressions of anger are more likely to label facial expressions as anger when those expressions are less exaggerated or intense (Pollak & Kistler, 2002).

These age-related changes may also reflect the degree of variability in the actual facial expressions children see from the stereotypical posed expressions used in the task. It may be that expressions of fear are more heterogenous and therefore lacking in cues that can be easily posed by someone not experiencing the emotion or discerned at lower levels of intensity. Indeed, within the present study, consensus judgments of labels of the facial expression stimuli were lowest for fear; thus, it is not surprising that these facial displays may have evoked a wider variety of labels and systematic confusions with other categories. Prototypical fear faces were most frequently mislabeled as sad, especially among younger children, suggesting that those children may have gleaned the consensus negative valence of the facial expression and perceived it as sadness because that is the most common negative emotional expression they encounter in daily life, while fear is among the least common (Calvo et al., 2014; Somerville & Whalen, 2006). Overall, the greater tendency to label anger, sadness, and fear in alignment with cultural consensus more frequently at lower intensity levels among older participants seems to be attributable to increases in perceptual sensitivity, as indicated by a lower proportion of "I don't know" responses, and, to a lesser degree, to decrease in the confusability with other labels. Older participants were less likely to label low intensity

Figure 5
Associations Between Fear Perception Threshold and Both Internalizing and Externalizing Psychopathology



Note. Significant linear association between a higher perceptual threshold for the stereotypical fear category and both internalizing and externalizing psychopathology, indicating that children, adolescents, and young adults who require more intense expressions of posed fear in order to consistently match it with the label "scared" tend to have more symptoms of psychopathology transdiagnostically. Error bars represent 95% confidence intervals. See the online article for the color version of this figure.

negative expressions as happiness, low intensity prototypical fear expressions with sadness, or low intensity prototypical anger expressions with fear than younger participants.

Responses to faces depicting canonical fearful expressions showed several interesting properties in this study. (a) Fear was the expression category requiring the most information to label consistently with cultural consensus overall. (b) Perceptual sensitivity for labeling the canonical fear expressions as fear demonstrated the most marked improvement from early childhood to adulthood. (c) Perceptual sensitivity for canonical fear was most strongly associated with psychopathology symptoms. Among older child and adolescent participants in the study, greater sensitivity to labeling canonical fear, but no other emotion expressions, was associated with lower self-reported internalizing problems and externalizing psychopathology symptoms, which mostly fell in the subclinical range. Difficulties inferring what other people are feeling based on their external cues may increase vulnerability to psychopathology by curtailing social support and adaptive emotion regulation in the context of interpersonal interactions (Williams et al., 2018). Thus, greater difficulty identifying cues that are culturally agreed on to denote fear may index increased risk for both internalizing and externalizing problems. Alternatively, because stereotypical fear was the most ambiguous and difficult expression to label, it is possible that those with lower emotional awareness and/or less granular emotion concepts—factors associated with both emotion perception and vulnerability to developing symptoms of psychopathology (Aaron et al., 2018; Barrett et al., 2001; Hemming et al., 2019; Lennarz et al., 2018; Nook et al., 2017; Starr et al., 2020; Weissman et al., 2020)—struggled most to identify canonical fear cues.

These patterns were observed despite—or potentially due to—the fact that these posed wide eyed, grimacing faces were not the naturalistic result of a person actually experiencing fear, and are not very consistently identified as fear by untrained observers (Tottenham et al., 2009). Thus, it is also possible that familiarity with stereotypical emotion expressions, but not necessarily sensitivity to perceiving others' fear in naturalistic settings, may be an index of low vulnerability to developing symptoms of psychopathology. In addition, it is important to note that internalizing and externalizing assessments differed in that externalizing estimates were based both on participant- and parent-report, whereas internalizing estimates were based only on participant-report, as parents naturally have much more limited ability to observe internalizing symptoms. However, we verified that this finding was significant even when only using participant-reported externalizing problems on the YSR (Weissman, 2023).

One major limitation of this work influences our interpretations and suggests potential future directions. The perceptual sensitivity task had a discrete set of emotion labels to choose from and only included a single positively valenced emotion expression. The age trajectories for perceptual sensitivity to different emotion categories are therefore vulnerable to bias introduced by a forced choice paradigm. Previous literature has established that limiting the response options available and forcing participant to choose between them limits the variability of potential responses and artificially inflates the consistency with which certain emotion labels are applied (DiGirolamo & Russell, 2017; Gendron et al., 2018; Hoemann et al., 2019). While this may be mitigated to some degree in our task by the presence of the "I don't know" and "nothing" options, the relative ease of assigning a happiness label, regardless of age, for example, may be specific to that

Table 3
Regression of Psychopathology on Age, Sex, and Perceptual Sensitivity Thresholds

Predictor	*B*	*SE*	β	*p*
Internalizing				
Age	.112	.0306	.400	<.001
Sex (female)	−.0490	.179	−.026	.784
Anger threshold	.00793	.00636	.131	.216
Fear threshold	.0173	.00611	.316	.006
Happiness threshold	.0146	.00813	.168	.076
Sadness threshold	−.00251	.00629	−.041	.691
R^2/adjusted R^2		0.217/0.166		
Externalizing				
Age	.0465	.0312	.173	.139
Sex (female)	−.272	.182	−.151	.138
Anger threshold	.00152	.00647	.026	.815
Fear threshold	.0133	.00827	.254	.035
Happiness threshold	.00802	.00827	.096	.335
Sadness threshold	−.00785	.00640	−.133	.223
R^2/adjusted R^2		0.121/0.064		

Note. *SE* = standard error.

specific category, or it may reflect that it is easier to differentiate expressions based on valence than to differentiate among stereotypically negatively valenced expressions. As another example, we find that perceptual sensitivity to canonical anger reaches a plateau at an earlier age than perceptual sensitivity to canonical sadness or canonical fear. While this may reflect a real developmental difference in the emergence of emotion concepts and associated their visual cues, it may alternatively reflect the absence of response options with high perceptual confusability with anger, in particular disgust (Dailey et al., 2002; Woodard et al., 2022). Similarly, the canonical fear poses in NimStim data set are most frequently confused with surprise (Tottenham et al., 2009), but that was not a response option in this task. If surprise were an option, it is plausible that the threshold for perceiving faces as "scared" would be even higher in this study and plateau even later. Future research using free response and/or a wider variety of positive and negative emotion expressions could therefore build upon this work. The use of the NimStim data set itself, in particular the fear expressions, could also be considered a limitation, both in terms external validity, in that the actors were not actually experiencing the emotions they were attempting to express, and in terms of how consistent the cultural consensus is about the emotions being expressed. As noted in the methods section, this is particularly true for the "fear expressions" in the NimStim data set, which are labeled as "fear" as little as 30% of the time for some stimuli. We mitigated this concern somewhat by selecting stimuli with the highest consistency in norming data, but this is an important concern for the validity of future research.

Some additional limitations of this work are its cross-sectional and correlational design. Age-related associations between individuals may or may not accurately characterize developmental trajectories within individuals. The cross-sectional design of this study therefore limits the extent to which these results can characterize the developmental trajectories of sensitivity to perceiving stereotypical emotion cues. Further, the causal direction of the association between sensitivity to perceiving fear and psychopathology also remains unclear. Difficulty perceiving facial expressions stereotypically judged as fear may be a cause, consequence, or symptom of psychopathology. Future work examining trajectories of sensitivity to perceiving emotion in facial cues and psychopathology longitudinally is essential to developing a more complete understanding of the development of perceptual sensitivity to emotion expressions and its role in mental health.

Constraints on Generality

This study was conducted in a U.S. sample of English-speaking children, adolescents, and young adults, and perceptual thresholds are based on English labels for facial emotion expressions. The generalizability of these findings therefore may be constrained to English speakers in the United States (Gendron et al., 2018; Hoemann et al., 2019). In addition, while the sample was racially diverse, it was majority white, with higher socioeconomic status than the U.S. population on average. Therefore, generalizability to more diverse groups requires further study. Finally, the sample was a community sample with relatively low levels of psychopathology symptoms. The associations between perceptual sensitivity thresholds and mostly subclinical externalizing and internalizing symptoms therefore may not imply that the same patterns will exist in people diagnosed with clinical levels of psychopathology.

Conclusion

In conclusion, this study provides evidence that perceptual sensitivity to labeling facial expressions of anger, fear, sadness, and happiness follow distinct, nonlinear developmental trajectories, all reaching plateaus by adolescence. Lower sensitivity to labeling fear—the most difficult emotion expression to categorize and the most slowly developing—may be a marker of vulnerability to developing symptoms of multiple forms of psychopathology.

References

Aaron, R. V., Snodgress, M. A., Blain, S. D., & Park, S. (2018). Affect labeling and other aspects of emotional experiences in relation to alexithymia following standardized emotion inductions. *Psychiatry Research*, *262*, 115–123. https://doi.org/10.1016/j.psychres.2018.02.014

Achenbach, T. M. (1991). *Manual for the child behavior checklist/4–18 and 1991 profile*. University of Vermont, Department of Psychiatry.

Aebi, M., Kuhn, C., Banaschewski, T., Grimmer, Y., Poustka, L., Steinhausen, H.-C., & Goodman, R. (2017). The contribution of parent and youth information to identify mental health disorders or problems in adolescents. *Child and Adolescent Psychiatry and Mental Health*, *11*(1), Article 23. https://doi.org/10.1186/s13034-017-0160-9

Angold, A., Weissman, M. M., John, K., Merikangas, K. R., Prusoff, B. A., Wickramaratne, P., Gammon, G. D., & Warner, V. (1987). Parent and child reports of depressive symptoms in children at low and high risk of depression. *Journal of Child Psychology and Psychiatry*, *28*(6), 901–915. https://doi.org/10.1111/j.1469-7610.1987.tb00678.x

Barisnikov, K., Thomasson, M., Stutzmann, J., & Lejeune, F. (2021). Sensitivity to emotion intensity and recognition of emotion expression in neurotypical children. *Children*, *8*(12), Article 1108. https://doi.org/10.3390/children8121108

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, *20*(1), 1–68. https://doi.org/10.1177/1529100619832930

Barrett, L. F., Gross, J., Christensen, T. C., & Benvenuto, M. (2001). Knowing what you're feeling and knowing what to do about it: Mapping the relation between emotion differentiation and emotion regulation. *Cognition and Emotion*, *15*(6), 713–724. https://doi.org/10.1080/02699930143000239

Becker, D. V., & Srinivasan, N. (2014). The vividness of the happy face. *Current Directions in Psychological Science*, *23*(3), 189–194. https://doi.org/10.1177/0963721414533702

Bird, H. R., Gould, M. S., & Staghezza, B. (1992). Aggregating data from multiple informants in child psychiatry epidemiological research. *Journal of the American Academy of Child & Adolescent Psychiatry*, *31*(1), 78–85. https://doi.org/10.1097/00004583-199201000-00012

Birmaher, B., Khetarpal, S., Brent, D., Cully, M., Balach, L., Kaufman, J., & Neer, S. M. K. (1997). The Screen for Child Anxiety Related Emotional Disorders (SCARED): Scale construction and psychometric characteristics. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(4), 545–553. https://doi.org/10.1097/00004583-199704000-00018

Boakes, J., Chapman, E., Houghton, S., & West, J. (2007). Facial affect interpretation in boys with attention deficit/hyperactivity disorder. *Child Neuropsychology*, *14*(1), 82–96. https://doi.org/10.1080/09297040701503327

Brotman, M. A., Skup, M., Rich, B. A., Blair, K. S., Pine, D. S., Blair, J. R., & Leibenluft, E. (2008). Risk for bipolar disorder is associated with face-processing deficits across emotions. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*(12), 1455–1461. https://doi.org/10.1097/CHI.0b013e318188832e

Bullock, M., & Russell, J. A. (1984). Preschool children's interpretation of facial expressions of emotion. *International Journal of Behavioral Development*, *7*(2), 193–214. https://doi.org/10.1177/016502548400700207

Calvo, M. G., Gutiérrez-García, A., Fernández-Martín, A., & Nummenmaa, L. (2014). Recognition of facial expressions of emotion is related to their frequency in everyday life. *Journal of Nonverbal Behavior*, *38*(4), 549–567. https://doi.org/10.1007/s10919-014-0191-3

Cantwell, D. P., Lewinsohn, P. M., Rohde, P., & Seeley, J. R. (1997). Correspondence between adolescent report and parent report of psychiatric diagnostic data. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(5), 610–619. https://doi.org/10.1097/00004583-199705000-00011

Carey, S., Diamond, R., & Woods, B. (1980). Development of face recognition: A maturational component? *Developmental Psychology*, *16*(4), 257–269. https://doi.org/10.1037/0012-1649.16.4.257

Collin, L., Bindra, J., Raju, M., Gillberg, C., & Minnis, H. (2013). Facial emotion recognition in child psychiatry: A systematic review. *Research in Developmental Disabilities*, *34*(5), 1505–1520. https://doi.org/10.1016/j.ridd.2013.01.008

Dailey, M. N., Cottrell, G. W., Padgett, C., & Adolphs, R. (2002). EMPATH: A neural network that categorizes facial expressions. *Journal of Cognitive Neuroscience*, *14*(8), 1158–1173. https://doi.org/10.1162/089892902760807177

DiGirolamo, M. A., & Russell, J. A. (2017). The emotion seen in a face can be a methodological artifact: The process of elimination hypothesis. *Emotion*, *17*(3), 538–546. https://doi.org/10.1037/emo0000247

Durand, K., Gallay, M., Seigneuric, A., Robichon, F., & Baudouin, J.-Y. (2007). The development of facial emotion recognition: The role of configural information. *Journal of Experimental Child Psychology*, *97*(1), 14–27. https://doi.org/10.1016/j.jecp.2006.12.001

Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Krause, R., LeCompte, W. A., Pitcairn, T., Ricci-Bitti, P. E., Scherer, K., Tomita, M., & Tzavaras, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, *53*(4), 712–717. https://doi.org/10.1037/0022-3514.53.4.712

Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science*, *164*(3875), 86–88. https://doi.org/10.1126/science.164.3875.86

Ellis, C. R., Lindstrom, K. L., Villani, T. M., Singh, N. N., Best, A. M., Winton, A. S. W., Axtell, P. K., Oswald, D. P., & Leung, J. P. (1997). Recognition of facial expressions of emotion by children with emotional and behavioral disorders. *Journal of Child and Family Studies*, *6*(4), 453–470. https://doi.org/10.1023/A:1025097429527

Gao, X., & 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/10.1016/j.jecp.2008.11.002

Gao, X., & Maurer, D. (2010). A happy story: Developmental changes in children's sensitivity to facial expressions of varying intensities. *Journal of Experimental Child Psychology*, *107*(2), 67–86. https://doi.org/10.1016/j.jecp.2010.05.003

Gendron, M., & Barrett, L. F. (2018). Emotion perception as conceptual synchrony. *Emotion Review*, *10*(2), 101–110. https://doi.org/10.1177/1754073917705717

Gendron, M., Crivelli, C., & Barrett, L. F. (2018). Universality reconsidered: Diversity in making meaning of facial expressions. *Current Directions in Psychological Science*, *27*(4), 211–219. https://doi.org/10.1177/0963721417746794

Grills, A. E., & Ollendick, T. H. (2003). Multiple informant agreement and the anxiety disorders interview schedule for parents and children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *42*(1), 30–40. https://doi.org/10.1097/00004583-200301000-00008

Grynberg, D., Chang, B., Corneille, O., Maurage, P., Vermeulen, N., Berthoz, S., & Luminet, O. (2012). Alexithymia and the processing of emotional facial expressions (EFEs): Systematic review, unanswered questions and further perspectives. *PLOS ONE*, *7*(8), Article e42429. https://doi.org/10.1371/journal.pone.0042429

Hemming, L., Haddock, G., Shaw, J., & Pratt, D. (2019). Alexithymia and its associations with depression, suicidality, and aggression: An overview of the literature. *Frontiers in Psychiatry*, *10*, Article 203. https://doi.org/10.3389/fpsyt.2019.00203

Herba, C. M., Landau, S., Russell, T., Ecker, C., & Phillips, M. L. (2006). The development of emotion-processing in children: Effects of age, emotion, and intensity. *Journal of Child Psychology and Psychiatry*, *47*(11), 1098–1106. https://doi.org/10.1111/j.1469-7610.2006.01652.x

Hoemann, K., Crittenden, A. N., Msafiri, S., Liu, Q., Li, C., Roberson, D., Ruark, G. A., Gendron, M., & Feldman Barrett, L. (2019). Context facilitates performance on a classic cross-cultural emotion perception task. *Emotion*, *19*(7), 1292–1313. https://doi.org/10.1037/emo0000501

Kauschke, C., Bahn, D., Vesker, M., & Schwarzer, G. (2019). The role of emotional valence for the processing of facial and verbal stimuli-positivity or negativity bias? *Frontiers in Psychology*, *10*, Article 1654. https://doi.org/10.3389/fpsyg.2019.01654

Kazak, A. E. (2018). Editorial: Journal article reporting standards. *American Psychologist*, *73*(1), 1–2. https://doi.org/10.1037/amp0000263

Kessler, R. C., Avenevoli, S., Costello, J., Green, J. G., Gruber, M. J., McLaughlin, K. A., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Merikangas, K. R. (2012). Severity of 12-month DSM-IV disorders in the national comorbidity survey replication adolescent supplement. *Archives of General Psychiatry*, *69*(4), 381–389. https://doi.org/10.1001/archgenpsychiatry.2011.1603

Kovacs, M. (2011). *Children's depression inventory (CDI2)*. Multi-Health Systems.

Lennarz, H. K., Lichtwarck-Aschoff, A., Timmerman, M. E., & Granic, I. (2018). Emotion differentiation and its relation with emotional well-being in adolescents. *Cognition and Emotion*, *32*(3), 651–657. https://doi.org/10.1080/02699931.2017.1338177

Linares, D., & López-Moliner, J. (2016). quickpsy: An R package to fit psychometric functions for multiple groups. *The R Journal*, *8*(1), 122–131. https://doi.org/10.32614/RJ-2016-008

Marsh, A. A., & Blair, R. J. R. (2008). Deficits in facial affect recognition among antisocial populations: A meta-analysis. *Neuroscience and Biobehavioral Reviews*, *32*(3), 454–465. https://doi.org/10.1016/j.neubiorev.2007.08.003

Martin-Key, N. A., Graf, E. W., Adams, W. J., & Fairchild, G. (2018). Facial emotion recognition and eye movement behaviour in conduct disorder. *Journal of Child Psychology and Psychiatry*, *59*(3), 247–257. https://doi.org/10.1111/jcpp.12795

Merikangas, K. R., He, J. P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., Benjet, C., Georgiades, K., & Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Survey Replication—Adolescent Supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(10), 980–989. https://doi.org/10.1016/j.jaac.2010.05.017

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

Moretti, M. M., Fine, S., Haley, G., & Marriage, K. (1985). Childhood and adolescent depression: Child-report versus parent-report information. *Journal of the American Academy of Child Psychiatry*, *24*(3), 298–302. https://doi.org/10.1016/S0002-7138(09)61090-6

Nook, E. C., Flournoy, J. C., Rodman, A. M., Mair, P., & McLaughlin, K. A. (2021). High emotion differentiation buffers against internalizing symptoms following exposure to stressful life events in adolescence: An intensive longitudinal study. *Clinical Psychological Science*, *9*(4), 699–718. https://doi.org/10.1177/2167702620979786

Nook, E. C., Lindquist, K. A., & Zaki, J. (2015). A new look at emotion perception: Concepts speed and shape facial emotion recognition. *Emotion*, *15*(5), 569–578. https://doi.org/10.1037/a0039166

Nook, E. C., Sasse, S. F., Lambert, H. K., McLaughlin, K. A., & Somerville, L. H. (2017). Increasing verbal knowledge mediates development of multidimensional emotion representations. *Nature Human Behaviour*, *1*(12), 881–889. https://doi.org/10.1038/s41562-017-0238-7

Nook, E. C., Stavish, C. M., Sasse, S. F., Lambert, H. K., Mair, P., McLaughlin, K. A., & Somerville, L. H. (2020). Charting the development of emotion comprehension and abstraction from childhood to adulthood using observer-rated and linguistic measures. *Emotion*, *20*(5), 773–792. https://doi.org/10.1037/emo0000609

Nowicki, S., & Duke, M. P. (1992). The association of children's nonverbal decoding abilities with their popularity, locus of control, and academic achievement. *The Journal of Genetic Psychology: Research and Theory on Human Development*, *153*(4), 385–393. https://doi.org/10.1080/00221325.1992.10753734

Nowicki, S., Jr., & Mitchell, J. (1998). Accuracy in identifying affect in child and adult faces and voices and social competence in preschool children. *Genetic, Social, and General Psychology Monographs*, *124*(1), 39–59. https://link.gale.com/apps/doc/A20945658/HRCA?u=anon~d5ec191&sid=googleScholar&xid=122f2979

Pollak, S. D., & Kistler, D. J. (2002). Early experience is associated with the development of categorical representations for facial expressions of emotion. *Proceedings of the National Academy of Sciences*, *99*(13), 9072–9076. https://doi.org/10.1073/pnas.142165999

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510–532. https://doi.org/10.1037/0033-2909.114.3.510

Rodger, H., Lao, J., & Caldara, R. (2018). Quantifying facial expression signal and intensity use during development. *Journal of Experimental Child Psychology*, *174*, 41–59. https://doi.org/10.1016/j.jecp.2018.05.005

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

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

Russell, J. A., & Bullock, M. (1985). Multidimensional scaling of emotional facial expressions: Similarity from preschoolers to adults. *Journal of Personality and Social Psychology*, *48*(5), 1290–1298. https://doi.org/10.1037/0022-3514.48.5.1290

Simonian, S. J., Beidel, D. C., Turner, S. M., Berkes, J. L., & Long, J. H. (2001). Recognition of facial affect by children and adolescents diagnosed with social phobia. *Child Psychiatry and Human Development*, *32*(2), 137–145. https://doi.org/10.1023/A:1012298707253

Simpson, G. (2014). *First derivative of a spline term* [Computer software]. https://gist.github.com/gavinsimpson/e73f011fdaaab4bb5a30

Singh, S. D., Ellis, C. R., Winton, A. S. W., Singh, N. N., Leung, J. P., & Oswald, D. P. (1998). Recognition of facial expressions of emotion by children with attention-deficit hyperactivity disorder. *Behavior Modification*, *22*(2), 128–142. https://doi.org/10.1177/01454455980222002

Somerville, L. H., & Whalen, P. J. (2006). Prior experience as a stimulus category confound: An example using facial expressions of emotion. *Social Cognitive and Affective Neuroscience*, *1*(3), 271–274. https://doi.org/10.1093/scan/nsl040

Starr, L. R., Hershenberg, R., Shaw, Z. A., Li, Y. I., & Santee, A. C. (2020). The perils of murky emotions: Emotion differentiation moderates the prospective relationship between naturalistic stress exposure and adolescent depression. *Emotion*, *20*(6), 927–938. https://doi.org/10.1037/emo0000630

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

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B. J., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. https://doi.org/10.1016/j.psychres.2008.05.006

Trampe, D., Quoidbach, J., & Taquet, M. (2015). Emotions in everyday life. *PLOS ONE*, *10*(12), Article e0145450. https://doi.org/10.1371/journal.pone.0145450

Vesker, M., Bahn, D., Degé, F., Kauschke, C., & Schwarzer, G. (2018). Developmental changes in the categorical processing of positive and negative facial expressions. *PLOS ONE*, *13*(8), Article e0201521. https://doi.org/10.1371/journal.pone.0201521

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence—Second edition (WASI-II)*. Pearson.

Weissman, D. G. (2023). *Perceptual sensitivity to emotion expressions: Associations with age and psychopathology*. PsyArXiv. https://osf.io/6jms7/

Weissman, D. G., Nook, E. C., Dews, A. A., Miller, A. B., Lambert, H. K., Sasse, S. F., Somerville, L. H., & McLaughlin, K. A. (2020). Low emotional awareness as a transdiagnostic mechanism underlying psychopathology in adolescence. *Clinical Psychological Science*, *8*(6), 971–988. https://doi.org/10.1177/2167702620923649

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

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

Williams, W. C., Morelli, S. A., Ong, D. C., & Zaki, J. (2018). Interpersonal emotion regulation: Implications for affiliation, perceived support, relationships, and well-being. *Journal of Personality and Social Psychology*, *115*(2), 224–254. https://doi.org/10.1037/pspi0000132

Wood, S. (2022). *mgcv: Mixed gam computation vehicle with automatic smoothness estimation* (Version 1.8-40) [Computer software]. https://CRAN.R-project.org/package=mgcv

Woodard, K., Zettersten, M., & Pollak, S. D. (2022). The representation of emotion knowledge across development. *Child Development*, *93*(3), e237–e250. https://doi.org/10.1111/cdev.13716

Received November 7, 2023
Revision received August 1, 2024
Accepted August 2, 2024 ■