Behavioural Processes 164 (2019) 193–200

Contents lists available at ScienceDirect

Behavioural Processes

journal homepage: www.elsevier.com/locate/behavproc

Children's perception of emotions in the context of live interactions: Eye movements and emotion judgements

Nicole L. Nelson[a,b], Catherine J. Mondloch[a,*]

[a] Brock University, Department of Psychology, 1812 Sir Isaac Brock Way, L2S3A1, St. Catharines, ON, United States
[b] University of Queensland, School of Psychology, Brisbane, QLD, Australia

ARTICLE INFO

Keywords:
Emotion
Development
Facial expressions
Eye-tracking
Interactive
Social cognition

ABSTRACT

Research examining children's understanding of emotional expressions has generally used static, isolated facial expressions presented in a non-interactive context. However, these tasks do not resemble children's experiences with expressions in daily life, where they must attend to a range of information, including others' facial expressions, movements, and the situation surrounding the expression. In this research, we examine the development of visual attention to another's emotional expressions during a live interaction. Via an eye-tracker, children (4–11 years old) and adults viewed an experimenter open a series of opaque boxes and make an expression (happiness, sadness, fear, or disgust) based on the object inside. Participants determined which of four possible objects (stickers, a broken toy, a spider, or dog poop) was in the box. We examined the proportion of the trial in which participants looked to three areas of the face (the eyes, mouth, and nose area), and the available contextual information (the box held by the experimenter, the four objects). Although children spent less time looking to the face than adults did, their pattern of visual attention within the face and to object AOIs did not differ from that of adults. Finally, for adults, increased accuracy was linked to spending less time looking to the objects whereas increased accuracy for children was not strongly linked to any emotion cue. These data indicate that although children spend less time looking to the face during live interactions than adults do, the proportion of time spent looking to areas of the face and context are generally adult-like.

1. Children's perception of emotions in the context of live interactions

1.1. Eye movements and emotion judgements

Several theories have been put forward to explain emotion and its development (see Camras, 2011 for a review). Izard's *differential emotion theory* proposes that discrete emotions develop according to a pre-programmed timetable (e.g., Izard and Malatesta, 1987); Campos' *functionalist view* proposes that emotions are not pre-programmed, but rather develop so as to maintain the individual's desired states/outcomes (Barrett and Campos, 1987); and Bridge's *differentiation view* proposes that discrete emotions emerge from undifferentiated states (e.g., fear and anger differentiate from a general negatively-valenced high-arousal state; Bridges, 1932). Researchers who investigate the development of emotion perception (i.e., the ability to recognise and/or label facial displays of emotion) typically contrast the discrete emotion theory, which posits universal and early recognition (Izard, 1994; Walle et al., 2017), with the differentiation view, which emphasises the slow development of emotion recognition during infancy and early childhood (Widen, 2013; Widen and Russell, 2008).

To measure children's emotion knowledge in laboratory tasks, children are generally presented with images of intense, posed, static facial expressions and asked to categorise or label them. Infants can discriminate expressions by their first birthday (Bornstein and Arterberry, 2003; Caron et al., 1982; Walker-Andrews, 1997) and can use others' expressions to guide their behaviour (e.g. *social referencing*; Repacholi and Meltzoff, 2007; Sorce et al., 1985; Walle et al., 2017). Nonetheless, a growing body of evidence suggests that children experience a prolonged development of emotion recognition. Notably, pre-school children do not differentiate among the negatively-valenced expressions (see Widen and Russell, 2008 and Widen, 2013 for a review), a pattern of results consistent with Bridges's (1932) differentiation view of early experiences of emotion (see also Sroufe, 1997). Children begin with broad emotion categories for "good" and "bad" and throughout childhood these categories are further differentiated into the emotion categories familiar to adults, such as anger, disgust, or fear (Camras and Allison, 1985; Durand et al., 2007; Kolb et al., 1992; Vicari

* Corresponding author.
E-mail address: cmondloch@brocku.ca (C.J. Mondloch).

https://doi.org/10.1016/j.beproc.2019.05.006
Received 26 November 2018; Received in revised form 1 May 2019; Accepted 3 May 2019
Available online 07 May 2019

et al., 2000; Widen and Russell, 2003; see Widen, 2013 for a review). It is not until children reach their teen years that they judge emotional expressions in the same way adults do (Widen, 2013; Widen et al., 2015).

Using static facial expressions to test emotion knowledge may not provide an accurate reflection of children's ability to use emotional expressions in daily life. Static expressions lack many of the emotional cues that are used to interpret others' emotions during real life interactions, such as movement, shifts in eye gaze direction, and contextual information (Aviezer et al., 2008; Nelson and Russell, 2011; Nelson and Mondloch, 2017, 2018; Vieillard and Guidetti, 2009). In daily life, expressions are dynamic, include a variety of cues to a person's emotion, and are part of a social, interactive experience. Research has shown that eliminating these cues particularly impacts children's performance in social and emotional tasks (e.g. Nelson and Mondloch, 2017, 2018; Chandler et al., 1989; Walle et al., 2017; Xiao et al., 2015; Xiao et al., 2017). It is possible that using expressions devoid of movement and contextual cues is the reason children show a protracted development of emotion knowledge in most studies of emotion recognition.

Such a view is entirely consistent with the Behaviour Systems View—a perspective that was championed by Bill Timberlake and colleagues (e.g., Gawley et al., 1987; Silva and Timberlake, 1997, 2005; Timberlake, 1983, 1984) and highlights the importance considering a functional account of learning. The dynamical system perspective provides an especially useful framework for emotion perception and its development. Rather than emphasizing prototypical expressions and one-on-one mappings between emotion and expression, the dynamical systems approach (DSA) posits that emotions are self-organizing dynamical systems influenced by both top-down and bottom-up (e.g., other ongoing behaviour) factors. Specific emotions (e.g., anger, fear) are best thought of as attractor states comprised of appraisal, physiological, cognitive, and expressive components. Notably, variability is an inherent component of such attractor states requiring children to recognise and understand emotions despite variability in their presentation (Camras and Witherington, 2005; Camras, 2011; Camras and Shuster, 2013). As noted by Fogel et al. (1992), developmental change will involve changes in the information to which infants and children are most sensitive. Given that emotions are not discrete entities controlled by emotion-specific control centres, but rather processes that emerge in the context of an individual's ongoing activity (Fogel et al., 1992), researchers must investigate children's perception of emotional expressions in the context of ongoing social interactions.

The aim of our study was not to provide a direct test of dynamical systems theory, but rather to examine children's sensitivity to facial expressions under more naturalistic, and yet controlled, conditions—an approach consistent with dynamical systems theory. We examined children's visual attention to facial expressions of emotion and their ability to accurately interpret facial expressions in the context of a controlled social interaction that included both dynamic expressions and a social context. Unlike static facial expressions, the controlled social interaction in this study presented children with dynamic emotional expressions. Previous research suggests that dynamic expressions are better recognised than static expressions by both adults and children (Calvo et al., 2016; Dukes et al., 2017; Krumhuber et al., 2013; Nelson and Mondloch, 2018; Nelson and Russell, 2015; Richoz et al., 2018). The benefits of dynamic expressions are particularly apparent when emotional expressions are subtle or degraded (e.g. point light displays) (Krumhuber et al., 2013). Importantly, children show adult-like visual processing (e.g., eye movements) of face and body emotional expressions when the expressions are dynamic but not static (Nelson and Mondloch, 2017, 2018) and are more influenced by the presentation of movement in complex scenes than are adults (Stoesz and Jakobson, 2014). Together, these studies suggest that using static stimuli interferes with children's recognition and visual scanning of a broad range of stimuli.

In addition to being dynamic, our emotional stimuli were presented in the context of a social interaction. Whereas the majority of research examining children's emotion knowledge has relied on tasks in which children passively view stimuli and make a judgement about another person's emotion, our task approximated daily interactions in which interpreting others' emotions is an inherently social and interactive process. When socio-cognitive tasks are presented to children in an interactive manner, children demonstrate more sophisticated skills than has been observed using passive tasks. For example, theory of mind tasks that require children to passively listen to a story present evidence for the emergence of explicit false belief understanding around four years of age (Perner et al., 1987; Wellman and Liu, 2004). In contrast, an interactive task that required children's active participation in "tricking" an experimenter showed evidence of false belief understanding as early as two and a half years (Chandler et al., 1989). Similarly, passive emotion recognition tasks suggest children's emotion knowledge develops slowly, with emotion category knowledge expanding greatly between four and 11 years (Izard, 1971; Widen, 2013; Widen and Russell, 2008; Vieillard and Guidetti, 2009). However, interactive tasks show that as early as 12 months, toddlers selectively approach and interact with objects based on others' emotional responses (Sorce et al., 1985; Repacholi and Meltzoff, 2007; Walle et al., 2017) and by 15 months they generalise another's expression to future interactions with that person (Repacholi et al., 2016). Furthermore both adults' (Laidlaw et al., 2011) and infants' (Franchak et al., 2011; Yu and Smith, 2013) visual attention is modulated by the nature of the interaction in which they are engaged (e.g., social vs. non-social; availability of information), suggesting that emotions presented in a controlled social interaction might be attended to differently than emotions presented in the form of static images. Children might demonstrate a prolonged development of emotion knowledge because passive emotion recognition tasks do not resemble their daily experiences in which interpreting another's emotions are most important.

In this study, we developed a novel protocol to examine participants' visual attention to another's emotional expression during a live interaction. Via a monitor linked to an eye-tracker, child (4–11 years old) and adult participants played a game with an experimenter who opened a series of boxes and produced a facial expression (happy, sad, disgusted, or fearful). Participants then guessed which of four possible objects (also visible on the screen) was inside the box (either a sticker, a broken toy, dog poop, or a spider). We recorded time spent looking at the experimenter's face, the box, and the four objects as well as the accuracy with which children provided the correct object label. This method allowed us to examine the development of visual attention to emotional expressions during a live interaction (a context that incorporates a wealth of cues), as well as children's accuracy in linking expressions to the causes of emotional expressions.

Because children interacted with a live experimenter, each expression displayed (n_{trials} = 368; 46 participants x 8 trials each) was unique. Areas of interest (AOIs) were marked for each trial and the dynamic nature of the AOIs required us to mark each AOI for each frame of the video. Our aim was to provide evidence as to whether children's scanning patterns, defined here as time spent looking to various areas of the face vs. the array of objects, and accuracy in labelling differed from those of adults and to identify any preliminary evidence of age-related change among children for future investigation. Thus, we elected to test a sample of children comprising a broad age range (4–11 years) and contrast them to a group of adults. We tested children aged 4–11 years because by 4 years of age children's allocation of visual attention is comparable to those of adults for dynamic whole-person stimuli when asked to provide a verbal label (Nelson and Mondloch, 2018) and they have also developed the theory of mind skills necessary to complete our task (Wellman et al., 2011; Wellman and Liu, 2004).

In addition to measuring the accuracy with which participants matched objects to each expression we measured time spent looking at three areas in the experimenter's face (eyes, nose, and mouth), each of the four objects in the array (stickers, broken car, spider, poop), and the box being held by the experimenter. We previously reported that proportion time spent looking at the face was comparable in 5-year-olds and adults when children were asked to label an expression (Nelson and Mondloch, 2018). Here we examined whether that finding held when the face competed with four objects and whether the allocation of looking within the face differed for children vs. adults. We examined looking at the four objects as past research has shown that children's looking to target and distractor objects can provide insight into how quickly they hone in on the correct object (Halberda, 2006). We predicted less differentiated looking at the four objects for emotions associated with lower accuracy (e.g., fear, disgust). We expected that children might scan more negative objects when the experimenter displayed a negatively valenced expression, perhaps using a process of elimination to determine which object was the correct one (e.g. Nelson & Russell, 2016a; Nelson and Russell, 2016b). We also examined children's looking to the box held by the experimenter as it was the referent of the experimenter's emotion. We expected that children might attend to the box longer when they were unsure of the experimenter's emotion (i.e., for emotional displays with lower accuracy).

Although our primary goal was to study children's emotional knowledge in a dynamic interactive task, we also included two control tasks. In one, participants labelled the emotion they thought would be elicited by each object (e.g., happy for the stickers, sad for the broken toy, disgusted for the dog poop and fearful for the spider); we did this to verify that children and adults both ascribed the appropriate label to each object. In the other control task, participants labelled a set of photographs of standardised facial expressions to provide a comparison for our interactive task.

2. Method

2.1. Participants

Participants were 25 children (4- to 11-years old, mean age: 7.8 years, SD: 1.7 years; 15 female), and 21 adults (mean age 19.5 years, SD: 2.1 years, 19 female). An additional 27 participants (15 children, 12 adults) were excluded from analysis due to technical failures (6 children, 12 adults), failure to follow instructions (2 children), or because the eye-tracker was unable to capture their fixation data throughout at least 70% of the testing protocol (7 children). An examination of participant accuracy in the interactive task showed no difference between participants who were included and excluded,[1] indicating that less accurate participants were not disproportionately excluded from the study.

2.2. Materials and procedure

Adults and children were tested with the identical protocol. Participants first participated in the Interactive task, followed by two control tasks: the Object Labeling and Expression Labeling tasks. Participants were seated at a 65 cm viewing distance from a 24-inch Tobii T60 XL eye-tracker (approximately 0.5° of precision, 60 Hz sampling rate, 1440 × 900 pixels resolution). On one half of the screen was the face of the experimenter and on the other half were pictures of the four possible objects that might be in the box opened by the experimenter in the Interactive task. The face of the experimenter with whom the participant interacted measured approximately 11.2° and 7.0° of vertical and horizontal visual angle respectively. For the face AOIs, the eyes measured approximately 2.2° and 5.7° of vertical and horizontal visual angle, the nose area measured approximately 2.2° and 3.5° of vertical and horizontal visual angle, and the mouth measured approximately 2.6° and 3.5° of vertical and horizontal visual angle. Each of the objects measured approximately 11.6° and 7.9° of vertical and horizontal visual angle.

2.2.1. Interactive task

The testing protocol was identical for adults and children. Participants were introduced to two experimenters (E1 and E2). Three different female experimenters served as E1 and E2. E1 then showed participants four example boxes that they were invited to open, one at a time. Each box contained a pair of objects that would evoke the same emotion. One box contained stickers and toy donuts (happiness), one contained a broken car and broken jacks (sadness), one contained a spider and a rat (fear), and one contained a toy eyeball and toy dog poop (disgust). Each object in the boxes was named for participants, but the emotion associated with the objects was not. Having participants open example boxes served several functions. It gave children the opportunity to interact with both E1 and E2, familiarised participants with the appearance and typical facial movements of E1 in particular, provided participants multiple exemplars of the kinds of objects associated with each emotion category, and gave them a first-person opportunity to experience emotions linked to each object while opening boxes themselves.

Next, E1 moved to the adjoining room to appear on the live video feed while E2 placed the participant in front of the eye-tracker monitor. The 24-inch monitor was divided vertically into two equal halves. In one half, participants saw E1 sitting in the adjacent room. In the other half of the monitor were images of four objects (a sticker, broken toy, spider, and dog poop), one from each of the boxes they had just opened (Fig. 1). We showed participants both the experimenter and the object array throughout the trial to reduce memory demands, to see whether time spent looking at distractor objects varied across emotions, and to see whether time spent looking at the face vs. the display of four objects varied across emotions. E2 explained that E1 would open a series of boxes identical to the four boxes they had just opened, and that one object would be inside each box. Participants were asked to guess what object was in the box based on the facial expression of E1, choosing one of the four objects visible on the monitor. To avoid participants relying on a process of elimination to guess the correct answer, E2 explained that E1 had many boxes she would be opening.

After a five-point eye tracking calibration procedure, the Interactive task trials started. Each trial began with E1 on the screen looking directly at the camera. E1 looked into the camera and asked the participants if they were ready for her to open the next box and waited until the participant responded in the affirmative. She then picked up a random box from a pile of boxes, opened it and looked inside, and then showed a facial expression in reaction to the object she saw. She then looked immediately to the camera for 1 s, glanced back at the object while still showing the expression and looked to the camera again for 1 s. Thus, the experimenter displayed each expression directly at the camera for 2 s per trial. E1 then closed the box and waited with a neutral expression while participants named the object they believed to be in the box. Responses were recorded by E2 who was in the room with the participant. To prevent anticipatory looking to the four objects, the location of each object in the array varied randomly across trials. Further, to avoid the participant using verbal cues the child could not hear the experimenter and to avoid the experimenter being influenced by the participant's behaviour the experimenter could not see the child. To ensure the task was interactive given these constraints, E1 could hear the participant and asked them if they were ready between boxes.

Participants viewed each of the four expressions seven times, which resulted in a total of 28 trials. The expressions were presented in a different random order for each participant with the following constraints: 1) within the first 8 trials, each expression was presented twice, and 2) no expression was presented more than two times in a row.

[1] A between-within 2 (age group) x 4 (emotion) x 2 (included vs excluded) ANOVA examining accuracy in the interactive task showed no main effects or interactions related to inclusion status (all $ps > .10$).

Fig. 1. Example of a happy trial with E1 on the left side of the screen and the four objects on the right. Objects are, clockwise from top left: broken toy, dog poop, spider, stickers.

2.2.2. Object labeling task

Next, both adult and child participants freely labelled the emotion associated with the four sets of objects. E2 presented participants with the four boxes they had opened previously and asked "What emotion would you feel if you opened a box and saw this?" Participants were free to label the objects with an emotion label of their choice. This task was designed to verify that participants associated the objects with the intended emotion category.

2.2.3. Expression labeling task

As a final task, participants freely labelled twelve 5 x 7 inch photographs of facial expressions drawn from the ADFES stimulus set (Van Der Schalk et al., 2011). The expressions were presented one at a time and intended to depict one of six emotions: happiness, anger, disgust, fear, sadness or surprise. The set of expressions include one male and one female poser for each expression, and each poser was shown only once. E2 presented participants the expression photographs one at a time and participants answered the question "What emotion is this person feeling?" Participants were free to label the expression with an emotion label of their choice. This task was designed to provide a baseline against which to compare accuracy of facial expression recognition in our interactive task.

2.3. Analyses

2.3.1. Accuracy

We analyzed accuracy in the interactive task, the object labeling task, and the expression labeling task by conducting separate 2 (Age group: children, adults) × 4 (Emotion: happy, sad, fearful, or disgusted) between-within ANOVAs for each task[2] . Preliminary analyses revealed practice effects in the interactive task; participants approached ceiling after the first 8 trials, with recognition across emotions exceeding 91% (85% for child participants and 96% for adults. See Figure S2 for more information). This change in performance across trials suggests that participants were simply matching expressions they had previously learned in the task (e.g. Nelson and Russell, 2016b), rather than spontaneously judging each expression for each trial. Thus, all subsequent analyses are based only on the first eight trials (two per emotion).

2.3.2. Areas of interest

We defined several areas of interest (AOI) including three areas in the face of E1 (eyes, nose area—which included the centre of the face, and mouth), each of the four objects (stickers, broken car, spider, poop) in the array, and the box being opened by E1. A new fixation was initiated when the eye moved more than 35 pixels. Because the expressions were performed live, for each participant the length of the trials varied slightly for each expression presented. Thus, our dependent variable for looking was the total time participants looked to each AOI as a proportion of the overall length of each trial. Looking behavior was analyzed from the frame in which the experimenter began opening the box to the frame in which she closed the box.

2.3.3. Scoring

Following a procedure outlined by Widen and Russell (2003), participants' free labeling responses for the Object labeling and Expression labeling tasks were sorted by three independent coders into one of the following categories: happiness, sadness, anger, fear, surprise, or disgust. Any responses for which coders could not agree were categorised as "other." Responses categorised as correct for happiness were: *happy, glad, excited*; for sadness were *sad, disappointed*; for disgust were *grossed out, gross, icky, yucky, disgusted*; and for fear were *scared, frightened, freaked out, nervous.*

2.3.4. Power

A post-hoc power analysis using G*Power (Faul et al., 2007) showed that with a sample size of 48, our primary analysis of a 2(age groups) × 4(emotions) repeated measures ANOVA was sensitive enough to detect a medium-sized effect ($\eta_p{}^2 =$ .06, $\alpha = 0.05$), with *power* reaching 0.99.

3. Results

3.1. Accuracy

Overall, participant's accuracy was high across the interactive task (80%), the object labeling task (87%) and the expression labeling task (91%). Bonferroni-corrected paired-samples *t*-tests (adjusted $\alpha = .05/$

[2] Initial ANCOVA analyses examining children's data for each task and using age as a continuous measure, revealed no effect of age (*p*s > .08). An examination of individual data points revealed no evidence supporting an effect of age within child participants (Figure S1). Thus, we collapsed children into a single age group for future analyses.

3 = .017) indicated that participants were less accurate in the interactive task than the object labeling task, $t(45) = 3.02$, $p = .004$, $d = 0.45$.

We examined participant's performance by conducting separate 2 (age group: children, adults) x 4 (emotion: happy, sad, fearful, or disgusted) mixed-design repeated measures ANOVAs for each task (interactive task, object labeling, expression labeling). Our DV was the proportion of trials in which participants produced the expected response. Main effects and interactions were followed up with Bonferroni-corrected post hoc tests.

For the Interactive task, the main effect of age indicated that adults were more accurate than children, $F(1, 44) = 6.30$, $p = .016$, $\eta_p^2 = .13$. The main effect of emotion, $F(3, 132) = 6.564$, $p < .001$, $\eta_p^2 = .13$, indicated that accuracy for happy was greater than for sad, fearful, and disgust expressions ($ps < 0.05$). The age x emotion interaction was not significant $F(3, 132) = .80$, $p = .50$ (For details see Figure S3).

For the Object task, there was no main effect of age, $F(1, 44) = 3.52$, $p = .07$, and no age x emotion interaction, $F(3, 132) = 1.01$, $p = .39$. The main effect of emotion, $F(3, 132) = 5.32$, $p = .002$, $\eta_p^2 = .11$, indicated that accuracy for happy and sad objects was greater than for disgusted objects ($ps < .023$); all other comparisons were not significant ($ps > .06$) (Figure S3).

For the Expression labeling task, there were main effects of age, $F(1, 44) = 5.01$, $p = .03$, $\eta_p^2 = .10$, and of emotion, $F(3, 132) = 9.99$, $p < .001$, $\eta_p^2 = .19$. These main effects were superseded by an age x emotion interaction, $F(3, 132) = 3.98$, $p = .009$, $\eta_p^2 = .08$. Adult's performance was higher than children's for disgust expressions only ($p = .002$, all other $ps = 1.00$), (Figure S3).

3.2. Visual attention during the interactive task

We next conducted a series of ANOVAs examining how allocation of visual attention varied with age group and the emotion expressed by the experimenter. Main effects and interactions were followed up with Bonferroni-corrected post hoc tests. As these analyses are focused on the AOIs examined, we report only the effects related to AOI.

3.2.1. Looking to the face, objects, and box

We first examined whether looking to the three main areas of the stimuli – the face, the objects, and the box held by the experimenter – varied with age and emotion. We conducted a 2 (age group: children, adults) × 4 (experimenter's expression: happy, sad, fearful, or disgusted) × 3 (Areas: Face, Objects, Box) mixed-design repeated measures ANOVA. Our DV was the proportion of the trial in which participants looked to each area.

The main effect of Area, $F(2, 88) = 192.36$, $p < .001$, $\eta_p^2 = .81$, showed that participants looked longer to the face (50.4%), followed by the objects (20.5%) and the box (4.9%) (all $p < .001$) (Fig. 2). The Age x Area interaction and the Emotion × Area interaction were superseded by an Age × Area × Emotion interaction, $F(6, 264) = 2.16$, $p = .05$, $\eta_p^2 = .05$, which indicated that adults spent more time than children looking to the face for happy, scared, and disgusted expressions ($ps < .001$). Children and adults looked to the sad face for a similar proportion of time ($p = .64$). Children and adults also spent a similar amount of time looking to the Objects ($ps = 1.00$) and the Box ($ps = 1.00$). In summary, participants spent the most time looking to the face, although children spent less time looking to the face than adults did. This was true for most expressions presented, with the exception of sadness.

3.2.2. Looking to the face AOIs

We next examined whether scanning of the experimenter's face areas – eyes, nose area and mouth – varied with age and emotion. We conducted a 2 (age group: children, adults) × 4 (experimenter's expression: happy, sad, fearful, or disgusted) × 3 (AOI: Face, Nose, Mouth) mixed-design repeated measures ANOVA. Our DV was the proportion of the trial in which participants looked to each AOI.

Although there was no main effect of AOI, $F(2, 88) = 1.27$, $p = .28$, the AOI × emotion effect was marginal, $F(2, 88) = 1.98$, $p = .07$, $\eta_p^2 = .04$. As AOI is our variable of interest, and to avoid type II error, we further examined this effect, which showed that for happy expressions, participants looked longer to the mouth than the eyes, $p = .003$. No other comparisons were significant ($ps > .39$). In summary, adults' and children's visual attention to face areas was similar and did not strongly vary with AOI or emotion. These data suggest that although children look less to the face than adults do, their allocation of looking within the face is adult-like.

3.2.3. Looking to the objects

Finally, we examined whether scanning of the objects differed across emotion and age. We conducted a 2 (age group: children, adults) × 4 (experimenter's expression: happy, sad, fearful, or disgusted) × 3 (Object AOI: spider, stickers, broken toy, poop) mixed-design repeated measures ANOVA. Our DV was the proportion of the trial in which participants looked to each object AOI.

The main effect of object AOIs indicated that participants looked longest to the broken toy ($ps < .04$), followed by the stickers and the spider, which were not different from each other ($ps = 1.00$), and least at the poop ($ps < .03$) (Fig. 3). The emotion x object interaction indicated that for expressions of happiness and sadness, participants looked longer to the target object (stickers for happiness, the broken toy for sadness) than any other object ($ps < .001$). For expressions of fear, participants looked longer to the spider than to the poop ($p = .002$) but looking to the spider was not different from the other two objects ($ps > .12$). Finally, for disgust expressions, participants looked longer to the poop than the stickers ($p = .04$) but looking to poop was not different from the other two objects ($ps = 1.00$).

In short, participants looked longer at the target object than (at least some) other objects, with a less differentiated pattern for fear and disgust expressions than for happy and sad expressions. In addition, there was no main effect or interaction related to age. Thus, adults and children showed similar visual attention to the objects, suggesting that, as with face AOIs, children's allocation of visual attention is adult-like.

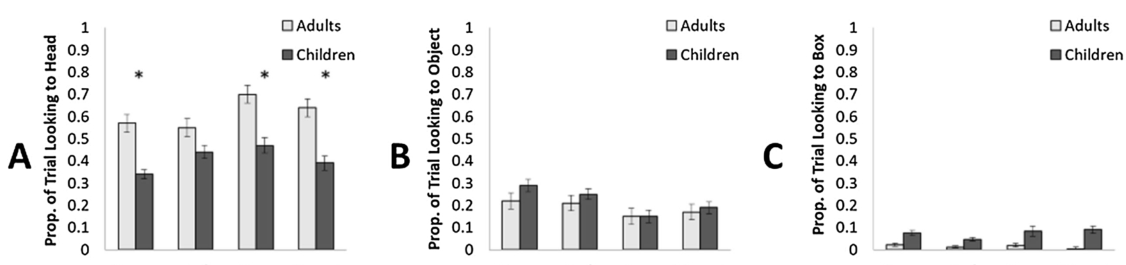

Fig. 2. Proportion of the trial in which adults and children looked to the (A) face, (B) objects, and (C) box, by emotion. Error bars are standard error.

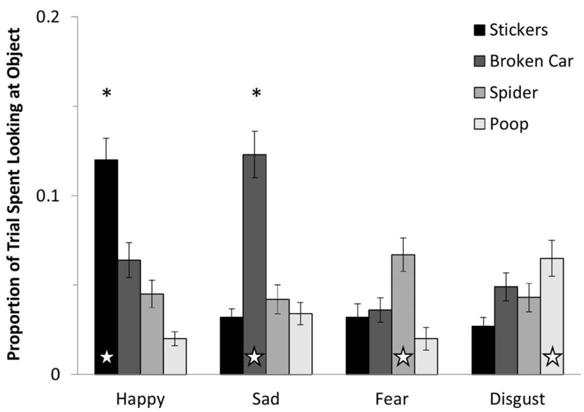

Fig. 3. Proportion of the trial participants looked to each object, by emotion. Bars with a star at the base are the target object for that trial. Error bars represent standard error.

3.3. Relationship between visual attention and accuracy

To determine whether visual attention to the AOIs was related to successfully matching the expression and the object in the box, we examined the correlation between participants' performance in the interactive task and the proportion of the trial spent looking to each of the eight AOIs. We used nonparametric Spearman's correlations and because this analysis was exploratory, we did not correct for multiple comparisons. We also examined these relationships for children and adults separately (Table 1) to determine whether children and adults showed similar patterns and to highlight any differential patterns that could influence future research.

For adults, visual attention was associated with higher performance in the interactive task only for disgust expressions. Specifically, looking more to the mouth ($p = .002$) and less to the eyes ($p = .008$) were both associated with increased accuracy for disgust expressions. For happy, sad, and fearful expressions there were no correlations with looking to any AOI ($ps > .073$).

For children, there was more evidence that allocation of looking was associated with accuracy in the interactive task, particularly in relation to their proportion of time looking at the objects and box. For disgust expressions, looking less to the spider ($p < .001$) and more to the box the experimenter was holding ($p = .006$) were associated with increased accuracy. For scared expressions, looking more to the spider ($p = .04$) was associated with increased accuracy. For happy expressions, looking less to the eyes ($p = .02$) and the poop ($p = .04$) predicted increased accuracy. For sad expressions, there were no correlations with looking to any AOI ($ps > .11$).

Overall, children's accuracy in the interactive task was generally associated with visual attention to the objects, suggesting that in cases where children's judgements were less accurate, they spent more time considering the different the objects. However, for adults, links between accuracy and allocation of looking were limited to face areas on the disgust expressions and suggest that looking to the mouth area increased accuracy.

4. Discussion

Our data demonstrate that passive emotion recognition tasks do not necessarily underestimate children's emotion knowledge. Participants were less likely to select the correct object in the interactive task than to correctly label the emotion associated with that object or to correctly label posed, standardised expressions. Furthermore, in the interactive task adults were more accurate than children regardless of emotion displayed, but in the expression labeling task this was true only when faces displayed disgust. Thus, it is possible that passive tasks actually overestimate both children's and adult's knowledge compared to how they use emotional information in daily life, perhaps because passive viewing tasks better allow children to use strategies such as a process of elimination to guess the correct answer (Nelson and Russell, 2016b). A richer understanding of how children decode facial displays of emotion might be obtained, then, in the context of ongoing social interactions.

We report two primary findings with respect to participants' looking at the Experimenter's face. First, children spent less of the trial looking to the face than adults did. This finding echoes other work indicating that children's visual attention to dynamic, complex scenes is broader than that of adults (Nelson and Mondloch, 2017, 2018; Stoesz and Jakobson, 2014). Second, when we examined children's visual attention to the face AOIs and to the object AOIs, children's looking patterns were generally similar to those of adults. These findings indicate that children were attending to the available emotion cues presented and show adult-like allocation of visual attention to complex, social scenes from as early as four years old. What appears to increase with age is the amount of time spent looking to the face in general rather than changes in visual attention to particular areas of the face or changes in how children attended to the objects or contextual information. Rather than being detrimental, children's attention to non-facial cues may be beneficial, assisting them in learning which facial expressions and situations tend to occur together (Nelson and Mondloch, 2018). Children's broad attention in our tasks also parallels research showing that preschool children demonstrate greater processing of task-irrelevant information than adults do, which results in decreased performance in assigned tasks, but a greater acquisition of information – both relevant and irrelevant – overall (e.g. Plebanek and Sloutsky, 2017).

Adults' and children's scanning of the four objects in the array also provides novel insights about the cognitive processes underlying emotion recognition and understanding. First, only for happy and sad expressions did participants look longer at the matching object (stickers and broken toy, respectively) than all non-matching objects. Less differentiated looking on fear and disgust trials might reflect greater difficulty recognizing fearful and disgusted expressions as compared to happy and sad expressions. This hypothesis is consistent with the negative correlation between children's accuracy and time spent looking at distractor objects. The tendency to look more at non-matching objects for fearful and disgust expressions is consistent with Halberda's (2006) report on visual scanning in a word learning task in which participants were asked which of two objects matched a verbal label.

Table 1
Spearman correlations between accuracy and looking to each AOI by emotion. Note. *** $p < 0.001$, ** $p < 0.010$, * $p < 0.05$.

Experimenter's Expression	Age Group	Eyes	Nose Area	Mouth	Sticker	Broken Toy	Poop	Spider	Box
Happy	Adults	0.231	0.258	0.148	−0.130	−0.391	−0.399	−0.399	0.184
	Children	−0.469*	0.194	−0.139	0.173	−0.361	−0.421*	−0.209	−0.076
Sad	Adults	−0.110	−0.140	0.180	0.118	0.298	−0.090	−0.220	0.126
	Children	−0.042	0.052	−0.095	−0.283	−0.328	−0.111	−0.259	−0.316
Fear	Adults	−0.042	−0.093	0.329	−0.399	−0.247	−0.189	0.023	0.040
	Children	−0.018	−0.219	0.152	0.349	−0.097	−0.24	0.418*	0.076
Disgust	Adults	−0.561**	−0.018	0.639**	−0.092	−0.021	0.245	−0.266	0.0100
	Children	0.195	−0.063	−0.258	−0.232	−0.299	0.255	−0.628***	0.531**

Upon hearing a novel label (e.g., *dax*) both adults and pre-school children looked towards a familiar object (i.e., double-checked that object) before selecting a novel object as the matching target. In contrast, upon hearing a familiar label (e.g., *cup*) neither group look towards the unfamiliar object in the pair before selecting the familiar object as the match. Thus, despite the object array in our study providing no cues as to which emotion the Experimenter displayed, both adults and children scanned those objects more when the task was arguably more difficult. The fact that children's scanning of the objects was correlated with accuracy might reflect that children found the task of linking expressions and emotional situations more challenging than did adults.

Abundant evidence suggests that children begin with very broad emotion categories (e.g., happy vs. not happy; see Bridges, 1932; see Widen, 2013 for a review) and gradually come to recognise more specific categories (e.g., fear, anger, and disgust). Our findings suggest that both children and adults rely on multiple sources of information when making emotion judgements, a pattern of results consistent with evidence that their perception of both static and dynamic facial expressions is influenced by the context (e.g., body posture, background scene) in which the faces are presented (Aviezer et al., 2008; Kret and de Gelder, 2010; Kret et al., 2013; Mondloch et al., 2013; Nelson and Mondloch, 2017, 2018). Our findings suggest that in responding to emotional information children (and adults) actively explore a wealth of cues, including potential causes of the expression such as objects.

Our research provides suggestions for future research examining the development of socio-emotional skills. Follow-up studies should increase task difficulty (e.g., by presenting subtle expressions and varying possible objects from trial-to-trial) in order to determine under what conditions children's attention to emotional cues and their ability to utilise emotional expressions in social interactions is adult-like. Future research might also examine a wider range of emotions and situations, particularly those that are closely linked to social interaction such as pride or embarrassment, with an emphasis on the temporal dynamics of emotion perception. For example, when emotional expressions are subtle do children scan the face more? Or, do they spend more time looking at non-face (e.g., postural or other contextual) cues?

One limitation of our study is that we tested a wide age range of children and a relatively small sample. The small sample simply reflects how labour-intensive it was to mark dynamic AOIs for each individual participant for each of eight trials. It would, of course, be valuable to test more children at each age. We note that we saw little evidence of age-related improvement in accuracy (see Figure S1). Whereas only younger children made any errors for happy and sad expressions, children at all ages made errors for angry and fearful expressions with no evidence of age-related improvement. However, our small sample size precluded looking at age-related changes in scanning patterns within our child participants. Future studies with a larger sample might investigate whether individual differences in accuracy correlate with differences in visual attention, perhaps using a protocol in which all children watch a pre-recorded adult opening a series of boxes. Future studies might also use larger facial stimuli in order to more closely examine whether children's visual attention to areas of faces is different from that of adults'. Doing so would also allow researchers to investigate whether subtle age-related changes in visual attention exist or whether details in scanning patterns (e.g., the order in which AOIs are scanned, duration of each look) vary with age. Such an approach would ensure that any scanning differences are not attributable to differences in the experimenter's behaviour, but would also compromise the generalizability of such a task, making the interaction less natural.

This research provides a novel protocol with which to examine the development of children's sensitivity to emotional cues. Our findings support recent literature showing that children demonstrate different patterns of performance during social, interactive tasks vs. tasks in which they are passive observers (Walle et al., 2017; Repacholi et al., 2016). Particularly when examining the development of a socio-emotional skill, incorporating social, interactive tasks provides children the most naturalistic opportunity to demonstrate their skill and also ensures that the cues they would use in everyday life when interacting with others are available. It is striking that a similar call to take the ecological context in which learning evolved and in which it develops has been made by researchers in the field of animal learning and behavior for several decades (Gawley et al., 1987; Silva and Timberlake, 1997, 2005; Timberlake, 1983, 1984, 2001). For example, Breland and Breland (1961) reported that animals' behaviour in an operant conditioning paradigm is impacted by instinctive behaviours, a phenomenon they dubbed instinctual drift. Findings like theirs ultimately led to researchers to consider the evolution and function of behaviour both in experimental protocols (e.g., conditioning chamber design) and in theoretical models. We echo that call for researchers who study the development of emotion recognition. Researchers must move to more testing paradigms that tap into how children perceive emotion outside of the laboratory in order to gain a comprehensive understanding of children's skills and how they interact with others in daily life. Such paradigms should involve dynamic stimuli and controlled social interactions.

Author note

This research was funded by a SSHRC Insight grant and Canada Foundation for Innovation grant to CJM and an NSF IRFP awarded to NN. We thank Breanna Elliott and Jasmine Bender for helping to pose faces, Jasmine Bender, Harmonie Chan, Lewis Lau, Allison Mondloch, and Gabriela Saldago, for helping to process data, Bryce Hunt for technical assistance, and the parents and children who participated.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.beproc.2019.05.006.

References

Aviezer, H., Hassin, R.H., Ryan, J., Grady, C., Susskind, J., Anderson, A., 2008. Angry, disgusted, or afraid? Studies on the malleability of emotion perception. Psychol. Sci. 19, 724–732. https://doi.org/10.1111/j.1467-9280.2008.02148.x.

Barrett, K., Campos, J., 1987. Perspectives on emotional development II: a functionalist approach to emotions. In: Osofsky, J. (Ed.), Handbook of Infant Development, 2nd ed. Wiley, New York, NY, pp. 555–578 Barrett, K.C., Zahn-Waxler, C., & Cole, P. M. (1993).

Bornstein, M.H., Arterberry, M.E., 2003. Recognition, discrimination and categorization of smiling by 5-month-old infants. Dev. Sci. 6, 585–599. https://doi.org/10.1111/1467-7687.00314.

Breland, K., Breland, M., 1961. The misbehavior of organisms. Am. Psychol. 16 (11), 681–684. https://doi.org/10.1037/h0040090.

Bridges, K.M.B., 1932. Emotional development in early infancy. Child Dev. 3, 324–341.

Calvo, M.G., Avero, P., Fernández-Martín, A., Recio, G., 2016. Recognition thresholds for static and dynamic emotional faces. Emotion 16, 1186–1200. https://doi.org/10.1037/emo0000192.

Camras, L.A., 2011. Differentiation, dynamical integration, and functional emotional development. Emot. Rev. 3 (2), 138–146. https://doi.org/10.1177/1754073910387944.

Camras, L.A., Allison, K., 1985. Children's understanding of emotional facial expressions and verbal labels. J. Nonverbal Behav. 9, 84–94. https://doi.org/10.1007/BF00987140.

Camras, L.A., Shuster, M.M., 2013. Current emotion research in developmental psychology. Emot. Rev. 5 (No. 3), 321–329. https://doi.org/10.1177/1754073913477516.

Camras, L.A., Witherington, D.C., 2005. Dynamical systems approaches to emotional development. Dev. Rev. 25, 328–350. https://doi.org/10.1016/j.dr.2005.10.002.

Caron, R.F., Caron, A.J., Myers, R.S., 1982. Abstraction of invariant face expressions in infancy. Child Dev. 53, 1008–1015.

Chandler, M., Fritz, A.S., Hala, S., 1989. Small scale deceit: deception as a marker of 2-, 3-, and 4-year-olds early theories of mind. Child Dev. 60, 1263–1277. https://doi.org/10.2307/1130919.

Dukes, D., Clément, F., Audrin, C., Mortillaro, M., 2017. Looking beyond the static face in emotion recognition: the informative case of interest. Vis. cogn. 25 (4–6), 575–588. https://doi.org/10.1080/13506285.2017.1341441.

Durand, K., Gallay, M., Seigneuric, A., Robichon, F., Baudouin, J.Y., 2007. The development of facial emotion recognition: the role of configural information. J. Exp. Child Psychol. 97, 14–27. https://doi.org/10.1016/j.jecp.2006.12.001.

Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A., 2007. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav.

Res. Methods 39 (2), 175–191. https://doi.org/10.3758/BF03193146.
Fogel, A., Nwokah, E., Dedo, J.Y., Messinger, D., Dickson, K.L., Matusov, E., Holt, A.A., 1992. Social process theory of emotion: a dynamic systems approach. Soc. Dev. 1, 122–142. https://doi.org/10.1111/j.1467-9507.1992.tb00116.x.
Franchak, J.M., Kretch, K.S., Soska, K.C., Adolph, K.E., 2011. Head-mounted eye tracking: a new method to describe infant looking. Child Dev. 82 (6), 1738–1750.
Gawley, D.J., Timberlake, W., Lucas, G.A., 1987. System-specific differences in behavior regulation: overrunning and underdrinking in molar nondeprived schedules. J. Exp. Psychol. Anim. Behav. Process. 13, 354–365.
Halberda, J., 2006. Is this a dax which I see before me? Use of the logical argument disjunctive syllogism supports word-learning in children and adults. Cogn. Psychol. 53, 310–344.
Izard, C.E., 1971. The Face of Emotion. Appleton-Century-Crofts, New York, NY.
Izard, C.E., 1994. Innate and universal facial expressions: evidence from developmental and cross-cultural research. Psychol. Bull. 115, 288–299. https://doi.org/10.1037/0033-2909.115.2.288.
Izard, C.E., Malatesta, C., 1987. Perspectives on emotional development I: differential emotions theory of early emotional development. In: Osofsky, J. (Ed.), Handbook of Infant Development, 2nd ed. Wiley, New York, NY, pp. 494–554.
Kolb, B., Wilson, B., Taylor, L., 1992. Developmental changes in the recognition and comprehension of facial expression: implications for frontal lobe function. Brain Cogn. 20, 74–84. https://doi.org/10.1016/0278-2626(92)90062-Q.
Kret, M.E., de Gelder, B., 2010. Social context influences recognition of bodily expressions. Exp. Brain Res. 203 (1), 169–180. https://doi.org/10.1007/s00221-010-2220-8.
Kret, M.E., Roelofs, K., Stekelenburg, J.J., de Gelder, B., 2013. Emotional signals from faces, bodies and scenes influence observers' face expressions, fixations and pupil-size. Front. Hum. Neurosci. 7, 810. https://doi.org/10.3389/fnhum.2013.00810.
Krumhuber, E.G., Kappas, A., Manstead, A.S.R., 2013. Effects of dynamic aspects of facial expressions: a review. Emot. Rev. 5 (1), 41–46. https://doi.org/10.1177/1754073912451349.
Laidlaw, K.E., Foulsham, T., Kuhn, G., Kingstone, A., 2011. Potential social interactions are important to social attention. Proc. Natl. Acad. Sci. 108 (14), 5548–5553.
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. J. Exp. Child Psychol. 114, 203–216. https://doi.org/10.1016/j.jecp.2012.06.003.
Nelson, N.L., Mondloch, C.J., 2017. Adults' and children's perception of facial expressions is influenced by body postures even for dynamic stimuli. Vis. cogn. 25, 563–574. https://doi.org/10.1080/13506285.2017.1301615.
Nelson, N.L., Mondloch, C.J., 2018. Children's visual attention to emotional expressions varies with stimulus movement. J. Exp. Child Psychol. 172, 13–24.
Nelson, N.L., Russell, J.A., 2011. Preschoolers' use of dynamic facial, bodily, and vocal cues to emotion. J. Exp. Child Psychol. 110, 52–61. https://doi.org/10.1016/j.jecp.2011.03.014.
Nelson, N.L., Russell, J.A., 2015. Children distinguish between positive pride and hubris. Dev. Psychol. 51 (11). https://doi.org/10.1037/dev0000048.
Nelson, N.L., Russell, J.A., 2016a. Building emotion categories: children use a process of elimination when they encounter novel expressions. J. Exp. Child Psychol. https://doi.org/10.1016/j.jecp.2016.02.012.
Nelson, N.L., Russell, J.A., 2016b. A facial expression of pax: assessing children's "recognition" of emotion from faces. J. Exp. Child Psychol. 141, 49–64. https://doi.org/10.1016/j.jecp.2015.07.016.
Perner, J., Leekam, S.R., Wimmer, H., 1987. Three-year-olds' difficulty with false belief. Br. J. Dev. Psychol. 5, 125–137. https://doi.org/10.1111/j.2044-835X.1987.tb01048.
Plebanek, D.J., Sloutsky, V.M., 2017. Costs of selective attention : when children notice what adults miss. Psychol. Sci. 28 (6), 723–732. https://doi.org/10.1177/0956797617711664.
Repacholi, B.M., Meltzoff, A.N., 2007. Emotional eavesdropping: infants selectively respond to indirect emotional signals. Child Dev. 78, 503–521. https://doi.org/10.1111/j.1467-8624.2007.01012.x.
Repacholi, B.M., Meltzoff, A.N., Hennings, T.M., Ruba, A.L., 2016. Transfer of social learning across contexts: exploring infants' attribution of trait-like emotions to adults. Infancy 21, 785–806. https://doi.org/10.1111/infa.12136.
Richoz, A.-R., Lao, J., Pascalis, O., Caldara, R., 2018. Tracking the recognition of static and dynamic facial expressions of emotion across the life span. J. Vis. 18, 1–27. https://doi.org/10.1167/18.9.5.
Silva, K.M., Timberlake, W., 1997. A behaviour systems view of conditioned states during long and short CS-US intervals. Learn. Motiv. 28, 465–490.
Silva, K.M., Timberlake, W., 2005. A behavior systems view of the organization of multiple responses during a partially reinforced or continuously reinforced interfood clock. Learn. Behav. 33, 99–110.
Sorce, J.F., Emde, R.N., Campos, J.J., Klinnert, M.D., 1985. Maternal emotional signaling: its effect on the visual cliff behavior of one-year-olds. Dev. Psychol. 21, 195–200. https://doi.org/10.1037/0012-1649.21.1.195.
Sroufe, L.A., 1997. Emotional Development: The Organization of Emotional Life in the Early Years. Cambridge University Press.
Stoesz, B.M., Jakobson, L.S., 2014. Developmental changes in attention to faces and bodies in static and dynamic scenes. Front. Psychol. 5 (193), 1–9. https://doi.org/10.3389/fpsyg.2014.00193.
Timberlake, W., 1983. Rats responses to a moving object related to food or water: a behavior systems analysis. Anim. Learn. Behav. 11, 309–320.
Timberlake, W., 1984. An ecological approach to learning. Learn. Motiv. 15, 321–333.
Timberlake, W., 2001. Integrating niche-related and general process approaches in the study of learning. Behav. Processes 54, 79–94.
Van Der Schalk, J., Hawk, S.T., Fischer, A.H., Doosje, B., 2011. Moving faces, looking places: validation of the Amsterdam Dynamic Facial Expression set (ADFES). Emotion 11, 907–920. https://doi.org/10.1037/a0023853.
Vicari, S., Reilly, J.S., Pasqualetti, P., Vizzotto, A., Caltagirone, C., 2000. Recognition of facial expressions of emotions in school-age children: the intersection of perceptual and semantic categories. Acta Paediatr. 89 (836), 845. https://doi.org/10.1111/j.1651-2227.2000.tb00392.x.
Vieillard, S., Guidetti, M., 2009. Children's perception and understanding of (dis)similarities among dynamic bodily/facial expressions of happiness, pleasure, anger, and irritation. J. Exp. Child Psychol. 102, 78–95. https://doi.org/10.1016/j.jecp.2008.04.005.
Walker-Andrews, A.S., 1997. Infants' perception of expressive behaviors: differentiation of multimodal information. Psychol. Bull. 121 (3), 437–456. https://doi.org/10.1037/0033-2909.121.3.437.
Walle, E.A., Reschke, P.J., Camras, L.A., Campos, J.J., 2017. Infant differential behavioral responding to discrete emotions. Emotion 17, 1078–1091. https://doi.org/10.1037/emo0000307.
Wellman, H.M., Liu, D., 2004. Scaling of theory-of-mind tasks. Child Dev. 75, 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x.
Wellman, H.M., Fang, F., Peterson, C.C., 2011. Sequential progressions in a Theory-of-Mind scale: longitudinal perspectives. Child Dev. 82, 780–792. https://doi.org/10.1111/j.1467-8624.
Widen, S.C., 2013. Children's interpretation of facial expressions: the long path from valence-based to specific discrete categories. Emot. Rev. 5, 72–77. https://doi.org/10.1177/1754073912451492.
Widen, S.C., Russell, J.A., 2003. A closer look at preschoolers' freely produced labels for facial expressions. Dev. Psychol. 39, 114–128. https://doi.org/10.1037/0012-1649.39.1.114.
Widen, S.C., Russell, J.A., 2008. Young children's understanding of other's emotions. In: Lewis, M., Haviland-Jones, J.M., Barrett, L.F. (Eds.), Handbook of Emotions, 3rd ed. Guilford Press, New York, NY, pp. 348–363.
Widen, S.C., Pochedly, J.T., Russell, J.A., 2015. The development of emotion concepts: a story superiority effect in older children and adolescents. J. Exp. Child Psychol. 131, 186–192. https://doi.org/10.1016/j.jecp.2014.10.009.
Xiao, N.G., Quinn, P.C., Liu, S., Ge, L., Pascalis, O., Lee, K., 2015. Eye tracking reveals a crucial role for facial motion in recognition of faces by infants. Dev. Psychol. 51, 744–757. https://doi.org/10.1037/dev0000019.
Xiao, N.G., Quinn, P.C., Ge, L., Lee, K., 2017. Facial movements facilitate part-based, not holistic, processing in children, adolescents, and adults. Dev. Psychol. 53, 1765–1776. https://doi.org/10.1037/dev0000360.
Yu, C., Smith, L.B., 2013. Joint attention without gaze following: human infants and their parents coordinate visual attention to objects through eye-hand coordination. PLoS One 8 (11) e79659.