Article:

Geangu, Elena orcid.org/0000-0002-0398-8398 and Vuong, Quoc (2020) Look up to the body:an eye-tracking investigation of 7-months-old infants' visual exploration of emotion body expressions. Infant Behavior and Development. 101473. ISSN 0163-6383

https://doi.org/10.1016/j.infbeh.2020.101473

Look up to the body. An eye-tracking investigation of 7-months-old infants' visual exploration of emotional body expressions.

*+Elena Geangu, University of York, UK
+Quoc C. Vuong, Newcastle University, UK

***Corresponding author**
Elena Geangu, PhD
Psychology Department
University of York
Heslington, York, YO10 5DD, UK
Email: elena.geangu@york.ac.uk

+ The authors contributed equally to this research output.

Abstract

The human body is an important source of information to infer a person's emotional state. Research with adult observers indicate that the posture of the torso, arms and hands provide important perceptual cues for recognising anger, fear and happy expressions. Much less is known about whether infants process body regions differently for different body expressions. To address this issue, we used eye tracking to investigate whether infants' visual exploration patterns differed when viewing body expressions. Forty-eight 7-months-old infants were randomly presented with static images of adult female bodies expressing anger, fear and happiness, as well as an emotionally-neutral posture. Facial cues to emotional state were removed by masking the faces. We measured the proportion of looking time, proportion and number of fixations, and duration of fixations on the head, upper body and lower body regions for the different expressions. We showed that infants explored the upper body more than the lower body. Importantly, infants at this age fixated differently on different body regions depending on the expression of the body posture. In particular, infants spent a larger proportion of their looking times and had longer fixation durations on the upper body for fear relative to the other expressions. These results extend and replicate the information about infant processing of emotional expressions displayed by human bodies, and they support the hypothesis that infants' visual exploration of human bodies is driven by the upper body.

Keywords: infancy, eye-tracking, body, emotion expressions

1. Introduction

Humans convey perceptual cues about their emotional states through a variety of sources. These sources include facial (Keltner, Tracy, Sauter, Cordaro, McNeil, 2016), bodily (de Gelder, 2006; de

Gelder 2009; Tamietto and de Gelder, 2006) and vocal expressions (Belin et al., 2012), as well as changes in body odour and temperature (Rosen et al., 2015; de Groot and Smeets, 2017; Salazar-López et al., 2015; Robinson et al., 2012). Observers rely on the perceptual cues from these sources for inferring people's emotional state which, in turn, is important for inferring how they are likely to act or appraise the environment (e.g., Walle and Campos, 2012; Walle, Dahl, & Campos, 2012; Shuman, Clark-Polner, Meuleman, Sander, & Scherer, 2017). In many instances, the cues from different sources work in concert to convey a person's emotions (Heck, Chroust, White, Jubran, & Bhatt, 2018; Stienen, Tanaka, de Gelder, 2011; Aviezer, Hassin, & Bentin, 2012; Aviezer, Trope, Todorov, 2012; Yeh, Geangu, Reid, 2016; de Gelder and Vroomen, 2000; Van den Stock et al., 2007; Campanella & Belin, 2007; Robbins & Coltheart, 2015). That said, perceptual cues from some sources are better than others for inferring emotions more accurately under different circumstances. For example, it may be easier to infer emotional states from body cues when facial cues are less visible, as when a person is at a distance.

Despite the relevance of the various sources for the non-verbal expression of emotions demonstrated in research with adults (e.g., Aviezer et al., 2012; Yeh et al., 2016; Jessen & Kotz, 2011), the majority of infancy research to date focus on understanding how infants develop the ability to process emotional facial expressions (Bayet & Nelson, 2019; Geangu et al., 2016a). Researchers have only recently begun to investigate which perceptual cues infants extract from the body during the critical first two years of life to infer emotions (Heck, Chroust, White, Jubran, Bhatt, 2018; Missana and Grossmann, 2014; Missana, Rajhans, Atkinson, & Grossmann, 2014; Missana, Atkinson, & Grossmann, 2014; Rajhans, Jessen, Missana, & Grossmann, 2016; Zieber, Kangas, Hock, & Bhatt, 2014). To address this issue, in the present study we used eye tracking to compare and contrast 7-month-old infants' visual exploration patterns when viewing static adult body postures expressing anger, fear, and happiness.

The body expressions of different emotions can be characterized in terms of signature combinations of postures, gestures, and muscle movements that help their recognition (Atkinson, Dittrich, Gemmel, & Young, 2004; Atkinson, Tunstall, & Dittrich, 2007; Atkinson, 2013; Dael, Mortillaro, & Scherer, 2012a; Dael, Mortillaro, & Scherer, 2012b; Glowinski, Camurri, Volpe, Dael, & Scherer, 2008; Glowinski, Dael, Camurri, Volpe, Mortillaro, & Scherer, 2011; Mancini, Varni, Glowinski, & Volpe, 2012). For example, when expressing happiness, adults adopt an upright posture with raised arms. By comparison when angry, adults tend to lean forward, and shake their fists or point. Fear is usually expressed by leaning backwards and raising the arms in front of the body. The most subtle of all body expressions seems to be body postures expressing sadness. When sad, adults tend to adopt a dropped position of the head, with the hands or arms brought close to the body. For static body postures, these characteristic signatures of emotional expressions tend to be mostly located in the upper part of the body (i.e., the torso, arms and hands). Adults rely greatly on these characteristic signatures for recognizing the emotions expressed by bodies. Disrupting these signatures can impair emotion recognition. For example, Ross and Flack (2019) showed that removing the hands impaired the recognition of fear and anger body expressions. Eye tracking studies with adults further show that much more time is spent looking at the upper body compared to the legs (Pollux, Craddock, & Guo, 2019; Poyo Solanas et al., 2020). Moreover, the pattern of visual fixations to the upper body shows subtle variations between expressions. For example, in contrast to sadness and anger, adults allocated more fixations to the arms of fearful and happy bodies. In particular, the contracted position of the arms has been associated with accurate recognition of fear and the activity of brain areas typically associated with rapid detection and response to fear (Poyo Solanas et al., 2020).

From the first days of life, infants frequently have people in their view, a pattern that remains constant throughout the first year of life (Jayaraman, Fausey & Smith, 2017). The persistent presence of people in the visual field provides infants with many opportunities to extract relevant perceptual cues about people's emotions and actions. The extent to which different parts of the human body are present in infants' view appears to change during infancy. While faces are more prevalent during the first four months after birth, other parts of people's bodies appear more frequently in infants' view afterwards (Jayaraman et al., 2017). For example, Fausey et al. (2016) showed an increase of the proportion of hands in the infant's field of view (as indicated by head-mounted camera recordings) with a corresponding decrease in the proportion of faces. These trends are observed across the first 2 years of life with a larger proportion of hands emerging between 6 and 9-months old. These changes increase infants' opportunities to extract characteristic signature postures associated with different emotional expressions (e.g., the extended arms for happiness or the contracted arms around the upper body for fear).

It remains yet unexplored whether infants use perceptual cues from different body regions and parts for discriminating different emotional states expressed by static body postures. Outside the emotional domain, there is evidence that infants are able to extract information from specific body parts in order to accomplish different social tasks (Smith, Jayaraman, Clerkin, & Yu, 2018; Geangu, Senna, Croci, & Turati, 2015). For example, from as early as 6-months, infants fixate on the hands of people who reach and grasp objects, and look less at other body parts that are in view (Geangu et al., 2015; Kochukhova & Gredebäck, 2010; Falck-Ytter, Gredebäck, & von Hofsten, 2006). Extracting visual information from different body parts (e.g., face, legs) was also shown to be related to the development of self-representations and to enable self-recognition in infants (Filippetti & Tsakiris, 2018; Nielsen, Suddendorf, & Slaughter, 2006; Geangu, 2008; Zmyj et al., 2011; Suddendorf & Butler, 2013). Seven- to 10-months-old infants use information about the visual appearance of their own legs in order to detect the visual-proprioceptive contingencies, which further allows them to differentiate between their own and another infant's legs movements (Zmyj et al., 2011). White et al (2018) showed that 3.5 and 6.5-months old were more likely to fixate different body regions for sex discrimination.

For emotion processing, from 5- to 8-months-old infants are able to discriminate body expressions of happiness from those of anger (Heck et al., 2018) and fear (Missana et al., 2015; Missana et al., 2014). Furthermore, they are also able to relate the emotional expressions displayed by bodies with that extracted from faces (Hock, Oberst, Jubran, White, & Bhatt, 2017) and voices (Rajhans et al., 2016; Heck et al., 2018; Zieber et al., 2014). Evidence from studies in which infants were presented with body-face stimuli that were emotionally congruent (e.g., happy body and face) and incongruent (e.g., happy body and sad face) suggest that infants within this age range may already include body postures in discrete emotion categories. Hock and colleagues (2017) showed that 6.5-months-old infants detected mismatches between face and body expressions not only when positive (e.g., happy) and negative (e.g., anger) emotion pairings are presented, but also when pairings of two negative emotions (e.g., sad and anger) are displayed to the participants.

The analysis of infants' electrical cortical responses to body images indicate the involvement of several processes in computing emotionally-relevant information (Missana & Grossmann, 2015; Missana et al., 2015; Missana, Rajhans, Atkinson, & Grossmann, 2014). In the study of Missana and colleagues (2014), differentiations between fear and happy body expressions were recorded as early as 290 ms after stimulus onset in 8-months-old infants. The N290 is the infant correspondent of the adult N170 (Gillmeister, Stets, Grigorova, & Rigato, 2019), a negative deflection in the event-related potentials (ERP) which has been previously related to processing information about face and body structure, as well as the activation of the perceptual representations of faces and bodies (Bernard, Content, Deltenre, & Colin, 2017; Bauser & Suchan, 2018; Eimer, 2000; Rossion & Jacques, 2008; Stekelenburg & de Gelder, 2004; Zion-Golumbic & Bentin, 2006). The increased amplitude of the N290 to fear compared to happy bodies in 8-months-old infants suggests that structural information relevant to expressing emotions may be extracted relatively fast and integrated in the perceptual representations of bodies (Missana & Grossmann, 2015; Missana, Atkinson, & Grossmann, 2015; Missana, Rajhans, Atkinson, & Grossmann, 2014). Later ERP components, such as the central negativity (Nc), suggest that 8-months-old infants also allocate more attentional resources to processing fearful bodies compared to the happy ones (Rajhans, Missana, Krol, & Grossmann, 2015; Krol, Rajhans, Missana, & Grossmann, 2015; Missana et al., 2014). The increased attention to bodies expressing fear is similar to the increased attention to negative facial and vocal expressions, and it may also be related to the higher attention orientation and allocation to threat-related information (Crespo-Llado, Vanderwert, Roberti, & Geangu, 2018; Geangu et al., 2016; Hoehl, 2014; Kaiser, Crespo-Llado, Turati, & Geangu, 2018; Leppänen & Nelson, 2009).

Thus by 7-months old, infants are able to preferentially fixate on body regions and parts which are the most relevant for different social tasks (Geangu et al., 2015; Kochukhova & Gredebäck, 2010;

Falck-Ytter, Gredebäck, & von Hofsten, 2006), and they can discriminate between emotional body postures at the behavioural and neural level (Missana & Grossmann, 2015; Missana, Atkinson, & Grossmann, 2015; Missana, Rajhans, Atkinson, & Grossmann, 2014; Rajhans, Missana, Krol, & Grossmann, 2015; Krol, Rajhans, Missana, & Grossmann, 2015; Missana et al., 2014; Heck et al., 2018; Hock et al., 2017; Zieber et al., 2014). What remains unknown is whether they differentially use perceptual cues from different body regions and parts, particularly the upper torso (including the arms and hands), as indicated by recent adult studies (Pollux et al., 2019; Poyo Solanas et al., 2020), for processing emotional body expressions.

The aim of the current study was to determine whether 7-months-old infants' visual exploration pattern of body expressions differed for angry, fear, happy and emotionally neutral static postures using eye tracking. Given the behavioural and neural evidence from the infant and adult studies reviewed above, we hypothesize that by 7-months-old infants are more likely to process visual information in the upper body to discriminate emotions expressed by body postures, particularly for fearful (or other negative) expressions that tend to be more attention grabbing. This hypothesis leads to two complementary predictions about infants' visual exploration patterns when presented with static body expressions. First, we predict that infants will be more likely to fixate (i.e., higher proportion of fixations) and dwell (i.e., longer proportion of looking time or fixation duration) on the upper body because the torso, arms and hands contain perceptual cues that discriminate between different body expressions, particularly fear and anger expressions (Atkinson, et al., 2004; Atkinson, et al., 2007; Atkinson, 2013; Dael et al., 2012a; Dael et al., 2012b; Glowinski et al., 2008; Glowinski et al., 2011; Mancini et al., 2012; Pollux et al., 2019; Ross & Flack, 2019). Second, we predict that infants' visual exploration patterns of the upper body region will show differences across emotion expressions. More specifically, given the previously documented attentional bias for fear body expressions (e.g., Missana et al., 2014) and that adults preferentially fixate the upper part of fearful bodies for accurate recognition (Pollux et al., 2019; Poyo Solanas et al., 2020), we predict that infants will be more likely to fixate and dwell upper body parts expressed during fear than other emotions.

2. Materials and Methods

2.1. Participants. Sixty 7-months-old infants (30 females; M = 227 days, SD = 8 days) were tested. They were recruited from urban and semi-urban areas in the North of England. The study was approved by the University's Ethics Committee. Participants' parents gave written informed consent before testing. The experiment was conducted according to the Declaration of Helsinki.

We used three criteria to include infants' eye-movement data in the final analyses. For each infant, (1) we first excluded trials if the infant looked at the body stimulus for a given trial less than 30% of the total body-stimulus duration; the remaining trials are referred to as valid trials. Following trial exclusion, we included infants' data for analyses (2) if they had 2 or more (out of 5) valid trials for each expression tested, and (3) if they had 10 or more (out of 20) valid trials in total. Forty-eight infants were included in the final analyses based on these criteria.

2.2. Stimuli. We selected female body stimuli from the Bodily Expressive Action Stimulus Test (BEAST, de Gelder et al., 2011). These included five body postures expressing anger, fear and happiness, as well as an emotionally neutral posture (20 stimuli in total). The inner face (eyes, nose, and mouth) of all the stimuli were uniformly filled with grey. In addition, for all of the five fear body postures, the hands and/or arms occluded to different extent the face and/or head. The body stimuli were presented as greyscale images. The bodies were roughly centred on a uniform grey background (grey level = 128), with the estimated location of the belly button at the centre of the screen (see Figure 1).

2.3. Apparatus. During the experiment, infants sat on their parent's lap in a cubicle which had thick black curtains to either side to minimise environmental distractions. There were approximately 60-70 cm between the infant head and the monitor screen (1920 x 1080 pixel resolution). Participants' eye movements were tracked binocularly by a Tobii X120 eyetracker (Tobii Technology, Sweden), sampling at 120 Hz with a spatial resolution of .5° and a drift of < .3 degrees. An infant-specific 5-points calibration procedure was used. All infants were tested with normal lighting conditions, with the exception of one infant that was tested with the lights turned off.

2.4. Design and procedure. Infants were presented with all 20 body stimuli on separate trials. The stimuli were presented in four randomised orders, and infants were randomly assigned to one of these orders. Figure 1 illustrates the sequence of events across trials. At the beginning of each trial, an attention grabber could be used to attract infants' attention if they were not looking at the screen. This attention grabber consisted of a colourful moving shape presented at the centre of a grey screen, along with a pleasant sound. Once the infant fixated on the attention grabber, the experimenter pressed the space bar to present a black cross at the centre of a grey screen. The experimenter pressed the space bar again when the infant fixated on the cross. This procedure ensured that infants began the body exploration from the same location. Following this, a body stimulus was presented for 4 s. The body's belly button was aligned with the fixation cross. The body stimulus was then replaced by a fixation cross (or an attention grabber if infants were not looking at the screen). The next trial began when the infant fixated on the cross; thus the inter-trial duration varied within and between infants.

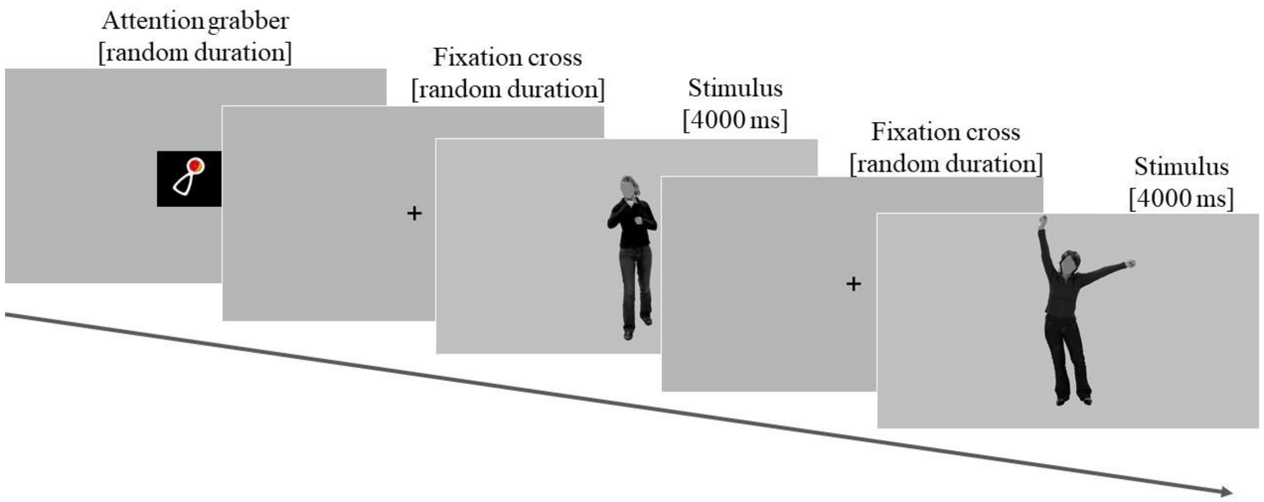


Figure 1. The sequence of events during the experiment. An attention grabber could be used to attract infants' attention to the screen. Once infants fixated on a fixation cross, the experimenter pressed the space bar to present the next stimulus. If infants looked away from the screen, the experimenter presented the attention grabber again.

2.5. Defining body areas of interest. We analysed eye-movement data with respect to the entire body and different areas of interests (AOIs; see White et al., 2018; Pollux et al., 2019): the head, upper body (torso, arms and hands) and lower body (hips, legs and feet). Figure 2 illustrates the three AOIs for the anger, fear, happy and neutral body expressions from the same female actor. We used the estimated location of each body's belly button to demarcate the upper and lower body portions. In cases when the arms/hands occluded the head, they were included as part of the upper AOI (see, for example, the fear posture in Figure 2).

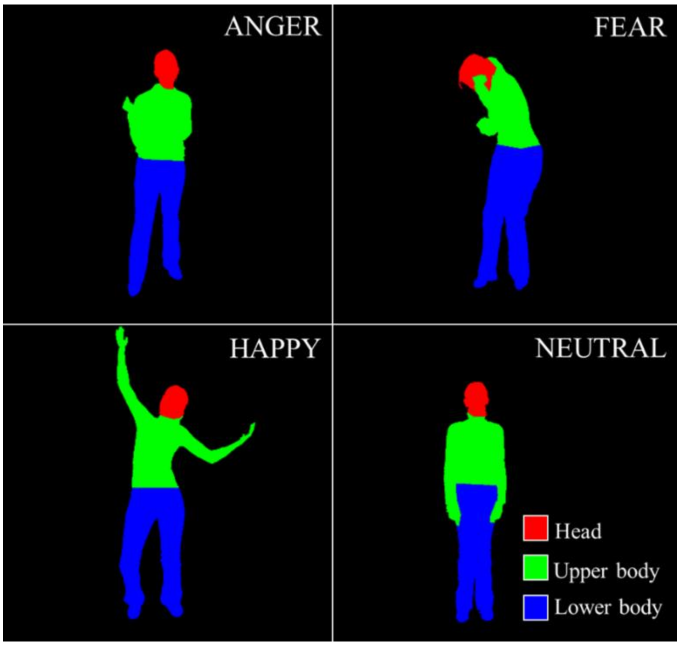


Figure 2. The three areas of interest (AOIs) used for the analysis of the eye-movement data. The AOIs are shown for expression from the same female actor. The approximate location of the belly button was used to demarcate the upper and lower body AOI. The arms and hands were assigned to the upper body even though it occluded the head for fear expressions.

<u>2.6. Comparison of low-level visual features as a function of AOI and expression.</u> There is the possibility that the stimuli could differ in terms of low-level visual features that do not directly contribute to the features that are diagnostic for emotion expressions. In this section we describe the approach we adopted to characterize the stimuli in terms of some of their low-level visual features, as well as the analyses performed to assess their possible contribution to infants' performance. For the whole body and each AOI, we computed several measures to characterise any potential low-level differences between body expressions in terms of surface area (size), luminance and visual saliency. Table 1 presents the percentage of the screen area and corresponding number of pixels occupied by

the body stimulus averaged across the five images for each expression. Note that the entire body stimulus occupied only ~6% of the screen on average. Table 2 presents the mean luminance averaged across the five image stimuli for each expression. Lastly, we used Harel et al.'s (2006) MATLAB implementation with default settings to compute visual saliency maps for each image (http://www.klab.caltech.edu/~harel/share/gbvs.php). The visual saliency at each image pixel was based on luminous intensity (at 8 different spatial frequency scales) and orientation (at 4 angles), normalized to a value between 0 and 1. The maps were then averaged across pixels as a function of AOI and expression. This measure of visual saliency better reflects strictly low-level visual features that define perceptually conspicuous regions in the image that can attract eye fixations and attentional processes, compared to measures such as size and mean luminance (Itti et al., 1998; Itti & Koch, 2001; Veale, Hafed, & Yoshida, 2017). Figure 3A illustrates the visual saliency map for example body images. Figure 3B shows the visual saliency of the different AOIs averaged across the five body images for each expression. Because eye fixations can be driven by the relative difference in visual saliency between AOIs, Figure 3C shows the visual saliency ratio between the upper and lower AOI (i.e., mean upper saliency/mean lower saliency), which are the two largest body regions.

We chose this approach, rather than manipulating the stimulus directly (e.g., stimulus inversion), for three main reasons. First, it allowed us to formulate testable predictions about which AOIs infants are likely to fixate more if their visual scanning were driven by low-level visual features without potentially affecting their perception of bodies that may occur if the stimuli are physically manipulated (e.g., Kibbe, Kàldy, and Blaser, 2017). Second, it provided information about the possible role of low-level visual features for the perceptual task that is more similar to what infants are familiar with and experience in everyday life, i.e., upright human bodies (e.g., Sugden & Moulson, 2017). Third, it reduces the number of conditions that is tested within-subject, which increases power and is less taxing for infants' limited attention span.

Table 1. The mean percentage of total screen area and total number of pixels occupied by the body stimulus as a function of AOI and expression. The number in parenthesis represents the standard deviation.

	Mean % Total Screen Area (SD)				**Pixels (SD)**			
	Body	**Head**	**Upper**	**Lower**	**Body**	**Head**	**Upper**	**Lower**
Anger	5.11	0.43	2.04	2.64	105902	8972	42258	54673
	(0.13)	(0.05)	(0.11)	(0.20)	(2717)	(955)	(2379)	(4097)
Fear	5.53	0.32	1.88	3.33	114741	6656	39020	69061
	(0.47)	(0.11)	(0.29)	(0.25)	(9776)	(2280)	(6015)	(5241)
Happy	6.24	0.40	2.73	3.11	129348	8390	56551	64406
	(0.19)	(0.08)	(0.07)	(0.19)	(3897)	(1566)	(1496)	(3876)
Neutral	5.66	0.41	2.64	2.61	117408	8545	54801	54060
	(0.45)	(0.10)	(0.22)	(0.28)	(9423)	(2075)	(4570)	(5892)
mean	5.64	0.39	2.32	2.92	116850	8141	48157	60550
	(0.52)	(0.09)	(0.42)	(0.38)	(10838)	(1879)	(8658)	(7933)

Table 2. The mean luminance (0 = black; 128 = mid-grey; 255 = white) of the body stimulus as a function of AOI and expression. The number in parenthesis represents the standard deviation. Note that we excluded the background in computing the mean luminance.

	Body	**Head**	**Upper**	**Lower**
Anger	56.3	115.7	50.2	52.1
	(16.7)	(18.3)	(27.0)	(17.0)
Fear	50.3	86.9	49.1	48.6
	(14.8)	(41.9)	(21.9)	(18.0)

Happy	45.3	114.5	39.7	41.8
	(6.2)	(18.2)	(12.8)	(9.7)
Neutral	72.6	135.7	75.3	60.0
	(21.0)	(26.8)	(20.8)	(23.9)
mean	56.1	113.2	53.6	50.6
	(17.8)	(31.3)	(23.7)	(17.7)

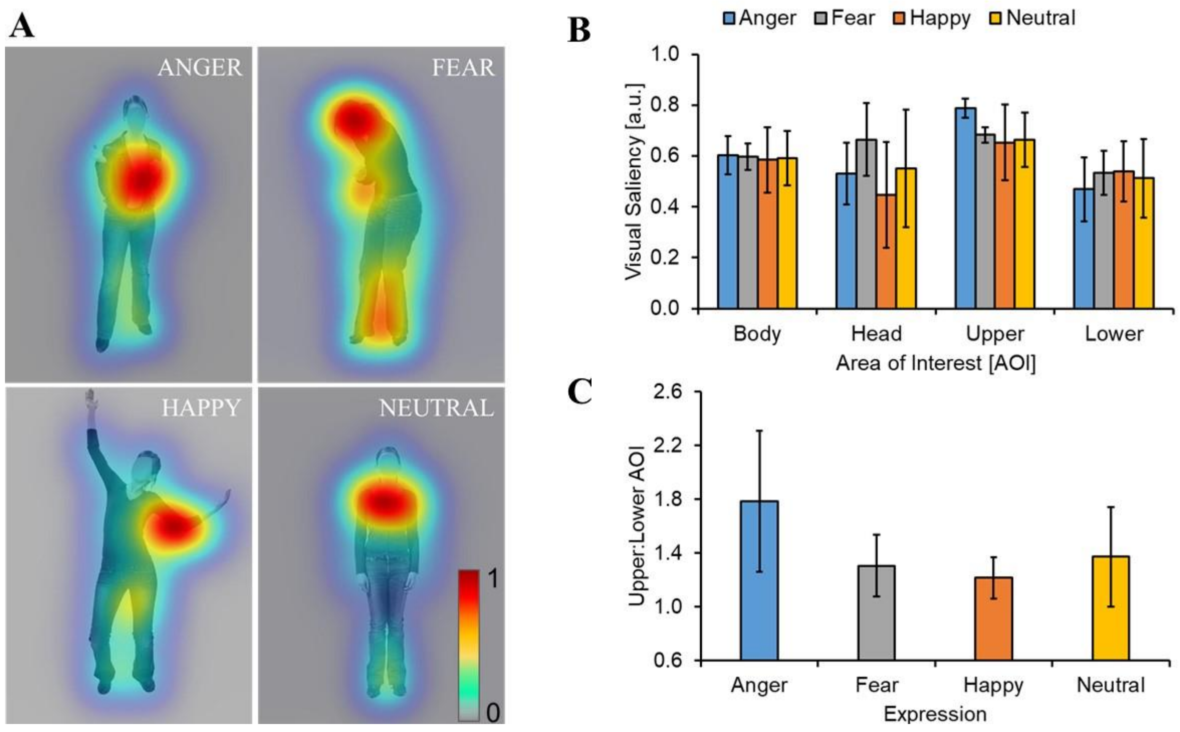


Figure 3. Visual saliency of body postures for the different expressions. A) The visual saliency map superimposed on the corresponding example body image for anger, fear, happy and neutral emotional

expressions based on Itti et al. (1998). The visual saliency at each image pixel is in arbitrary units (a.u.) normalized between 0 and 1. The images match those in Figures 2 and 4 (see Supplementary Figures 5 – 8 for the visual saliency map of all 20 body images). B) The mean visual saliency as a function of AOI and expression. C) The visual saliency ratio between the upper and lower AOI (i.e., mean upper saliency/mean lower saliency). Error bars reflect the standard deviation. Values greater than 1 reflect higher visual saliency (on average) in the upper AOI relative to the lower AOI.

2.7. Assigning fixations to AOIs. Tobii Studio (version 3.3.1; Tobii Technology, Sweden) was first used to extract fixations from each infant's eye-movement data. For the fixation filter, we enabled interpolation with a maximum temporal gap length of 75 ms; used a velocity threshold of 30°/s; and merged temporally adjacent fixations if they were within 75 ms of each or the spatial angle between fixations were within 0.5°. Valid fixations were defined as fixations that fell within the screen (i.e., irrespective of whether the fixation fell within the body stimulus or not) during the 4-s stimulus presentation on each trial. We excluded the first valid fixation on each trial, given that the onset of the body stimulus was initiated when infants fixated the central cross (Geangu et al., 2016a).

We used the following steps to determine whether a valid fixation fell within the body and within a specific AOI. First for each fixation, a circle with a radius of 10 pixels centred on the fixation's x- and y-coordinate was defined. Second, we assigned the fixation based on the intersection between the circle and body or AOI area. For the body, if 10 or more pixels of the fixation circle intersected with pixels from the body stimulus, then the fixation was considered to be on the body. For each AOI, if 5 or more pixels of the fixation circle intersected with pixels of that AOI, then that fixation was assigned to that AOI. A smaller threshold was used given the AOIs smaller area relative to the bodies' area. If a fixation circle intersected with more than one AOI, the fixation was assigned to the AOI which had the larger pixel overlap. In the case of a tie in overlap, the fixation was randomly assigned to one of the AOI. These conditions ensured that fixations were assigned to only one AOI.

2.8. Statistical data analyses. To compare infants' pattern of looking behaviour to different body regions as a function of body expressions, we computed three complementary fixation parameters on each trial for each infant. First, we calculated the *proportion fixations* by counting the total number of valid fixations falling within the whole body or AOI on a given trial, and then dividing it by the total number of valid fixations on that trial. This parameter measures how likely infants are to fixate on the whole body or regions of the body (i.e., AOIs) and how much they are scanning the visual stimulus (e.g., Gredeback et al., 2012; Helo et al., 2016). Second, the *proportion looking time* was computed by summing the fixation duration of all valid fixations falling within the whole body or AOI on a given trial and dividing by 4-s (i.e., the duration of the body stimulus). This parameter measures the relative amount of time infants fixate on the whole body or regions of the body and is regarded as index of perceptual preference and attention allocation (Bronson, 1991; Colombo & Mitchell, 2009). Finally, the *fixation duration* was computed by averaging the fixation duration of all valid fixations falling within the whole body or AOI on a given trial. For this parameter, we eliminated outliers by removing durations that were less than 80 ms (2.0% of the data) or greater than 3.5SD above the grand mean (averaged across all participants and fixations; 1.5% of the data). The *fixation duration* measures how much time it takes observers to process visual information at the point of fixation, and has been associated with speed of processing and cognitive effort (e.g., Helo et al., 2016; Mayer et al., 2015; Wass & Smith, 2014).

There were no fixations in some of the expressions for each AOI. Thus, there were missing *fixation-duration* data for some infants as zero is not a meaningful *fixation duration*. Across all infants, there were 192 cells (4 expressions x 48 infants). There was a small percentage of missing values for the whole body (1.04%, 2 cells) and upper AOI (1.56%; 3 cells). Given the small percentage, we replaced missing values with the infant's mean *fixation duration* averaged across the remaining expressions. By comparison, there was a large percentage of missing values for the head (25.52%, 49 cells) and the lower AOI (27.60%, 53 cells). Given the large percentage of missing values, *fixation duration* was not analysed for these AOIs.

For each infant, AOI (including the whole body) and expression, we averaged the three parameters across valid trials (as defined above). The trial-averaged mean fixation parameters were submitted to different repeated measures analyses of variance (ANOVAs; see below for details). We conducted post-hoc simple ANOVAs or paired *t*-tests to further interrogate significant main effects and interactions in the omnibus ANOVAs. We report uncorrected *p*-values for pairwise tests. For all analyses, an $\alpha = .05$ was adopted as the level of statistical significance ($\alpha = .008$ for Bonferroni-corrected significance level for 6 pairwise tests).

3. Results

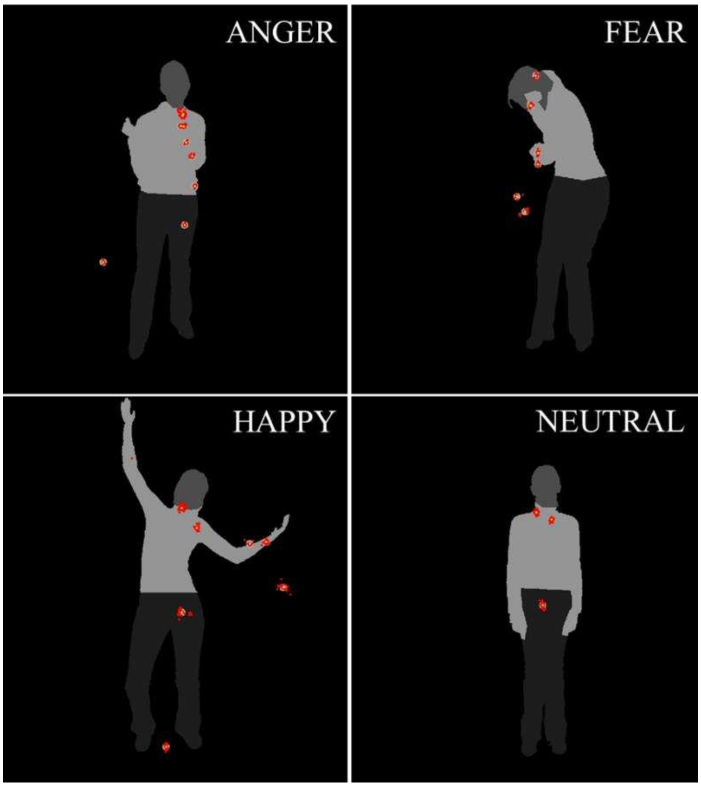


Figure 4. Examples of visual exploration patterns for anger, fear, happy and neutral body images. For each expression, the fixations from an individual infant and a single trial are superimposed onto the body image. The red points represent gaze points and the yellow point represents the fixation point (averaged across all gaze points). The head, upper body and lower body AOI are presented as different grey levels so that the gaze and fixation points are visible.

3.1. *Looking behaviour for the whole body for different body expressions*

Figure 4 shows individual infants' visual exploration pattern for an image from body expression (see Supplementary Figures 1-4 for all infants' fixations on all 20 body images). Figure 5 shows the mean of the three fixation parameters averaged across infants as a function of AOI (including the whole body) and expression. We first investigated how infants' pattern of looking behaviour on the whole body differed for the four expressions. The three fixation parameters were submitted to a repeated-measure ANOVA with expression as a within-subjects factor. For the whole body, there was only a main effect of expression for *proportion fixations*, $F(3,141) = 8.697$, $p < .001$, partial-eta = .156. Post-hoc *t*-tests showed that the *proportion fixations* was significantly lower for the happy expression than the anger, $t(47) = -5.261$, $p < .001$; fear, $t(47) = -2.912$, $p = .005$; and neutral expressions, $t(47) = -3.230$, $p = .002$.

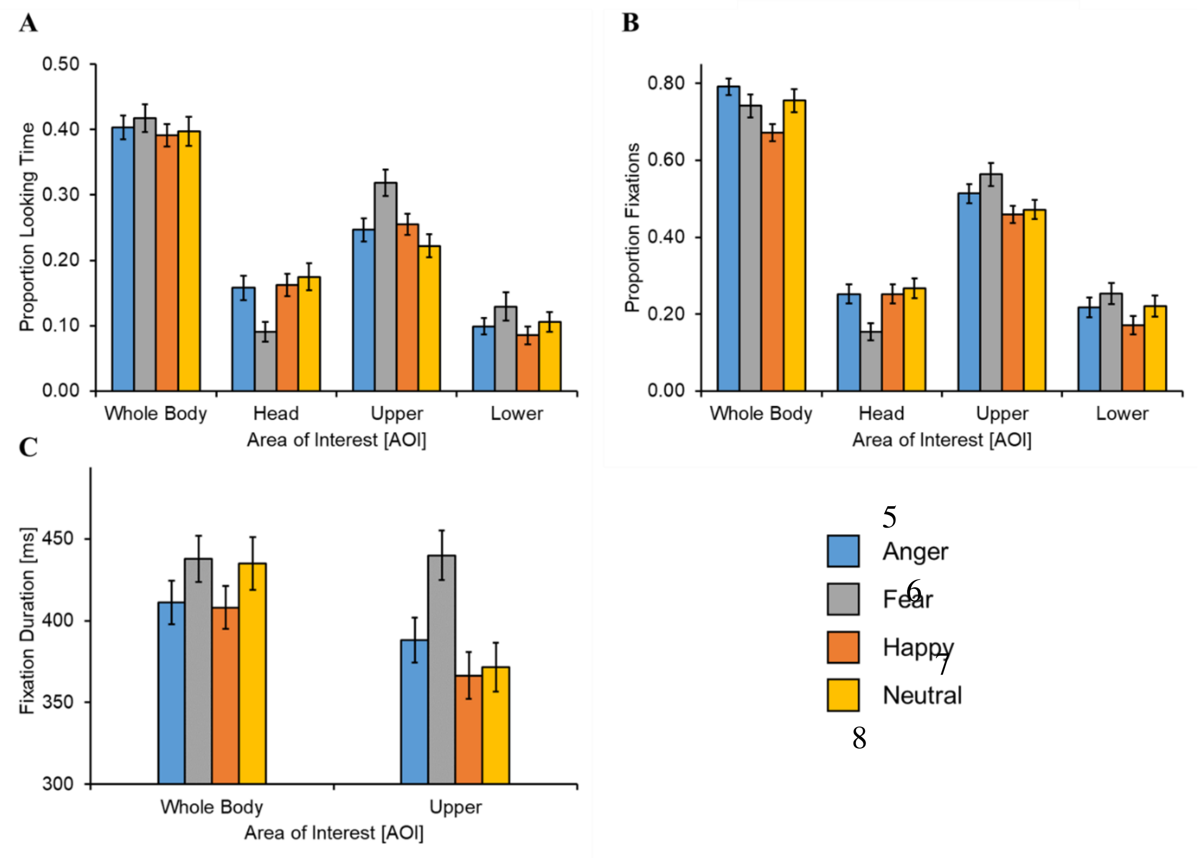


Figure 5. Mean and standard error of the mean (SE) of the three fixation parameters averaged across infants as a function of AOI (including the whole body) and expression. A) The proportion looking time was computed by summing the fixation duration of all valid fixations falling within the whole body or AOI on a given trial and dividing by 4-s (i.e. the duration of the body stimulus). B) The proportion fixations was computed by dividing the number of fixations by the total number of valid fixations on a specific trial. C) The fixation duration was computed by averaging the fixation duration of all valid fixations falling within the whole body or AOI on a given trial.

3.2. *Looking behaviour for the head AOI for different body expressions*

The *proportion looking time* and *proportion fixations* were submitted to a repeated-measures ANOVA with expression as a within-subjects factor. There was a main effect of expression for *proportion looking time*, $F(3,141) = 7.776$, $p < .001$, partial-eta = .142. Post-hoc *t*-tests showed that infants manifested a smaller *proportion looking time* on the head for the fear expression compared to the anger, $t(47) = -3.237$, $p = .002$; happy, $t(47) = -3.575$, $p = .001$; and neutral expressions, $t(47) = -4.228$, $p < .001$. For *proportion fixations*, there was also a main effect of expression, $F(3,141) = 8.888$, $p < .001$, partial-eta = .159. Post-hoc *t*-tests showed that infants manifested a smaller *proportion fixations* for the fear expression compared to the anger, $t(47) = -3.438$, $p = .001$; happy, $t(47) = -3.845$, $p < .001$; and neutral expressions, $t(47) = -4.383$, $p < .001$.

3.3. *Looking behaviour for the upper and lower AOIs for different body expressions*

The upper and lower AOIs had approximately the same size (see Table 1). Therefore, the fixation parameters were submitted to a 2 AOI (upper, lower) x 4 expression (anger, fear, happy, neutral) repeated-measures ANOVA, with both AOI and expression as within-subjects factors.

<u>Proportion looking time.</u> Figure 4A illustrates the mean *proportion looking time* as a function of AOI and expression. There was a main effect of both AOI, $F(1,47) = 81.075$, $p < .001$, partial-eta = .633 (Upper: $M = .261$, $SE = .013$; Lower: $M = .105$, $SE = .011$); and expression, $F(3,141) = 6.981$, $p < .001$, partial-eta = .129 (Anger: $M = .173$, $SE = .011$; Fear: $M = .224$, $SE = .015$; Happy: $M = .170$, $SE = .010$; and Neutral: $M = .164$, $SE = .013$). These main effects were qualified by a significant interaction between the two factors, $F(3,141) = 2.854$, $p = .039$, partial-eta = .057. Simple ANOVAs for each AOI showed that there was a main effect of expression for the upper AOI, $F(3,141) = 7.977$, $p < .001$, partial-eta = .145; but not for the lower AOI, $F(3,141) = 1.865$, $p = .138$, partial-eta = .038. Infants showed a higher *proportion of looking time* for the upper AOI of the fear expression compared to the upper AOI of the anger, $t(47) = 3.230$, $p = .002$; happy, $t(47) = 2.770$, $p = .008$; and neutral expressions, $t(47) = 4.397$, $p < .001$.

Proportion fixations. Figure 4B illustrates the mean *proportion fixations* as a function of AOI and expression. There was a main effect of AOI, $F(1,47) = 109.511$, $p < .001$, partial-eta = .700, with a higher proportion of fixations to the upper AOI ($M = .502$, $SE = .019$) compared to the lower AOI ($M = .216$, $SE = .018$). There was also a main effect of expression, $F(3,141) = 6.885$, $p < .001$, partial-eta = .128 (Anger: $M = .366$, $SE = .016$; Fear: $M = .409$, $SE = .023$; Happy: $M = .315$, $SE = .013$; and Neutral: $M = .346$, $SE = .019$). Across both upper and lower body regions, post-hoc *t*-tests revealed that the *proportion fixations* was lower for the happy expression compared to the anger, $t(47) = -3.144$, $p = .003$; and fear expressions, $t(47) = -4.288$, $p < .011$. The *proportion fixations* on fear expression was marginally more than on neutral expression, $t(47) = 2.666$, $p = .01$. There was no interaction between the two factors.

Fixation duration. Figure 4C illustrates the mean *fixation duration* for the whole body and upper body AOI as a function of expression. There were too many missing data values for the head and lower-body AOIs. Given the large percentage of missing data for the lower AOI, we submitted the *fixation-duration* data to a repeated-measures ANOVA with expression as a within-subjects factor only for the upper AOI. There was a main effect of expression for the upper AOI, $F(3,141) = 7.080$, $p < .001$, partial-eta = .131. Post-hoc *t*-tests showed that *fixation duration* was significantly longer for the fear expression compared to the anger, $t(47) = 3.036$, $p = .004$, happy, $t(47) = 4.111$, $p < .001$, and neutral expressions, $t(47) = 3.589$, $p = .001$.

4. Discussion

The aim of the current study was to investigate whether 7-months-old infants' visual exploration pattern of body expressions differed for angry, fearful, happy and emotionally neutral static body postures using eye tracking. In particular, we predicted that infants will be more likely to fixate and dwell on the upper body because the torso, arms and hands contain perceptual cues that discriminate between different body expressions, particularly fear and anger expressions (e.g., Atkinson, et al., 2004; Pollux et al., 2019; Poyo Solanas et al., 2020). Given the previously documented attentional bias for fear body expressions (e.g., Missana et al., 2014) and that adults preferentially fixate the upper part of fearful bodies for accurate recognition (Pollux et al., 2019; Poyo Solanas et al., 2020), we also predicted that infants will be more likely to fixate and dwell on upper body parts expressed during fear than other emotions. Towards this aim, infants were presented randomly with static images of adult female bodies with happy, fearful, angry, and emotionally neutral expressions. Considering the whole body, infants tended to allocate a lower proportion of fixations to happy bodies compared to the fearful and the angry ones. That said, we found that infants at this age spent a larger proportion of looking time on the upper relative to the lower part of the bodies. They also had a higher proportion of fixations on the upper relative to the lower body.

Importantly, we provide the first evidence that 7-months-old infants fixate differently on different body regions depending on the expression of the body posture. In line with our predictions, infants dwelled longer on the upper body for fear relative to the upper body of the other expressions as reflected by the proportion looking time and fixation duration. This increased looking time appears to derive from a similar proportion of fixations as for other expressions, but with longer duration. Infants' fixations on the head region also differed for the different body expressions, despite this region's relatively small area compared to the upper and lower body regions. In contrast to the upper body regions, infants spent less time and made proportionally fewer fixations to the head region for fear expressions compared to the other expressions. Lastly, infants did not explore the lower body differently for the different body expressions. Overall, these results extend and replicate the information about infant processing of emotional expressions displayed by human bodies.

Our findings support the hypothesis that the upper body, which includes the torso, arms and hands contain perceptual cues that can discriminate between emotional expressions from static body postures (Atkinson et al., 2004; Atkinson et al., 2007; Dael et al., 2012a,b; Atkinson, 2013; Glowinski et al., 2012). In line with this, we found no clear differences in infants' visual exploration behaviours between expressions for the lower body. Previous research has shown that infants discriminate between happy, fearful, angry, and sad static body postures when the emotion expressions are presented in pairs (Hock et al., 2017; Krol et al., 2015; Missana et al., 2014; Rajhans et al., 2015). The present study shows that this discriminatory pattern extends to some extent to a task where infants see more than two emotional body expressions. Importantly, in this case, the discrimination is predominantly driven by the upper body region which has been shown to be the most informative (Atkinson et al., 2004; Atkinson et al., 2007; Dael et al., 2012a,b; Atkinson, 2013; Glowinski et al., 2012).

Seven-months-old infants in our study looked proportionally longer at the upper fear bodies compared to the other conditions. Several previous ERP studies have shown that fearful bodies elicit increased ERP amplitude of the component functionally associate with attention (i.e., the Nc - Krol et al., 2015; Missana et al., 2014; Rajhans et al., 2015), thus it is possible that the visual preference observed in our study reflects increased allocation of attention. Interestingly, the increased dwell time to the upper region of fearful bodies resulted from longer fixation durations rather than more proportion of fixations compared to the other expressions. Longer fixation durations are thought to reflect detailed processing and increased cognitive effort (Bronson, 1991; Helo et al., 2016; Wass & Smith, 2014). This increased allocation of cognitive resources to the upper part of fearful bodies could also result in a richer representation for fear expression. The present data further indicates that this bias for fear is not particular to specific pairwise contrasts (most previous studies contrasted fear with happiness), but it is also elicited when fear is presented in the context of several other expressions, including emotionally neutral body postures. Fear expressions are salient cues for the presence of threat in our environment, and from this perspective the allocation of cognitive resources for accurate identification and representation of this expression could be adaptive (Atkinson, 2013; Emery & Amaral., 2000; de Gelder, 2006; de Gelder et al., 2004; Grezes, Pichon, & De Gelder, 2007; Leppänen & Nelson, 2009).

Rather surprisingly, some of the infants' looking behaviours on the head also differed between fear and the other expressions despite its much smaller size compared to the other body regions. In contrast to the upper body, infants looked less at fearful heads: they spent less time looking and made less proportion of fixations for fear relative to the other expressions. Other studies with infants have shown that they can discriminate expressions from static facial and body cues (e.g., Bayet & Nelson, 2019; Heck et al., 2018; Missana et al., 2014), and can even match expressions between faces and bodies (Hock et al., 2017). Our results extend these findings by suggesting that infants fixate on the head *per se* independently of any emotional expressions from facial cues (as the facial features were removed).

It is possible that some of the differences in infants' visual exploration behaviours between expressions for the upper body and head may reflect a bias to scan the upper half of the screen, which can lead to a more superficial processing of the lower compared to the upper body and head (Van Renswoude et al., 2016). However, several previous findings suggest that this is less likely to be the case. First, 7-months-old infants have well developed abilities for endogenous control of visual attention which can support the exploration of complex visual stimuli (Johnson et al., 1991). White and colleagues (2018) showed that infants preferentially scan either the upper or the lower part of human bodies dependent on the stimulus sex, looking longer at the upper half for female bodies and at the lower half for male bodies. Cross-cultural investigations also indicate that looking at the upper half of emotional facial expressions is not necessarily an automatic and generic perceptual strategy (Geangu et al., 2018). Western-Caucasian infants fixate both the upper and the lower part of happy and fearful faces, while East-Asian infants tend to fixate predominantly the upper half of faces. These findings converge with those showing that the type of emotion expression significantly influences how susceptible is infants' visual attention to distractors: they are more likely to maintain fixation on fearful compared to happy faces when distractors are present (Peltola et al., 2008). It is thus possible that infants' preferential looking to the upper body in the present study is driven by the information relevant for discriminating emotion expressions rather than pre-existing biases. The fact that the stimuli used here depict female bodies may have further consolidated this visual preference. However, it would be interesting to investigate in future studies whether the infants' visual exploration of emotional bodies is modulated by the stimulus sex (White et al., 2018).

It is possible that some of the differences that we found in infants' visual exploration behaviours between expressions may be driven by low-level visual features of the body images (White et al., 2018; Frank et al., 2009; van Renswoude et al., 2019; Amso et al., 2014; Kwon et al., 2016; Itti et al., 1998; Itti & Koch, 2001). Previous studies have attempted to control for this by various stimulus manipulations, such as stimulus inversion in the picture plane. However, there is mounting evidence to suggest that it is not clear how inversion affects stimulus processing. Consequently, it is not clear what inversion would reveal about the contribution of low-level features to perception (e.g., Burton, Schweinberger, Jenkins, & Kaufmann, 2015; Murphy & Cook, 2017. In order to determine if low-level differences could drive infants' visual exploration behaviours, rather than inverting the body images, our approach was to compare low-level visual features across AOIs and expressions using established measures (see Tables 1 and 2; and Figure 3). More specifically, we computed different low-level visual features for the stimuli, including mean surface area (size), luminance and visual saliency as function of AOI and expression. The visual saliency is a particularly informative measure, as salient regions are more likely to be fixated on or attended for a longer duration during free viewing of natural images (Itti et al., 1998; Itti & Koch, 2001; van Renswoude et al., 2019; Amso et al., 2014; Veale et al., 2017).

Some of the low-level visual features could partly explain differences in infants' visual exploration behaviours between expressions for the upper body and head. The size, spatial extent and visual saliency of the outstretched arms in the case of the happy body expression could have partially driven the lower proportion of fixations for the body in contrast to the fearful and angry expressions. As illustrated in the Supplementary Figure 5, many fixations were located in the vicinity of the arms' boundaries of the happy upper AOI. This pattern of fixation did not necessarily characterize other body parts that are relatively small (e.g., the head) or other body parts with similar shape (e.g., the legs). Thus, it could be that infants have difficulties fixating certain spatial configurations of body parts, like the wider extension of the arms in expressing happiness, with implications for how they perceive the emotion expressions. The slightly smaller size of the head region for fear compared to the other expressions (due to the hands often covering the face) could also be partially responsible for the differences in dwell time and proportion of fixations to different body parts.

Importantly, however, visual saliency was unlikely to drive differences in infants' visual exploration behaviours for upper and lower body regions across the different body expressions. In our study if visual saliency solely accounted for infants' visual exploration behaviours, then we would expect increased looking towards the upper body AOI of the stimuli relative to the lower body AOI *and* increased looking towards the upper body AOI of angry body postures compared to the other emotional body postures. Although our results showed increased looking towards the upper AOI, we found a larger proportion of looking time and longer fixation durations on the upper body for fear relative to the other expressions rather than for anger. Furthermore, the number of fixations recorded for the upper part of fearful bodies indicate that they were explored as much as the other expressions, despite having a smaller area (Gredebäck et al., 2012). Overall, it is more likely that diagnostic features for this emotion expression such as the position of the arms, rather than low-level visual features, represent the focus of infants' attention (Atkinson et al., 2004; Atkinson et al., 2007; Atkinson, 2013; Dael et al., 2012a; Dael et al., 2012b; Glowinski et al., 2008; Glowinski et al., 2011; Mancini et al., 2012; Poyo Solanas et al., 2020). Taken together, these results suggest 7-months-old infants' visual exploration of emotional body postures is unlikely to be entirely reliant on bottom-up processes and exogenous control of attention, and that top-down processes related to emotion representations are likely to be involved as well. Furthermore, it is important to emphasise that visual saliency alone cannot completely account for eye movement data, and this is more so for older participants. For example, Frank et al. (2009) showed that visual saliency was a good predictor for fixations to faces for 3-months-old infants, but was a poor predictor for 6- and 9-months-old infants and adults (see also Kwon et al., 2016). In our study, there were no significant correlations between visual saliency and 7-months-old infants' visual exploration of emotional body postures as a function of AOIs (see Supplementary Table 1).

Our study highlights the need for future research about the relationship between low-level visual features of body-part images, scanning biases and infants' visual exploration behaviours, which will be highly valuable for understanding the development of emotional body expression processing. Methodological approaches that include the analysis of visual saliency and other low-level visual features will be particularly informative in this respect (Amso et al., 2014; Frank et al., 2009; Itti et al., 1998; Itti & Koch, 2001; Kibbe et al., 2018; Kwon et al., 2016; Mayer et al., 2015, 2017, 2020; van Renswoude et al., 2019; Veale et al., 2017). The processing of low-level features can be affected by top-down feedback (e.g., Garrido-Vásquez et al., 2018; Hochstein & Ahissar, 2002), which means that for different perceptual tasks the same low-level features could contribute differently to participants' performance. For example, Kibbe and colleagues (2018) have shown that the same low-level visual features of the stimuli can lead to different patterns of looking in infants when the stimuli are presented in different contexts. In comparison to other methodological options (e.g., stimulus inversion in the picture plane), the analysis of visual saliency can provide more information about the potential contribution of low-level features for the perceptual task that is more similar to what infants are familiar with and experience in everyday life - upright human bodies (e.g., Sugden & Moulson, 2017). Lastly on a pragmatic note, the use of visual saliency means that experimental paradigms do not have to include additional manipulations to control for low-level stimulus properties (e.g., stimulus inversion), unless these are the direct focus of the study, and thus could be less taxing for infants' limited attention span.

The present infant eye-tracking study helps to bring together different lines of research regarding infants' ability to process perceptual cues from different body parts to accomplish a variety of social tasks. Naturalistic sampling of infants early visual experiences using head-mounted cameras have shown that while bodies remain constant in infants' field of view during the first 2 years of life, the proportion of faces in the field of view decreases over the same period (Fausey et al., 2016;

Jayaraman et al., 2017; Smith et al., 2018). Perhaps most critically for the overall findings here, the proportion of non-face body parts (e.g., hands) becomes larger than the proportion of faces between 6-9 months (Fausey et al., 2016), and the proportion of visible bodies with faces decreases substantially around this period (from ~0.60 to ~0.40; Jayaraman et al., 2017). These visual experiences coincide with infants' abilities to, for example, attend to goal-directed actions of hands (Geangu et al., 2015) or build self-representations of their own body parts (Fillippetti & Tsakiris, 2018; Nielsen et al., 2006; Geangu, 2008; Zmyj et al., 2011; Suddendorf & Butler, 2013). They also approximately coincide to changes to the neural responses to faces and bodies. For example, ERP studies with infants have shown an increase in the N290 component to fear compared to happy body expressions (Missana et al., 2014) that becomes evident at around 8 months. This component is analogous to the N170 component for faces and bodies in the mature adult (Gillmeister et al., 2019; Itier & Taylor, 2002; Minnebusch et al., 2008; Reed et al., 2003; Righart & de Gelder, 2008; Stekelenburg & de Gelder, 2004; Taylor et al., 2010; Thierry et al., 2006; Watanabe et al., 2003). The increase in the N290 component for different expressions can reflect extraction of structural information, including different body parts (Gillmeister et al., 2019; Stekelenburg & de Gelder, 2004). Infants' early visual experiences, their increasing ability to process different body parts for different tasks, and changes at the neural level may lead them to use different body regions for discriminating different emotional expressions.

Adult studies on the perception of expressions from body postures highlight the importance of characteristic poses for the different expressions (Atkinson et al., 2004; Atkinson et al., 2007; Dael et al., 2012a,b; Pollux et al., 2019). For static body postures, these characteristics signature tend to be mostly located in the upper body. Indeed, recent eye-tracking studies with adults found that adults' visual exploration behaviours concentrated around upper body parts including the arms and hands (e.g., Pollux et al., 2019; Poyo Solanas et al., 2020; Ross & Flack, 2019). Interestingly, fear and anger were the expressions that were best distinguished by fixations on the upper body. However, in contrast to infants in our study, adults in Pollux's et al.'s study spent a relatively larger proportion of their looking times on faceless heads (~0.3-0.4, compared to ~0.1-0.2 in this study). Interestingly, adults' proportion of viewing times to faceless heads were least for fear body expressions, similar to our infant data.

Despite the similarity between infant and adult visual exploration behaviours to static body expressions, it is most likely that infants have a protracted developmental trajectory for recognising emotional expressions from the body without the face, as they do for faces (e.g., Taylor, Batty & Itier, 2004). For example, 3-years-old children's facial responses do not yet converge with the communicative value of emotional body expressions (Geangu et al., 2016b). Furthermore, the activity of the neural network sensitive to this type of social information is not yet fully matured and adult-like before late childhood and early adolescence (Ross et al., 2012, 2014, 2019). Our results only show that 7-months-old infants' visual exploration behaviours can discriminate between different expressions, but not whether the infants recognise the emotion *per se*. Building on our findings, future studies can integrate eye-tracking with other measures, such as ERPs, to better establish the relation between differential visual exploration of bodies, the underlying cognitive processes, and emotion recognition (e.g., Ayneto & Sebastian-Galles, 2017; Poulin-Dubois et al., 2018; Vanderwert et al., 2015). Future eye-tracking studies can also use more naturalistic images, such as those in which the body is embedded in a scene, or videos containing dynamic body cues (e.g., Heck et al., 2018; Kaiser et al., 2018; Missana & Grossmann, 2015; Poulin-Dubois et al., 2018; Quiroz, Geangu, & Hooi Yong, 2018). The effect of the presence of auditory information, generated by voice (e.g., ., Crespo-Llado et al., 2018; de Gelder & Vroomen, 2000; Yeh et al., 2016) or by the moving body (Geangu, Quadrelli, Lewis, Macchi Cassia, & Turati, 2015; Quadrelli, Geangu, & Turati, 2019), may also be informative in this case. Adults, for instance, can find humans more quickly and more efficiently than they can find machines in both static natural images and videos (Mayer et al., 2015, 2017). Moreover, visual-saliency models also did not predict adults' visual search behaviours. Similar studies with infants, using naturalistic images or videos and contrasting biological and mechanical categories, can further help us understand infants' ability to process bodies for different social tasks and the developmental trajectory of this ability.

Conclusion. Our study provides the first evidence that 7-months-old infants use perceptual cues from upper body regions and parts for discriminating emotional states expressed by static body postures. The upper body and head (without the face) are particularly important for discriminating fear from the other expressions, which is surprisingly similar to the regions adults fixate on for recognising different emotional expressions from static body postures (Pollux et al., 2019, Poyo Solanas et al., 2020). These differences in visual exploration behaviours for the different body regions we tested coincide with period when infants' visual experience increases for the body parts (e.g., hands) and concurrently decreases for faces (Fausey et al., 2016; Jayaraman et al., 2017; Smith et al., 2018). Overall, these findings suggest that visual experience and informativeness of different body parts may both play important roles in shaping how infants perceive and process static body postures for discriminating emotional expressions during the critical first 2-years of life.

Acknowledgements

We would like to express our gratitude to all families who dedicated their time to participate in this research project with their children. Without their continuous interest in our work and desire to help, these findings would have not been possible. The authors would also like to thank Camilla Tognato for her assistance with data collection. The authors wish to declare no conflict of interests.

References

Amso, D., Haas, S., & Markant, J. (2014). An eye tracking investigation of developmental change in bottom-up attention orienting to faces in cluttered natural scenes. *PLoS ONE*, 9(1), e85701. doi:10.1371/journal.pone.0085701

Atkinson, A. P. (2013). Bodily expressions of emotion: Visual cues and neural mechanisms. In J. Armony and P. Vuilleumier (Eds.). "The Cambridge handbook of human affective neuroscience". Cambridge University Press. https://doi.org/10.1017/CBO9780511843716

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, *33*(6), 717-746.

Atkinson, A. P., Tunstall, M. L., & Dittrich, W. H. (2007). Evidence for distinct contributions of form and motion information to the recognition of emotions from body gestures. *Cognition*, *104*(1), 59-72.

Aviezer, H., Hassin, R. R., & Bentin, S. (2012). Impaired integration of emotional faces and affective body context in a rare case of developmental visual agnosia. *Cortex*, *48*(6), 689-700.

Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science*, *338*(6111), 1225-1229.

Bauser, D. S., & Suchan, B. (2018). The role of the head in configural body processing. *British Journal of Psychology*, *109*(3), 564-582.

Bayet, L., & Nelson, C. A. (2019). The perception of facial emotion in typical and atypical development. In *Handbook of Emotional Development* (pp. 105-138). Springer, Cham.

Bernard, P., Content, J., Deltenre, P., & Colin, C. (2018). When the body becomes no more than the sum of its parts: the neural correlates of scrambled versus intact sexualized bodies. *Neuroreport*, *29*(1), 48-53.

Bronson, G. W. (1991). Infant differences in rate of visual encoding. *Child Development*, *62*(1), 44-54.

Campanella, S., & Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences*, *11*(12), 535-543.

Crespo-Llado, M. M., Vanderwert, R. E., & Geangu, E. (2018). Individual differences in infants' neural responses to their peers' cry and laughter. *Biological psychology*, *135*, 117-127.

Dael, N., Mortillaro, M., & Scherer, K. R. (2012a). Emotion expression in body action and posture. *Emotion*, *12*(5), 1085.

Dael, N., Mortillaro, M., & Scherer, K. R. (2012b). The body action and posture coding system (BAP): Development and reliability. *Journal of Nonverbal Behavior*, *36*(2), 97-121.

De Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, *7*(3), 242.

De Gelder, B. (2009). Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3475-3484.

De Gelder, B., Snyder, J., Greve, D., Gerard, G., & Hadjikhani, N. (2004). Fear fosters flight: a mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proceedings of the National Academy of Sciences*, *101*(47), 16701-16706.

De Gelder, B., & Van den Stock, J. (2011). The bodily expressive action stimulus test (BEAST). Construction and validation of a stimulus basis for measuring perception of whole body expression of emotions. *Frontiers in Psychology*, *2*, 181.

De Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition & Emotion*, *14*(3), 289-311.

Garrido-Vásquez, P., Pell, M. D., Paulmann, S., & Kotz, S. A. (2018). Dynamic facial expressions prime the processing of emotional prosody. *Frontiers in human neuroscience*, *12*, 244.

Grezes, J., Pichon, S., & De Gelder, B. (2007). Perceiving fear in dynamic body expressions. *Neuroimage*, *35*(2), 959-967.

de Groot, J. H., & Smeets, M. A. (2017). Human fear chemosignaling: evidence from a meta-analysis. *Chemical Senses*, *42*(8), 663-673.

Eimer, M. (2000). The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport*, *11*(10), 2319-2324.

Emery, N. J., & Amaral, D. G. (2000). The role of the amygdala in primate social cognition. *Cognitive neuroscience of emotion*, 156-191.

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience*, *9*(7), 878.

Fausey, C. M., Jayaraman, S., & Smith, L. B. (2016). From faces to hands: Changing visual input in the first two years. *Cognition*, *152*, 101-107.

Filippetti, M. L., & Tsakiris, M. (2018). Just before I recognize myself: The role of featural and multisensory cues leading up to explicit mirror self-recognition. *Infancy*, *23*(4), 577-590.

Geangu, E. (2008). Notes on self-awareness development in early infancy. *Cognition, Brain, Behavior*, *12*(1), 103.

Geangu, E., Ichikawa, H., Lao, J., Kanazawa, S., Yamaguchi, M. K., Caldara, R., & Turati, C. (2016a). Culture shapes 7-month-olds' perceptual strategies in discriminating facial expressions of emotion. *Current Biology*, *26*(14), R663-R664.

Geangu, E., Quadrelli, E., Conte, S., Croci, E., & Turati, C. (2016b). Three-year-olds' rapid facial electromyographic responses to emotional facial expressions and body postures. *Journal of Experimental Child Psychology*, *144*, 1-14.

Geangu, E., Quadrelli, E., Lewis, J. W., Macchi Cassia, V., & Turati, C. (2015). By the sound of it: an ERP investigation of human action sound processing in 7-month-old infants. *Developmental Cognitive Neuroscience*, *12*, 134-144. https://doi.org/10.1016/j.dcn.2015.01.005

Geangu, E., Senna, I., Croci, E., & Turati, C. (2015). The effect of biomechanical properties of motion on infants' perception of goal-directed grasping actions. *Journal of Experimental Child Psychology*, *129*, 55-67.

Gillmeister, H., Stets, M., Grigorova, M., & Rigato, S. (2019). How do bodies become special? Electrophysiological evidence for the emergence of body-related cortical processing in the first

14 months of life. *Developmental Psychology, 55*(10), 2025-2038. https://doi.org/10.1037/dev0000762

Glowinski, D., Camurri, A., Volpe, G., Dael, N., & Scherer, K. (2008). Technique for automatic emotion recognition by body gesture analysis. In *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* (pp. 1-6). IEEE.

Glowinski, D., Dael, N., Camurri, A., Volpe, G., Mortillaro, M., & Scherer, K. (2011). Toward a minimal representation of affective gestures. *IEEE Transactions on Affective Computing*, *2*(2), 106-118.

Gredebäck, G., Eriksson, M., Schmitow, C., Laeng, B., & Stenberg, G. (2012). Individual differences in face processing: Infants' scanning patterns and pupil dilations are influenced by the distribution of parental leave. *Infancy*, *17*(1), 79-101.

Harel, J., Koch, C., & Perona, P. (2006). Graph-based visual saliency. *NIPS*, 1-8.

Heck, A., Chroust, A., White, H., Jubran, R., & Bhatt, R. S. (2018). Development of body emotion perception in infancy: From discrimination to recognition. *Infant Behavior and Development*, *50*, 42-51.

Heck, A., Chroust, A., White, H., Jubran, R., & Bhatt, R. S. (2018). Development of body emotion perception in infancy: From discrimination to recognition. *Infant Behavior and Development*, *50*, 42-51.

Helo, A., Rämä, P., Pannasch, S., & Meary, D. (2016). Eye movement patterns and visual attention during scene viewing in 3-to 12-month-olds. *Visual neuroscience*, *33*.

Hock, A., Oberst, L., Jubran, R., White, H., Heck, A., & Bhatt, R. S. (2017). Integrated emotion processing in infancy: Matching of faces and bodies. *Infancy*, *22*(5), 608-625.

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron, 36,* 791-804.

Hoehl, S. (2014). Emotion processing in infancy. In *Children and Emotion* (Vol. 26, pp. 1-12). Karger Publishers.

Itier, R.J., & Taylor, M.J. (2002). Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: a repetition study using ERPs. *Neuroimage*, 15, 352–372.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature reviews neuroscience*, *2*(3), 194-203.

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 1254-1259.

Jayaraman, S., Fausey, C. M., & Smith, L. B. (2017). Why are faces denser in the visual experiences of younger than older infants? *Developmental Psychology*, *53*(1), 38.

Jessen, S., & Kotz, S. A. (2011). The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. *Neuroimage*, *58*(2), 665-674.

Johnson, M. H. (1990). Cortical maturation and the development of visual attention in early infancy. *Journal of cognitive neuroscience*, *2*(2), 81-95.

Kaiser, J., Crespo-Llado, M. M., Turati, C., & Geangu, E. (2017). The development of spontaneous facial responses to others' emotions in infancy: An EMG study. *Scientific reports*, *7*(1), 1-10.

Keltner, D., Tracy, J., Sauter, D. A., Cordaro, D. C., & McNeil, G. (2016). Expression of emotion. *Handbook of emotions*, 467-482.

Kibbe, M. M., Kàldy, Z., & Blaser, E. (2018). Rules infants look by: Testing the assumption of transitivity in visual salience. *Infancy*, *23*(2), 156-172.

Kochukhova, O., & Gredebäck, G. (2010). Preverbal infants anticipate that food will be brought to the mouth: An eye tracking study of manual feeding and flying spoons. *Child Development*, *81*(6), 1729-1738.

Krol, K. M., Rajhans, P., Missana, M., & Grossmann, T. (2015). Duration of exclusive breastfeeding is associated with differences in infants' brain responses to emotional body expressions. *Frontiers in Behavioral Neuroscience*, *8*, 459.

Kwon, M-K., Setoodehnia, M., Baek, J., Luck, S. J., & Oakes, L. M. (2016). The development of visual search in infancy: Attention to faces versus salience. *Developmental Psychology*, 52(4), 537-555.

Leppänen, J. M., & Nelson, C. A. (2009). Tuning the developing brain to social signals of emotions. *Nature Reviews Neuroscience*, *10*(1), 37.

Mancini, M., Varni, G., Glowinski, D., & Volpe, G. (2012, October). Computing and evaluating the body laughter index. In *International Workshop on Human Behavior Understanding* (pp. 90-98). Springer, Berlin, Heidelberg.

Mayer, K. M., Vuong, Q. C., & Thornton, I. M. (2015). Do people "pop out"?. *PloS One*, *10*(10), e0139618.

Mayer, K. M., Vuong, Q. C., & Thornton, I. M. (2017). Humans are detected more efficiently than machines in the context of natural scenes. *Japanese Psychological Research*, *59*(2), 178-187.

Mayer, K. M., Thornton, I. M., & Vuong, Q. C. (2020). Comparable search efficiency for human and animal targets in the context of natural scenes. *Attention, Perception & Psychophysics, 82(3),* 954-965.

Missana, M., & Grossmann, T. (2015). Infants' emerging sensitivity to emotional body expressions: Insights from asymmetrical frontal brain activity. *Developmental Psychology*, *51*(2), 151.

Missana, M., Atkinson, A. P., & Grossmann, T. (2015). Tuning the developing brain to emotional body expressions. *Developmental Science*, *18*(2), 243-253.

Missana, M., Rajhans, P., Atkinson, A. P., & Grossmann, T. (2014). Discrimination of fearful and happy body postures in 8-month-old infants: an event-related potential study. *Frontiers in Human Neuroscience*, *8*, 531.

Nielsen, M., Suddendorf, T., & Slaughter, V. (2006). Mirror self-recognition beyond the face. *Child Development*, *77*(1), 176-185.

Peltola, M. J., Leppänen, J. M., Palokangas, T., & Hietanen, J. K. (2008). Fearful faces modulate looking duration and attention disengagement in 7-month-old infants. *Developmental science*, *11*(1), 60-68.

Pollux, P. M., Craddock, M., & Guo, K. (2019). Gaze patterns in viewing static and dynamic body expressions. *Acta psychologica*, *198*, 102862.

Poulin-Dubois, D., Hastings, P., Chiarella, S., Geangu, E., Hauf, P., Ruel, A., & Johnson, A. (2018). The eyes know it: Toddlers' visual scanning of sad faces is predicted by their theory of mind skills. *PLOS one*. https://doi.org/10.1371/journal.pone.0208524

Poyo Solanas, M., Vaessen, M., & de Gelder, B. (2020). Limb contraction drives fear perception. *bioRxiv*.

Quiroz, J. C., Geangu, E., & Hooi Yong, M. (2018). Emotion-recognition using smart watch sensor data: Mixed-design study. *JMIR Mental Health*, *5*(3), [e10153]. https://doi.org/10.2196/10153

Quadrelli, E., Geangu, E., & Turati, C. (2019). Human action sounds elicit sensorimotor activation early in life. *Cortex*, *117*, 323-335. https://doi.org/10.1016/j.cortex.2019.05.009

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

Reed, C.L., Stone, V.E., Bozova, S., & Tanaka, J. (2003). The body-inversion effect. *Psychological Science,* 14, 302–308.

Robinson, D. T., Clay-Warner, J., Moore, C. D., Everett, T., Watts, A., Tucker, T. N., & Thai, C. (2012). Toward an unobtrusive measure of emotion during interaction: Thermal imaging techniques. In *Biosociology and Neurosociology* (pp. 225-266). Emerald Group Publishing Limited.

Rosen, J. B., Asok, A., & Chakraborty, T. (2015). The smell of fear: innate threat of 2, 5-dihydro-2, 4, 5-trimethylthiazoline, a single molecule component of a predator odor. *Frontiers in Neuroscience*, *9*, 292.

Ross, P. D., de Gelder, B., Crabbe, F., & Grosbras, M. H. (2014). Body-selective areas in the visual cortex are less active in children than in adults. *Frontiers in Human Neuroscience*, *8*, 941.

Ross, P. D., Polson, L., & Grosbras, M. H. (2012). Developmental changes in emotion recognition from full-light and point-light displays of body movement. *PLoS One*, *7*(9), e44815.

Ross, P., & Flack, T. R. (2019). Removing hand information specifically impairs emotion recognition for fearful and angry body stimuli. *Perception*. https://doi.org/10.1177/0301006619893229.

Ross, P., de Gelder, B., Crabbe, F., & Grosbras, M. H. (2019). Emotion modulation of the body-selective areas in the developing brain. *Developmental Cognitive Neuroscience*, *38*, 100660.

Rossion, B., & Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *Neuroimage*, *39*(4), 1959-1979.

Salazar-López, E., Domínguez, E., Ramos, V. J., De la Fuente, J., Meins, A., Iborra, O., ... & Gómez-Milán, E. (2015). The mental and subjective skin: Emotion, empathy, feelings and thermography. *Consciousness and Cognition*, *34*, 149-162.

Shuman, V., Clark-Polner, E., Meuleman, B., Sander, D., & Scherer, K. R. (2017). Emotion perception from a componential perspective. *Cognition and Emotion*, *31*(1), 47-56.

Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The developing infant creates a curriculum for statistical learning. *Trends in Cognitive Sciences*, *22*(4), 325-336.

Stekelenburg, J. J., & de Gelder, B. (2004). The neural correlates of perceiving human bodies: An ERP study on the body-inversion effect. *Neuroreport, 15*(5), 777-780.

Stienen, B. M., Tanaka, A., & de Gelder, B. (2011). Emotional voice and emotional body postures influence each other independently of visual awareness. *PLoS One*, *6*(10), e25517.

Suddendorf, T., & Butler, D. L. (2013). The nature of visual self-recognition. *Trends in Cognitive Sciences*, *17*(3), 121-127.

Sugden, N. A., & Moulson, M. C. (2017). Hey baby, what's "up"? One-and 3-month-olds experience faces primarily upright but non-upright faces offer the best views. *Quarterly Journal of Experimental Psychology*, *70*(5), 959-969.

Tamietto, M., & De Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews Neuroscience*, *11*(10), 697.

Van den Stock, J., Righart, R., & De Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, *7*(3), 487.

van Renswoude, D. R., Johnson, S. P., Raijmakers, M. E. J., & Visser, I. (2016). Do infants have the horizontal bias? *Infant Behavior and Development,* 44, 38-48.

van Renswoude, D. R., Visser, I., Raijmakers, M. E. J., Tsang, T., & Johnson, S. P. (2019). Real-world scene perception in infants: What factors guide attention allocation? *Infancy*, 24, 693-717.

Veale, R., Hafed, Z. M., & Yoshida, M. (2017). How is visual salience computed in the brain? Insights from behaviour, neurobiology and modelling. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1714), 20160113.

Walle, E. A., & Campos, J. J. (2014). The development of infant detection of inauthentic emotion. *Emotion*, *14*(3), 488.

Walle, E. A., Dahl, A., & Campos, J. J. (2012). How can one piece together emotion when a crucial piece is missing? *Emotion Review*, *4*(3), 299-300.

Wass, S. V., & Smith, T. J. (2014). Individual differences in infant oculomotor behavior during the viewing of complex naturalistic scenes. *Infancy*, *19*(4), 352-384.

Watanabe, S., Kakigi, R., & Puce, A. (2003). The spatiotemporal dynamics of the face inversion effect: a magneto- and electro-encephalographic study. *Neuroscience*, 116(3), 879–895.

White, H., Jubran, R., Heck, A., Chroust, A., & Bhatt, R. S. (2019). Sex-specific scanning in infancy: Developmental changes in the use of face/head and body information. *Journal of Experimental Child Psychology*, *182*, 126-143.

Yeh, P. W., Geangu, E., & Reid, V. (2016). Coherent emotional perception from body expressions and the voice. *Neuropsychologia*, *91*, 99-108.

Zieber, N., Kangas, A., Hock, A., & Bhatt, R. S. (2014). Infants' perception of emotion from body movements. *Child Development*, *85*(2), 675-684.

Zion-Golumbic, E., & Bentin, S. (2006). Dissociated neural mechanisms for face detection and configural encoding: evidence from N170 and induced gamma-band oscillation effects. *Cerebral Cortex*, *17*(8), 1741-1749.

Zmyj, N., Jank, J., Schütz-Bosbach, S., & Daum, M. M. (2011). Detection of visual–tactile contingency in the first year after birth. *Cognition*, *120*(1), 82-89.

Figures with Captions

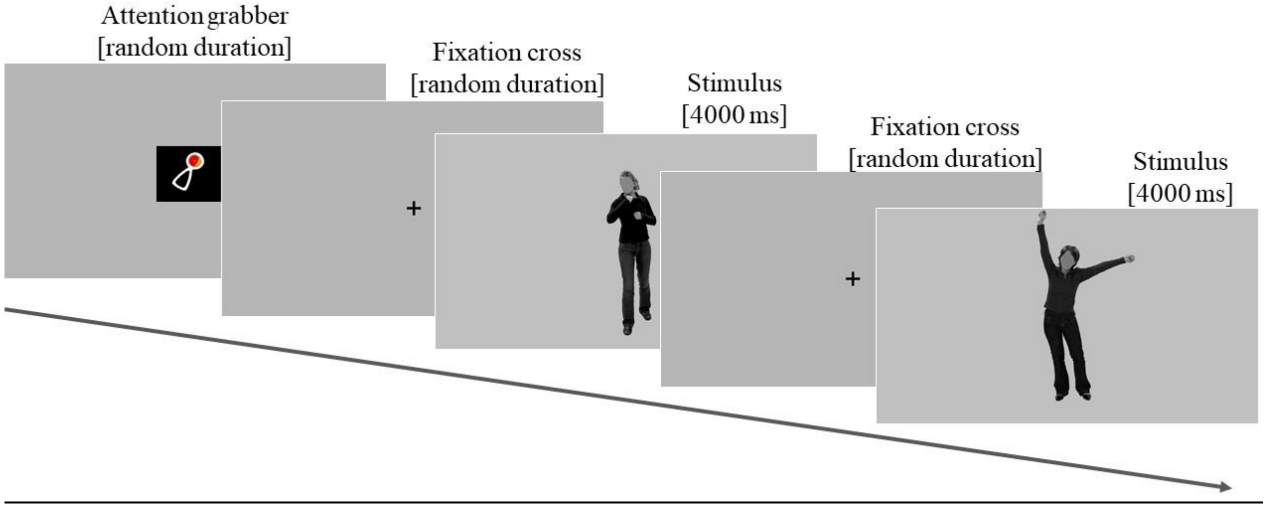


Figure 1. The sequence of events during the experiment. An attention grabber could be used to attract infants' attention to the screen. Once infants fixated on a fixation cross, the experimenter pressed the space bar to present the next stimulus. If infants looked away from the screen, the experimenter presented the attention grabber again.

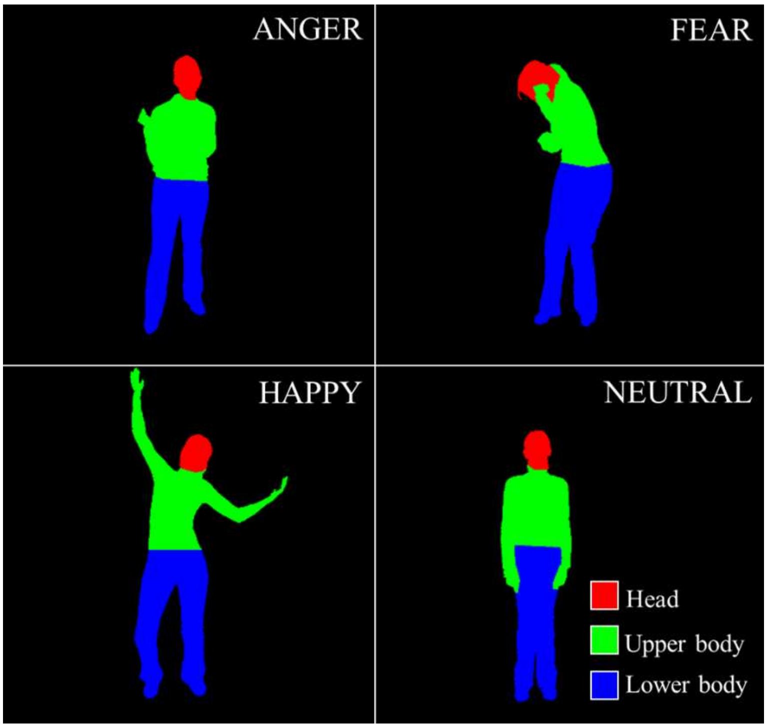


Figure 2. The three areas of interest (AOIs) used for the analysis of the eye-movement data. The AOIs are shown for each body expression from the same female actor. The approximate location of the belly button was used to demarcate the upper and lower body AOI. The arms and hands were assigned to the upper body even though it occluded the head for fear expressions.

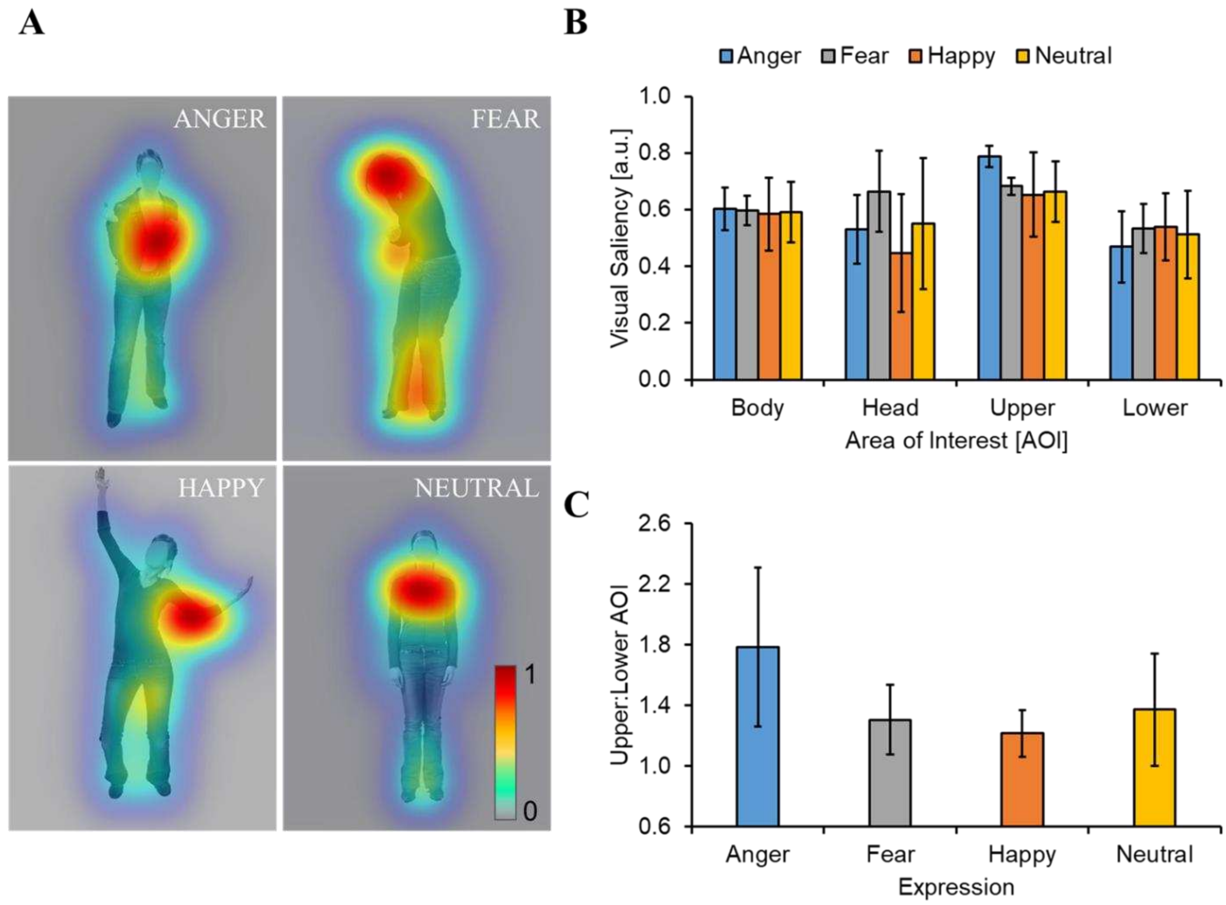


Figure 3. Visual saliency of body expressions. A) Examples of visual saliency map superimposed on the body image for anger, fear, happy and neutral emotional expressions based on Itti et al. (1998). The images match those in Figures 2 and 4 (please see Supplemental Figures 5 – 8 for the visual saliency map of all 20 body images). The visual saliency at each image pixel was based on grayscale intensity (at 8 image scales) and orientation of edges in the image (at 4 angles), in arbitrary units [a.u.] normalised to a value between 0 and 1. Harel et al.'s (2006) MATLAB implementation with default settings was used to compute visual saliency (http://www.klab.caltech.edu/~harel/share/gbvs.php). Warm colours represent image regions of high saliency (closer to 1) where as cool colours represent regions of low saliency (closer to 0 = no colour). B) The mean visual saliency as a function of AOI and expression. The error bars represents the standard deviation. C) The visual saliency ratio between the upper and lower AOI (i.e., mean upper saliency/mean lower saliency). Values greater than 1 reflect higher visual saliency (on average) in the upper AOI relative to the lower AOI.

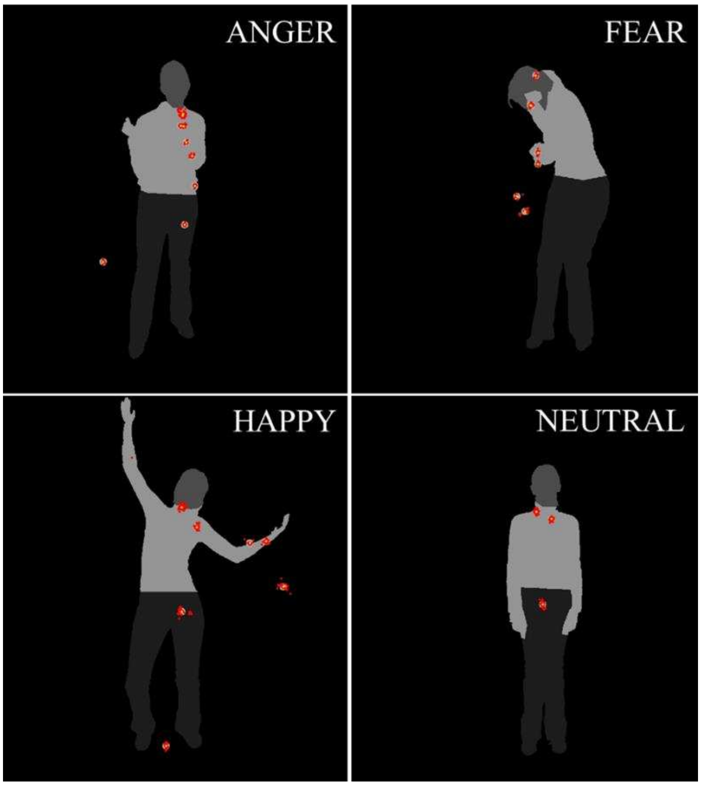


Figure 4. Examples of visual exploration patterns for anger, fear, happy and neutral body images. For each expression, the fixations from an individual infant are superimposed onto the body image. The red points represent gaze points and the yellow point represents the fixation point (averaged across all gaze points). The head, upper body and lower body AOI are presented as different grey levels here so that the fixation points are visible.

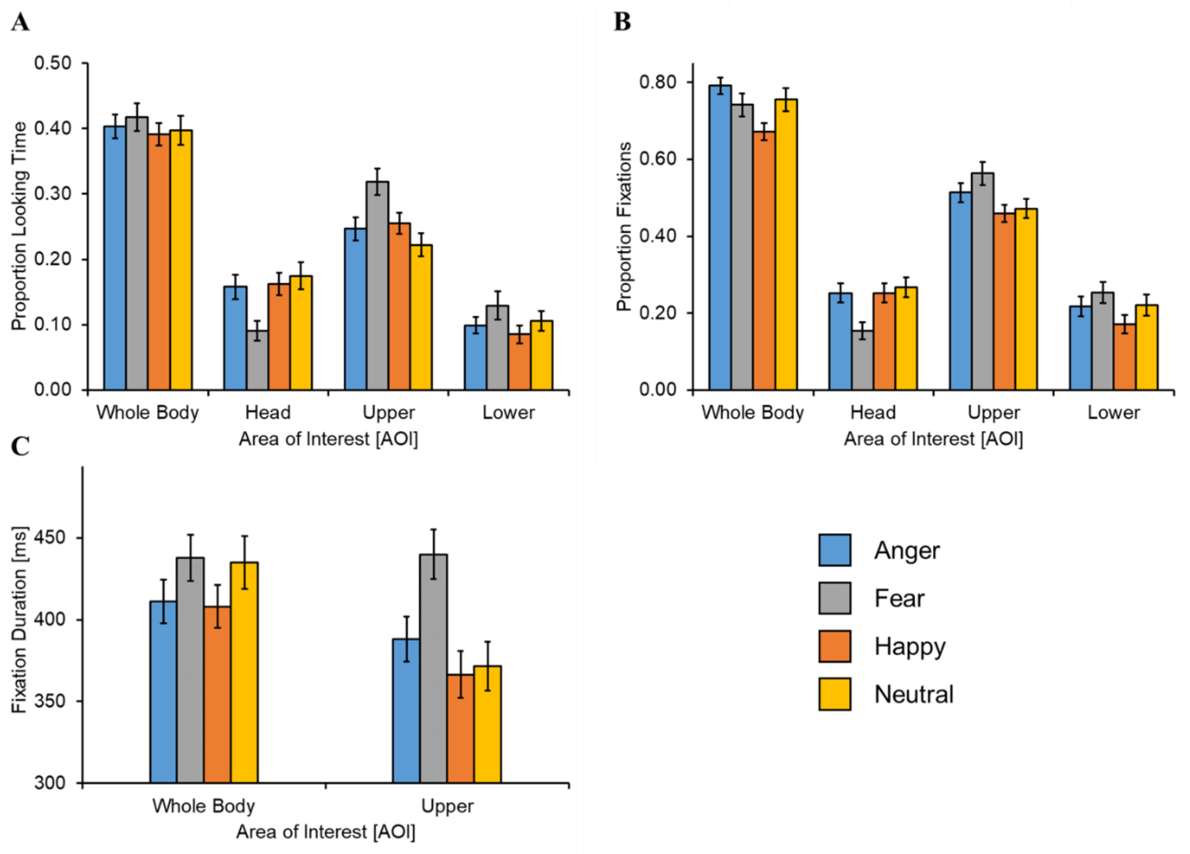


Figure 5. Mean and standard error of the mean (SE) of the three fixation parameters averaged across infants as a function of AOI (including the whole body) and expression. A) The proportion looking time was computed by summing the fixation duration of all valid fixations falling within the whole body or AOI on a given trial and dividing by 4-s (i.e. the duration of the body stimulus). B) The proportion fixations was computed by dividing the number of fixations by the total number of valid fixations on a specific trial. C) The fixation duration was computed by averaging the fixation duration of all valid fixations falling within the whole body or AOI on a given trial.

Tables with Captions

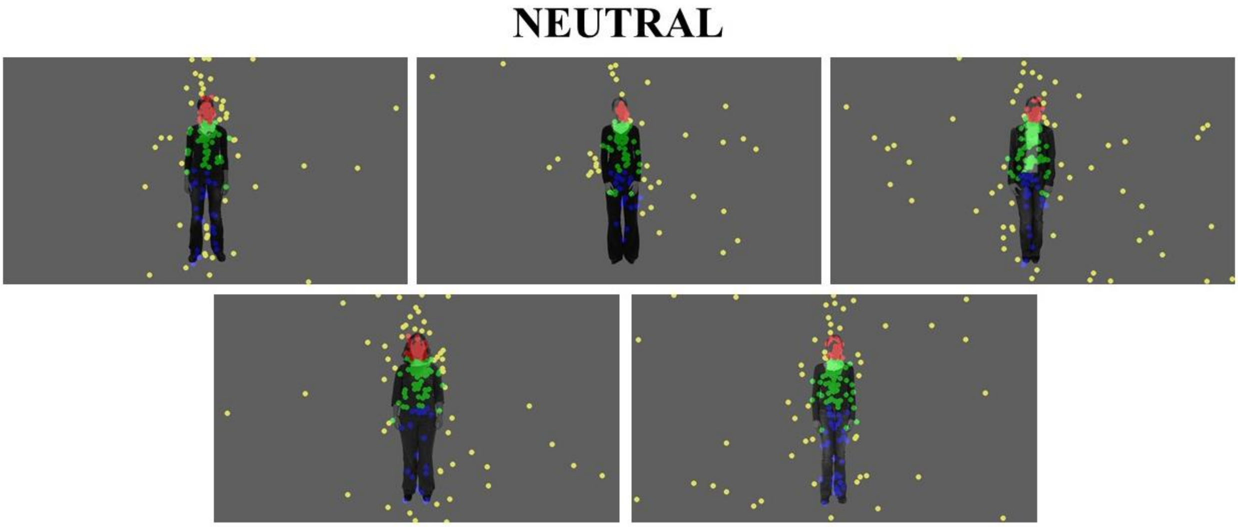
Table 1. The mean percentage of total screen area and total number of pixels occupied by the stimulus as a function of AOI and body expression. The number in parenthesis represent the standard deviation.

	Mean % Total Screen Area (SD)				**Pixels (SD)**			
	Body	**Head**	**Upper**	**Lower**	**Body**	**Head**	**Upper**	**Lower**
Anger	5.11 (0.13)	0.43 (0.05)	2.04 (0.11)	2.64 (0.20)	105902 (2717)	8972 (955)	42258 (2379)	54673 (4097)
Fear	5.53 (0.47)	0.32 (0.11)	1.88 (0.29)	3.33 (0.25)	114741 (9776)	6656 (2280)	39020 (6015)	69061 (5241)
Happy	6.24 (0.19)	0.40 (0.08)	2.73 (0.07)	3.11 (0.19)	129348 (3897)	8390 (1566)	56551 (1496)	64406 (3876)
Neutral	5.66 (0.45)	0.41 (0.10)	2.64 (0.22)	2.61 (0.28)	117408 (9423)	8545 (2075)	54801 (4570)	54060 (5892)
mean	5.64 (0.52)	0.39 (0.09)	2.32 (0.42)	2.92 (0.38)	116850 (10838)	8141 (1879)	48157 (8658)	60550 (7933)

Table 2. The mean luminance (0 = black; 128 = mid-grey; 255 = white) of the stimulus as a function of AOI and body expression. The number in parenthesis represent the standard deviation. Note that we excluded the background in computing the mean luminance.

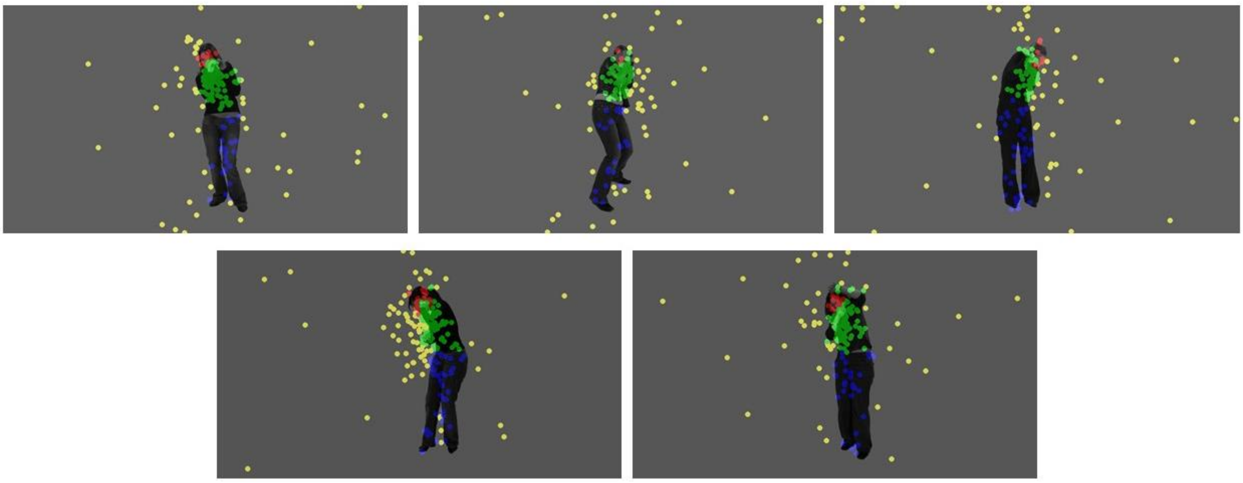
	Body	**Head**	**Upper**	**Lower**
Anger	56.3	115.7	50.2	52.1
	(16.7)	(18.3)	(27.0)	(17.0)
Fear	50.3	86.9	49.1	48.6
	(14.8)	(41.9)	(21.9)	(18.0)
Happy	45.3	114.5	39.7	41.8
	(6.2)	(18.2)	(12.8)	(9.7)
Neutral	72.6	135.7	75.3	60.0
	(21.0)	(26.8)	(20.8)	(23.9)
mean	56.1	113.2	53.6	50.6
	(17.8)	(31.3)	(23.7)	(17.7)

Supplementary Information



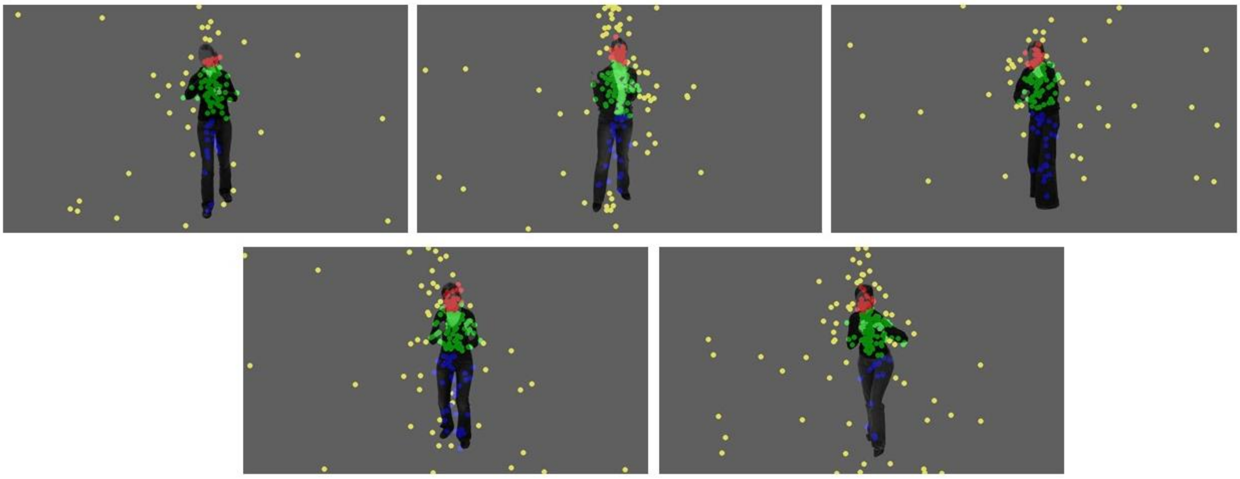
Supplementary Figure 1. The position of valid fixations across all infants for the five emotionally neutral body images. Red: fixations assigned to the Head ROI; Blue: fixations assigned to the Lower Body ROI; Green: fixations assigned to the Upper Body ROI; Yellow: valid fixations that were not assigned to any of the body AOIs.

FEAR



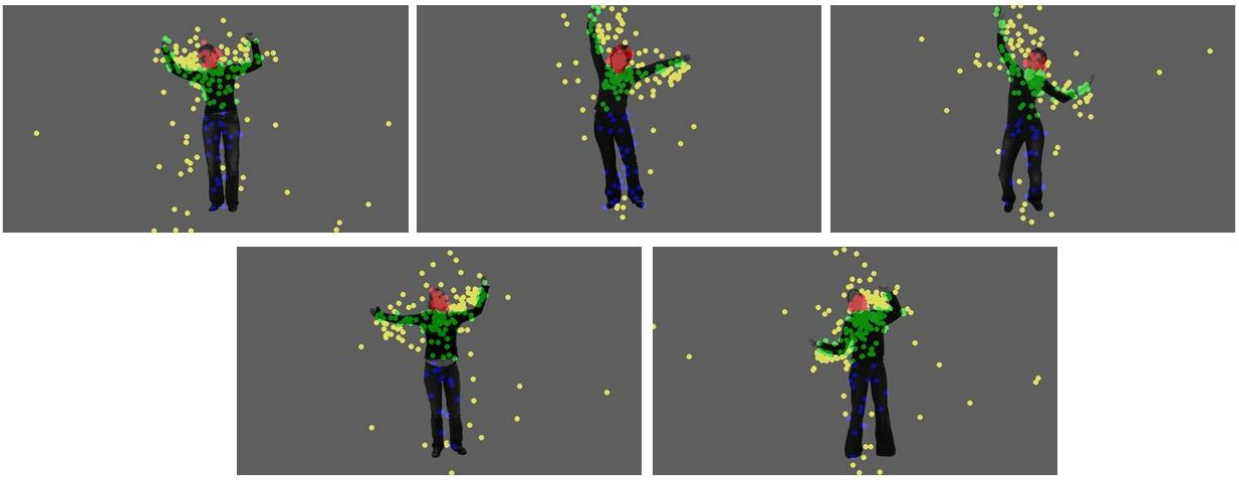
Supplementary Figure 2. The position of valid fixations across all infants for the five fear body images. Red: fixations assigned to the Head ROI; Blue: fixations assigned to the Lower Body ROI; Green: fixations assigned to the Upper Body ROI; Yellow: valid fixations that were not assigned to any of the body AOIs.

ANGER



Supplementary Figure 3. The position of valid fixations across all infants for the five angry body images. Red: fixations assigned to the Head ROI; Blue: fixations assigned to the Lower Body ROI; Green: fixations assigned to the Upper Body ROI; Yellow: valid fixations that were not assigned to any of the body AOIs.

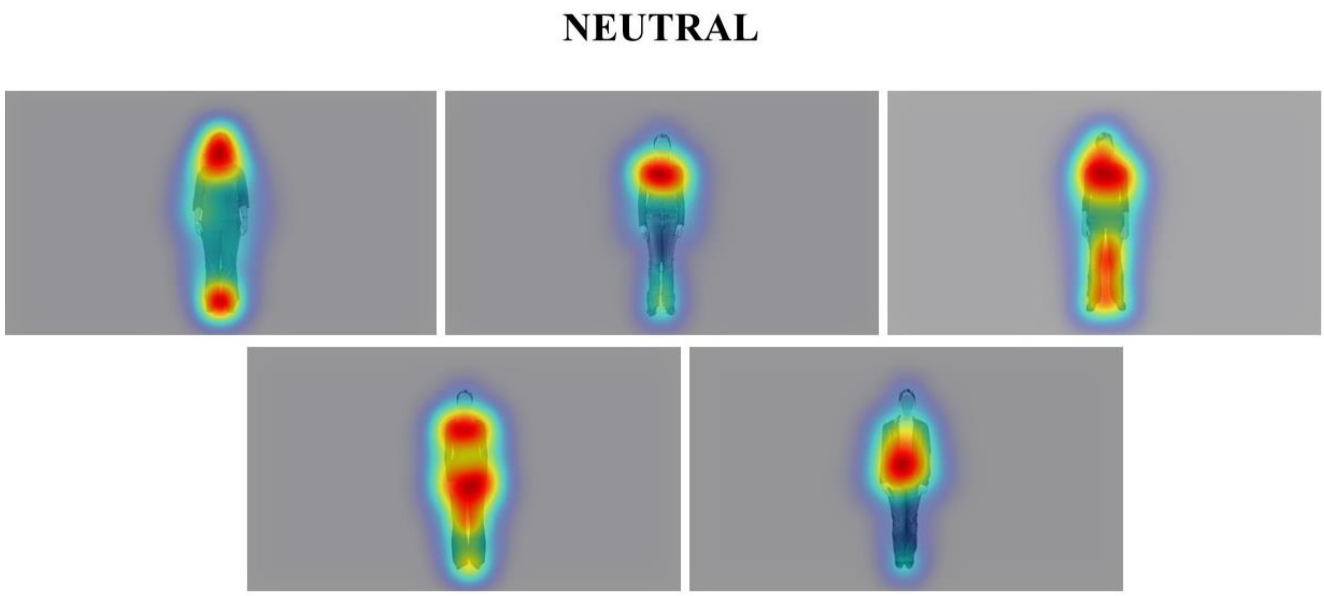
HAPPY



Supplementary Figure 4. The position of valid fixations across all infants for the five happy body images. Red: fixations assigned to the Head ROI; Blue: fixations assigned to the Lower Body ROI; Green: fixations assigned to the Upper Body ROI; Yellow: valid fixations that were not assigned to any of the body AOIs.

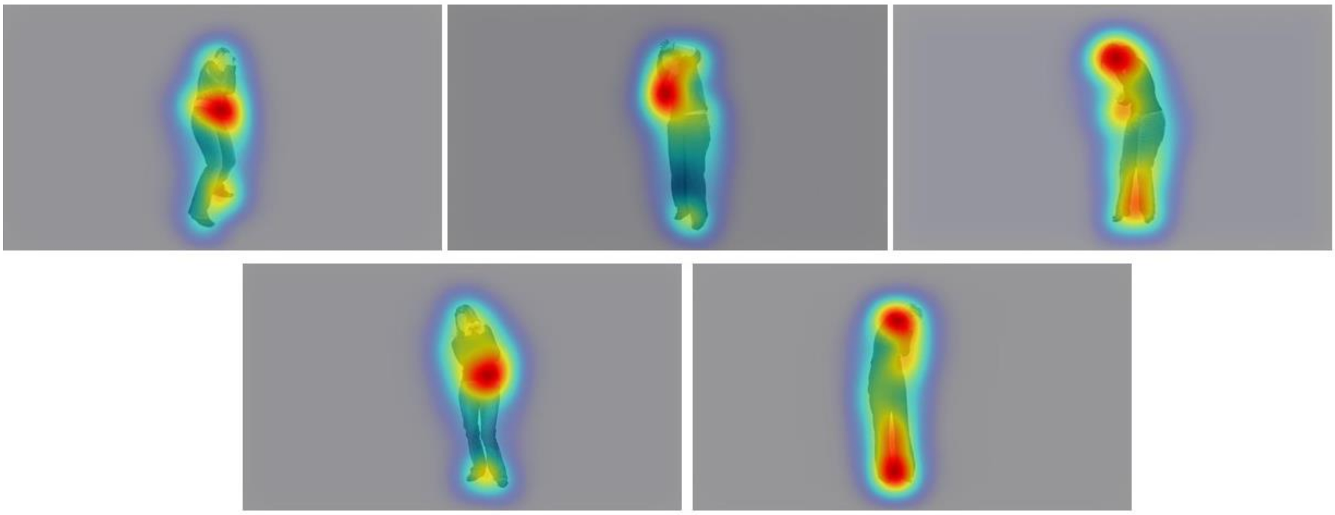
Visual Saliency Maps

The visual saliency of different body parts in an image can be calculated based on luminance and orientation of edges in the image. The resulting visual saliency map is a two-dimensional map, with the value at each image pixel representing how perceptually conspicuous the corresponding region is in the image (Veale et al., 2017). Based on the model proposed by Itti and colleagues (Itti, Koch, & Niebur, 1998; Itti & Koch, 2001), we calculated the visual saliency map for each body image (*Supplementary Figures 5 to 8*).



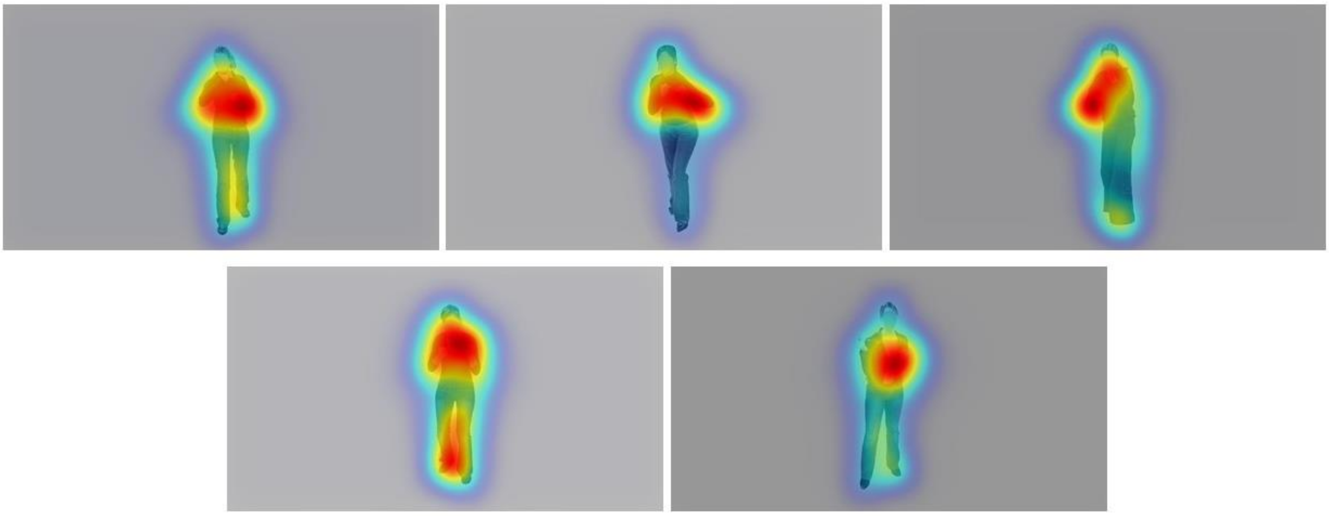
Supplementary Figure 5. Visual saliency maps for the five emotionally neutral body images. Warm colours represent image regions of high saliency (closer to 1) where as cool colours represent regions of low saliency (closer to 0 = no colour).

FEAR



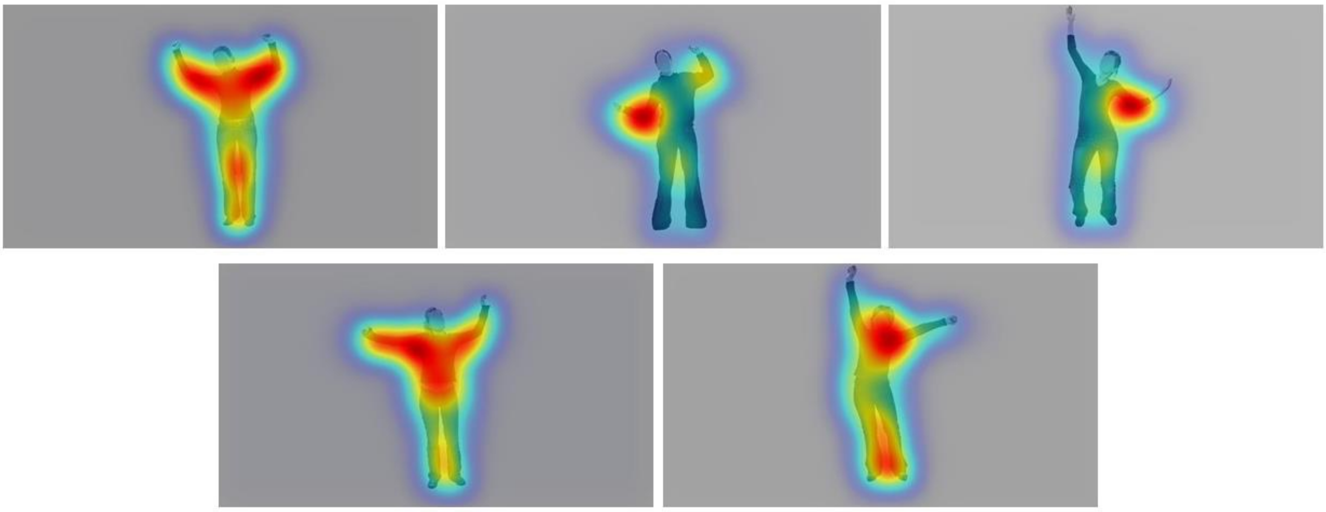
Supplementary Figure 6. Visual saliency maps for the five fear body images. Warm colours represent image regions of high saliency (closer to 1) where as cool colours represent regions of low saliency (closer to 0 = no colour).

ANGER



Supplementary Figure 7. Visual saliency maps for the five angry body images. Warm colours represent image regions of high saliency (closer to 1) where as cool colours represent regions of low saliency (closer to 0 = no colour).

HAPPY



Supplementary Figure 8. Visual saliency maps for the five happy body images. Warm colours represent image regions of high saliency (closer to 1) where as cool colours represent regions of low saliency (closer to 0 = no colour).

Supplementary Table 1. Correlation between eye-movement parameters and visual saliency as a function of AOI. For each of the 20 body images (5 images per expression), we calculated the average proportion fixation, proportion looking times and fixation duration across valid fixations pooled across infants. The correlation between each parameter and the visual saliency averaged across all pixels within each AOI was then computed (i.e., there were 20 pairs for each parameter and each AOI). There were no significant correlations (all *p*s > .19). The U:L is the visual saliency ratio of the upper to lower AOI.

	Prop Fixation	**Prop Looking Time**	**Fixation Duration (ms)**
Body	-.19	-.26	.00
Head	-.22	-.06	.31
Upper	.08	-.09	.02
Lower	.18	.03	.01
U:L	.27	-.05	-.10