Infants’ sensitivity to the congruence of others’ emotions and actions

Robert Hepach,⇑, Gert Westermann

⇑Corresponding author. 
E-mail address: hepach@eva.mpg.de (R. Hepach).

Introduction

Emotions form an integral part of human interactions. Our ability to detect the mood and emotion of the person with whom we are interacting allows us to swiftly adjust our behavior (e.g., the topics we raise, the requests we make). As children, we quickly learn that our request for a new toy has
greater chances of approval if our mother is smiling and humming a happy tune. Although such an example may appear to be trivial, the underlying ability is sophisticated, namely, using a person’s emotional display (e.g., facial expression, vocalizing) to interpret her actions. The mechanisms contributing to infants’ emerging understanding of social phenomena such as others’ emotions are being extensively studied in the wider context of the developing “social brain” (Grossmann & Johnson, 2007).

Infants can discriminate a variety of different emotional facial expressions. During their first year of life, infants discriminate surprise and happiness at the age of 3 months (Young-Browne, Rosenfeld, & Horowitz, 1977), anger and happiness at 4 months (Barrera & Maurer, 1981), and fear and happiness at 7 months (Kotsoni, de Haan, & Johnson, 2001; Ludemann & Nelson, 1988). However, such discriminatory abilities do not necessarily indicate that infants understand the communicative function of emotional facial expressions in the same way as adults do (de Haan & Groen, 2006). More evidence along the lines of what infants understand of others’ emotions comes from studies on cross-modal matching, which show that 7-month-olds become sensitive to the congruence of an emotional facial expression and the emotional tone of a voice (Soken & Pick, 1992). It has been argued that this ability indicates the development of an invariant representation of emotions that is independent of specific modalities (Walker-Andrews, 1997).

More sophisticated abilities in emotion processing have been addressed in social referencing studies that investigate infants’ abilities to use others’ emotions to gain information about objects or events (e.g., Klinnert, Campos, Sorce, Emde, & Svejda, 1983; Walden & Ogan, 1988). In a typical setup, an infant faces an adult with an object placed between them. The adult then emotes either positively or negatively toward the object. After 10 months of age, infants will avoid or approach objects toward which the adult emoted either negatively or positively, respectively (Saarni, Mumme, & Campos, 1998). However, it is unclear whether such results depend on the infant’s understanding of the adult’s emotion or whether they are caused by emotional contagion by which the infant merely adopts the emotion of the adult in response to an object.

Stronger evidence for an understanding of emotions comes from a study that investigated whether infants can use others’ emotions to infer desires even when they are in conflict with their own (Repacholi & Gopnik, 1997). In this study, 14- and 18-month-olds witnessed an adult expressing a food preference that either matched or mismatched the infant’s own preference. The adult expressed a positive emotion (pleasure) when tasting one food and expressed a negative emotion (disgust) when tasting the other food. In one condition the adult’s preference matched the (previously established) preference of the infant, and in the other condition it was the opposite of the infant’s preference. Finally, the adult would reach out her hand and request “some more.” The results showed that 18-month-olds were more likely to hand the adult the food item to which the adult had emoted positively, independent of their own preference. In contrast, 14-month-olds were more likely to give the adult the food they themselves preferred in both conditions.

To address infants’ understanding of the link between others’ emotional displays and actions on a more abstract level, Barna and Legerstee (2005) presented 9-month-olds with either a happy actor or an unhappy actor looking at an ambiguous object and saying that she either liked or did not like the object, respectively. Infants were habituated to a stimulus where one actor was holding the object in her hand, but her face (indicating whether she was the happy actor or the angry actor) was occluded. Subsequently, the emotion of the actor was revealed again, and infants looked longer in the case of the unhappy actor holding the object compared with the happy actor holding it. Similarly, Phillips, Wellman, and Spelke (2002) examined whether 12- and 14-month-olds form assumptions about an actor’s subsequent behavior based on the actor’s gaze and facial emotional expression toward an object. In their study, an experimenter looked at and emoted positively toward one of two objects placed in front of her. In the test trial, either the actor was shown holding the object toward which she had emoted positively or she was shown holding the other object. The authors found that 14-month-olds looked longer at test trials in which the actor held the other object. However, the results of this study did not rule out the possibility that infants used the actor’s attention rather than the displayed emotion to form their expectations because the actor emoted to only one object and ignored the other.
To address this concern through separating the effects of emotion and attention, Vaish and Woodward (2010) presented 14-month-olds with an actor who looked into one of two cups while showing either a happy or disgusted emotional expression. During a subsequent test phase, infants showed greater novelty preference when the actor reached for the unattended cup regardless of the actor’s previous emotional display. These results suggest that in the study by Phillips and colleagues (2002), infants’ looking behavior was guided not by the actor’s positive emotion toward an object but rather by the directing of attention to one object but not the other, leaving open the question of when infants begin to be able to link emotional displays with expectations for actions.

In sum, all previous studies have focused on infants’ sensitivity to emotions as referents to others’ desires and preferences related to objects toward which an emotion was displayed. It remains unclear when infants become able to link emotions and actions that are congruent with these emotions on a more abstract conceptual level, when they are not interacting with the person displaying an emotion, and when no cause of the emotion is provided and actions and emotions are presented as being independent of one another. Here we address this question.

In the current study, we showed 10- and 14-month-old infants videos of two actors: one introducing himself or herself to the infants with a happy (positive) emotional facial expression and vocal tone and the other with an angry (negative) emotional display (e.g., Barrera & Maurer, 1981). Unlike in social referencing studies, the causes of their respective emotions were not revealed. Subsequently, the actors performed two actions toward a toy tiger that was not present during the introductory phase. One action consisted of the actor patting and gently stroking the toy, an approach behavior that is generally associated with a positive emotion (e.g., Saarni et al., 1998). The second action showed the actor thumping the toy with a closed fist, a behavior with a negative connotation similar to pushing an object down after it followed an upward trajectory (Kuhlmeier, Wynn, & Bloom, 2003). Crucially, an action was therefore either congruent or incongruent with the actor’s displayed emotion. Our study used a violation of expectation paradigm to assess infants’ responses to congruent and incongruent emotion–action pairings. We used measures of pupil diameter as an indicator of changes in sympathetic activity in response to unexpected events.

In a world that is uncertain but predictable, humans need to constantly update prior held beliefs and knowledge with incoming information. Individuals need to constantly predict, and thus form expectations of their social and nonsocial surroundings. Unexpected outcomes or events will trigger an orienting response through which the individual will be in a state of heightened attention to allow for efficient information processing (Sokolov, 1963). One physiological component of this orienting toward novel and unexpected information is systematic changes in pupil dilation (Nassar et al., 2012; Preuschoff, Hart, & Einhäuser, 2011). Pupillary changes indicate autonomic reactivity through heightened sympathetic activity that puts an individual in a state of heightened alertness and attention. Therefore, unexpected events can trigger sympathetic arousal that can be measured in changes of pupil dilation. Although pupil dilation has long been used to measure heightened sympathetic arousal in adults, more recently it has begun to be employed in infancy research as well. Results have indicated that infants show increased pupil dilation in response to unexpected social and physical events (Gredebäck & Melinder, 2010; Jackson & Sirois, 2009; see Laeng, Sirois, & Gredebäck, 2012, for a recent review) and that this measure can be more sensitive than the traditionally used looking time (Jackson & Sirois, 2009).

In the current study, we hypothesized that infants would show increased pupil dilation for trials in which an actor was shown performing an action incongruent with his or her emotional display.

Method

Participants

The participants were 10- and 14-month-old, full-term healthy infants recruited from a database and invited to the department’s research facility. A total of 57 children participated in the study, with 30 (16 male and 14 female) participants in the younger age group (median age = 10 months 3 days, range = 9 months 5 days to 11 months 10 days) and 27 (15 female and 12 male) participants in the
older age group (median age = 14 months 10 days, range = 13 months 14 days to 15 months 2 days). An additional 8 participants were excluded because of fussiness.

Materials and setup

Stimuli consisted of movie clips showing a female or male actor. All clips were recorded with actors showing either a neutral, happy, or angry emotional expression. The types of clips used were as follows. An introductory clip showed the actor waving and saying “Hello baby, I’m [actor’s name] and I’m happy/angry” with the appropriate emotional expression and tone (in the neutral expression condition, the actor only mentioned his or her name). Two action clips showed the actor performing either a positive action (patting a toy animal) or a negative action (thumping the toy animal with a closed fist) with either a neutral, happy, or angry facial expression (see Fig. 1 for illustrations). Crucially, the toy was not present during the introductory clips so as not to be linked to the actors’ emotions. The duration of each action clip was kept short at 4 s because the measure of pupil dilation relies on participants looking at the stimulus for as long as possible. On the other hand, the measure of looking time relies on participants looking away. In this sense, designs including pupil dilation or looking time as the main dependent measure are diametrically opposed.

![Fig. 1. Examples of the stimuli sequence used in the study. The first column presents still frames from the introductory clips during which actors showed their respective emotions and waved to the children (see text for details). The three remaining columns present still frames from the actual action clips to illustrate the patting action (top and bottom rows) and the thumping action (middle two rows). The top two rows show the female actor displaying a happy facial expression performing the congruent patting action (top row) and the incongruent thumping action (second row from top). The bottom two rows show the male actor displaying the angry emotion performing the congruent thumping action (second row from bottom) and the incongruent patting action (bottom row). In the actual study, the videos were presented in color. The proportions of the above images are different from those of the videos used in the study. For illustration purposes, close-ups are provided in this figure.](image-url)
A trial consisted of the following sequence: introductory clip → first action (either patting or thumping) → second action (patting or thumping, different from the first action). Before each trial, an attention-grabber movie was shown to keep participants engaged (e.g., Frank, Vul, & Johnson, 2009).

The stimuli (720 × 576 pixels) were presented centrally on a 22-inch computer screen using Tobii Studio. Eye movements and the pupil size of both eyes were recorded simultaneously with a Tobii X120 eye-tracker positioned below the screen.

Design and procedure

During the experiment, each child sat on the parent’s lap approximately 65 cm from the computer screen. Parents were instructed not to interact with their children during the experiment. The light in the experimental room was dimmed and kept constant for all participants. From where participants sat in the room, the luminosity was measured at 41 lux when facing the illuminated screen and 15 lux when facing the surrounding walls in the room. The actual procedure took approximately 5.5 min and consisted of a familiarization block, which was always shown first, and an experimental block. The purpose of the familiarization block was to introduce participants to the actors and actions and to familiarize them with the order of stimuli. The familiarization block consisted of two trials: one with the male actor and one with the female actor (counterbalanced) displaying a neutral facial expression and performing both actions in a counterbalanced order (e.g., if the first actor performed the patting action before the thumping action, then the second actor would perform the two actions in the reverse order). The experimental block consisted of two trials, with one actor displaying one emotion and the other actor displaying the other emotion while each performed the actions in the same order as he or she had done during the familiarization block. Each participant saw both actors perform an incongruent action and a congruent action. Congruence was defined as a happy actor patting the toy and an angry actor thumping the toy. Therefore, each participant was shown four test trials in total. Half of the participants saw the congruent action first. Each actor displayed the happy and angry emotions equally often. Finally, the order of appearance of the actors was also counterbalanced. This led to 2 (Emotion) × 2 (Order of Actor Appearance) × 2 (Order of Action Appearance) = 8 different experimental blocks. Participants were randomly assigned to the experimental blocks.

Manipulation check

A manipulation check was conducted to ensure that the video clips actually depicted a negative emotion and a positive emotion (angry and happy) as well as a negative action and a positive action (thumping and patting). Two coders, blind to conditions and hypotheses, rated all of the movie clips with regard to actors’ emotions and actions. The videos were edited such that the action was obscured for the rating of emotions and the actor’s face was obscured for the rating of actions. The videos were played without any sound. For each video clip, the raters were asked to rate the valence of the actor’s emotion or action on a 9-point scale using the self-assessment manikin (SAM; Bradley & Lang, 1994). Values of 1 and 9 referred to very positive and very negative, respectively.

For rating the emotions, reliability for the raters’ coding was high (intraclass correlation coefficient [ICC] = .95, p < .001). The videos showing the actors with a happy emotion were rated as positive (M = 3.17, Min = 1, Max = 4), and the videos showing the actors being angry were rated as negative (M = 7.5, Min = 7, Max = 8). This suggested that the videos portrayed the hypothesized valence with respect to the emotion.

For the rating of the actions, reliability for the raters’ coding was also high (ICC = .92, p < .001). The videos showing the actors performing the patting action were rated as positive (M = 3.34, Min = 3, Max = 4), and the videos showing the actors performing the thumping action were rated as negative (M = 7, Min = 6, Max = 9). This suggested that the action clips portrayed the hypothesized valence as well.
Preliminary analysis of luminance

To ensure that there were no substantial differences in the luminance of the test stimuli, we computed the average luminance for each frame (100 frames in total) and each video (8 videos in total) employing the following formula (as suggested by Jackson & Sirois, 2009):

$$ Y = 0.2126 \times R + 0.7152 \times G + 0.0722 \times B. $$

More specifically, for each image, we obtained the average value of the red channel ($R$), the green channel ($G$), and the blue channel ($B$). For each frame of the videos, the values were entered in the above formula to obtain the average luminance per image. The calculations were carried out using Matlab (Version 7.4, R2007a). For the action clips showing the female actor, there were only minor differences in luminance among the happy emotion–patting action ($M = .58, SD = .0026$), the happy emotion–thumping action ($M = .59, SD = .0028$), and the angry emotion–patting action ($M = .59, SD = .0028$). Similarly, the differences for the videos portraying the male actor were also minor: happy emotion–patting action ($M = .60, SD = .0006$), happy emotion–thumping action ($M = .60, SD = .0015$), angry emotion–patting action ($M = .60, SD = .0008$), and angry emotion–thumping action ($M = .60, SD = .001$). Crucially, one would not conclude from looking at the luminance values that the incongruent emotion action pairings have systematically lower luminance values. Any systematic increase in pupil dilation, therefore, would not be due to the physical properties of the stimuli but rather would be due to an effect of their conceptual content.

It is important to note that the videos showing the female actor had systemically lower luminance values than those showing the male actor. This makes sense given that the female actor wore dark clothes, whereas the male actor wore bright clothes (see also Fig. 1). Therefore, we used the luminance values for each individual clip as a control variable in the statistical analyses (see below).

Pupil diameter analyses

The sampling rate for the majority of participants was 60 Hz. A number of participants were recorded with a higher sampling rate of 120 Hz, in which case the data were down-sampled to match recordings of 60 Hz. The raw data files were exported from Tobii Studio and analyzed in Matlab (Version 7.4, R2007a). Data were recorded for both the left and right eyes, estimating the pupil diameter in millimeters. In case a participant was not looking at the screen, the values for the left and right eyes were replaced with a “NaN” (not a number) value. Initially, the data for an entire recording were filtered by excluding those absolute sample-to-sample differences that exceeded the 90th percentile of all differences. Subsequently, missing samples were linearly interpolated only if the gap between two samples was not greater than four samples. With regard to the sampling frequency of 60 Hz, this was approximately 70 ms. The data for both eyes were then averaged, filtered, and interpolated again to smooth out extreme values resulting from averaging across both eyes (see Hepach, Vaish, & Tomasello, 2012, for a similar analysis). The analyses were carried out on the action clips when the actors were performing their respective actions. The duration of each clip was 4 s and consisted of a pre-period during which the actor was shown sitting at a table looking at a toy tiger. This period ended when the actor started moving his or her hand toward the toy demonstrating either a patting or thumping action. This marked the point of action differentiation, which varied in time between clips ($M = 883$ ms, $SD = 267$). A second coder blind to the hypotheses of the study coded the time point of action differentiation. The correlation between the primary and secondary coders’ coding was high (Spearman’s rho = .99, $p < .001$). With reference to each clip’s individual time point of action differentiation, all succeeding individual values were baseline-corrected by subtracting the mean of all preceding values. Finally, the mean of all baseline-corrected values was computed for the statistical analyses.
Statistical analyses

The data were analyzed by fitting generalized linear mixed models (GLMMs; Baayen, 2011) with Gaussian error distribution and identity link function using the function `lmer` (lme4 package; Bates, Maechler, & Bolker, 2011) in R (Version 2.15.1; R Core Team, 2012). GLMMs have recently been applied in developmental research to investigate age effects using more complex designs (Rossano, Rakoczy, & Tomasello, 2011). The overall influence of a fixed effect (independent variable) was investigated by comparing a full model including the effect with a reduced model excluding the effect. Model comparisons of full and reduced models were carried out with a likelihood ratio test (Dobson, 2002) using the function `anova` with a “Chisq” test argument in R. Therefore, a significant difference between models would indicate that the independent variable under investigation had an overall effect on the dependent measure (in this case pupil dilation). Furthermore, to investigate the specific influence of the independent variable of interests, p values were computed with the function `pvals.fnc` from the R package language, which is based on Markov chain Monte Carlo sampling (Baayen, 2011).

Pairwise comparisons were carried out using post hoc, paired-sample Wilcoxon exact tests. The pairwise analyses were within-participant for each actor. Therefore, differences in luminance did not affect the pairwise results. In the following, details regarding the individual models later referred to in Results are provided. If a fixed effect did not have an overall influence, it was removed from all subsequent model comparisons. The dependent measure was always participants’ average baseline-corrected pupil diameter. Subject was always entered as a random effect. The fixed effect of luminance was transformed (to arrive at a uniform distribution) and z-standardized. The transformation assigned each luminance value its respective position in the list of sorted values (i.e., given that there were 8 luminance values [for each action clip], the lowest value of .58 would be assigned a 1, etc.).

Analysis of overall effects of luminance

First, two models were fitted to investigate whether the luminance of the video clips had an overall effect on participants’ pupil dilation. The full model included the fixed effects of participants’ gender, block order (test trial: 1 or 2), action order (order within trial: 1 or 2), and luminance (see “Preliminary analysis of luminance” section above) as well as the hypothesized three-way interaction of emotion (happy vs. angry), action (patting vs. thumping), and age (14 vs. 10 months). The reduced model included all of the above effects except luminance.

Second, two models were fitted to assess whether the luminance of the video clips had an overall effect on the hypothesized three-way interaction of emotion, action, and age. The full model comprised the fixed effects of participants’ gender, block order, and action order as well as the four-way interaction of emotion, action, age, and luminance. The reduced model included all of the above effects except luminance.

Analysis of overall order effects

First, two models were fitted to investigate whether there was an effect of block order or action order on participants’ pupil dilation. The full model included gender, block order, action order, and luminance as well as the hypothesized interaction of emotion, action, and age as fixed effects. Two reduced models were fitted excluding block order and action order, respectively.

Second, it was investigated whether block order had an overall effect on the hypothesized three-way interaction of emotion, action, and age. The full model comprised the fixed effects of participants’ gender, luminance, and action order as well as the four-way interaction of emotion, action, age, and block order. The reduced model included all of the above effects with the exception that the four-way interaction was replaced with all possible three-way interactions of age, emotion, action, and luminance.

Analysis of overall order effects

First, two models were fitted to investigate whether there was an effect of block order or action order on participants’ pupil dilation. The full model included gender, block order, action order, and luminance as well as the hypothesized interaction of emotion, action, and age as fixed effects. Two reduced models were fitted excluding block order and action order, respectively.

Second, it was investigated whether block order had an overall effect on the hypothesized three-way interaction of emotion, action, and age. The full model comprised the fixed effects of participants’ gender, luminance, and action order as well as the four-way interaction of emotion, action, age, and block order. The reduced model included all of the above effects with the exception that the four-way interaction was replaced with all possible three-way interactions of age, emotion, action, and block order. No separate analysis was carried for action order because there was no overall effect of this variable (see Results). Therefore, action order was dropped from all subsequent analyses. Furthermore, we investigated a potential interaction of action order and block order by fitting two additional models. The full model included the fixed effects of participants’ gender, luminance, and
the three-way interaction of age, emotion, and action as well as the interaction of action order and block order. The reduced model was similar to the full model with the important difference that action order and block order were entered as individual main effects.

Analysis of an overall effect of age

Two models were fitted to investigate whether there was an overall effect of age on the interaction of emotion and action. In other words, these models investigated whether the interaction of emotion and action varied between the two age groups. The full model included participants’ gender, block order, and luminance as well as the hypothesized three-way interaction of emotion, action, and age. The reduced model included all of the above effects except that the three-way interaction was removed and all possible two-way interactions of emotion, action, and age were included.

Separate analyses for 10- and 14-month-olds

The following steps were carried out for each age group individually. First, two models were fitted to investigate whether there was a hypothesized interaction of emotion and action. The full model included participants’ gender, block order, and luminance as well as the interaction of emotion and action. The reduced model included all of the above except that the interaction was removed and emotion and action were entered as two separate main effects.

Second, it was investigated whether luminance had on overall effect on the hypothesized interaction of emotion and action. A full model included as fixed effects participants’ gender and block order as well the as three-way interaction of emotion, action, and luminance. The reduced model was similar with the exception that the three-way interaction was removed and all possible two-way interactions of emotion, action, and luminance were included.

Third, to investigate whether participants’ gender had an overall effect on the interaction of emotion and action, a full model was fitted including luminance and block order as fixed effects as well as the three-way interaction of emotion, action, and gender. The reduced model was similar with the only exception that the three-way interaction was replaced by all possible two-way interactions of emotion, action, and gender.

Fourth, to investigate whether block order interacted with emotion and action, a full model was fitted including luminance and gender as fixed effects as well as the three-way interaction of emotion, action, and block order. The reduced model was similar with the only exception that the three-way interaction was replaced by all possible two-way interactions of emotion, action, and block order.

Preliminary analysis of looking time

We fitted two GLMMs to investigate whether the hypothesized three-way interaction for pupil dilation also affected children’s overall looking behavior, that is, whether the interaction of emotion and action led to looking times that varied between the two age groups. For these models, the dependent measure was children’s looking time. More specifically, we divided the amount of time children looked at the video after the action differentiated by the remaining time of the video (from action differentiation to the end). Therefore, the dependent measure was a proportion score, with 1 indicating that children looked at the video for its entire time after the action took the form of either patting or thumping. The full model included participants’ gender, block order, and luminance as well as the hypothesized three-way interaction of emotion, action, and age. The reduced model included all of the above effects except that the three-way interaction was removed and all possible two-way interactions of emotion, action, and age were included.

In addition, we fitted two further models to investigate whether the interaction of emotion and action had an effect on children’s looking time. The full model included age, participants’ gender, block order, and luminance as well as the interaction of emotion and action. The reduced model was similar to the full model with the difference that emotion and action were entered as individual fixed effects.
Results

Analysis of looking time

There was no overall effect of age on the interaction of emotion and action, $LR(1) = 1.75, p = .19$. That is, the time children spent looking at the stimuli did not systematically vary with the hypothesized overall effect of age on the congruence of emotion and action. Furthermore, there was no interaction effect of emotion and action on children's looking behavior, $LR(1) = 0.23, p = .63$. Therefore, participants' looking time did not significantly vary with the congruence or incongruence of the emotion–action pairings.

Analysis of overall luminance effects on pupil dilation

The luminance of the individual video clips had a statistically significant overall influence on participants' pupil dilation, $LR(1) = 61.83, p < .001$. Specifically, as luminance increased, pupil dilation decreased ($-0.16 \pm 0.019, t = -8.66, p < .001$). This is the common relationship given that the pupil constricts with increased levels of luminance (Loewenfeld, 1993). Crucially, there was no overall effect of luminance on the hypothesized interaction of age, emotion, and action, $LR(1) = 0.024, p = .88$.

Analysis of overall order effects

There was an overall effect of block order on participants' pupil dilation, $LR(1) = 7.51, p = .0061$, but not of action order on participants' pupil dilation, $LR(1) = 3.0026, p = .083$. More specifically, pupil dilation increased from the first test trial to the second ($0.095 \pm 0.034, t = 2.78, p = .0208$). However, there was no overall effect of block order on the hypothesized interaction of age, emotion, and action, $LR(1) = 0.087, p = .77$. Furthermore, there was no overall effect of the interaction of action order and block order on children's pupil dilation, $LR(1) = 0.39, p = .53$.

Fig. 2. Box plots illustrating the baseline-corrected changes in pupil diameter. The scale of measurement is the absolute value of pupil diameter following the target action minus the average of all absolute values preceding the action (i.e., baseline-corrected). The boxes represent the distribution of values between the first and fourth quartiles. The solid lines inside the boxes indicate the group medians. The dashed lines with the whiskers at their respective end points represent the locations of extreme values. Outliers that exceed the interquartile distance by more than 1.5 are represented by dots and were not removed from the statistical analysis. Significance: * $p < .05$; ** $p < .001$. Graphics were created using the ggplot2-package in R (Wickham, 2009).
Analysis of an overall effect of age

As hypothesized, there was an overall effect of age on the interaction of emotion and action, $LR(1) = 4.107$, $p = .043$. That is, the interaction of emotion and action varied between age groups. Therefore, in a subsequent step, individual analyses were carried out separately for the age groups of 10 and 14 months. The results for both age groups are illustrated in Figs. 2 and 3.

Analysis of 10-month-olds

In accordance with our hypotheses, there was an overall interaction of emotion and action on children’s pupil dilation, $LR(1) = 19.6$, $p < .001$. Furthermore, there was a main effect for action ($-.42 \pm .079$, $t = -5.32$, $p < .001$) and for emotion ($-.53 \pm .072$, $t = -7.30$, $p < .001$). Infants showed more pupil dilation when an actor performed the patting action ($M = -.21$, $SD = .45$) compared with the thumping action ($M = -.29$, $SD = .46$). In addition, infants showed more pupil dilation when an actor displayed the angry emotion ($M = -.12$, $SD = .41$) compared with the happy emotion ($M = -.37$, $SD = .46$). Furthermore, there was a significant overall effect of luminance ($-.19 \pm .027$, $t = -6.87$, $p < .001$) but neither an overall effect of gender nor an overall effect of block order (all $p$s > .14). In addition, none of the three fixed effects of luminance, gender, and block order interacted with the hypothesized interaction of emotion and action (all $p$s > .11). Pairwise comparisons revealed that there was no statistically significant difference in 10-month-olds’ pupil dilation when a happy actor performed the incongruent thumping action ($M = -.36$, $SD = .47$) compared with the congruent patting action ($M = -.39$, $SD = .45$), $T = 231$, $p = .54$. However, 10-month-olds showed significantly more pupil dilation when an angry actor performed the incongruent patting action ($M = -.016$, $SD = .35$) compared with the congruent thumping action ($M = -.22$, $SD = .46$), $T = 220$, $p = .046$.

Analysis of 14-month-olds

Similar to 10-month-olds, there was an overall interaction of emotion and action on children’s pupil dilation, $LR(1) = 36.21$, $p < .001$. Furthermore, there was a main effect for action ($-.55 \pm .075$, $t = -7.30$, $p < .001$) and for emotion ($-.42 \pm .072$, $t = -5.85$, $p < .001$). Infants showed more pupil dilation when an actor performed the patting action ($M = -.04$, $SD = .36$) compared with the thumping action ($M = -.15$, $SD = .39$). In addition, infants showed more pupil dilation when an actor displayed the angry emotion ($M = -.086$, $SD = .35$) compared with the happy emotion ($M = -.11$, $SD = .40$).

Fig. 3. Time course of participants’ pupil dilation following action differentiation while watching the action clips. The y axis represents the baseline-corrected change in pupil diameter.
addition, there was a significant overall effect of luminance (\(-.14 \pm .027, t = -5.17, p < .001\)) but neither an overall effect of gender nor an overall effect of block order (\(ps > .09\)). Furthermore, none of the three fixed effects of luminance, gender, and block order interacted with the hypothesized interaction of emotion and action (all \(ps > .11\)). The 14-month-old infants showed more pupil dilation when a happy actor performed the incongruent thumping action \((M = -.043, SD = .40)\) compared with the congruent patting action \((M = -.17, SD = .41)\), \(T = 236, p = .048\), and they showed more pupil dilation when an angry actor performed the incongruent patting action \((M = .091, SD = .23)\) compared with the congruent thumping action \((M = -.27, SD = .36)\), \(T = 232, p < .001\).

**Discussion**

The aim of the current study was to investigate the emergence of infants’ sensitivity to the congruence of other people’s emotional facial expressions and actions. When presenting two age groups with actors performing an action either congruent or incongruent with their emotional display, we found that both 10- and 14-month-olds showed increased pupil dilation when presented with an angry actor performing a positive patting action compared with the same actor performing a negative thumping action. Furthermore, only 14-month-olds also showed increased pupil dilation when a happy actor performed the incongruent thumping action compared with the congruent patting action. Therefore, at 10 months of age, infants are sensitive to the congruence of others’ emotions and actions when actors display an angry emotion. By 14 months, this ability has expanded to include sensitivity to the incongruence of others’ emotions and actions when actors display a happy facial expression. Because, unlike in previous studies (e.g., Repacholi & Gopnik, 1997), infants here were not presented with any information about the cause of the actors’ emotions (the toy tiger was not present during the introductory clips), the only source of information the 14-month-olds could have used was their prior knowledge of positive and negative emotional displays.

Our interpretation of the pupil dilation is that infants show increased levels of sympathetic arousal because viewing incongruent emotion–action pairings runs counter to their expectation about how people with positive and negative emotions behave. In adults, the pupils will systematically dilate when participants are engaged in cognitively demanding tasks (Beatty & Lucero-Wagoner, 2000; Kahneman, 1973) and when presented with emotionally arousing stimuli (Bradley, Miccoli, Escrig, & Lang, 2008). Both the more cognitive and more emotional aspects of induced pupillary changes are part of the same autonomic response system triggered by novel outcomes of events and actions that are both cognitively and affectively engaging.

The question of what infants actually understand of other people’s emotional displays, as well as the question of when this ability emerges, has stimulated much research (see Thompson & Lagatutta, 2008, and Widen & Russell, 2008, for recent overviews). There are different interpretations of what it means to understand emotions, and indeed this term has not been well specified in the literature and is often used merely to refer to the processing of emotions in others as opposed to infants’ own production of emotions (Widen & Russell, 2008). Previous studies have found that with the onset of the second year of life, infants are able to regulate their own emotions on the basis of emotional displays in others (Klinnert et al., 1983; Repacholi & Gopnik, 1997) and that both 15- and 18-month-olds use emotive cues to evaluate others’ actions (Repacholi, 2009) when the cause of the emotion was made obvious to them (e.g., if the actor’s emotion was introduced as being associated with a particular object). However, although social referencing studies show that infants can modify their own behavior on the basis of emotional displays in others, it is not clear whether this ability depends on an understanding of these emotions or whether infants merely adopt the same emotion toward the object or event as the adult through a process of emotional contagion (Widen & Russell, 2008). Therefore, it is difficult to argue that infants’ success in social referencing studies implies an understanding of the meaning of emotions. Arguably, one essential component of an understanding of emotions emerges when infants can use another person’s emotional state to form expectations about that person’s behavior. The results of the current study speak to this issue by using a design that controlled for the possibility that infants’ behavior could be influenced by emotional contagion toward objects; in our study, the actors’ emotion was not caused by a specific object. This study also avoided the
confound between emotion and attention toward an object (see also Vaish & Woodward, 2010) because the actor attended to the object equally in all situations. Our results suggest that infants’ more sophisticated understanding of emotions might be more flexible (i.e., independent of the presence of the cause of an emotion) and emerge earlier than previously assumed, namely between 10 and 14 months of age.

It is possible that infants’ own changes in sympathetic arousal were caused by a process of affect sharing through which they experienced the actors’ emotion as part of their own. This would align with results by Geangu, Hauf, Bhardwaj, and Bentz (2011), who demonstrated that infants’ own pupillary responses change as a function of viewing emotional stimuli. This sensitivity to others’ emotional displays and behaviors might serve as a precursor for the sorts of abilities infants and toddlers demonstrate in social referencing studies later in development.

A point worth discussing refers to the question of what the actions of patting and thumping actually mean to infants at 10 and 14 months of age and to what extent they are prototypically associated with positive and negative emotions, respectively. The social referencing literature suggests that infants view a positive happy emotion to signal approach behavior (e.g., a person who emotes positively toward an object is expected to subsequently reach for it; Repacholi & Gopnik, 1997). A negative emotion generally signals avoidance in social referencing studies (e.g., Klinnert et al., 1983). Although those studies used negative emotions such as fear (Mumme, Fernald, & Herrera, 1996) or disgust (Repacholi & Gopnik, 1997), we chose anger because the action we associated with it, thumping with a closed fist, includes similar movements (e.g., extended arm) as the positive patting action. We aimed to keep the perceptual difference between the two actions as small as possible to reduce the effects of low-level movement on infants’ pupillary responses. As for the positive emotion, it is interesting that for 10-month-olds there was no difference between the situations when an actor was shown thumping and when the actor was shown patting. This could reflect the fact that for infants at this age, the thumping action is not negative per se because it might be involved in playful behavior as well. However, by 14 months of age, infants are sensitive to the incongruence of the thumping action in the context of the positive emotion because they showed increased pupil dilation when the positive actor was thumping the toy. It could be that 14-month-olds’ sensitivity to the incongruence of both the happy–thumping actions and the angry–patting action pairing reflects the more adult-like perception of the video stimuli (given the results from the manipulation check). Furthermore, the actions of patting and thumping are only examples of approach and avoidance behaviors, and future studies could assess infants’ understanding of others’ actions and emotions using varying approaches and avoidance behaviors with different sets of emotions.

Furthermore, the fact that both 10- and 14-month-olds are sensitive to the incongruence when they view the angry actor, but that only older infants are also sensitive to the congruence when viewing the happy actor, is interesting in its own right in the light of a negativity bias in children’s emotional development (Vaish, Grossmann, & Woodward, 2008). This bias may serve the important function of directing infants’ attention to negative stimuli in the environment that are prone to hold valuable information as to how to adjust one’s own behavior to learn to avoid negative outcomes. It is possible that such heightened attention to negative stimuli includes increases in sympathetic arousal similar to our results that infants showed more pupil dilation when watching the actors display an angry emotion (see also Geangu et al., 2011). Later in development, a similar negativity bias toward negative information is present when preschoolers learn the display rules for negative emotions earlier than those for positive emotions (Banerjee, 1997) and when parent–child conversation between 2 and 5 years of age includes more reports of negative emotional episodes compared with positive emotional episodes (Lagattuta & Wellman, 2002). Our results suggest that infants’ emergent sensitivity to the congruence of emotional facial expression and actions in others’ emotive behavior may develop from within their natural inclination to be more attentive and vigilant toward adults displaying a negative emotional expression.

With regard to interpreting the magnitude of changes in infants’ pupil dilation (i.e., interpreting the scale of measurement) (see Fig. 2), it is important to note that negative values do not indicate systematic pupillary constriction and a value of 0 does not indicate a null effect. The fact that the results contained negative changes in pupil dilation is a direct consequence of the baseline correction technique; the action clips (4 s in duration) were always preceded by a black screen such that upon presentation
of the clip the pupils would be maximally dilated (as a consequence of the dark screen) and would subsequently maximally constrict (i.e., showing the pupillary light reflex to a bright stimulus). Following constriction, pupillary movements will adjust to the brightness of the presentation surface, generally resulting in smaller values compared with those at stimulus onset. Therefore, the mean of the baseline period (i.e., time before action differentiation) included the large pupil diameter, and given that this mean was subtracted from all subsequent values (i.e., time after action differentiation), there was a general tendency to obtain negative values of pupil dilation.

It is worth pointing out that the experimental manipulation in our study was rather subtle. We did not habituate participants to an actor’s emotion; rather, we only presented them with one-trial brief introductions. Our results suggest that the measure of pupil dilation is sensitive to infants’ processing of subtle experimental manipulations without the need for habituation phases. This provides further evidence for the argument that pupil dilation provides a sensitive measure for tapping into time-locked changes in infants’ processing of unexpected events (Jackson & Sirois, 2009). Therefore, in addition to assessing infants’ sympathetic activity in response to impossible physical events (Jackson & Sirois, 2009) and uncommon social interactions (Gredebäck & Melinder, 2010) our results encourage the investigation of infants’ understanding of social cognitive phenomena, such as others’ emotions, using measures of pupil dilation (see also Geangu et al., 2011).

Conclusion

Successfully identifying emotions and the ability to link these emotions, with typical actions is an important part of successfully, engaging in social interaction. Our finding that this ability emerges at the end of the first year of life dovetails with work showing that after 9 months infants begin to display a variety of social skills such as jointly attending to objects with others, following the pointing and gaze of others, and imitating the actions of others (Carpenter, Nagell, & Tomasello, 1998). The more sophisticated processing of emotions shown here allows infants to go beyond viewing emotions in a specific context. Instead, the emotions themselves can establish a context in which subsequent actions are interpreted. Our results suggest that by 14 months of age, infants come to view emotions as concepts themselves that generalize across contexts. As adults, it comes naturally to us to view emotions as having a multitude of possible causes but predictable consequences. Here we have presented evidence that this ability may emerge earlier than previously assumed at the beginning of the second year of life.

Acknowledgments

We would like to thank all the children and parents who participated in this study and S. Wallace-Hadrill and N. Althaus for their assistance with data collection. Furthermore, we thank N. Ruh for his help with Matlab; A. Kirtley and J. Domoney for their help with recording the videos; L. Baldock for her help recruiting participants; E. Klonowski, N. Heussner, and R. Buechner for their help with coding. This research was supported in part by an ESRC grant (RES-000-22-3394) to GW.

References
