Cognitive Psychology 102 (2018) 1–20

Contents lists available at ScienceDirect

Cognitive Psychology

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

Young infants expect an unfamiliar adult to comfort a crying baby: Evidence from a standard violation-of-expectation task and a novel infant-triggered-video task

Kyong-sun Jin[a,*], Jessica L. Houston[b], Renée Baillargeon[c], Ashley M. Groh[d], Glenn I. Roisman[e]

[a] Department of Psychology, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea
[b] Center of Excellence for Children in State Custody, University of Tennessee Health Sciences Center, Memphis, TN 38103, United States
[c] Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign, IL 61820, United States
[d] Department of Psychological Sciences, University of Missouri, 204B McAlester Hall, Columbia, MO 65211, United States
[e] Institute of Child Development, University of Minnesota, 51 East River Parkway, Minneapolis, MN 55455, United States

ARTICLE INFO

Keywords:
Infancy
Social cognition
Prosociality
Expectations about comforting actions
Crying baby
Violation-of-expectation methods

ABSTRACT

Do infants expect individuals to act prosocially toward others in need, at least in some contexts? Very few such expectations have been uncovered to date. In three experiments, we examined whether infants would expect an adult alone in a scene with a crying baby to attempt to comfort the baby. In the first two experiments, 12- and 4-month-olds were tested using the standard violation-of-expectation method. Infants saw videotaped events in which a woman was performing a household chore when a baby nearby began to cry; the woman either comforted (*comfort* event) or ignored (*ignore* event) the baby. Infants looked significantly longer at the ignore than at the comfort event, and this effect was eliminated if the baby laughed instead of cried. In the third experiment, 8-month-olds were tested using a novel forced-choice violation-of-expectation method, the infant-triggered-video method. Infants faced two computer monitors and were first shown that touching the monitors triggered events: One monitor presented the comfort event and the other monitor presented the ignore event. Infants then chose which event they wanted to watch again by touching the corresponding monitor. Infants significantly chose the ignore over the comfort event, and this effect was eliminated if the baby laughed. Thus, across ages and methods, infants provided converging evidence that they expected the adult to comfort the crying baby. These results indicate that expectations about individuals' actions toward others in need are already present in the first year of life, and, as such, they constrain theoretical accounts of early prosociality and morality.

1. Introduction

Over the past two decades, a wealth of research has explored two broad issues in the development of prosociality in infancy. One issue has to do with infants' *own* prosocial actions (i.e., overt actions intended to benefit others, such as helping or comforting others in need). Around their first birthday, infants begin to produce prosocial actions, and these actions steadily increase in frequency and

* Corresponding author.
E-mail addresses: kyongsun.jin@gmail.com (K.-s. Jin), jhousto2@gmail.com (J.L. Houston), rbaillar@illinois.edu (R. Baillargeon), groha@missouri.edu (A.M. Groh), roism001@umn.edu (G.I. Roisman).

https://doi.org/10.1016/j.cogpsych.2017.12.004
Accepted 28 December 2017

variety as infants grow older (for reviews, see Brownell, 2013; Davidov, Vaish, Knafo-Noam, & Hastings, 2016; Decety & Howard, 2014; Dunfield, 2014; Dunn, 2014; Eisenberg, Fabes, & Spinrad, 2006; Eisenberg, Spinrad, & Knafo-Noam, 2015; Hastings, Zahn-Waxler, & McShane, 2006; Martin & Olson, 2015; Paulus, 2014; Vaish & Tomasello, 2014; Warneken, 2015). In addition to documenting the emergence and development of early prosocial actions, investigations have shed light on the situational factors that influence normative patterns of prosocial responding as well as on the dispositional and environmental factors that contribute to individual differences in infants' prosocial repertoires.

The other issue has to do with infants' reasoning about *others'* prosocial actions (for reviews, see Baillargeon, Setoh, Sloane, Jin, & Bian, 2014; Baillargeon et al., 2015; Bloom & Wynn, 2016; Hamlin, 2013b; Premack, 2007; Van de Vondervoort & Hamlin, 2016; Wynn & Bloom, 2014). As we discuss in more detail in the next sections, investigations of early prosocial reasoning have addressed three main questions: whether infants can distinguish between prosocial and antisocial actions and correctly evaluate the moral valences of these actions, whether infants can identify appropriate targets for prosocial actions, and whether infants expect individuals, at least in some contexts, to act prosocially toward others in need. This last question was the focus of the present research. As will become clear, there have been very few demonstrations to date of social contexts in which infants expect individuals to engage in prosocial actions. Yet knowing when infants expect prosocial actions would seem to be critical for understanding how they interpret and evaluate interactions in their social environments.

Building on prior findings, the present research examined one possible context in which young infants might expect individuals to produce prosocial actions. In three experiments, we asked whether infants would expect an unfamiliar adult alone in a scene with a crying baby to attempt to comfort the baby, and would detect a violation if the adult ignored the baby instead. The first two experiments tested 12- and 4-month-olds using the standard violation-of-expectation (VOE) method; the third experiment tested 8-month-olds using a novel forced-choice VOE method, the infant-triggered-video (ITV) method. We reasoned that positive evidence across ages and across methods would be important for several reasons. First, such evidence would bring to light one particular context in which infants, from a very young age, expect individuals to act prosocially toward others. As noted above, very few prosocial expectations have been uncovered to date, and none in the first year of life; positive results would thus make a significant empirical contribution to our understanding of early prosocial reasoning.

Second, from a methodological standpoint, positive evidence in our ITV experiment would introduce to the field of infancy research a novel method for assessing infants' expectations about events in the social domain as well as in other domains of cognition.

Finally, positive evidence that young infants expect an unfamiliar adult to comfort a crying baby would pave the way for future inquiries into the mechanisms responsible for this expectation. One possibility might be that from a young age, infants feel, and expect others to feel, empathic concern for the distress of others (e.g., Davidov, Zahn-Waxler, Roth-Hanania, & Knafo, 2013; de Waal, 2008; Decety, Michalska, & Kinzler, 2011; Eisenberg et al., 2006; Hastings et al., 2006; Hoffman, 2007; Knafo, Zahn-Waxler, Van Hulle, Robinson, & Rhee, 2008; Vaish & Warneken, 2012). Another possibility might be that most infants detect a regularity, in the first months of life, in how adults typically act toward crying babies (e.g., Killen & Cooley, 2014; Killen, Rutland, Abrams, Mulvey, & Hitti, 2013; Paulus, 2014; Paulus & Moore, 2012; Perner, 2010; Ruffman, Taumoepeau, & Perkins, 2012; Sripada & Stich, 2006). Yet another possibility might be that this expectation is sociomoral in nature. According to some accounts of morality, infants are born with a small set of moral principles that guides their expectations about how individuals will act toward others; this "first draft" of moral cognition (which is later revised by culture; Graham et al., 2013) includes not only principles of harm avoidance and fairness, which capture obligations that apply broadly to all individuals within the same moral circle, but also principles of ingroup support and authority, which capture obligations that are more limited in scope and depend on group affiliations or roles (e.g., Baillargeon et al., 2015; Brewer, 1999; Cosmides & Tooby, 2013; Graham et al., 2013; Jin & Baillargeon, 2017; Rai & Fiske, 2011; Shweder, Much, Mahapatra, & Park, 1997; Spokes & Spelke, 2017). From this sociomoral perspective, finding that infants as young as 4 months of age already expect caregivers to respond to crying babies might be taken as evidence of early sensitivity to harm avoidance and/or ingroup support. We return to these possibilities in the General Discussion.

1.1. Infants' expectations about helping actions

Much of the research on early prosocial reasoning has focused on helping actions. Initial investigations examined whether infants could distinguish between helping and hindering actions and correctly judge the moral valences of these actions (e.g., Hamlin, 2013a, 2014; Hamlin & Wynn, 2011; Hamlin, Wynn, & Bloom, 2007, 2010; Hamlin, Wynn, Bloom, & Mahajan, 2011; Kuhlmeier, Wynn, & Bloom, 2003; Premack & Premack, 1997). In a series of experiments, for example, Hamlin and her colleagues familiarized infants ages 3–10 months to a live scenario involving three non-human agents (blocks with eyes or animal puppets; e.g., Hamlin, 2014; Hamlin & Wynn, 2011; Hamlin et al., 2007, 2010). In a typical scenario, a character tried in vain to achieve a goal (e.g., reach the top of a steep hill or open a clear box to retrieve a toy); on alternate trials, a helper helped the character achieve its goal (*help* event), and a hinderer prevented the character from achieving its goal (*hinder* event). Following the familiarization trials, infants tested with a forced-choice affiliative-preference task consistently preferred the helper over the hinderer (e.g., Hamlin, 2014; Hamlin & Wynn, 2011; Hamlin et al., 2007, 2010), and (beginning at about 10 months) infants tested with a VOE task expected the character to show the same preference and detected a violation when the character approached the hinderer instead (Hamlin et al., 2007; see also Fawcett & Liszkowski, 2012; Lee, Yun, Kim, & Song, 2015). These results suggest that from a young age, infants can correctly evaluate a variety of helping and hindering actions and can use these evaluations to generate affiliative attitudes: Infants prefer, and expect others to prefer, individuals who have produced helping actions over individuals who have produced hindering actions.

Additional investigations examined whether infants could identify appropriate targets for individuals' helping actions. Köster, Ohmer, Nguyen, and Kärtner (2016) asked whether infants ages 9–18 months would selectively expect a protagonist to help a

character who needed assistance over one who did not. In the test portion of each computer-animated scenario, a human protagonist stood centered behind two non-human characters (shapes with eyes and limbs). One character's ball was within reach, and the other character's ball was out of reach. The protagonist first leaned forward (anticipatory phase) and then helped either the character who needed assistance (*expected* event) or the one who did not (*unexpected* event) (VOE phase). Infants looked first at the character in need during the anticipatory phase, and they looked longer at the unexpected event during the VOE phase. Thus, in line with claims that early psychological reasoning is guided by a principle of rationality, with corollaries of consistency and efficiency (e.g., Baillargeon, Scott, & Bian, 2016; Gergely, Nádasdy, Csibra, & Bíró, 1995), infants viewed it as less rational when the protagonist chose to help the character who did not need assistance. Kanakogi et al. (2017) focused on sociomoral as opposed to psychological considerations and examined whether 6-month-old infants would selectively expect a protagonist to help a victim as opposed to a wrongdoer. Infants were first familiarized to a computer-animated event involving three non-human agents (shapes with eyes): A wrongdoer repeatedly hit a victim while a protagonist watched from inside an enclosure. In the test events, the wrongdoer and the victim attempted in vain to enter the enclosure, and the protagonist helped either one of them. Infants looked significantly longer when the protagonist chose to help the wrongdoer as opposed to the victim, and this pattern reversed when the protagonist engaged in hitting actions instead. Together, these results suggest that from a young age, infants selectively expect helping actions to be targeted at individuals who need or deserve assistance over individuals who do not.

Finally, investigations have also examined whether infants ever view helping a character in need as expected rather than as optional and hence detect a violation when a protagonist fails to help the character. The initial investigations of helping actions cited above yielded largely negative results: Across a wide range of scenarios, infants ages 3–19 months tended to look equally at the help and hinder events shown in the familiarization trials (e.g., Hamlin, 2013a, 2014; Hamlin & Wynn, 2011; Hamlin et al., 2011; Lee et al., 2015; Premack & Premack, 1997). These negative results indicated that infants held no particular expectation that the helper and the hinderer would help the character; had infants possessed such an expectation, they would have looked significantly longer when the hinderer not only failed to help but actually hindered the character.

In light of these results, Jin and Baillargeon (2017) asked whether infants might be more likely to expect a protagonist to help a character in need when given clear evidence that the two belonged to the same social group. This speculation was derived in part from the well-established finding that adults and older children are more likely to help ingroup than non-ingroup individuals (e.g., Balliet, Wu, & De Dreu, 2014; Fehr, Bernhard, & Rockenbach, 2008; Killen & Turiel, 1998; Levine, Prosser, Evans, & Reicher, 2005; Olson & Spelke, 2008; Renno & Shutts, 2015). Jin and Baillargeon tested 17-month-olds using a VOE task with a minimal-group manipulation. Infants watched interactions among three unfamiliar female experimenters, E1-E3, who occupied windows around three sides of a puppet-stage apparatus. In the familiarization trials, the women announced their group memberships using novel labels (e.g., "I'm a topid!", "I'm a topid, too!", "I'm a jaybo!"). In the test trial, only E1 and E2 were present; while E2 watched, E1 selected discs of decreasing sizes from a box next to her window and stacked the discs, one by one, on a base. The final, smallest disc rested across the apparatus from E1, out of her reach, but within E2's reach. E1 tried in vain to reach the final disc until a bell rang; at that point, she announced that she would return and then left. In her absence, E2 picked up the final disc and either placed it in E1's box so that she could complete her stack when she returned (*help* event) or returned it to its original position on the apparatus floor, out of E1's reach (*ignore* event). When E1 and E2 belonged to the same minimal group, infants looked significantly longer at the ignore than at the help event. However, when E1 and E2 belonged to different minimal groups, or when their group memberships were not specified, infants looked equally at the help and ignore events.

These results, together with those of the initial investigations cited above, suggest that when a protagonist belongs to the same group as a character in need of instrumental assistance, infants expect the protagonist to provide the necessary help, and they detect a violation if the protagonist either ignores or hinders the character instead. When the protagonist and the character are not clearly identified as members of the same group, however, infants hold no particular expectation as to whether the protagonist will help, ignore, or hinder the character.

1.2. Expectations about others' comforting actions

Johnson and her colleagues (Johnson, Dweck, & Chen, 2007; Johnson et al., 2010) investigated whether infants would view comforting actions as expected in the context of an interaction between an adult and a crying baby. In one experiment, 13-month-olds were first habituated to a computer-animated event involving two faceless ovals, a large red one (the "adult") and a small blue one (the "baby"). At the start of each habituation trial, the adult and the baby stood at the bottom of a steep hill; the adult then climbed half-way up the hill, and the baby began to cry (with a pre-recorded human infant cry). Following habituation, infants saw two test events. At the start of each event, the adult stood half-way up the hill, and the baby cried at the bottom of the hill; the adult either returned to the baby (*return* event) or continued to the top of the hill (*leave* event). After completing this task, infants' attachment status was assessed using the Strange Situation procedure (Ainsworth, Blehar, Waters, & Wall, 1978).

Overall, infants tended to look equally at the return and leave events. When grouped by their attachment status, however, infants who were securely attached looked significantly longer at the leave than at the return event, whereas infants who were insecurely attached tended to show the reverse looking pattern (the latter difference was not statistically significant). Johnson et al. (2007, 2010) concluded that the adult-baby separation enacted in the habituation event triggered infants' internal working models of attachment to their primary caregivers (e.g., Bowlby, 1969). These models, in turn, influenced infants' expectations about how the adult would act in the test events. Thus, securely attached infants expected the adult to return to the crying baby, but insecurely attached infants did not. Biro, Alink, Huffmeijer, Bakermans-Kranenburg, and Van IJzendoorn (2015) subsequently examined how 12-month-olds allocated their attention during a similar adult-baby separation scenario, and their results provided partial support for

those of Johnson and her colleagues. When the baby character cried during the test events, infants with more sensitive mothers looked significantly longer at the adult character, as though expecting it to return to the baby, whereas infants with less sensitive mothers monitored both characters equally.

At first glance, these findings would appear to be inconsistent with those of Jin and Baillargeon (2017) reviewed in the preceding section. One might have predicted that infants would represent an interaction between an adult and a baby oval character alone in a scene as an ingroup interaction, and that they would expect the adult to respond prosocially to the baby if it became distressed. However, only infants who were classified as secure or who experienced more sensitive parenting (henceforth only infants with more supportive parenting) expected the adult to return to the crying baby. How can we reconcile these discrepant findings?

One possibility is that expectations about helping and comforting actions follow different developmental timetables, perhaps because these expectations are differentially affected by early caregiving experiences (e.g., Dunfield, 2014; Dunfield & Johnson, 2015; Dunfield & Kuhlmeier, 2013; Gross, Stern, Brett, & Cassidy, 2017). Another possibility is that infants generally do expect an adult to comfort a crying baby, but those who receive less supportive parenting demonstrate this expectation only under limited, optimal conditions. From this perspective, there were at least two possible (and complementary) explanations for the negative responses of the infants with less supportive parenting in the experiments described above.

A first explanation was that because of their negative attachment-related experiences, these infants perceived the leave event to be not merely unexpected but also emotionally distressing, and they therefore tended to look at it *less*. In a comprehensive overview of research on attachment and its impact on the processing of social information from infancy to adulthood, Dykas and Cassidy (2011) concluded that "potentially painful negative attachment-relevant social information (e.g., information about the absence of attachment figures when needed)" (p. 23) is processed differently by secure and insecure individuals. Specifically, secure individuals process the information relatively fully, whereas insecure individuals tend to defensively block it or exclude it from further processing—for example, by turning away from it. This notion of defensive exclusion supports the possibility that the insecure infants tested by Johnson et al. (2007, 2010) did detect a violation in the leave event, but tended to look away because their prior experiences with their caregivers led them "to suppress their attention to distressing infant-mother attachment related information" (Dykas & Cassidy, 2011, p. 26).

A second explanation was that the infants with less supportive parenting had difficulty processing the abstract, stylized animations they were shown. Evidence for this possibility comes from a physical-reasoning experiment by Hohenberger et al. (2012). They presented 10-month-olds with abstract animations depicting expected and unexpected collision events, adapted from naturalistic live events devised by Kotovsky and Baillargeon (1994, 1998). Infants with more sensitive mothers looked significantly longer at the unexpected than at the expected event (consistent with the results obtained by Kotovsky and Baillargeon with infants age 6.5 months and older), but infants with less sensitive mothers did not. These results suggest that infants with less supportive parenting may experience more stress in a novel laboratory situation and as such may be less able to process and interpret abstract animations, whether these depict physical interactions between objects or social interactions between individuals (e.g., Bornstein, Mash, Arterberry, & Manian, 2012; Sohr-Preston & Scaramella, 2006).[A]

Both of these explanations suggested that regardless of their attachment status, infants might expect an unfamiliar adult to respond to a crying baby if presented with a scenario that was both milder (i.e., less distressing) and more natural (i.e., less difficult to interpret) than the separation scenario devised by Johnson et al. (2007, 2010).[A] Accordingly, infants in the present experiments were shown videos of natural, everyday events depicting a mild distress scenario without separation from an adult. In everyday life, infants cry for many reasons other than fear of separation from caregivers, including hunger, cold, pain, boredom, frustration, and fatigue (e.g., Ainsworth et al., 1978; Bowlby, 1969). In our scenario, a baby who lay out of view in a stroller began to cry for no visible cause (e.g., the baby might have woken up or become hungry or bored and was now seeking attention); an adult standing nearby either responded to or ignored the baby, but in neither case did the adult move away, leaving the baby alone.

1.3. The present research

Our experiments examined whether 12- (Experiment 1), 4- (Experiment 2), and 8-month-olds (Experiment 3) would expect an unfamiliar woman to comfort a crying baby in a mild distress scenario. All experiments involved typical samples of infants recruited from university-maintained databases of parents interested in participating in child development research; attachment status and parenting sensitivity were not assessed. In each experiment, infants saw a comfort and an ignore test event. At the start of each event, a woman folded laundry at a table on one side of a room; behind her, at the back of the room, were a chair with more laundry and a large stroller. Next, a baby began to cry (the stroller shook slightly to signal that the baby was lying out of sight in the stroller). The

[A] dditional evidence that infants with less supportive mothers may have difficulty processing stimuli drawn from either the social or the physical domain comes from experiments conducted by Bornstein and his colleagues with 5-month-olds (Bornstein, Arterberry, Mash, & Manian, 2011; Bornstein et al., 2012). Following familiarization, infants of non-depressed mothers looked significantly longer at a novel than at a familiar facial expression (Bornstein et al., 2011) and at a novel than at a familiar view of a complex object (Bornstein et al., 2012), whereas infants of depressed mothers did not. In the case of the complex object, the infants of depressed mothers tended to look longer at the familiar view, and this familiarity effect suggests that they could discriminate between the two views but had difficulty processing them (Bornstein et al., 2012).

[A] nother facet of the separation scenario used by Johnson et al. (2007, 2010) that might have been distressing to infants involved the habituation event: Recall that the adult character moved away from the baby character, who then began to cry. On average, infants received 9.4 habituation trials, and across trials this repeated separation might have become puzzling or distressing: Since the baby cried each time the adult left it alone at the bottom of the hill, why would the adult continue to do so? Perhaps for this reason, Johnson reported that in addition to the 33 infants included in their final sample, another 20 infants (38%) were tested but excluded due to failure to habituate (n = 13) or to complete the procedure (n = 7).

woman then walked toward the back of the room and either bent over the stroller and rocked it gently, as though trying to comfort the baby (*comfort* event), or bent over the chair and picked up more laundry, as though ignoring the baby (*ignore* event). Two different women comforted or ignored the baby in the experiments with 12- and 8-month-olds, but the same woman did so in the experiment with 4-month-olds, to make the events simpler for them.

To assess infants' expectations, we used two different VOE methods. In Experiments 1 and 2, we used the traditional VOE method: Infants saw the comfort and ignore events in separate trials, and we measured how long they looked at each event. If infants expected the woman to respond to the crying baby, then they should detect a violation in the ignore event, and they should therefore look significantly longer at that event than at the comfort event. In Experiment 3, we introduced a novel forced-choice VOE method, the infant-triggered-video (ITV) method. Infants faced two computer monitors and were first shown that touching the monitors triggered events: One monitor presented the comfort event and the other monitor presented the ignore event. Infants then chose which event they wanted to watch again by touching the corresponding monitor. The rationale was that if infants viewed one of the events as unexpected, they might choose to watch it again in order to process it further and try to make sense of it. Thus, if infants expected the woman to respond to the crying baby, they should detect a violation in the ignore event, and they should therefore be significantly more likely to choose that event than the comfort event. Like the traditional VOE method, the ITV method thus takes advantage of infants' natural tendency to inspect further those events that deviate from their expectations.

As discussed earlier, we reasoned that converging evidence across ages and across methods that infants expected the woman to comfort the crying baby (a) would bring to light one early prosocial expectation, (b) would give rise to new questions about the nature of the mechanisms responsible for this expectation, and (c) would introduce to the field of infant cognition a new method for assessing early expectations about events.

2. Experiment 1

Twelve-month-old infants were randomly assigned to a crying or a laughing condition and received two familiarization trials and two test trials (Fig. 1). Each pair of trials in the crying condition involved two different women; for ease of description, we refer to the

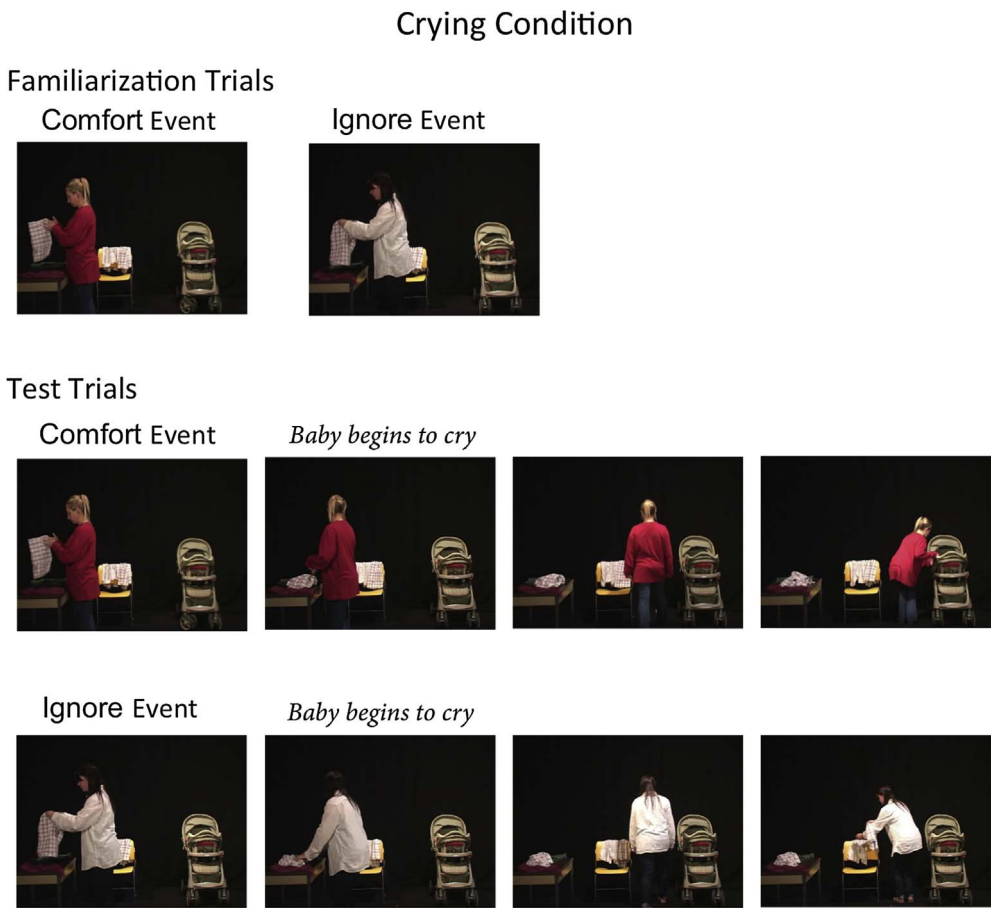

Fig. 1. Familiarization and test events in the crying condition of Experiment 1. For half the infants, the blonde woman comforted and the brunette woman ignored the crying baby; for the other infants, these roles were reversed. Events in the laughing condition were identical except that the baby laughed instead of cried in the test trials.

familiarization event involving the woman who would later appear in the comfort test event as the comfort familiarization event, and to the familiarization event involving the woman who would later appear in the ignore test event as the ignore familiarization event. For half the infants, the woman in the comfort familiarization and test events was a blonde in a red shirt, and the woman in the ignore familiarization and test events was a brunette in a white shirt; for the other infants, these roles were reversed. Half the infants saw the comfort event first in both the familiarization and test trials, and half the infants saw the ignore event first.

The two familiarization trials simply served to introduce the two women and the setting of the events. In each trial, a 25-s event loop was presented. After a 1.5 s blank screen, the woman could be seen folding kitchen towels at a small table on the left side of a room; behind her, at the back of the room, were a chair draped with three more unfolded towels, on the left, and a large stroller, on the right (it was not possible to see inside the stroller). The event loop ended as the woman folded a fourth towel, and it was repeated until the trial ended (see Procedure for criteria).

In each of the two test trials, a 22-s event loop was presented. After a 1.5 s blank screen, the woman folded kitchen towels for about 9.5 s. At that point, a baby began to cry (this was a pre-recorded infant cry from Groh & Roisman, 2009), and the stroller shook slightly to help infants infer that the baby was occupying it. The woman then set down the towel she was folding and walked toward the back of the room, following a course midway between the chair and the stroller (4 s). Because the woman turned toward the back of the room when the baby began to cry, her face (which remained neutral in any case) became hidden from the camera; thus, infants could not use the presence or absence of affective concern in her facial expression to predict her actions. When the woman arrived at the back of the room, her responses differed in the two events. In the *comfort* event, the woman bent over the stroller and rocked it gently while looking inside it, as though trying to comfort the crying baby. In the *ignore* event, the woman bent over the chair and picked up the three unfolded towels on the back of the chair one by one, ignoring the crying baby. These comforting or ignoring actions lasted about 7 s, and then the event loop ended. The whole 22-s loop then repeated until the trial ended. In each event loop, the baby cried throughout the second half of the event (11 s), so that the soundtracks of the comfort and ignore events were identical.

Infants in the *laughing* condition received familiarization and test trials identical to those in the crying condition, with one exception: During the second half of the 22-s event loop in the test trials, the baby laughed instead of cried (this was a pre-recorded infant laugh from Groh & Roisman, 2009).

Our predictions were as follows. First, if the infants in the crying condition expected the woman in each test event to respond to the crying baby, then they should detect a violation in the ignore event, and they should therefore look significantly longer at that event than at the comfort event. Second, if the infants in the laughing condition held no particular expectation about how the woman would respond to the laughing baby (who was clearly happily entertained), then they should look about equally at the comfort and ignore events. Such a negative result would help rule out low-level alternative interpretations of a positive result in the crying condition (e.g., infants looked longer at the ignore event because they preferred seeing the woman pick up towels as opposed to rock the stroller). Finding the expected results in the crying and laughing conditions would thus provide evidence that by their first birthday, infants already expect an unfamiliar adult alone in a scene with a crying baby to attempt to comfort the baby.

2.1. Method

2.1.1. Participants

Participants were 44 healthy term infants, 25 male (11 months, 18 days to 13 months, 1 day, M = 12 months, 10 days). Another 13 infants were excluded (5 in the crying condition and 8 in the laughing condition), 8 because they were distracted (e.g., by their shoes or a dropped pacifier) (4), fussy (3), or overly active (1), 3 (2 in the crying condition and 1 in the laughing condition) because the difference in their looking times at the two test events was over 2.5 standard deviations from the condition mean, and 2 because they looked for the maximum allowed (45 s) in both test trials. All of these criteria are similar to ones used in prior VOE tasks (e.g., Scott & Baillargeon, 2009; Setoh, Wu, Baillargeon, & Gelman, 2013). Half the infants (n = 22) were randomly assigned to the crying or laughing condition. Parents were offered reimbursement for their transportation expenses but were not compensated for their participation. Each infant's parent gave written informed consent, and the protocol was approved by the Institutional Review Board of the University of Illinois at Urbana-Champaign.

2.1.2. Apparatus and events

The apparatus consisted of a display booth (201 cm high × 164 cm wide × 78 cm deep) with a large opening (46 cm × 158 cm) in its front wall; between trials, a supervisor lowered a curtain in front of this opening. A 20-inch Panasonic television monitor rested on the apparatus floor, flush with its front edge; foam board covered with adhesive paper surrounded the monitor and filled the opening of the apparatus. Behind the apparatus, a DVD player was used to show the video appropriate for each trial. The soundtracks of the test videos were presented from the television monitor's internal speakers, at a comfortable listening level of about 65 dB.

Each testing session took place in a brightly lit room. During the session, a camera captured an image of the video shown on the television monitor, and another camera captured an image of the infant. The two images were combined, projected onto a monitor located behind the apparatus, and checked by the supervisor to confirm that the correct video was shown in each trial. The mixed images were also recorded and checked offline for accuracy.

2.1.3. Adults' descriptions of the test events

To confirm our descriptions of the women's actions in the test events, we showed 23 undergraduate students (4 male, mean age = 21.22 years) the ignore and comfort events from the crying and laughing conditions (order was counterbalanced across participants). After watching each event, participants completed the sentence, "The woman is ___ the baby." For the ignore event in the

crying condition, all 23 (100%) participants stated that the woman was ignoring or not paying attention to the baby. For the comfort event in the same condition, 21/23 (91%) participants indicated that the woman was responding to the crying baby in some fashion (e.g., comforting, consoling, rocking, or trying to soothe the baby). For the ignore event in the laughing condition, 18/23 (78%) participants stated that the woman was ignoring the baby, and 4/23 (17%) used descriptors that hinted at some acknowledgement of the baby (e.g., letting the baby be). Finally, for the comfort event in the same condition, 12/23 (52%) participants described a playful interaction between the woman and the laughing baby (e.g., playing with or entertaining the baby), and 9/23 (39%) described a more neutral interaction (e.g., checking on or rocking the baby).

Thus, as expected, adults generally perceived the woman in the ignore events of the crying and laughing conditions to be ignoring the baby; they perceived the woman in the comfort event of the crying condition to be attempting to comfort the baby; and they perceived the woman in the comfort event of the laughing condition to be playing with or checking on the baby.

2.1.4. Procedure

Each infant sat on a parent's lap in front of the television monitor; parents were instructed to remain silent and close their eyes during the test trials. At the start of each trial, a brightly colored attention-getter was displayed on the television monitor. Once the curtain was fully raised, the supervisor started the video for the trial. Each infant's looking behavior was monitored by two naïve observers hidden behind large frames on either side of the apparatus; the primary observer's responses were used to determine the ending of the trials. From their viewpoints, the observers could not determine whether the women in the test trials comforted or ignored the baby (in each condition, the soundtracks of the comfort and ignore events were identical). Agreement between the two observers was calculated for each trial by dividing the number of 100-ms intervals in which the two observers agreed on whether the infant was or was not looking at the event by the total number of intervals in the trial. Inter-observer agreement was calculated for all 44 infants and averaged 96% per trial per infant.

Each familiarization trial ended when infants either (1) looked away for 2 consecutive seconds after having looked for at least 25 cumulative seconds or (2) looked for a maximum of 45 s. The minimum value of each trial corresponded to the duration of one event loop and (as in prior VOE research, e.g., Setoh et al., 2013; Wang, Zhang, & Baillargeon, 2016) was selected to ensure that infants had the opportunity to observe at least one loop before the trial could end.

Each test trial consisted of a pre-trial followed by a main-trial. The pre-trial corresponded to the first 15 s of the first 22-s event loop shown in the trial. During this time, the women performed identical actions in the comfort and ignore events: They folded towels until the baby began to cry, and then they walked toward the back of the room following a course midway between the chair and the stroller. The main-trial began when the women turned toward the stroller (comfort event) or the chair (ignore event), and it continued on as the first event loop ended and the next event loop began. Looking times during the pre-trial and main-trial of each test trial were computed separately; this ensured that brief looks away during the pre-trial (e.g., while the women were folding towels) could not result in the trial ending prematurely, before the women had produced their comforting or ignoring actions. Across conditions and events, infants were highly attentive during the pre-trials and looked, on average, for 12.75/15 s. The main-trial of each test trial ended when infants (1) looked away for 0.5 consecutive seconds after having looked for at 4 cumulative seconds or (2) looked for a maximum of 45 s. The 4-s minimum value ensured that infants had the opportunity to see some of the woman's actions at the stroller or chair before the trial could end. Finally, because infants' attention was easily recaptured as the event loop repeated, the 0.5-s look-away value made it possible to detect when infants had sufficiently processed the event to determine whether it was expected (for similar criteria, see e.g., Scott, He, Baillargeon, & Cummins, 2012; Stavans & Baillargeon, 2018).

2.1.5. Ratings of infants' fussiness, concern, and amusement

In an experiment on early sensitivity to others' distress, Roth-Hanania, Davidov, and Zahn-Waxler (2011) presented 8-, 10-, and 12-month-olds with three distress scenarios: two 60-s live scenarios in which the mother simulated pain (i.e., pretended to hurt her finger or knee), and one 60-s video scenario of an unfamiliar baby crying. Results were similar across ages and scenarios. Although infants rarely became distressed themselves, they showed modest levels of affective concern in their facial expressions or vocalizations (means of 2.17–2.75 on a 4-point scale). In line with these results (see also Geangu, Hauf, Bhardwaj, & Bentz, 2011; Hay, Nash, & Pedersen, 1981; Liddle, Bradley, & Mcgrath, 2015), we also examined whether infants in Experiment 1 became fussy or showed concern or amusement during the test trials.

Fussiness. At the end of each trial, the two observers noted the infant's state (drowsy, quiet and alert, active, or fussy) during the trial.[A] As mentioned earlier, three infants (one in the crying condition and two in the laughing condition) were excluded from the experiment due to fussiness; they all began to cry during the familiarization trials and received no test trials. Of the 22 infants included in the crying condition, no infant was described as fussy during the test trials; infants thus showed no obvious distress when the baby cried.[I] In the laughing condition, only one of the 22 infants was described by the two observers as somewhat fussy during the

[A] t the end of each trial, observers also noted any events that might have affected the infant's looking responses during the trial, such as the infant becoming distracted (e.g., by items of clothing or a dropped pacifier) or actively trying to get off the parent's lap. These comments were later used for making decisions about exclusions, as reported in the Participants sections of Experiments 1 and 2.

[I] n the study of Roth-Hanania et al. (2011) and in the present experiments, infants exhibited little or no distress themselves during the distress scenarios they observed. These findings may seem inconsistent with prior reports of early "contagious crying": Newborns (e.g., Dondi, Simion, & Caltran, 1999; Sagi & Hoffman, 1976) and infants ages 1 to 9 months (e.g., Geangu, Benga, Stahl, & Striano, 2010, 2011) have been found to show facial and vocal expressions of distress in response to another infant's cry vocalizations. In all of these reports, however, the distress scenario lasted several minutes (e.g., 4 to 6 min), and infants were tested in an infant seat or crib with the parent absent or out of view, which might have added to their distress, especially at the older ages.

test trials.

Concern and amusement. To examine whether infants exhibited concern or amusement in their facial expressions, vocalizations, or gestures during the test trials, two independent coders watched infants' recorded testing sessions. For each test trial, they rated the infant's concern and amusement on separate 4-point scales (1 = absent, 2 = slight, 3 = moderate, and 4 = substantial). Because the coders had to listen to the infant's vocalizations, they were not naïve about the condition to which the infant was assigned; however, they were naïve about the order in which the comfort and ignore test events were displayed. Four infants (one in the crying condition and three in the laughing condition) could not be rated due to technical difficulties (e.g., the infant's face was not visible in one or both test trials). None of the 40 infants who were rated were fussy prior to or during the test trials. Inter-coder reliability was high, intraclass correlation coefficient = 0.98, $p < .001$. Accordingly, the two coders' ratings were averaged for each scale, and these composite ratings were used in the analyses of infants' concern and amusement.

2.2. Results

2.2.1. Looking times

Preliminary analyses of the test data revealed no interaction of condition and event with infants' sex, test order, or which woman comforted, all $Fs(1, 40) < 1.31$, $ps > .250$; the data were therefore collapsed across the latter three factors.

Infants' looking times during the familiarization trials were subjected to an analysis of variance (ANOVA) with condition (crying or laughing) as a between-subject factor and event (comfort or ignore) as a within-subject factor. The main effects of condition and event were not significant, both $Fs(1, 42) < 0.46$, $p > .250$, nor was the Condition × Event interaction, $F(1, 42) = 1.95$, $p = .170$, suggesting that infants in the two conditions tended to look equally when the two women folded towels during the familiarization trials (crying condition: ignore, $M = 34.63$, $SD = 7.86$, comfort, $M = 31.57$, $SD = 6.03$; laughing condition: ignore, $M = 33.03$, $SD = 7.32$, comfort, $M = 34.10$, $SD = 7.24$).

Infants' looking times during the main-trials of the test trials (Fig. 2) were analyzed as above. The main effects of condition and event were again not significant, both $Fs(1, 42) < 0.18$, $p > .250$, but the Condition × Event interaction was significant, $F(1, 42) = 6.07$, $p = .018$, $\eta_p^2 = 0.13$. Planned comparisons revealed that infants in the crying condition looked significantly longer at the ignore event ($M = 21.74$, $SD = 14.67$) than at the comfort event ($M = 16.56$, $SD = 11.42$), $F(1, 42) = 4.10$, $p = .049$, $d = 0.39$, whereas infants in the laughing condition looked about equally at the ignore ($M = 18.67$, $SD = 11.69$) and comfort ($M = 22.40$, $SD = 12.57$) events, $F(1, 42) = 2.13$, $p = .152$, $d = -0.31$. An analysis of covariance using as covariates infants' looking times at the comfort and ignore familiarization events again revealed a significant Condition × Event interaction, $F(1, 40) = 5.04$, $p = .030$, $\eta_p^2 = 0.11$.

Inspection of infants' individual responses yielded similar results. In the crying condition, 16/22 (73%) infants looked longer at the ignore than at the comfort event, $p = .026$ (cumulative binomial probability), and a nonparametric Wilcoxon signed-ranks test

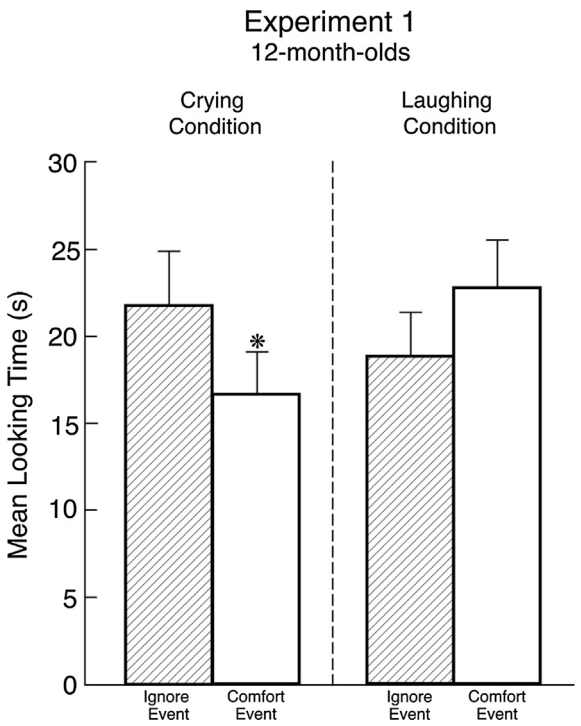

Fig. 2. Mean looking times of infants in Experiment 1 during the main-trials of the test trials as a function of condition and event. Errors bars represent SEs, and an asterisk denotes a significant difference between the events within a condition ($p < .05$).

confirmed that looking times at the two events differed significantly, $Z = 2.55$, $p = .011$ (all tests in this report are two-tailed). In the laughing condition, 10/22 (45%) infants looked longer at the ignore event, $p > .250$, and looking times at the two events did not differ significantly, $Z = -0.99$, $p > .250$.

2.2.2. Concern and amusement

The coders' averaged ratings of infants' *concern* during the test events were analyzed using an ANOVA with condition (crying or laughing) as a between-subject factor and event (comfort or ignore) as a within-subject factor. The analysis yielded a significant main effect of condition, $F(1, 38) = 11.56$, $p = .002$, indicating that infants in the crying condition ($M = 1.58$, $SD = 0.76$) showed significantly more concern than did those in the laughing condition ($M = 1.05$, $SD = 0.23$). The ratings of infants' *amusement* were analyzed in the same manner, and the analysis again yielded a significant main effect of condition, $F(1, 38) = 4.16$, $p = .048$, indicating that infants in the laughing condition ($M = 1.51$, $SD = 0.76$) showed significantly more amusement than did those in the crying condition ($M = 1.18$, $SD = 0.47$). In each analysis, the main effect of event and the Condition × Event interaction were not significant, all Fs$(1, 38) < 2.56$, ps $> .117$, indicating that within each condition, infants had similar affective responses to the comfort and ignore events.

Finally, we conducted correlational analyses to examine whether the concern shown by infants in the crying condition when watching either or both test events was related to their differential looking times at the events (ignore minus comfort event). None of these correlations were significant; rs ranged from 0.03 to 0.07, ps $> .250$, suggesting that infants who showed more concern were no more likely to look differentially at the two events. Inspection of the individual infants' responses yielded similar results. Of the 14 infants who received a concern rating greater than 1 (= absent) in at least one test trial, 9 (64%) looked longer at the ignore than at the comfort event. Of the remaining 7 infants, who showed no concern in either test trial, 6 (86%) looked longer at the ignore event. The two groups' responses did not differ significantly, $p > .250$ (Fisher's exact test); given the small sample sizes involved, however, these results should be viewed with caution.

In sum, infants in the crying and laughing conditions exhibited distinct and appropriate affective responses: Overall, infants were more likely to show concern when the baby cried and to show amusement when the baby laughed. In both cases, however, these responses were very slight (all means were between 1 = absent and 2 = slight). Finally, infants' concern for the crying baby did not appear to be related to their test looking times, as infants who did not show concern were as likely as those who did to look longer at the ignore event.

2.3. Discussion

When watching the test events, infants showed mild signs of concern when the baby cried and mild signs of amusement when the baby laughed; these responses did not differ whether infants were watching the comfort or the ignore event. In their looking times, however, infants in the crying condition looked significantly longer at the ignore than at the comfort event, whereas infants in the laughing condition looked equally at the events. Thus, when the baby began to cry, infants expected the woman to interrupt her work and comfort the baby, and they detected a violation if she continued her work and ignored the baby instead. When the baby began to laugh, in contrast, infants held no particular expectation about whether the woman would respond to or ignore the baby, who seemed happily entertained.

3. Experiment 2

The results of Experiment 1 indicated that by their first birthday, infants expect an unfamiliar adult to comfort a crying baby. Does this expectation gradually emerge over the course of the first year, or is it already present early in life? To address this question, we next tested 4-month-olds using events similar to those shown the crying and laughing conditions of Experiment 1, except for a few changes instituted to render the events simpler, shorter, and more attractive for very young infants. First, infants saw only one woman, the blonde, throughout the testing session (Fig. 3); they received two identical familiarization trials in which she folded towels and two test trials in which she either comforted or ignored the baby. Second, a chime melody was added to the familiarization event to help attract and maintain infants' attention. Finally, the test events were edited to make them shorter: After a 1-s blank screen, the woman spent only a few seconds folding towels before the baby began to cry or laugh, so that each event loop now lasted 15 s instead of 22 s. In each event loop, the baby cried or laughed during the last 11 s of the loop, as in Experiment 1, so that the soundtracks of the comfort and ignore events were identical within each condition.

3.1. Method

3.1.1. Participants

Participants were 36 healthy term infants, 16 male (3 months, 21 days to 5 months, 27 day, $M = 4$ months, 23 days). Another 6 infants were excluded (2 in the crying condition and 4 in the laughing condition), 3 because they were fussy (2) or distracted (1), 2 (1 in each condition) because the difference in their looking times at the two test events was over 2.5 standard deviations from the condition mean, and 1 because she looked for the maximum allowed (60 s) in both test events. Half the infants ($n = 18$) were randomly assigned to the crying or laughing condition.

Experiment 2 was conducted in two different locations: 26 infants were tested at the University of Illinois Infant Cognition Laboratory, and 10 infants (5 in each condition) were tested at the Yonsei University Human Development Laboratory (where the first

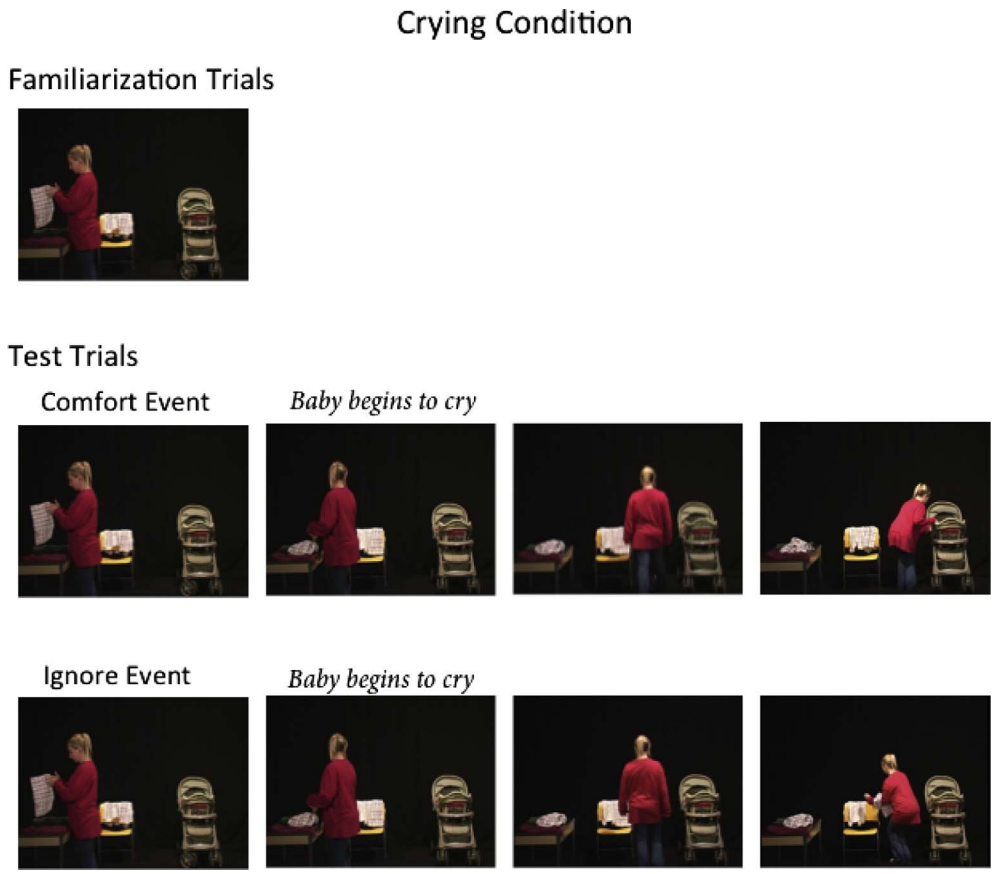

Fig. 3. Familiarization and test events in the crying condition of Experiment 2. Infants received two identical familiarization trials in which the blonde woman folded towels; in the test trials, she either comforted or ignored the crying baby. Events in the laughing condition were identical except that the baby laughed instead of cried in the test trials.

author went to pursue postdoctoral studies after completing her doctorate at Illinois). The protocol was approved by the Institutional Review Boards at both institutions. Comparison of the test data from the two countries revealed no interaction of condition and event with country, $F(1, 32) = 0.01$, $p > .250$; subsequent analyses therefore collapsed across country as a factor.

3.1.2. Apparatus and procedure

At the University of Illinois, the apparatus used in Experiment 2 was identical to that in Experiment 1 except that a 37-inch Panasonic television monitor was used. At Yonsei University, the apparatus again consisted of a display booth (200 cm × 95 cm × 64 cm) with a large opening (53 cm × 88 cm) in its front wall. A 22-inch LG computer monitor rested on the apparatus floor, flush with its front edge; foam board covered with adhesive paper surrounded the monitor and filled the opening of the apparatus. Behind the apparatus, a laptop was used to show the video appropriate for each trial. The soundtracks of the test videos were presented from the laptop's internal speakers, at a comfortable listening level of about 65 dB.

The procedure used in Experiment 2 was identical to that in Experiment 1 except that slightly different criteria were used to end the trials. First, each familiarization and test trial had a maximum value of 60 s, instead of 45 s, to allow our very young participants more time to watch and process the events. Second, because the event loop shown in each test event was now shorter (15 s instead of 22 s), with very little incidental folding of towels at the start of each event, the test trials were no longer divided into a pre-trial and main-trial; instead, the minimum value of each test trial simply corresponded to the duration of one event loop, as in the familiarization trials. Thus, each familiarization trial ended when infants (1) looked away for 2 consecutive seconds after having looked for at least 25 cumulative seconds (the duration of one event loop) or (2) looked for a maximum of 60 s, and each test trial ended when infants (1) looked away for 0.5 consecutive seconds after having looked for at least 15 cumulative seconds (the duration of one event loop) or (2) looked for a maximum of 60 s. Interobserver agreement in each trial was calculated for 35 infants (only one observer was available for 1 infant) and averaged 93% per trial per infant.

3.1.3. Ratings of infants' fussiness, concern, and amusement

As in Experiment 1, the two observers noted the infant's state during each trial. As mentioned earlier, two infants were excluded from the experiment due to fussiness; both were in the laughing condition and became too fussy to complete the second test trial. Of

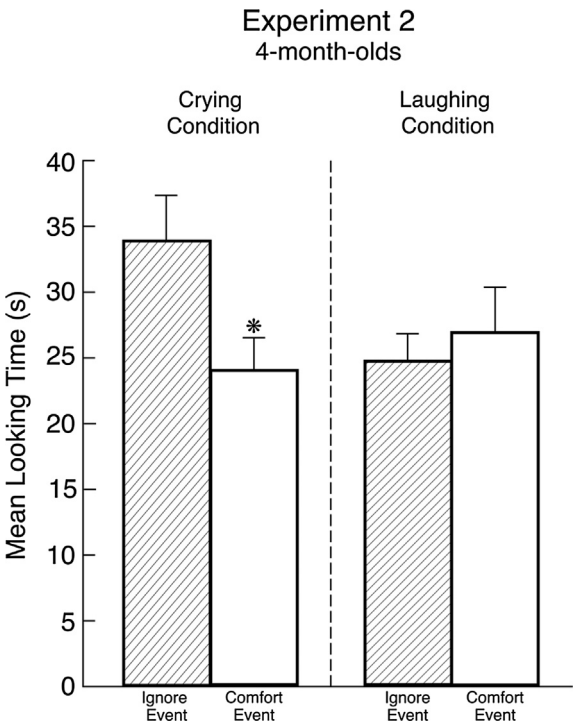

Fig. 4. Mean looking times of infants in Experiment 2 during the test trials as a function of condition and event. Errors bars represent SEs, and an asterisk denotes a significant difference between the events within a condition ($p < .05$).

the 18 infants included in the crying condition, only three were described by one or both observers as somewhat fussy during one or both test trials; as in Experiment 1, infants thus showed little distress themselves when the baby cried. In the laughing condition, only one of the 18 infants was described by the observers as somewhat fussy in the test trials.

Finally, infants' concern and amusement during the test trials were assessed as in Experiment 1. Two infants (one in each condition) were excluded from these analyses because they became somewhat fussy before the test trials began, making it difficult to determine whether their affective responses to the test events were due to their prior emotional states or to the contents of the events. Inter-coder reliability was high, intraclass correlation coefficient = 0.91, $p < .001$. Accordingly, the two coders' ratings were averaged for each scale, and these composite ratings were used in the analyses of infants' concern and amusement.

3.2. Results

3.2.1. Looking times

Preliminary analyses of the test data revealed no interaction of condition and event with infants' sex or test order, both Fs(1, 32) < 0.45, $p > .250$; the data were therefore collapsed across the latter two factors.

Infants' looking times during the two familiarization trials were averaged and subjected to an ANOVA with condition (crying or laughing) as a between-subject factor. The main effect of condition was not significant, $F(1, 34) = 1.73$, $p = .198$, suggesting that the infants in the two conditions tended to look equally when the woman folded towels during the familiarization trials (crying condition: $M = 47.12$, $SD = 11.20$, laughing condition: $M = 41.76$, $SD = 13.22$).

Infants' looking times during the test trials (Fig. 4) were subjected to an ANOVA with condition (crying or laughing) as a between-subject factor and event (comfort or ignore) as a within-subject factor. The main effects of condition and event were not significant, both Fs(1, 34) < 2.53, $p > .121$, but the Condition × Event interaction was significant, $F(1, 34) = 6.32$, $p = .017$, $\eta_p^2 = 0.16$. Planned comparisons revealed that infants in the crying condition looked significantly longer at the ignore event ($M = 33.78$, $SD = 14.52$) than at the comfort event ($M = 23.93$, $SD = 10.40$), $F(1, 34) = 8.41$, $p = .006$, $d = 0.78$, whereas infants in the laughing condition looked about equally at the ignore ($M = 24.58$, $SD = 8.78$) and comfort ($M = 26.80$, $SD = 15.00$) events, $F(1, 34) = 0.43$, $p > .250$, $d = -0.18$. An analysis of covariance using as a covariate infants' averaged looking times during the familiarization trials again revealed a significant Condition × Event interaction, $F(1, 33) = 7.75$, $p = .009$, $\eta_p^2 = 0.19$.

Inspection of infants' individual responses yielded similar results. In the crying condition, 14/18 (78%) infants looked longer at the ignore than at the comfort event, $p = .015$ (cumulative binomial probability), and a Wilcoxon signed-ranks test confirmed that looking times at the two events differed significantly, $Z = 2.20$, $p = .028$. In the laughing condition, 9/18 (50%) infants looked longer at the ignore event, $p > .250$, and looking times at the two events did not differ significantly, $Z = -0.20$, $p > .250$.

3.2.2. Concern and amusement

The coders' averaged ratings of infants' *concern* during the test events were analyzed as in Experiment 1. The analysis yielded a significant main effect of condition, $F(1, 32) = 4.35$, $p = .045$, indicating that infants in the crying condition ($M = 1.37$, $SD = 0.54$) showed significantly more concern than did those in the laughing condition ($M = 1.09$, $SD = 0.26$). The ratings of infants' *amusement* were analyzed in the same manner; the main effect of condition was not significant, $F(1, 32) = 0.96$, $p > .250$, though mean differences were in the predicted direction (crying condition: $M = 1.18$, $SD = 0.47$; laughing condition: $M = 1.35$, $SD = 0.65$). In each analysis, the main effect of event and the Condition × Event interaction were not significant, all $Fs(1, 32) < 0.87$, $p > .250$, indicating that within each condition, infants had similar affective responses to the comfort and ignore events.

Finally, we conducted correlational analyses to examine whether the concern shown by infants in the crying condition when watching either or both test events was related to their differential looking times at the events (ignore minus comfort event). None of these correlations were significant; rs ranged from -0.04 to 0.24, $ps > .250$, suggesting that infants who showed more concern were no more likely to look differentially at the two events. Inspection of the individual infants' responses yielded similar results. Of the 7 infants who received a concern rating greater than 1 (= absent) in at least one test trial, 5 (71%) looked longer at the ignore than at the comfort event. Of the remaining 10 infants, who showed no concern in either test trial, 9 (90%) looked longer at the ignore event. The two groups' responses did not differ significantly, $p > .250$ (Fisher's exact test); given the small sample sizes involved, however, these results should be viewed with caution.

In sum, infants exhibited more concern when the baby cried as opposed to laughed, and they also tended to show more amusement when the baby laughed as opposed to cried, though the latter difference was not statistically significant. In both cases, affective responses were very slight (all means were between 1 = absent and 2 = slight). Finally, as in Experiment 1, infants' concern for the crying baby did not appear to be related to their test looking times, as infants who did not show concern were as likely as those who did to look longer at the ignore event.

3.3. Discussion

Infants in the crying condition looked significantly longer at the ignore than at the comfort event, whereas infants in the laughing condition looked equally at the events. These results indicate that by 4 months of age, infants already expect an unfamiliar woman to respond to a crying baby, and they detect a violation if she ignores the baby instead. Our results thus extend those of Experiment 1 downward in age from 12 to 4 months and suggest that beginning early in life, infants hold at least one expectation related to individuals' prosocial actions toward others in need.

The results of Experiment 2 are also consistent with a few prior reports that young infants who are faced with a crying baby sometimes look to nearby adults (e.g., Hay et al., 1981; Liddle et al., 2015; Roth-Hanania et al., 2011). In a study by Hay et al., for example, pairs of mothers and their 6-month-olds sat on the floor, facing each other, during two 10-min trials. When one of the infants became distressed, the mother provided comfort as she normally did, and the responses of the other infant were examined. Results indicated that 63% of the infants looked at the peer's mother and 42% looked at their own mothers during at least one distress episode. Although these responses are somewhat difficult to interpret (e.g., infants might have looked at the peer's mother because she had already begun comforting the peer, or because they were trying to understand why the peer was crying), one interpretation consistent with our findings is that infants looked at the peer's mother or their own at least in part because they expected them to intervene and comfort the crying baby.

4. Experiment 3

Experiments 1 and 2 demonstrated, using the traditional VOE method, that young infants expect an unfamiliar adult to comfort a distressed baby. Experiment 3 sought to confirm this expectation using a novel VOE method. We reasoned that obtaining converging evidence across methods would not only buttress our conclusions about the early emergence of this expectation, but would also make available to developmental researchers a new method for exploring expectations about prosociality, morality, and other facets of infant cognition.

To date, at least three different VOE methods have been used to study infants' expectations in studies of cognitive development. The oldest and most prevalent of these methods compares infants' looking times at unexpected and expected events (e.g., Baillargeon, Spelke, & Wasserman, 1985). Another method measures changes in infants' pupil diameter (a measure of sympathetic activity) when watching unexpected and expected events (e.g., Gredebäck & Melinder, 2010). Yet another method compares infants' exploration of objects featured in unexpected as opposed to expected events (e.g., Stahl & Feigenson, 2015). In Experiments 1 and 2, we used the first of these methods. In Experiment 3, we introduce a novel forced-choice VOE method, the ITV method. In this method (as was explained in the Introduction), infants first see an expected and an unexpected event, each displayed on a separate computer monitor; next, infants choose which event they want to see again by touching the corresponding monitor. Like other VOE tasks, ITV tasks thus take advantage of infants' natural tendency to explore further those events that deviate from their expectations.

Participants in Experiment 3 were 8-month-old infants. We chose this particular age group because we wanted to test young infants who would be comfortable reaching for and tapping our computer monitors to select events; pilot data with other events suggested that infants in this age group would have no difficulty with our novel ITV method.

As in Experiments 1 and 2, infants were assigned to a crying or a laughing condition. Each infant received two demonstration trials and four test trials; each trial had a choice phase and a viewing phase. During the choice phase of the first demonstration trial, the infant sat on a parent's lap between two angled computer monitors; the parent's wheeled chair was positioned a short distance

from the monitors so that the parent could reach them, but the infant could not. The monitors displayed still pictures from the videotaped test events used in Experiment 1; recall that two different women, a blonde and a brunette, were used in these events, making the two still pictures more perceptually distinct and hence more suitable for a forced-choice task with young infants. In one picture, the woman was bent over the stroller (*comfort* event); in the other picture, the woman was bent over the chair (*ignore* event). Which woman comforted and which monitor displayed the comfort event were counterbalanced across infants. Each still picture was surrounded by a flashing, beeping yellow band, to help attract attention to both pictures and signal that a choice was needed. A hidden online assistant asked the parent to briefly touch the picture on the right monitor (the parent was instructed before the test session to first wave his or her hand gently in front of the infant, so as to capture the infant's attention and ensure that the infant saw the parent touch the monitor). When this touch occurred, the assistant activated the corresponding event, ending the choice phase and beginning the viewing phase. The picture on the left monitor disappeared (it was replaced by a grey rectangle), and the event on the right monitor played once. The comfort and ignore events shown in Experiment 3 were identical to those in Experiment 2 (with little folding of towels at the start of each event), except that there was no blank screen at the start of each event loop; each loop thus lasted 14 s (instead of 15 s as in Experiment 2). When the event ended, the two still pictures (with their flashing, beeping bands) returned, beginning the choice period of the second demonstration trial. The parent was then asked to touch the picture on the left monitor, and the corresponding event was activated. Thus, by the end of the two demonstration trials, infants had had the opportunity to watch both events and had also seen that each event was triggered by touching the corresponding monitor.

After the demonstration trials, a screen-saver depicting animals appeared on the two monitors, to allow the parent to wheel forward so that the infant came within reach of the two monitors. The four test trials then began (in contrast to affiliative-preference tasks, in which infants typically receive a single choice trial, infants in our ITV task received as many as four choice trials, thus substantially increasing the data from which conclusions could be drawn). Each test trial had a *choice* period, in which the infant could choose an event by touching a monitor (Fig. 5), followed by a *viewing* period, in which the selected event played. We reasoned that if infants in the crying condition expected the woman in each event to respond to the crying baby, then they should detect a violation when the woman ignored the baby instead. Infants should thus be significantly more likely to trigger the ignore event, to process it further and try to make sense of the woman's behavior. In contrast, infants in the laughing condition should hold no particular expectation about whether the woman in each event would respond to or ignore the laughing baby, and they should thus be equally likely to trigger either event.

4.1. Method

4.1.1. Participants

Participants were 36 healthy full-term infants, 18 male (7 months, 6 days to 9 months, 17 days, M = 8 months, 13 days). Another 7 infants were excluded (4 in the crying condition and 3 in the laughing condition), 1 because of parental interference, 1 because of fussiness, 1 because of inattentiveness during the demonstration phase, 1 because the infant never triggered test events, and 3 because the online assistant made an error in registering what events the infants had selected. Equal numbers of infants (n = 18) were randomly assigned to the crying or laughing condition.

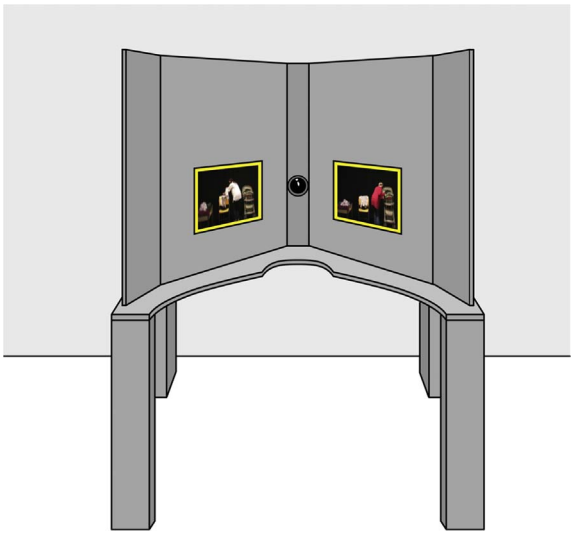

Fig. 5. Schematic drawing of the infant-triggered-video (ITV) apparatus used in Experiment 3. In the choice period at the start of each test trial, the infant saw still pictures from the comfort and ignore events on the two monitors. Which woman comforted and which monitor displayed the comfort event were counterbalanced across infants.

4.1.2. Apparatus and events

The apparatus consisted of an inverted-V-shaped table (64 cm × 122 cm × 122 cm) supporting two angled 19-inch computer monitors with a wide-angle-lens camera between them. The soundtrack of each video was presented via a speaker placed behind the corresponding monitor. A tall divider made of Plexiglas (76 cm × 102 cm × 0.5 cm) stood on each arm of the table and was affixed to the front of the monitor. These dividers isolated the infant from the rest of the testing room and also protected the monitors so that the infant could tap on them without damaging them. Black adhesive paper covered the dividers everywhere except for a rectangular area (22 cm × 36 cm) in front of each monitor. When the parent's chair was wheeled forward, the infant fit into a small circular cut-out area in the front of the table, between the two monitors; this ensured that the infant was secure and able to reach freely with little parental supervision.

Testing sessions were conducted in semi-darkness, with only the light from the monitors. During each session, the wide-angle-lens camera captured an image of the infant, and an overhead camera captured an image of the monitors. The two images were combined and projected on a television monitor in a curtained-off area of the testing room. The mixed images were also recorded and checked offline for accuracy.

4.1.3. Procedure

Each infant sat on a parent's lap, centered in front of the ITV apparatus; parents were instructed to remain silent and to close their eyes or look down during the test trials. The online assistant watched the computer screen in the curtained-off area of the testing room and pressed the left or right button on a wireless mouse when the infant touched the corresponding monitor (touches to either monitor produced a lit-up reflection in the wide-angle-lens camera output, while touches to the dark areas outside the monitors did not). The assistant was naïve about which event was presented on which monitor: The image from the overhead camera was covered up, and the comfort and ignore events in each condition were the same duration and had identical soundtracks so that no auditory cues distinguished them. A special-purpose computer program used the assistant's button presses to trigger the events: In each choice period, the infant's first touch to a monitor triggered the corresponding event; while this event played, additional touches to either monitor had no effect. A naïve offline assistant watched all of the recorded testing sessions without their soundtracks, to check the online assistant's accuracy in registering what event was selected in each choice period. The offline assistant detected errors for only three infants, who were excluded from the analyses (as reported above).

Offline frame-by-frame coding established that infants were attentive during the viewing phases of the two demonstration trials and watched, on average, for 76% of each viewing phase. Of the 36 infants tested, 31 completed all four test trials and 5 completed only three test trials, for a total of 139 triggers. For each infant, we computed the proportion of triggers to the ignore event, out of the total number of triggers to either event.

4.1.4. Ratings of infants' fussiness

To determine whether infants became distressed as they watched the events, an offline assistant watched each recorded testing session and noted the infant's state during each demonstration and test trial. As mentioned earlier, one infant in the laughing condition was excluded from the experiment due to fussiness. Of the 36 infants included in the experiment, one could not be assessed due to technical difficulties; the remaining 35 infants did not show any fussiness during the demonstration or test trials.

Because infants were tested in semi-darkness and tended to orient toward the left or right computer monitor (instead of facing the wide-angle-lens camera between the monitors), it was not possible to code their affective concern or amusement during the events.

4.2. Results

Preliminary analyses of the test data revealed no interaction of condition with infants' sex, which woman comforted, or which monitor presented the comfort event, all Fs(1, 32) < 2.27, ps $> .141$; the data were therefore collapsed across the latter three factors.

Infants' proportions of triggers to the ignore event during the test trials were subjected to an ANOVA with condition as a between-subject factor (Fig. 6). The analysis yielded a significant main effect of condition, $F(1, 34) = 5.47$, $p = .025$, indicating that infants in the crying condition were significantly more likely to choose the ignore event ($M = 0.72$, $SD = 0.24$) than were infants in the laughing condition ($M = 0.49$, $SD = 0.33$), $d = 0.80$. One-sample t-tests against chance (0.50) indicated that infants in the crying condition selected the ignore event significantly more often than would be expected by chance, $t(17) = 3.83$, $p = .001$, whereas the choices of infants in the laughing condition did not differ from chance, $t(17) = -0.12$, $p > .250$.

Overall, the 18 infants in the crying condition produced 50 ignore and 20 comfort triggers, $p < .001$ (cumulative binomial probability), and 13 (72%) infants produced more ignore than comfort triggers. In contrast, the 18 infants in the laughing condition produced 34 ignore and 35 comfort triggers, $p > .250$, and only 6 (33%) infants produced more ignore than comfort triggers. The distributions of ignore and comfort triggers in the two conditions were significantly different by a Fisher's exact test, $p = .009$.

4.3. Discussion

Infants in the crying condition were significantly more likely to select the ignore than the comfort event, whereas infants in the laughing condition selected either event at random. These results suggest that 8-month-olds expect an unfamiliar woman who is alone with a crying baby to comfort the baby, and they detect a violation if she ignores the baby instead. The results are thus consistent with those obtained with 12- and 4-month-olds in the preceding experiments, and provide converging evidence that infants in the first year

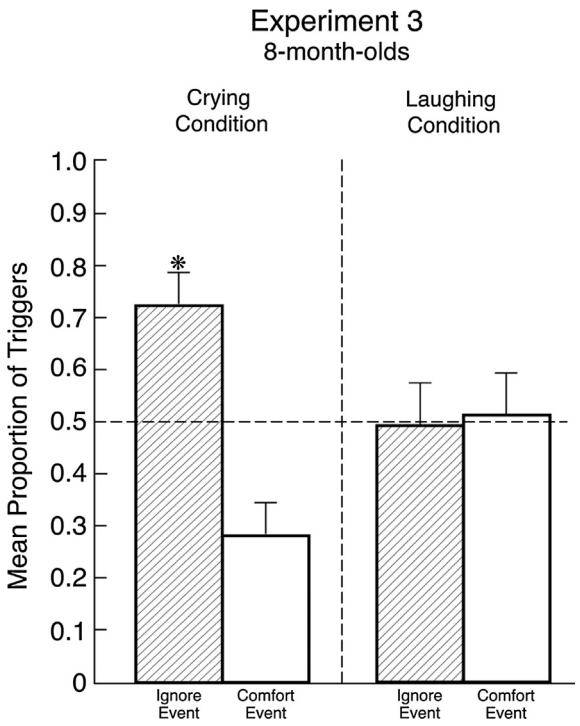

Fig. 6. Mean proportions of triggers by infants in Experiment 3 as a function of condition and event. Errors bars represent SEs, and an asterisk denotes a significant difference from chance level (0.50) for the condition ($p < .05$).

of life already hold at least one expectation related to individuals' prosocial actions.

In addition, the results of Experiment 3 demonstrate that the ITV method offers a useful tool for assessing young infants' expectations about events: Infants in the crying condition chose to watch again an event that violated as opposed to confirmed their expectations. At first glance, these results might appear inconsistent with those from forced-choice affiliative-preference tasks (discussed earlier) in which infants ages 5–10 months consistently chose a helper over a hinderer (e.g., Hamlin, 2013a; Hamlin & Wynn, 2011; Hamlin et al., 2007). Given these results, it might be asked why infants in the crying condition of Experiment 3 did not choose the responsive woman who had attempted to comfort the baby over the unresponsive woman who had ignored the baby. The answer to this question lies in the fact that our ITV task is a VOE task, not an affiliative-preference task. Infants in Experiment 3 were not choosing which *woman* they preferred, but rather which *event* they wanted to see again. Thus, infants in the crying condition were significantly more likely to trigger the ignore event not because they preferred the woman in that event (it is difficult to imagine on what basis they would do so) but because they were puzzled by her behavior and wanted to see it again, in an attempt to make sense of it. Their responses were thus similar to those of the 12- and 4-month-olds in the crying conditions of Experiments 1 and 2, who were tested with a standard VOE task and looked significantly longer at the ignore than at the comfort event.

5. General discussion

In three experiments, 12-, 8-, and 4-month-old infants watched videotaped events in which an unfamiliar woman was performing a household task when a baby resting nearby began to cry. At all ages, infants expected the woman to comfort the crying baby, and they detected a violation if she continued her work and ignored the baby instead. At 12 and 4 months, infants were tested using a traditional VOE task, and they looked significantly longer when the woman ignored as opposed to comforted the baby. At 8 months, infants were tested with a novel ITV task, and they significantly chose to watch again the event in which the woman ignored as opposed to comforted the baby. In all three experiments, the effect was eliminated if the baby laughed instead of cried. These results thus bring to light one context in which young infants expect prosocial actions: They expect an adult who is alone with a crying baby to comfort the baby, and they detect a violation if the adult ignores the baby instead.

5.1. How do infants attain their expectation about comforting actions?

What mechanisms might lead infants, by as early as 4 months of age, to expect adults to comfort crying babies? There are several possibilities, derived from different—though not necessarily mutually exclusive—approaches to the development of prosociality and morality more generally.

Empathic concern. According to the empathic-concern approach, sympathetic or empathic concern—the capacity to understand and feel concern for the plights of others—plays a key role in the development of morality (e.g., Davidov et al., 2013; de Waal, 2008; Decety et al., 2011; Eisenberg et al., 2006; Hastings et al., 2006; Hoffman, 2007; Knafo et al., 2008; Vaish & Warneken, 2012). Common assumptions are that the capacity to experience concern for the welfare of others is innate; that it evolved from natural selection pressure for the care of offspring and other family members; and that with development, empathy-related responding incorporates increasingly sophisticated emotional, cognitive, and behavioral components.

In line with this approach, it might be suggested that infants in the present experiments (a) felt empathic concern for the crying baby, (b) assumed that the woman near the baby felt the same concern, (c) expected her to act on this concern, and hence (d) detected a violation when she chose to ignore the baby instead.

Although this interpretation is plausible, there is some reason for doubting it. Recall that infants in Experiments 1 and 2 showed very little affective concern when they heard the baby cry in the test events (all means were between 1 = absent and 2 = slight). Moreover, correlations between these concern ratings and infants' differential looking times at the ignore and comfort events were not significant. Across the two experiments, 14/21 (67%) infants who showed some concern in one or both test trials looked longer at the ignore than at the comfort event, whereas 15/17 (88%) infants who showed no visible concern at all did so. These results provide little support for the suggestion that infants' expectations about the woman's actions were driven primarily by their empathic concern for the crying baby.

Acquired expectation. According to the acquired-expectation approach, infants detect regularities, as they observe and interact with their social environments, in how individuals typically behave toward others in different contexts. Learning may occur through domain-general statistical-learning mechanisms (e.g., Paulus, 2014; Paulus & Moore, 2012; Perner, 2010; Ruffman et al., 2012), through domain-specific mechanisms designed to facilitate the acquisition of the social rules that prevail in the local environment and to motivate compliance with these rules (e.g., Sripada & Stich, 2006), or through socialization processes that enable children to detect, internalize, and adhere to social rules (e.g., Killen & Cooley, 2014; Killen et al., 2013).

In line with this approach, it might be suggested that infants in the present experiments had already detected a particular social rule, "adults attempt to comfort babies in distress", based on the caring ministrations they had received from their own caregivers and on similar ministrations they had observed directed at other infants. As a result, infants detected a violation in the ignore event when the woman deviated from this rule.

Although this explanation is plausible, it does raise questions about how it might apply to very young infants. The percentages of infants who, when the baby cried, detected a violation in the ignore event (i.e., looked longer at that event or triggered it more often) were very similar across ages: 73% (16/22) of the 12-month-olds in Experiment 1, 78% (14/18) of the 4-month-olds in Experiment 2, and 72% (13/18) of the 8-month-olds in Experiment 3 did so. This suggests that by 4 months of age, most infants had detected the regularity tested here and had generalized it sufficiently to extend it to the unfamiliar woman and baby in the events. To support this suggestion, evidence would be needed that infants age 4 months or younger can already acquire and generalize social rules.

Sociomoral expectation. According to the sociomoral-expectation approach, the "first draft" of human moral cognition (Graham et al., 2013) includes a small set of abstract sociomoral principles that guides early reasoning about how individuals will act toward others (e.g., Baillargeon et al., 2015; Brewer, 1999; Cosmides & Tooby, 2013; Graham et al., 2013; Krebs, 2008; Premack, 2007; Rai & Fiske, 2011; Shweder et al., 1997). Common assumptions are that these principles evolved during the millions of years our ancestors lived in small groups of hunter-gatherers, where survival depended on cooperation, and that different cultures implement, stress, and rank-order the principles in different ways, resulting in the diverse moral landscape that exists in the world today.

Although there is much controversy over the number, nature, and contents of these sociomoral principles, two candidates seem particularly relevant to our research. One principle is *harm avoidance*: Individuals will avoid harming other members of the same broad moral circle and will assist those who are suffering, especially the young and the weak (e.g., Graham et al., 2013). In line with this principle, infants in the present experiments might have expected the woman to care for the crying baby because it was a fellow human, young and vulnerable, and clearly distressed. The other principle is *ingroup support*: Individuals will act in ways that support other members of their social groups (e.g., Jin & Baillargeon, 2017). Because our test events depicted interactions between a woman who was alone with a baby, and thus was likely to belong to the same social group as the baby (e.g., to be the baby's mother, aunt, or other caregiver), infants might have expected the woman to care for the crying baby because it was a member of her own social group and in need of comfort.

Future research can evaluate these and related possibilities by showing infants mild distress scenarios involving comfort and ignore events and varying whether the characters in the events are human or not, whether they belong to the same social group or not, and whether the distressed character is a baby or not.[W] Such research would allow us to determine, for example, whether infants view comforting as expected, in the case of humans, whenever an adult finds him- or herself alone with a crying baby, or only when there is some reason to suppose that the adult and the baby belong to the same social group.

Recent findings by Spokes and Spelke (2017), obtained with non-human characters, already support the idea that infants do not generally expect any adult to comfort any crying baby. In a series of experiments, 16-month-olds watched interactions among two adult and three baby characters who all differed in shape and color. Infants detected a violation if a baby approached and danced with a baby who had been soothed by a different adult as opposed to a baby who had been soothed by the same adult. These results

[W] ith older infants, it may also be important to vary the nature of the comforting responses presented. In adult-baby distress scenarios involving human characters, for example, infants may come with age to expect more sophisticated responses from the adult (e.g., picking up and talking to the crying baby) than the simple responses portrayed in the present research (e.g., rocking the baby's stroller).

suggest that infants assumed that each adult was selectively comforting the baby or babies in its own moral circle or social group, and they therefore inferred an affiliation between the babies soothed by the same adult.[A]

5.2. Prior findings on comforting actions

Recall that when the 12- and 13-month-olds tested by Johnson, Biro, and their colleagues were shown an intense adult-baby separation scenario (with faceless oval characters), only infants who experienced more supportive parenting gave evidence that they expected the adult to return to the crying baby (Biro et al., 2015; Johnson et al., 2007, 2010). One interpretation of these results—consistent with the acquired-expectation approach described in the preceding section—is that infants with more supportive parenting had learned from their social environments that caregivers typically respond to crying babies, whereas infants with less supportive parenting had not.

As was suggested in the Introduction, another interpretation of these results—consistent with the sociomoral-expectation approach described above—is that *both* infants with more supportive parenting and infants with less supportive parenting expected the adult to return to the crying baby, in line with abstract and early-emerging expectations of harm avoidance and/or ingroup support. However, infants with less supportive parenting found the violation in the leave event distressing, due to their negative attachment-related experiences, and they therefore tended to look away from it. This interpretation predicted that if shown a milder adult-baby distress scenario, typical samples of infants might detect a violation when the adult failed to respond to the baby. The present experiments supported this prediction: 12-, 8-, and 4-month-olds gave robust evidence that they detected a violation when an adult ignored as opposed to comforted a crying baby. Moreover, infants rarely became distressed and showed at most very slight concern when the baby cried, supporting the idea that we did succeed in creating a relatively mild adult-baby distress scenario.

There are, of course, other ways of reconciling our results with those of Johnson, Biro, and their colleagues (Biro et al., 2015; Johnson et al., 2007, 2010). For example, it could be that the percentages of infants with more supportive parenting were consistently higher in our experiments than in theirs. Thus, more of our infants had learned that caregivers typically comfort crying babies, resulting in positive findings overall. Future research can examine this and other possibilities by presenting infants with more or less supportive parenting with a range of distress scenarios, using both behavioral and physiological measures.

This research will be crucial for furthering our understanding of infants with less supportive parenting. In particular, if it is the case that (a) from a very early age, infants expect individuals to care about their ingroup members and to support their welfare via comforting, helping, and other prosocial actions, and (b) this abstract expectation persists even in the face of unsupportive parenting, then this could help explain why lower quality parental care, which repeatedly violates this expectation, has enduring negative consequences for social competence and mental health (e.g., Fearon, Bakermans-Kranenburg, Van IJzendoorn, Lapsley, & Roisman, 2010; Fraley, Roisman, & Haltigan, 2013; Groh, Roisman, Van IJzendoorn, Bakermans-Kranenburg, & Fearon, 2012; Groh et al., 2014; Jaffari-Bimmel, Juffer, Van IJzendoorn, Bakermans-Kranenburg, & Mooijaart, 2006; Raby, Roisman, Fraley, & Simpson, 2015; Roisman & Fraley, 2012).

5.3. Conclusions

Do infants expect individuals, in some contexts at least, to care for others in need and to act prosocially to restore or maintain their welfare? Or do infants hold no particular expectation about whether individuals will act prosocially or antisocially toward others in need? As was discussed in the Introduction, prior reports had revealed only one prosocial expectation in normative samples of infants: 17-month-olds expected a woman to help another woman from the same social group who needed instrumental assistance, and they detected a violation if she ignored her instead (Jin & Baillargeon, 2017). The present experiments brought to light one additional prosocial expectation: 4-, 8- and 12-month-olds expected a woman alone with a crying baby to comfort the baby, and they detected a violation if she ignored the baby instead.

Our results provide the first demonstration that infants in the first year of life already possess at least one prosocial expectation and, as such, they raise important questions for future research about the mechanisms responsible for the emergence of this expectation. In particular, does this expectation reflect a social rule that is generally acquired early in life, or does it stem from abstract and early-emerging sociomoral principles, such as harm avoidance and/or ingroup support? Our results also call for additional research on the effects of more or less supportive parenting on the prosocial expectation demonstrated here and related expectations. Evidence that such expectations emerge early in life and persist, at least to some degree, in the face of unsupportive parenting, would support a characterization of early morality that goes well beyond the acquisition of local social rules.

Finally, our results introduce to the field of infancy research a new forced-choice VOE method for assessing infants' expectations about events, the ITV method. As shown in Experiment 3, infants chose to watch again an event they viewed as unexpected over an event they viewed as expected. This new method provides investigators with another valuable tool for exploring early prosocial and moral expectations and hence for shedding light on the infant roots of human moral cognition.

[A] nother interpretation of our findings might be this: Perhaps young infants merely expect adults who hear an aversive noise (be it a crying baby or a smoke alarm) and find it distressing to do something about it. Given the results of Johnson, Biro, Spokes and their colleagues, however, this interpretation seems unlikely (Biro et al., 2015; Johnson et al., 2007, 2010; Spokes & Spelke, 2017). In the experiments of Spokes and Spelke, for example, it is unclear why infants would have expected a baby to prefer another baby who was stopped from producing an aversive crying noise by the same adult as opposed to a different adult. In our own experiments, infants also had no evidence that the adult was in any way distressed by the baby's crying, as she turned toward the back of the room as soon as the crying began.

Acknowledgments

This research was supported by a grant from the John Templeton Foundation to R. B. and by a postdoctoral fellowship from the National Research Foundation of Korea (NRF-2015S1A3A2046711) to K.J. The opinions expressed in this article are those of the authors and do not necessarily reflect the views of the John Templeton Foundation. We are very grateful to three individuals who helped us develop the infant-triggered-video task used in Experiment 3: David A. Rosenbaum first suggested the idea for this method (in a conversation with R. B.); Amélie Bernard designed and oversaw the construction of the ITV apparatus; and Scott Fraundorf designed and implemented the Matlab computer program used to run the task (with suggestions from Amélie Bernard, K. J., and R. B.). Finally, we wish to thank Fernando Sanchez Hernandez and Fransisca Ting for suggestions on the manuscript, and the parents and infants who participated in the research.

References

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation.* Hillsdale, NJ: Erlbaum.

Baillargeon, R., Setoh, P., Sloane, S., Jin, K., & Bian, L. (2014). Infant social cognition: Psychological and sociomoral reasoning. In M. S. Gazzaniga & G. R. Mangun (Eds.-in-chief), *The cognitive neurosciences V* (pp. 7–14). Cambridge, MA: MIT Press.

Baillargeon, R., Scott, R. M., He, Z., Sloane, S., Setoh, P., Jin, K., & Bian, L. (2015). Psychological and sociomoral reasoning in infancy. In M. Mikulincer, P. R. Shaver (Eds.), E. Borgida, & J. A. Bargh (Assoc. Eds.), *APA handbook of personality and social psychology: Vol. 1. Attitudes and social cognition* (pp. 79–150). Washington, DC: American Psychological Association.

Baillargeon, R., Scott, R. M., & Bian, L. (2016). Psychological reasoning in infancy. *Annual Review of Psychology, 67*, 159–186.

Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition, 20*, 191–208.

Balliet, D., Wu, Y., & De Dreu, C. K. W. (2014). In-group favoritism and cooperation: A meta-analysis. *Psychological Bulletin, 140*, 1556–1581.

Biro, S., Alink, L. R., Huffmeijer, R., Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2015). Attachment and maternal sensitivity are related to infants' monitoring of animated social interactions. *Brain and Behavior, 5*, e00410.

Bloom, P., & Wynn, K. (2016). What develops in moral development? In D. Barner, & A. S. Baron (Eds.). *Core knowledge and conceptual change* (pp. 347–364). New York: Oxford University Press.

Bornstein, M. H., Arterberry, M. E., Mash, C., & Manian, N. (2011). Discrimination of facial expression by 5-month-old infants of nondepressed and clinically depressed mothers. *Infant Behavior and Development, 34*, 100–106.

Bornstein, M. H., Mash, C., Arterberry, M. E., & Manian, N. (2012). Object perception in 5-month-old infants of clinically depressed and nondepressed mothers. *Infant Behavior and Development, 35*, 150–157.

Bowlby, J. (1969). Attachment and loss: Vol. 1. Attachment. New York, NY: Basic Books.

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues, 55*, 429–444.

Brownell, C. A. (2013). Early development of prosocial behavior: Current perspectives. *Infancy, 18*, 1–9.

Cosmides, L., & Tooby, J. (2013). Evolutionary psychology: New perspectives on cognition and motivation. *Annual Review of Psychology, 64*, 201–229.

Davidov, M., Vaish, A., Knafo-Noam, A., & Hastings, P. D. (2016). The motivational foundations of prosocial behavior from a developmental perspective–Evolutionary roots and key psychological mechanisms: Introduction to the special section. *Child Development, 87*, 1655–1667.

Davidov, M., Zahn-Waxler, C., Roth-Hanania, R., & Knafo, A. (2013). Concern for others in the first year of Life: Theory, evidence, and avenues for research. *Child Development Perspectives, 7*, 126–131.

de Waal, F. B. M. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology, 59*, 279–300.

Decety, J., & Howard, L. H. (2014). A neurodevelopmental perspective on morality. In M. Killen, & J. G. Smetana (Eds.). *Handbook of moral development* (pp. 454–474). (2nd ed.). New York, NY: Psychology Press.

Decety, J., Michalska, K. J., & Kinzler, K. D. (2011). The contribution of emotion and cognition to moral sensitivity: A neurodevelopmental study. *Cerebral Cortex, 22*, 209–220.

Dondi, M., Simion, F., & Caltran, G. (1999). Can newborns discriminate between their own cry and the cry of another newborn infant? *Developmental Psychology, 35*, 418–426.

Dunfield, K. (2014). A construct divided: Prosocial behavior as helping, sharing, and comforting subtypes. *Frontiers in Psychology, 5*, 958.

Dunfield, K., & Johnson, S. C. (2015). Variability in social reasoning: The influence of attachment security on the attribution of goals. *Frontiers in Psychology, 6*, 1487.

Dunfield, K., & Kuhlmeier, V. A. (2013). Classifying prosocial behavior: Children's responses to instrumental need, emotional distress, and material desire. *Child Development, 84*, 1766–1776.

Dunn, J. (2014). Moral development in early childhood and social interaction in the family. In M. Killen, & J. G. Smetana (Eds.). *Handbook of moral development* (pp. 135–159). (2nd ed.). New York, NY: Psychology Press.

Dykas, M. J., & Cassidy, J. (2011). Attachment and the processing of social information across the life span: Theory and evidence. *Psychological Bulletin, 137*, 19–46.

Eisenberg, N., Fabes, R. A., Spinrad, T. L. (2006). Prosocial behavior. In W. Damon & R. M. Lerner (Series Eds.) & N. Eisenberg (Vol. Ed.), Handbook of Child psychology: Vol. 3. Social, emotional, and personality development (6th ed., pp. 646–718). New York, NY: Wiley.

Eisenberg, N., Spinrad, T. L., & Knafo-Noam, A. (2015). Prosocial development. In M. E. Lamb (Vol. Ed.) & R. M. Lerner (Series Ed.), Handbook of child psychology and developmental science, Vol. 3: Social, emotional and personality development (7th ed., pp. 610–656). New York, NY: Wiley.

Fawcett, C., & Liszkowski, U. (2012). Infants anticipate others' social preferences. *Infant and Child Development, 21*, 239–249.

Fearon, R. P., Bakermans-Kranenburg, M. J., Van IJzendoorn, M. H., Lapsley, A., & Roisman, G. I. (2010). The significance of insecure attachment and disorganization in the development of children's externalizing behavior: A meta-analytic study. *Child Development, 81*, 435–456. http://dx.doi.org/10.1111/j.1467-8624.2009.01405.x.

Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature, 454*, 1079–1083.

Fraley, R. C., Roisman, G. I., & Haltigan, J. D. (2013). The legacy of early experiences in development: Formalizing alternative models of how early experiences are carried forward over time. *Developmental Psychology, 49*, 109–126.

Geangu, E., Benga, O., Stahl, D., & Striano, T. (2010). Contagious crying beyond the first days of life. *Infant Behavior and Development, 33*, 279–288.

Geangu, E., Benga, O., Stahl, D., & Striano, T. (2011). Individual differences in infants' emotional resonance to a peer in distress: Self–other awareness and emotion regulation. *Social Development, 20*, 450–470.

Geangu, E., Hauf, P., Bhardwaj, R., & Bentz, W. (2011). Infant pupil diameter changes in response to others' positive and negative emotions. *PLoS ONE, 6*, e27132.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56*, 165–193.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology, 47*, 55–130.

Gredebäck, G., & Melinder, A. (2010). Infants' understanding of everyday social interactions: A dual process account. *Cognition, 114*, 197–206.

Groh, A. M., Fearon, R. P., Bakermans-Kranenburg, M. J., Van IJzendoorn, M. H., Steele, R. D., & Roisman, G. I. (2014). The significance of attachment security for children's social competence with peers: A meta-analytic study. *Attachment & Human Development, 16*, 103–136.

Groh, A. M., & Roisman, G. I. (2009). Adults' autonomic and subjective emotional responses to infant vocalizations: The role of secure base script knowledge. *Developmental Psychology, 45*, 889–893.

Groh, A. M., Roisman, G. I., Van IJzendoorn, M. H., Bakermans-Kranenburg, M. J., & Fearon, R. P. (2012). The significance of insecure and disorganized attachment for children's internalizing symptoms: A meta-analytic study. *Child Development, 83*, 591–610.
Gross, J. T., Stern, J. A., Brett, B. E., & Cassidy, J. (2017). The multifaceted nature of prosocial behavior in children: Links with attachment theory and research. *Social Development, 2017*, 1–18.
Hamlin, J. K. (2013a). Failed attempts to help and harm: Intention versus outcome n preverbal infants' social evaluations. *Cognition, 18*, 451–474.
Hamlin, J. K. (2013b). Moral judgment and action in preverbal infants and toddlers evidence for an innate moral core. *Current Directions in Psychological Science, 22*, 186–193.
Hamlin, J. K. (2014). Context-dependent social evaluation in 4.5-month-old human infants: The role of domain-general versus domain-specific processes in the development of social evaluation. *Frontiers in Psychology, 5*, 614.
Hamlin, J. K., & Wynn, K. (2011). Young infants prefer prosocial to antisocial others. *Cognitive Development, 26*, 30–39.
Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature, 450*, 557–559.
Hamlin, J. K., Wynn, K., & Bloom, P. (2010). Three-month-olds show a negativity bias in their social evaluations. *Developmental Science, 13*, 923–929.
Hamlin, J. K., Wynn, K., Bloom, P., & Mahajan, N. (2011). How infants and toddlers react to antisocial others. *Proceedings of the National Academy of Sciences USA, 108*, 19931–19936.
Hastings, P. D., Zahn-Waxler, C., & McShane, K. (2006). We are, by nature, moral creatures: Biological bases of concern for others. In M. Killen, & J. G. Smetana (Eds.). *Handbook of moral development* (pp. 483–516). Mahwah, NJ: Erlbaum.
Hay, D. F., Nash, A., & Pedersen, J. (1981). Responses of six-month-olds to the distress of their peers. *Child Development, 51*, 1071–1075.
Hoffman, M. L. (2007). The origins of empathic morality in toddlerhood. In C. A. Brownell, & C. B. Kopp (Eds.). *Socioemotional development in the toddler years: Transitions and transformations* (pp. 132–145). New York, NY: Guilford Press.
Hohenberger, A., Elsabbagh, M., Serres, J., de Schoenen, S., Karmiloff-Smith, A., & Aschersleben, G. (2012). Understanding goal-directed human actions and physical causality: The role of mother–infant interaction. *Infant Behavior and Development, 35*, 898–911.
Jaffari-Bimmel, N., Juffer, F., Van IJzendoorn, M. H., Bakermans-Kranenburg, M. J., & Mooijaart, A. (2006). Social development from infancy to adolescence: Longitudinal and concurrent factors in an adoption sample. *Developmental Psychology, 42*, 1143–1153.
Jin, K., & Baillargeon, R. (2017). Infants possess an abstract expectation of ingroup support. *Proceedings of the National Academy of Sciences, 114*, 8199–8204.
Johnson, S. C., Dweck, C. S., & Chen, F. S. (2007). Evidence for infants' internal working models of attachment. *Psychological Science, 18*, 501–502.
Johnson, S. C., Dweck, C. S., Chen, F. S., Stern, H. L., Ok, S. J., & Barth, M. (2010). At the intersection of social and cognitive development: Internal working models of attachment in infancy. *Cognitive Science, 34*, 807–825.
Kanakogi, Y., Inoue, Y., Matsuda, G., Butler, D., Hiraki, K., & Myowa-Yamakoshi, M. (2017). Preverbal infants affirm third party interventions that protect victims from aggressors. *Nature Human Behaviour, 1*, 0037.
Killen, M., & Cooley, S. (2014). Morality, exclusion, and prejudice. In M. Killen, & J. Smetana (Eds.). *Handbook of moral development* (pp. 340–360). (2nd ed.). New York, NY: Psychology Press.
Killen, M., Rutland, A., Abrams, D., Mulvey, K. L., & Hitti, A. (2013). Development of intra- and intergroup judgments in the context of moral and social-conventional norms. *Child Development, 84*, 1063–1080.
Killen, M., & Turiel, E. (1998). Adolescents' and young adults' evaluations of helping and sacrificing for others. *Journal of Research on Adolescence, 8*, 355–375.
Knafo, A., Zahn-Waxler, C., Van Hulle, C., Robinson, J. L., & Rhee, S. H. (2008). The developmental origins of a disposition toward empathy: Genetic and environmental contributions. *Emotion, 8*, 737–752.
Köster, M., Ohmer, X., Nguyen, T. D., & Kärtner, J. (2016). Infants understand others' needs. *Psychological Science, 27*, 542–548.
Kotovsky, L., & Baillargeon, R. (1994). Calibration-based reasoning about collision events in 11-month-old infants. *Cognition, 51*, 107–129.
Kotovsky, L., & Baillargeon, R. (1998). The development of calibration-based reasoning about collision events in young infants. *Cognition, 67*, 311–351.
Krebs, D. (2008). Morality: An evolutionary perspective. *Perspectives on Psychological Science, 3*, 149–172.
Kuhlmeier, V., Wynn, K., & Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological Science, 14*, 402–408.
Lee, Y., Yun, J., Kim, E., & Song, H. (2015). The development of infants' sensitivity to behavioral intentions when inferring others' social preferences. *PLoS ONE, 10*, e0135588.
Levine, M., Prosser, A., Evans, D., & Reicher, S. (2005). Identity and emergency intervention: How social group membership and inclusiveness of group boundaries shape helping behavior. *Personality and Social Psychology Bulletin, 31*, 443–453.
Liddle, M. J. E., Bradley, B. S., & Mcgrath, A. (2015). Baby empathy: Infant distress and peer prosocial responses. *Infant Mental Health Journal, 36*, 446–458.
Martin, A., & Olson, K. R. (2015). Beyond good and evil: What motivations underlie children's prosocial behavior? *Perspectives on Psychological Science, 10*, 159–175.
Olson, K. R., & Spelke, E. S. (2008). Foundations of cooperation in young children. *Cognition, 108*, 222–231.
Paulus, M. (2014). The emergence of prosocial behavior: Why do infants and toddlers help, comfort, and share? *Child Development Perspectives, 8*, 77–81.
Paulus, M., & Moore, C. (2012). Producing and understanding prosocial actions in early childhood. *Advances in Child Development and Behavior, 42*, 275–309.
Perner, J. (2010). Who took the cog out of cognitive science? Mentalism in an era of anti-cognitivism. In P. A. Frensch, & R. Schwarzer (Vol. Eds.), *Cognition and neuropsychology: International perspectives on psychological science: vol. 1*, (pp. 241–261). Hove, England: Psychology Press.
Premack, D., & Premack, A. J. (1997). Infants attribute value +/- to the goal-directed actions of self-propelled objects. *Journal of Cognitive Neuroscience, 9*, 848–856.
Premack, D. (2007). Foundations of morality in the infant. In O. Vilarroya, & F. I. Argimon (Eds.). *Social brain matters: Stances on the neurobiology of social cognition* (pp. 161–167). Amsterdam, The Netherlands: Rodopi.
Raby, K. L., Roisman, G. I., Fraley, R. C., & Simpson, J. A. (2015). The enduring predictive significance of early sensitivity: Social and academic competence through age 32 years. *Child Development, 86*, 695–708.
Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review, 118*, 57–75.
Renno, M. P., & Shutts, K. (2015). Children's social category-based giving and its correlates: Expectations and preferences. *Developmental Psychology, 51*, 533–543.
Roisman, G. I., & Fraley, R. C. (2012). A behavior–genetic study of the legacy of early caregiving experiences: Academic skills, social competence, and externalizing behavior in kindergarten. *Child Development, 83*, 728–742.
Roth-Hanania, R., Davidov, M., & Zahn-Waxler, C. (2011). Empathy development from 8 to 16 months: Early signs of concern for others. *Infant Behavior and Development, 34*, 447–458.
Ruffman, T., Taumoepeau, M., & Perkins, C. (2012). Statistical learning as a basis for social understanding in children. *British Journal of Developmental Psychology, 30*, 87–104.
Sagi, A., & Hoffman, M. L. (1976). Empathic distress in the newborn. *Developmental Psychology, 12*, 175–176.
Scott, R. M., & Baillargeon, R. (2009). Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development, 80*, 1172–1196.
Scott, R. M., He, Z., Baillargeon, R., & Cummins, D. (2012). False-belief understanding in 2.5-year-olds: Evidence from two novel verbal spontaneous-response tasks. *Developmental Science, 15*, 181–193.
Setoh, P., Wu, D., Baillargeon, R., & Gelman, R. (2013). Young infants have biological expectations about animals. *Proceedings of the National Academy of Sciences, 110*, 15937–15942.
Shweder, R. A., Much, N. C., Mahapatra, M., & Park, L. (1997). The "big three" of morality (autonomy, community and divinity) and the "big three" explanations of suffering. In A. M. Brandt, & P. Rozin (Eds.). *Morality and health* (pp. 119–169). New York: Routledge.
Sohr-Preston, S. L., & Scaramella, L. V. (2006). Implications of timing of maternal depressive symptoms for early cognitive and language development. *Clinical Child and Family Psychology Review, 9*, 65–83.
Spokes, A. C., & Spelke, E. S. (2017). The cradle of social knowledge: infants' reasoning about caregiving and affiliation. *Cognition, 159*, 102–116.
Sripada, C., & Stich, S. (2006). A framework for the psychology of norms. In P. Carruthers, S. Laurence, & S. Stich (Vol. Eds.), *Culture and cognition: . vol. 2. The innate mind* (pp. 280–301). Oxford, England: Oxford University Press.

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science, 348*, 91–94.
Stavans, M., & Baillargeon, R. (2018). 4-month-old infants individuate and track simple tools following functional demonstrations. *Developmental Science, 21*, e12500.
Vaish, A., & Tomasello, T. (2014). The early ontogeny of human cooperation and morality. In M. Killen, & J. G. Smetana (Eds.). *Handbook of moral development* (pp. 279–298). (2nd ed.). New York, NY: Psychology Press.
Vaish, A., & Warneken, F. (2012). Social-cognitive contributors to young children's empathic and prosocial behavior. In J. Decety (Ed.). *Empathy: From bench to bedside* (pp. 131–146). Cambridge, MA: MIT Press.
Van de Vondervoort, J. W., & Hamlin, J. K. (2016). Evidence for intuitive morality: Preverbal infants make sociomoral evaluations. *Child Development Perspectives, 10*, 143–148.
Wang, S., Zhang, Y., & Baillargeon, R. (2016). Young infants view physically possible support events as unexpected: New evidence for rule learning. *Cognition, 157*, 100–105.
Warneken, F. (2015). Precocious prosociality: Why do young children help? *Child Development Perspectives, 9*, 1–6.
Wynn, K., & Bloom, P. (2014). The moral baby. In M. Killen, & J. G. Smetana (Eds.). *Handbook of moral development* (pp. 435–453). (2nd ed.). New York, NY: Psychology Press.