Received: 20 November 2020 | Accepted: 15 November 2021

DOI: 10.1111/infa.12447

RESEARCH ARTICLE

INFANCY THE OFFICIAL JOURNAL OF THE INTERNATIONAL CONGRESS OF INFANT STUDIES WILEY

Mechanisms of social evaluation in infancy: A preregistered exploration of infants' eye-movement and pupillary responses to prosocial and antisocial events

Enda Tan | **Jane Kiley Hamlin**

University of British Columbia, Vancouver, British Columbia, Canada

Correspondence
Enda Tan, Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, BC, V6T 1Z4, Canada.
Email: enda.tan@psych.ubc.ca

Funding information
Natural Sciences and Engineering Research Council of Canada, Grant/Award Number: RGPIN-2016-03775

Abstract

Past research shows infants selectively touch and look longer at characters who help versus hinder others (Social evaluation by preverbal infants. *Nature*, 2007, 450, 557; Three-month-olds show a negativity bias in their social evaluations. *Developmental Science*, 2010, 13, 923); however, the mechanisms underlying this tendency remain underspecified. The current preregistered experiment approaches this question by examining infants' real-time looking behaviors during prosocial and antisocial events, and exploring how individual infants' looking behaviors correlate with helper preferences. Using eye-tracking, 34 five-month-olds were familiarized with two blocks of the "hill" scenario originally developed by Kuhlmeier et al. (Attribution of dispositional states by 12-month-olds. *Psychological Science*, 2003, 14, 402), in which a climber tries unsuccessfully to reach the top of a hill and is alternately helped or hindered. Infants' visual preferences were assessed after each block of 6 helping and hindering events by proportional looking time to the helper versus hinderer in an image of the characters side by side. Results showed that, at the group level, infants looked longer at the helper after viewing 12 (but not after viewing 6) helping and hindering videos. Moreover, individual infants' average preference for the helper was predicted by their looking behaviors, particularly those suggestive of an understanding of the climber's unfulfilled goal. These results shed light on how infants process helping/hindering scenarios, and suggest that goal understanding is important for infants' helper preferences.

1 | INTRODUCTION

The tendency to evaluate others based on how they interact with third parties, even from the perspective of an independent bystander, is a human universal (Brown, 1991), and there exists high levels of agreement as to what kinds of interactions are considered positive around the world (Curry et al., 2019). Where does this tendency come from? Answering this question requires understanding the developmental origins of sociomoral evaluation. Recently, a growing literature suggests that preverbal infants may evaluate others based on their prosocial and antisocial behaviors, preferring those who engage in prosocial versus antisocial acts over those who engage in antisocial ones (for recent reviews, see Hamlin & Sitch, 2020; Ting et al., 2020). Although this work suggests that some tendency to engage in social evaluation emerges early in life, the nature of these evaluations has been hotly contested (Scarf et al., 2012a, 2012b; Tafreshi et al., 2014). This debate may stem from the fact that, at a mechanistic level, it is insufficiently clear what mental processes support these evaluations. The current preregistered study aimed to address this issue, using eye-tracking to explore the mechanisms underlying infants' responses to sociomoral scenarios.

Infants' evaluation of sociomoral events was first documented using the "hill" scenario (Hamlin et al., 2007; adapted from Kuhlmeier et al., 2003). In these studies, infants watched a live puppet show in which a climber repeatedly tried but failed to reach the top of a steep hill. On his third attempt, the climber was consistently helped (pushed up the hill) by a prosocial character and hindered (pushed down the hill) by an antisocial character. After being habituated to alternating helping/hindering events, 6- and 10-month-old infants preferentially reached toward the helper versus hinderer (Hamlin et al., 2007), and both 6- and 3-month-olds (who are too young to reach) looked longer at the helper versus hinderer (Hamlin et al., 2010). These findings were taken to suggest that young infants evaluate prosocial characters more positively than antisocial ones.

Infants' preference for prosocial characters has been documented via scenarios depicting additional forms of help/harm (Buon et al., 2014; Hamlin & Wynn, 2011; Kanakogi et al., 2017; Scola et al., 2015) and prosocial/antisocial acts in other sociomoral domains such as fairness (e.g. DesChamps et al., 2015; Geraci & Surian, 2011; Lucca et al., 2018). These studies suggest that infants' prosocial preferences emerge across a variety of physically distinct interactions that adults would also positively and negatively evaluate. In addition, various studies of infants' prosocial preferences have included nonsocial control conditions, in which the same physical acts (e.g., pushing up versus down hills) are directed toward inanimate entities incapable of possessing unfulfilled goals and so unworthy of being treated prosocially or antisocially (e.g., Hamlin et al., 2007; Hamlin et al., 2010; see also Geraci & Surian, 2011; Hamlin & Wynn, 2011). In these conditions, infants have routinely chosen randomly between characters, suggestive that preferences for prosocial over antisocial characters in the original conditions were in response to social features of the displays. This social interpretation is also supported by a recent event-related potential (ERP) study, in which 6-month-olds were familiarized to helping/hindering events and subsequently shown iterated images of the helper and the hinderer (Gredebäck et al., 2015). Infants' neural responses to helper versus hinderer images revealed reliable differences in the P400 component, thought to index social perception in the infant brain. Together, these findings have been interpreted by some as evidence for social (and indeed moral) evaluations in infancy (Hamlin, 2013b; Ting et al., 2020; Wynn & Bloom, 2014).

On the contrary, there has been considerable debate about both the robustness and the nature of infants' responses in these studies. First, some studies have failed to find that infants prefer prosocial to antisocial characters at all (Cowell & Decety, 2015; Nighbor et al., 2017; Schlingloff et al., 2020), including when methods are kept very close to the original studies (Salvadori et al., 2015; for review and meta-analyses, see Holvoet et al., 2016; Margoni & Surian, 2018). Second, some researchers have argued that although infants may show a general tendency to reach toward helpers, this tendency need not reflect evaluations of the *social* value of others' acts. Instead, infants' choices could be based on various appealing and aversive perceptual features of the displays (Scarf et al., 2012a, 2012b; cf. Hamlin, 2015; Hamlin et al., 2012). For instance, Scarf et al. (2012b) ran several conditions, suggestive that infants' helper preferences in the hill scenario may be due to the climber bouncing upon reaching the top of the hill, a positive perceptual event, rather than a preference for prosociality. Indeed, in a condition wherein the climber bounced at both the top and the bottom of the hill, infants chose randomly between the helper and hinderer, suggestive that bouncing may influence infants' character choices. In response, Hamlin (2015) ran several additional conditions, hypothesizing that infants' preferences in Scarf et al. (2012b) study stemmed from ambiguous cues to goal-directedness in their climber whose eyes were unfixed and so did not point up the hill toward its goal. Indeed, infants in Hamlin's (2015) studies selectively chose the helper anytime the climber's eyes were fixed looking upward, whether or not bouncing occurred (for related evidence see Lee et al., 2020).

As the previous paragraphs illustrate, although some studies purport to demonstrate high-level tendencies for social evaluation in the first year of life, others suggest either that such tendencies do not exist or that they are rooted in perceptual rather than social processes. How might we adjudicate between these possibilities? One common approach in the social evaluation literature has involved running experiments in which different conditions pit competing accounts against each other, as in the work by Scarf et al. (2012b) and Hamlin (2015) described above. Unfortunately, to date this approach has not been entirely fruitful, given that the different papers have found evidence for both higher-level social and lower-level perceptual interpretations. Further, it is exceedingly difficult to ensure that distinct conditions are matched along every single possibly relevant dimension, which limits the effectiveness of this approach. Another potentially fruitful approach is to explore long-term correlates of individual differences in infants' preferences, examining whether or not differences in sociomoral evaluation during infancy predict meaningful differences in later social development. A recent paper demonstrating individual-level correlations between infants' evaluations and preschool measures of social and moral functioning supports the hypothesis that infants' responses are meaningful (e.g., Tan et al., 2018); however, this study involved only a very small sample of infants who participated in a wide variety of infant measures.

The current study adopts a third approach to determining the processes underlying infants' helper preferences, attempting to draw inferences about the nature of infants' preferences based on a detailed analysis of individual infants' real-time responses throughout a study. Specifically, by examining whether and how individual infants' real-time responses during both helper/hinderer events and preference measurements relate to differences in the tendency to prefer helpful characters, one can attempt to better elucidate the mechanisms underlying infants' social evaluations. This approach is facilitated by the use of eye-tracking, which provides moment-to-moment measures of infants' fixations (maintenance of eye gaze on specific locations), fixation shifts (gaze shifts between fixations), and pupil size. Indeed, previous studies have productively utilized eye-tracking measures to provide information about infants' preferences, expectations, relational processing, and social evaluations (Bornstein et al., 2011; Fawcett & Liszkowski, 2012; Geraci & Surian, 2011; Holvoet et al., 2019; Keemink et al., 2019; for review, see Aslin, 2012; Gredebäck et al., 2009; Oakes, 2012), and individual differences in how infants process scenes predict meaningful differences in performance on cognitive tasks (Johnson et al., 2004; Yu & Smith, 2011).

To our knowledge, only one paper to date has examined the links between infants' real-time looking behaviors and their preferences for helpers versus hinderers. In this study, Cowell and Decety (2015) recorded eye movements, while 12- to 24-month-old infants watched an animated version of the hill scenario (and pictures depicting various other prosocial/antisocial interactions). They found that infants' overall attention to the hill display did not differ between helping/hindering events, nor did attentional differences to helping versus hindering events predict infants' tendency to reach for the helper versus the hinderer (which, at a group level, infants did not do). However, an analysis of particular areas of interest (AOIs) revealed that during helping events, infants looked longer at the agent (the helper) than at the recipient (the climber), but during hindering events, infants looked equally at the agent (the hinderer) and the recipient (the climber). These results suggest that infants may be more attentive toward agents of positive versus negative acts and that attention to helping/hindering events may differ in kind. These results call for further study to examine how differences in patterns of attention may relate to infants' helper preferences.

The current study aimed to provide a more comprehensive account of how infants process and respond to helping/hindering scenarios. To this end, we made several changes to the paradigms used by both Cowell and Decety (2015) and Hamlin et al. (2007), Hamlin et al. (2010), in order to maximize both procedural standardization and meaningful individual differences. These design features were preregistered on Open Science Framework (OSF)—see Supplemental Materials for OSF link and 4 changes to the preregistration. First, we utilized an animated version of the hill scenario to maximize control over stimulus luminance, perceptual features, and event timing. Second, we familiarized (vs. habituated) babies to the stimuli (2 rounds of 6 events) and fixed the amount of time infants were given to process each event after the movements ended (5s, versus 60s in early studies of the hill paradigm). We chose familiarization rather than habituation because habituation procedures and long trial duration are typically utilized in order to ensure that infants have fully processed an event (e.g., Colombo & Mitchell, 1990); we reasoned that familiarization and short trial duration might lead infants to process our events to different extents, and therefore maximize our changes of observing meaningful relationships between looking behaviors and preferences. Third, we assessed infants' preferences visually instead of manually: Visual preference measures are well suited for eye-tracking and provide richer information than would a one-shot reaching preference measure. Given these changes, this study should not be interpreted as a direct replication of either Cowell and Decety (2015) or Hamlin et al. (2007), Hamlin et al. (2010).

In addition to paradigm changes, we also recruited relatively young participants (5-month-olds) and a narrow age range (1 month). Past research has demonstrated that infants around this age not only reach for but also look longer at helpers versus hinderers in the hill paradigm (Hamlin et al., 2010), providing evidence for links between visual and manual preferences in this paradigm at this age. Further, a recent meta-analysis suggests that infants' preferences are stable across infancy from 4 months of age (Margoni & Surian, 2018). Thus, we reasoned that it would be possible to observe meaningful individual differences in preferences for helpers in this age group and that it would be appropriate to relate 5-month-olds' visual responses to the broader sociomoral evaluation literature.

Analysis strategies for exploring infants' looking behaviors were also preregistered. First, we divided the hill scenario into different phases to assess how infants responded to different aspects of the prosocial/antisocial events, in order to better delineate the dynamic process of infants' reactions. Second, we analyzed infants' looking behaviors toward and between several AOIs, including the moving characters (the climber, helper, and hinderer), as well as the top of the hill (the climber's goal; see Figure 1). We were particularly interested in looks from the climber to the hilltop, as fixation shifts have been utilized to explore infants' relational processing (Bornstein et al., 2011; Rennels & Cummings, 2013), and because past work suggests that infants visually predict agents' action goals in some contexts from around 6 months of age (Adam & Elsner, 2020; Gredebäck et al., 2018; Kanakogi & Itakura, 2011; Kochukhova & Gredebäck, 2010; Monroy et al., 2020; for review, see Elsner & Adam, 2020). Specifically, we reasoned that whereas fixations on the climber could be driven by perceptual features of the displays (i.e., attention being drawn to a moving object), anticipatory looking from the climber to the hilltop during the climber's ascent would be more likely to reflect high-level social processing (i.e., anticipation of the climber's goal; though see Discussion for alternative interpretations). Based on past work suggestive that the ability to perceive the climber's goal is critical to infants' group-level helper preferences (Hamlin, 2015), we hypothesized that there might be a link between climber–hilltop fixation shifts and individual infants' helper preferences. Finally, we examined infants' fixation shifts between the helper and hinderer during preference trials.

In addition to measuring how infants distributed their attention around the displays, we also measured infants' pupil size during and after prosocial/antisocial scenarios. Pupil size is influenced by the activity of the autonomous nervous system (Hepach & Westermann, 2016). Past research reveals that children show an increase in pupil size when seeing surprising or inconsistent events (Krüger et al., 2020), as well as when seeing others in need of help (notably, pupil size decreases after children either help or observe a third-party help, the needy person; for a review, see Hepach, 2017). Thus, in the current study we explored changes in infants' pupil diameter during and after helping and hindering events, and whether these changes are associated with infants' social preferences.

Overall, the study was designed to address three main questions. First, we compared eye-tracking measures between event types (helping/hindering) to explore whether infants process prosocial/antisocial scenarios differently. Second, we examined whether infants, as a group, show visual preferences for prosocial (vs. antisocial) characters during preference trials. Finally, and critically, we correlated various eye-tracking measures with visual preferences for the prosocial character, to examine links between individual infants' looking behaviors, pupillary responses, and social preferences.

2 | METHOD

2.1 | Participants

Participants were 34 full-term and healthy infants (20 females; mean age = 5.13 months, range = 4.53–5.73) living in a North American city. Most of the participants were from Caucasian

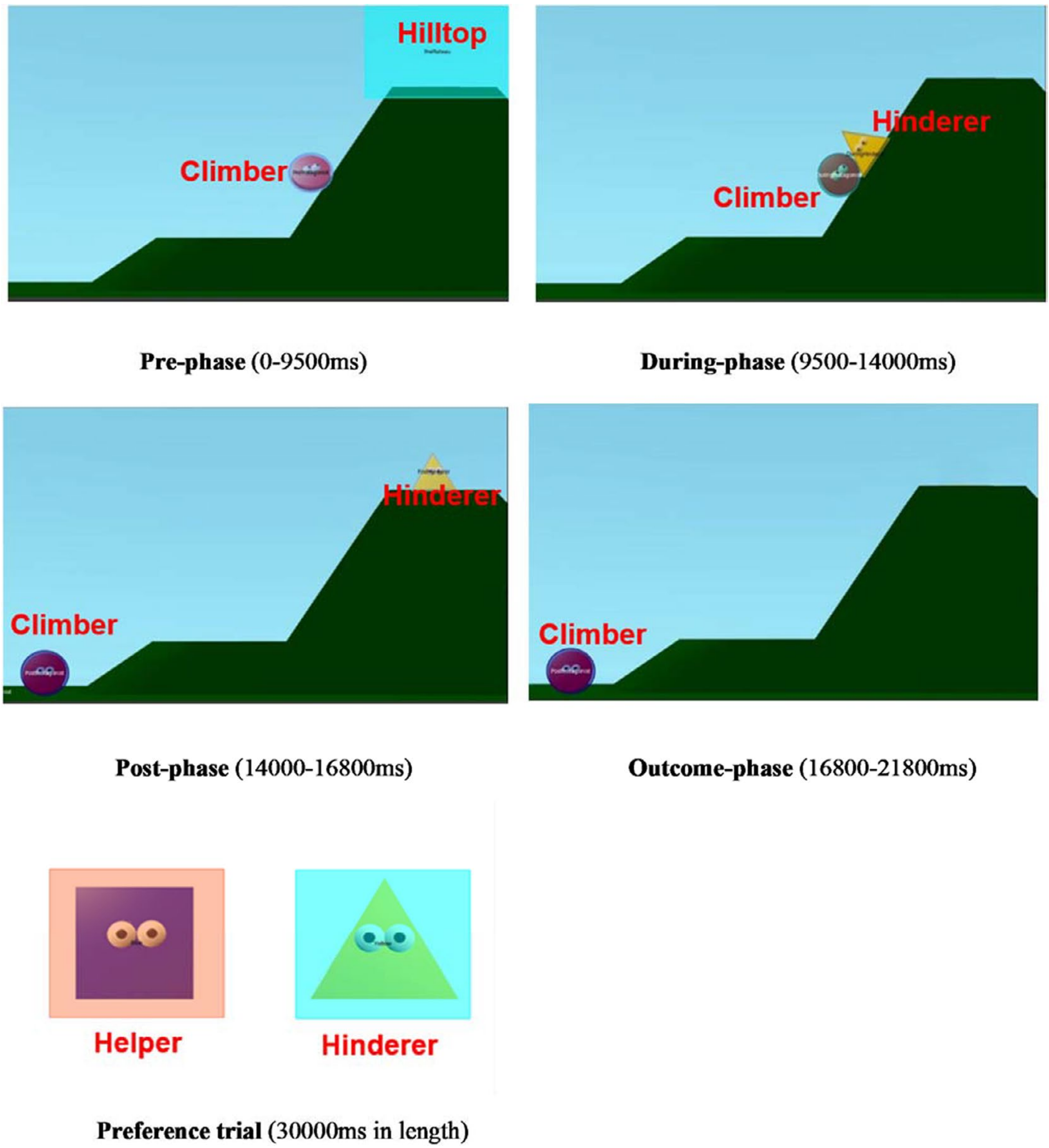

FIGURE 1 Study stimuli and areas of interest (AOIs). *Note*. For videos, AOIs were created covering the contours of the climber, the helper, and the hinderer with 0° margin

and Asian backgrounds, representative of the population in the area. An additional 17 infants who reached calibration were excluded due to poor tracking ratio ($N = 7$; see below for details), inaccurate calibration ($N = 6$; see below for details), fussiness ($N = 3$), and procedure errors ($N = 1$). The sample size was determined based on sample sizes of past studies that found a group-level preference for the helper using the hill paradigm ($N = 16$ in Hamlin et al., 2007; $N = 24$ in Hamlin et al., 2010 and Hamlin, 2015), and increased given our interest in individual differences. The sample size, inclusion criteria, procedure, and analysis plan of the study were preregistered (see Supplemental Material 1). Two deviations from preregistration are described and justified in Supplemental Materials 2.1–2.2. The study was conducted according to guidelines laid down in the Declaration of Helsinki, with written informed consent obtained from a parent or guardian for each child before any data collection. All procedures were approved by the Behavioral Research Ethics Board at the University of British Columbia.

2.2 | Apparatus and room setup

The stimuli were presented on a 24" LCD screen (Dell E2414H; width: 1920 pixels, 44° visual angle; height: 1080 pixels, 29° visual angle). Infants sat either in a highchair or on parent's lap (if they felt uncomfortable with the highchair) approximately 60cm from the screen. Infants' eye gaze and pupil size were recorded using an SMI REDn Scientific remote eye-tracker (sampling rate = 60 Hz) attached underneath the screen. SMI reports a gaze position accuracy of 0.4°, a spatial resolution (root mean square) of 0.05°, and a tracking range (head box) of 50 × 30 cm (SensoMotoric Instruments, 2014). Parents either stood/sat behind the infant or (if the infant felt uncomfortable not being able to see their parent) sat beside the screen looking away from the infant and the screen, and were instructed not to interfere with their infants. The position of the parent (behind the infant vs. beside the screen) did not affect tracking ratio or any other key eye-tracking measures, *p*s > .05. Stimulus presentation was controlled by the experimenter in an adjacent room using the SMI Experiment Center (version 3.7). Infants' behaviors were videotaped and monitored by the experimenter throughout the study. Room lighting was kept constant across participants.

2.3 | Calibration

Prior to stimulus presentation, the experimenter played a short video to attract infants' attention to the screen while making sure that infants' eyes were detected by the eye-tracker. The eye-tracker was calibrated using a 5-point calibration procedure, with a looming red dot showing at the center and four corners of the screen. Calibration was deemed successful if, in a minimum of four out of five locations, the offset was less than 1° for both eyes. For infants who were inattentive, the calibration was repeated until satisfactory results were obtained. Calibration accuracy was further confirmed after testing through visual inspection of gaze replay videos; calibration was deemed successful if infants' eye gaze coordinates showed no consistent deviations from salient stimuli (e.g., moving characters). Participants who showed consistent deviations were excluded from analyses ($N = 6$). After calibration, stimulus presentation began.

2.4 | Procedure and stimuli

Infants viewed two blocks of familiarization and preference trials presented in alternation; participants in the final sample completed all trials, except for one participant who became fussy and so viewed only one block of trials. This participant was included in the final sample based on preregistered exclusion criteria, and the participant's data were adjusted for trial number differences.

Each block (about 3 min in length) consisted of three helping and three hindering videos presented in a counterbalanced order, followed by a 30-s preference trial. Each familiarization trial started with an attention getter. The helping/hindering videos (21,800 ms in length) began after infants looked at the attention getter for 2000 ms consecutively. Videos were created using

Blender (www.blender.org; see Figure 1) and had equal luminance (see Supplemental Material 3). At the outset of each video, a climber (a red circle with googly eyes; diameter = 4° visual angle) appeared on the left side of a steep hill. The climber climbed the first half of the hill, reaching a small plateau in the middle of the hill (0–1000 ms), and jiggled up and down twice (1000–2000 ms). The climber then made two unsuccessful attempts to climb the second half of the hill (2000–9000 ms), each time falling back to the middle plateau. On his third attempt, a helpful/unhelpful character (a blue square or a yellow triangle with googly eyes; diameter ≈ 4° visual angle) entered the scene. In the helping event, the helper emerged from the bottom of the hill (9000–9500 ms), assisted the climber by pushing him up the hill (9500–13,000 ms), and exited the scene from the bottom of the hill (13,000–16,800 ms). In the hindering event, the hinderer emerged from the top of the hill (9000–9500 ms), hindered the climber by pushing him down the hill (9500–13,000 ms), and exited the scene from the top of the hill (13,000–16,800 ms). The climber did not bounce upon reaching the hilltop during helping events, nor did he roll end over end to the bottom of the hill during hindering events (see Scarf et al., 2012b). The videos ended with a still image depicting the climber on the top of the hill (in the helping event) or on the bottom of the hill (in the hindering event) for 5000 ms (16,800–21,800 ms). Throughout each video, the climber's eyes were fixated at the top of the hill during his failed attempts, and the helper and hinderer's eyes were fixated at the climber. The emergence of the climber and helper/hinder was accompanied by a "ding" sound, and the helper/hinderer's pushing movements were accompanied by a knocking sound. During preference trials, infants viewed a still image depicting the helper and hinderer placed side by side against a white background for 30,000 ms (diameter of each character ≈ 11.31° visual angle).

Within participants, the position of the helper and the hinderer was counterbalanced across the two preference trials. The following factors were counterbalanced across participants: the color/shape of the helper and the hinderer (blue square or yellow triangle), the order of helping and hindering videos, and the position of the helper and hinderer during preference trials. Participants showed no group-level visual preferences for the blue square or the yellow triangle across preference trials, $p = .412$, or within the first/second preference trial, $ps > .377$.

2.5 | Eye-movement analyses

Data processing was performed using the SMI BeGaze software (3.7) and R software (3.6.2; packages in Supplemental Material 4). Raw data of the study are deposited on OSF (see Supplemental Material 1 for OSF link). We used tracking ratio (the proportion of time the eye-tracker registered eye gaze over the entire study) to assess data quality. Participants with tracking ratio below 30% were deemed unreliable (see Amso et al., 2014; Havy & Zesiger, 2017; Korhonen et al., 2018) and excluded from analyses. The final sample had an average tracking ratio of 47.17% with an SD of 11.97%. Considering the young age of our participants and the long durations of each of our videos (21.8 s) and picture stimuli (30 s), we considered a tracking ratio of 47.17% to be satisfactory (see, for comparison, Gulz et al., 2020, in which preschoolers had an average tracking ratio of 50.4%, and LoBue et al., 2017, who used a tracking ratio cut-off of 15% with 4- to 24-month-olds; average tracking ratio was not reported). We also note that the main findings of the study remained largely unchanged when infants with tracking ratios above and below the median are analyzed separately (see Supplemental Material 5) and that no correlations were found between tracking ratio and any main measures of the study. These results suggest that tracking ratio neither moderated nor mediated reported effects.

The helping/hindering videos were divided into several phases (one deviation from preregistration is described and justified in Supplemental Material 2.3), including Baseline (attention-getter period immediately before the video; −2000 to 0 ms), Pre-phase (video outset to helper/hinderer appearance, includes climber's failed attempts; 0–9500 ms), During-phase (helper/hinderer appearance to end of helping/hindering action; 9500–14,000 ms), Post-phase (end of helping/hindering action to helper/hinderer's exit; 14,000–16,800 ms), and Outcome (climber resting at the top/bottom of hill; 16,800–21,800 ms). AOIs were created covering the contours of the climber, the helper, and the hinderer with 0° margin (see Figure 1). We decided to use smaller AOIs (i.e., no buffer margin) for characters due to lack of space between AOIs; this practice balances the ratio of true- and false-positive detections of fixations (Orquin et al., 2016). An additional AOI covered the top of the hill during helping/hindering videos, and during preference trials, AOIs covered the helper and the hinderer. Fixations were detected using SMI's default algorithm. Data from the left eye were used to estimate gaze coordinates. Fixations were defined as gazes within a 100-pixel radius for at least 80ms. Fixation shifts were defined as gaze shifts between fixations on AOIs. Specifically, fixation shifts between X and Y (AOIs) would be calculated as the number of fixation changes from X to Y plus the number of fixation changes from Y to X.

Eye-movement analyses addressed three main questions. First, to explore whether infants processed prosocial/antisocial scenarios differently, we compared infants' looking behaviors between conditions, including (a) overall looking to the videos, (b) looking at each character in each phase, (c) fixation shifts between characters during the During- and Post-phases, and (d) hilltop looking behaviors, including fixation shifts between the climber and hilltop (preregistered measure) and looking time at the hilltop (non-preregistered measure; see Supplemental Material 2.4 for justification) during the Pre-phase. Second, to assess visual preferences, we compared infants' looking times on the helper and hinderer during preference trials. Finally, to explore what aspects of looking behaviors predicted infants' preferences for the prosocial character, we extracted individual difference measures based on looking times and fixation shifts (see Table 1 for measures), and correlated these measures with infants' proportional looking times to the helper during preference trials. We calculated measures of looking behaviors separately for helping/hindering events (except for the Pre-phase, where stimuli were identical) to explore whether looking behaviors to different events showed distinct correlations with visual preferences.

2.6 | Pupillary analyses

For pupillary analyses, preprocessing procedures, parameters, and R scripts were adapted from Hepach et al. (2012). We first removed samples wherein the pupil was obscured by blinking. Next, pupil size data for each eye were filtered by excluding sample-to-sample differences that exceeded the 90th percentile of the data sequence, and gaps between samples that were no greater than 4 samples (66.67 ms) were linearly interpolated. The processed data from the left and right eyes were then averaged, and the averaged values were filtered and interpolated again using the same procedure.

After preprocessing, average pupil sizes and relative change scores (pupil size [phase] – pupil size [Baseline]) were calculated for different conditions and phases (see Supplemental Material 6 for supplementary analysis using the Pre-phase as the baseline, and Supplemental Material 7 for supplementary analysis using time windows adjusted for infants' slower pupil reactions). Pupillary analyses addressed two questions: First, to explore whether infants showed different sympathetic responses to prosocial/antisocial scenarios, we compared relative pupil size changes

TABLE 1 Spearman's correlations between looking behaviors and visual preferences for the helper

Type	Measure	Correlation with proportional looking time on the helper
Looking time	Helping scenario (overall)	.07
	Hindering scenario (overall)	.00
	Preference trials (overall)	.41*
	Pre: climber	.02
	Pre: hilltop	.52**
	During: climber (helping)	.10
	During: climber (hindering)	.14
	During: helper	.17
	During: hinderer	.25
	Post: climber (helping)	−.04
	Post: climber (hindering)	.12
	Post: helper	.00
	Post: hinderer	.06
	Outcome: helpee	.07
	Outcome: hinderee	−.07
Fixation shift	Pre: climber–hilltop	.46**
	During: climber–helper	.16
	During: climber–hinderer	.24
	Post: climber–helper	.31
	Post: climber–hinderer	.10
	Preference trial: helper–hinderer	.37*

Note: $^*p < .05$. $^{**}p < .01$ (2-tailed).

in different phases between prosocial/antisocial events. Second, to explore whether infants' pupillary responses predicted visual preferences for the prosocial character, we extracted individual difference measures based on relative pupil size changes in different phases, and correlated these measures with infants' visual preferences for the helper during preference trials.

3 | RESULTS

3.1 | Did infants process helper/hinderer scenarios differently?

We compared infants' total looking times and number of fixation shifts across trials between prosocial/antisocial scenarios using paired-samples t tests. When Shapiro–Wilk tests showed violations of normality, we performed Wilcoxon signed-rank tests. The results showed that, during the Post-phase, infants looked longer at the climber ($M_{hindering} = 2749.39$ ms, $M_{helping} = 1268.98$ ms), Wilcoxon exact test, $V = 118$, $Z = 2.67$, $p = .007$, and showed more fixation shifts between the climber and the actor ($M_{hindering} = 1.85$, $M_{helping} = 0.56$), Wilcoxon exact test, $V = 36$, $Z = 3.39$, $p < .001$, during hindering versus helping events. Following Cowell and Decety (2015), we also compared infants' looking time on the climber and the helper/hinderer within each condition.

The results showed that infants looked longer at the climber (hinderee) than at the hinderer ($M_{hinderee}$ = 6394.24 ms, $M_{hinderer}$ = 3365.00 ms), Wilcoxon exact test, $V = 83$, $Z = 3.67$, $p < .001$, but did not look differently between the climber (helpee) and the helper. Taken together, these findings suggest that infants pay more attention to the climber and to the relationship between the climber and the actor during antisocial (vs. prosocial) events.

To compare pupillary responses between prosocial/antisocial events, we entered baseline-corrected pupil size changes into a 2 (condition: prosocial vs. antisocial) × 4 (phase: Pre, During, Post, Outcome) repeated-measures ANOVA. As shown in Figure 2, there was a main effect of phase, $F(2.17, 71.63) = 133.60$, $p < .001$, $\eta^2 = .80$, qualified by a significant interaction between condition and phase, $F(2.53, 83.51) = 26.35$, $p < .001$, $\eta^2 = .44$. Post hoc tests revealed that pupil size was smaller for hindering (versus helping) scenarios during the Post-phase ($M_{hindering}$ = 0.06 mm, $M_{helping}$ = 0.23 mm), $t(33) = 5.12$, $p < .001$, $d = .88$, 95% CI [0.098, 0.227], but not during the Pre-, During-, and Outcome phases. This effect is unlikely to be attributable to general luminosity differences because the luminance of the screen and the room was kept constant across conditions (see Discussion below). These findings suggest that infants showed lower levels of arousal immediately following the hindering (versus the helping) act, as the character left the scene.

3.2 | Did infants show visual preferences for the helper during preference trials?

We compared infants' total looking times on the two characters using paired-samples *t* tests. When Shapiro–Wilk tests showed violations of normality, we performed Wilcoxon signed-rank tests. The results showed that across the two preference trials, infants did not look longer at the helper over the hinderer, $t(33) = 0.80$, $p = .43$, $d = .12$, 95% CI [−1112.41, 2553.71]. Infants' individual patterns of response showed the same result: 17 of 34 infants preferred the helper across both events (binomial test, $p > .99$).

As a non-preregistered/exploratory analysis, we performed block-specific tests to examine infants' visual preferences during each of the two preference trials. We reasoned that infants might show different looking patterns between the two preference trials due to different amounts of exposure to helping/hindering videos (6 trials versus 12 trials). Specifically, infants might have gotten bored during the second block of presentation and hence shown a weaker effect during the second preference trial; alternatively, infants might not have had sufficient exposure to the stimuli during the first block and therefore shown a weaker effect during the first preference trial. Results showed that although infants did not look longer at the helper during the first preference trial, $p = .67$, they did look longer at the helper (versus the hinderer) during the second preference trial (M_{helper} = 4390.79 ms, $M_{hinderer}$ = 3386.51 ms; see Figure 3), Wilcoxon exact test, $V = 410$, $Z = 2.31$, $p = .02$; individual patterns of response: 23 of 34 infants preferred the helper (binomial test, $p = .058$). These results suggest that infants' visual preferences for the helper (versus hinderer) may require more than 6 exposures to the scenarios (i.e., more than 3 of each event). Indeed, an exploratory analysis of infants' average looking times to the (still) Outcome phase across events (see Figure 1) suggests that although infants' attention decreased over the course of the experiment, the overall level of decrease from the first to the last events was rather low: Infants looked an average total of 6107.46 ms following the first three familiarization trials and an average total of 4107.81ms following the last three familiarization trials, a decrease of only 32.74%. This result suggests that infants may have still been processing the events at the end of the second block, and the null result of the first preference trial (and overall) could have been

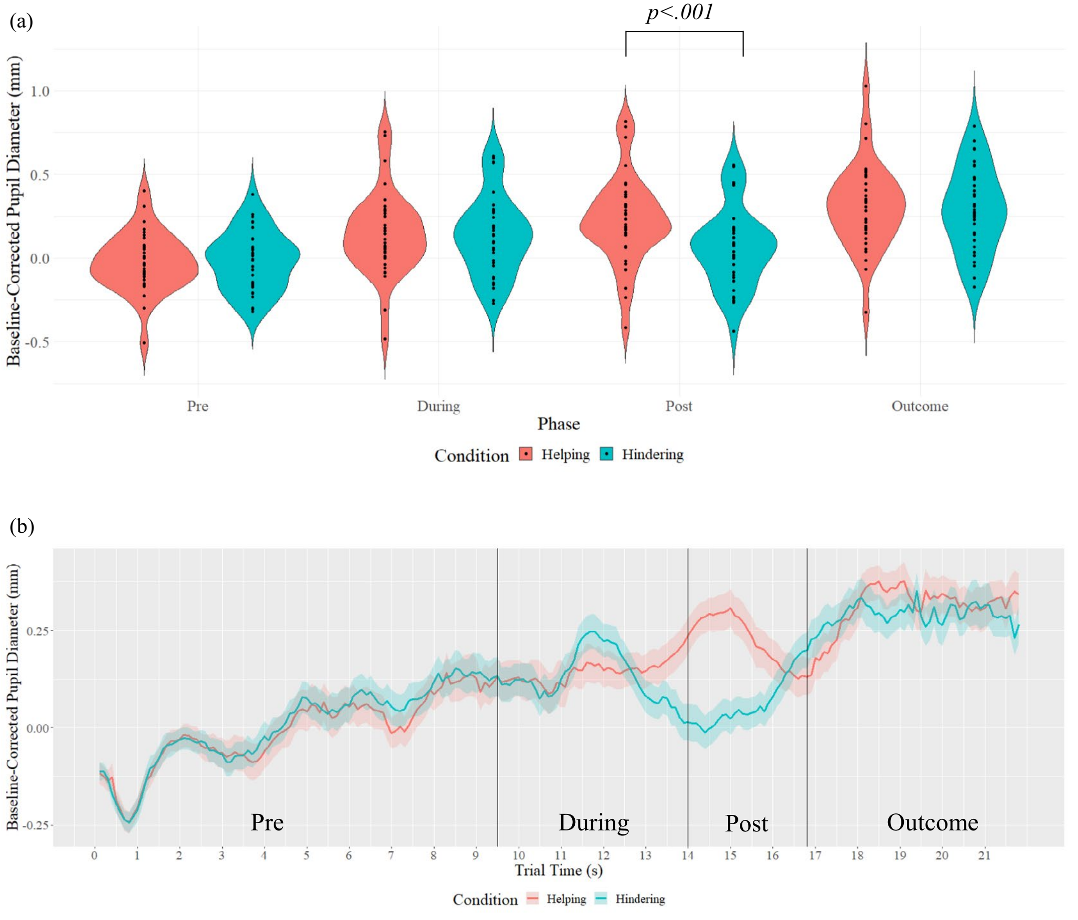

FIGURE 2 Infants' baseline-corrected pupillary responses to different phases of helping and hindering scenarios. *Note*. (B) Error envelopes represent standard errors

caused by insufficient exposure to the stimuli. Given this block-specific analysis was exploratory, this possibility should be explored in future work.

3.3 | What predicted visual preferences for the helper?

To explore whether infants' looking behaviors predicted visual preferences in preference trials, we calculated individual difference measures based on looking times and fixation shifts, and correlated these measures with infants' proportional looking time to the helper (of total looking time on the helper and the hinderer) during preference trials (see Table 1). Consistent with our preregistration, these measures were calculated across trials/blocks, in order to best tap reliable individual differences. Because proportional looking time has a restricted range of 0–1 (which causes ceiling and floor effects), and because Shapiro–Wilk's tests and visual inspection revealed that several eye-movement measures violated normality, we report Spearman's correlation coefficients. Results showed that visual preferences for the helper were predicted by specific looking behaviors during helper and hinderer events. Specifically, significant positive relationships emerged between number of Pre-phase climber–hilltop fixation shifts and helper preferences (see Figure 4), $r_s(32) = .46$,

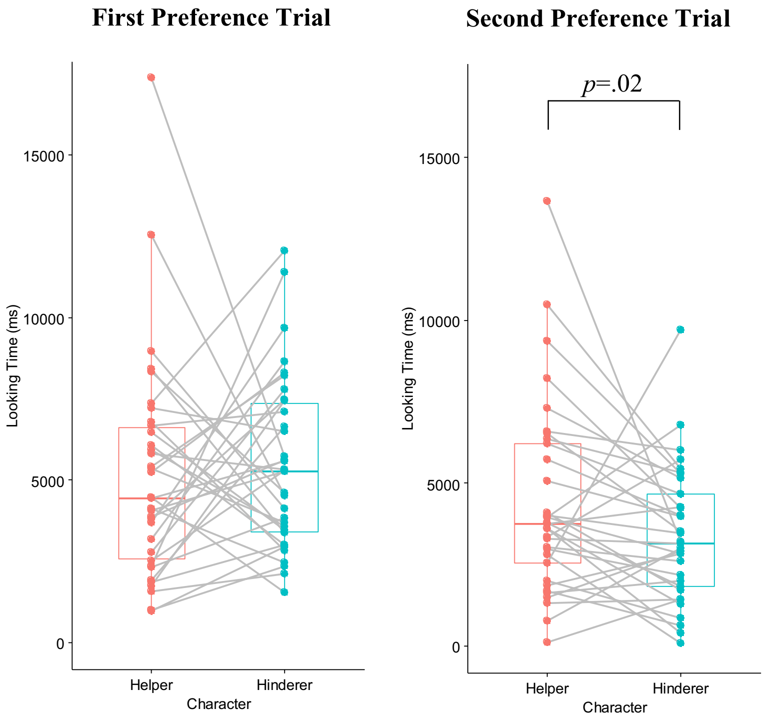

FIGURE 3 Looking times during preference trials

p = .006, and between amount of Pre-phase looking toward the hilltop and helper preferences, $r_s(32) = .52$, $p = .001$ (note, the second analysis was not preregistered). These correlations were robust across helping and hindering events. In addition, visual preferences for the helper correlated with looking behaviors during the preference trials themselves: Infants who looked longer overall, $r_s(32) = .41$, $p = .01$, and who showed more fixation shifts between the helper and hinderer, $r_s(32) = .37$, $p = .03$, also showed greater proportional attention to the helper. In sum, attending/fixating more to the top of the hill during the climber's attempts, attending/fixating more during preference trials, predicted reliable differences in visual preferences for the helper.

Because many infants did not show climber–hilltop fixation shifts ($N = 21$), nor look at the hilltop at all ($N = 16$) during the Pre-phase (see Figure 4), we decided to perform exploratory analyses solely using infants who did show these looking behaviors, reasoning based on past work (Elsner & Adam, 2020) that these might be the infants who understood the Climber's unfulfilled goal to reach the hilltop, and that understanding of the Climber's goal may be critical for social evaluation (Hamlin, 2015). Indeed, t tests against chance (50%) revealed significant preferences for the helper (across both preference trials) in those infants who showed climber–hilltop fixation shifts ($M = 57.51\%$, $t(12) = 2.72$, $p = .019$, $d = .75$, 95% CI [0.51, 0.64]), and who looked at the hilltop at all ($M = 55.80\%$, $t(17) = 2.53$, $p = .021$, $d = .60$, 95% CI [0.51, 0.61]) during the Pre-phase (see Figures 4-5). These findings provide further evidence that attending to the hilltop (i.e., the climber's goal) during the Pre-phase is critical for infants' social evaluations.

We also examined correlations between infants' pupillary responses in different phases and their visual preferences for the helper during preference trials. These analyses showed that infants' pupil size changes and pupil size differences between helper and hinderer events in different phases did not relate to visual preferences for the helper.

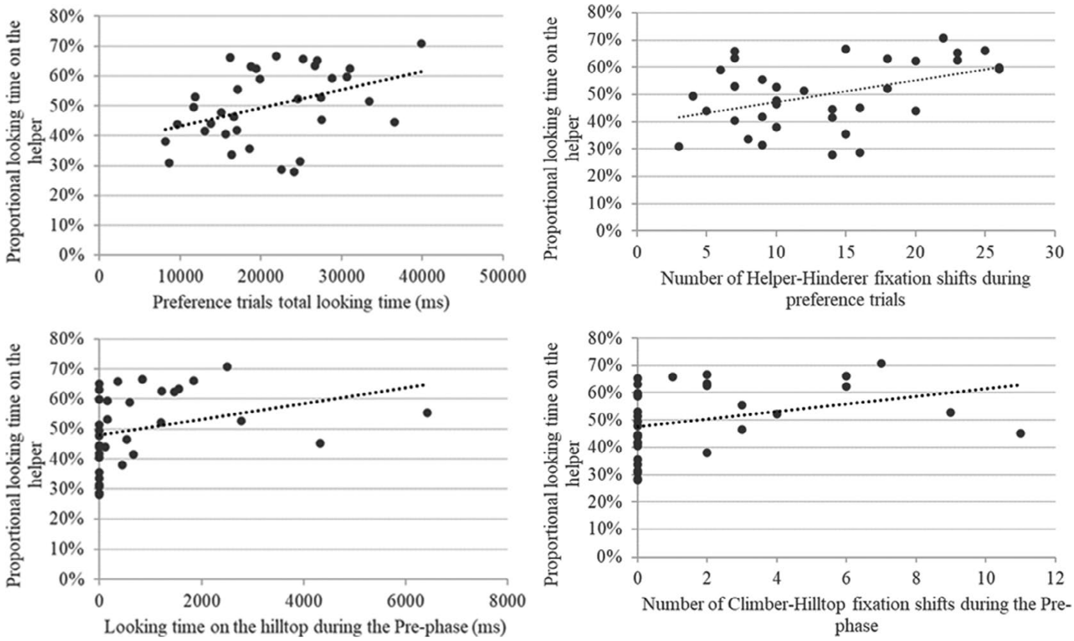

FIGURE 4 Relations between looking behaviors and helper preferences. *Note*. Dots falling on the y axis in the bottom-left graph represent Non-Hilltop Lookers (infants who did not look at the hilltop during the Pre-phase)

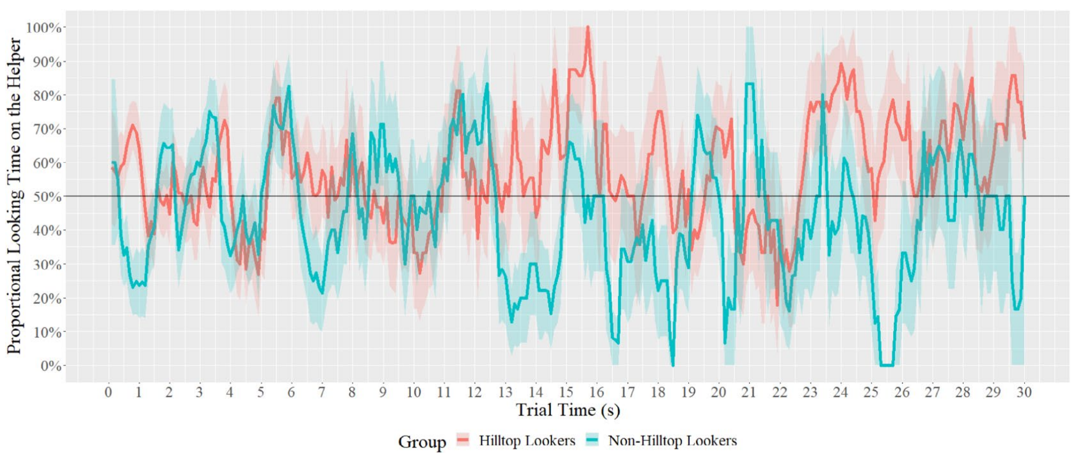

FIGURE 5 Time course of infants' visual preferences during preference trials. *Note*. Error envelopes represent standard errors. Hilltop Lookers and Non-Hilltop Lookers represent infants who did and did not look at the hilltop during the Pre-phase, respectively

4 | DISCUSSION

The current preregistered study used eye-tracking to explore the mental processes underlying infants' responses to the hill scenario. We found that infants, at the group level, showed a preference for the helper during preference trials after viewing 12 (but not 6) helping and hindering videos. In addition, individual infants' preferences for the helper were predicted by their looking behaviors, both during preference trials and during helping/hindering events. Specifically, more attention toward the top of the hill and more back-and-forth fixations between the climber and the hilltop during the climber's ascent was positively related to looking toward the helper, as was more attention and more back-and-forth fixations between the characters during the preference trials. Finally, the group of infants who showed hilltop looking behaviors during the climber's ascent did reliably prefer the helper to the hinderer across preference trials. Together, these findings provide evidence that infants' social evaluations are related to how they process sociomoral scenarios.

The fact that infants, as a group, showed a visual preference for the helper after viewing 12 (but not 6) helping and hindering videos suggests that sufficient exposure to helper/hinderer's action is important for social preferences toward the helper. We note that past studies that found a group-level helper preference with the hill scenario have primarily used habituation (versus familiarization; Hamlin, 2015; Hamlin et al., 2007, 2010), and Outcome phases were kept onscreen for a considerably longer period (30–60 s versus 5 s); notably, infants in these studies tended to take more than 6 trials to habituate (on average about 9). Indeed, our looking time analysis revealed that the decrease in infants' attention over the course of the current study was relatively low, suggestive that infants may have still been processing the scenarios at the end of the presentation. We also note that the animated videos used in the current study might be less engaging and/or more difficult to process than are live puppet shows (see Barr, 2010; but see Margoni & Surian, 2018, for evidence that stimulus type does not influence social evaluations). This finding may help to inform interpretation of a recent study using animated stimuli and 6 familiarization trials, in which 15-month-old infants showed no manual preferences for the helper in the hill paradigm (Schlingloff et al., 2020), and suggests that future research should further explore the role of both animated stimuli and amount of stimulus exposure in infants' social preferences.

What explains the correlations between hilltop looking and infants' social preferences? One possibility is that hilltop looking reflects infants' anticipatory looking toward the climber's goal. Although from past work it remains unclear whether, at the group level, infants under 9 months of age understand the goal of unsuccessful actions—some studies suggest they do (Daum et al., 2008; Hamlin et al., 2008, 2009; Marsh et al., 2010) while others suggest they do not (Brandone et al., 2014; Brandone & Wellman, 2009; Csibra et al., 2003)—past research shows that infants can *anticipate* others' action goals at a group level within certain contexts from 6 months (Adam & Elsner, 2020; Kanakogi & Itakura, 2011; Kochukhova & Gredebäck, 2010; Monroy et al., 2020; for review, see Elsner & Adam, 2020). Further, individual differences in action prediction at 6 months have been associated with action evaluation, defined as surprise when social interactions do not unfold as expected (Gredebäck et al., 2018). Therefore, the hilltop looking behaviors observed in the current study with 5-month-olds could (in principle) reflect individual differences in action and goal understanding; that is, it is possible that only some infants understood the goal of the climber's failed attempts to climb the hill, and only these infants showed preferences for the helper. This interpretation, if true, would suggest that higher-level (e.g., goal-related) measures of attention related to whether or not infants demonstrated social evaluations in our task, consistent with past work suggesting that understanding the climber's goal plays a crucial role in infants' interpretation of the hill scenario (Hamlin, 2015). Critically, if infants' preferences for helpers rely on an understanding of a protagonist's unfulfilled goal, this would support the idea that infants' manual and visual preferences for helpers are based on evaluations of the social value of prosocial/antisocial acts.

Alternatively, it is possible that hilltop looking does not reflect goal understanding per se, but instead relatively lower-level processes that may lead to goal understanding. Specifically, hilltop looking might serve to maximize the uptake of information about the movement of the climber (e.g., the direction of the movement, his target location) and the challenge he faces (e.g., the physics of the hill). Attention toward these features might have facilitated infants' understanding of the climber's goal. This possibility aligns with research demonstrating that 3-month-olds' object perception is predicted by individual infants' visual scanning patterns: Infants who perceived object unity attended more to stimulus features that were relevant to perceptual completion than infants who did not (Johnson et al., 2004), and the perceivers of object unity showed more efficient selective attention in a separate search task than did non-perceivers (Amso & Johnson, 2006). These findings suggest that being able to attend to task-relevant features is important for infants' task success. In the current study, the moving climber (during the Pre-phase) was presumably more perceptually salient than the static hilltop. Hence, infants who looked at the hilltop might have been better at attending to task-relevant features (beyond perceptually salient features) than were infants who did not show such looking behaviors.

Another interpretation of our results is that the relations between hilltop looking and helper preferences were driven solely by perceptual features. For example, infants might simply prefer "upness" (e.g., moving upward, being in a higher place), and infants who preferred to look at the hilltop might also like agents who moved up (versus down). We argue that this "upness" preference account is not supported by our data. Specifically, if infants' preferences were solely driven by a preference for upward movements, we would also expect to find a correlation between infants' helper preferences and their fixation time on the helper in the During-phase, when the helper was moving up. The fact that helper preferences were unrelated to looking time to the helper during its upward motion suggests that helper preferences were not simply driven by a preference for upward movements; further, the fact that we found associations between infants' helper preferences and climber–hilltop fixation shifts suggests that it is attending to the *relationship* between the climber and the hilltop that is important for helper preferences.

We also observed correlations between attention during the preference trials themselves and infants' social preferences. It is possible that infants who attended more during preference trials were simply more engaged; this engagement may have made these infants better able to express any evaluations they made during the helper/hinderer events. Indeed, infants' proportional looking at the helper was relatively consistent throughout the entire length of the preference trials (see Figure 5), suggestive that infants' visual preferences may reflect evaluations they previously made. Further, the positive relationship between back-and-forth fixations between the helper and hinderer during preference trials and preference for the helper suggests that engaging in attentional behaviors thought to facilitate relational processing (Bornstein et al., 2011; Rennels & Cummings, 2013) might also have facilitated infants' evaluations, perhaps by allowing infants to actively compare the two characters onscreen. Together, these results suggest that attentional engagement and relational processing were related to whether or not infants demonstrated social evaluations in our task.

At the group level, infants also responded differently to prosocial/antisocial events. Specifically, during the Post-phase, infants looked longer at the climber and showed more fixation shifts between the climber and the actor (either helper or hinderer) during antisocial events than during prosocial events. Although we did not find longer looking time on the helper versus the climber during prosocial event as in Cowell and Decety (2015), we did find that infants looked longer at the climber than at the hinderer during the antisocial event. This result is consistent with other work from Decety and colleagues (Decety et al., 2012) showing longer looking time at victims versus perpetrators in children and adults aged 4–37 years. Relatedly, Shimizu et al. (2021) found a general attentional bias to recipients versus agents of (positive and negative) sociomoral acts in 3- to 4-year-olds. Although the reason for this pattern is currently unclear, that differences were found during the Post-phase (but not during other phases) might indicate that these looking behaviors reflect infants' empathic concern for the climber after he receives antisocial actions, and/or a general aversion to a character who has just behaved antisocially. Future studies should test this possibility by examining infants' attention to the victims (vs. beneficiaries) in other prosocial/antisocial scenarios.

In addition to eye movements, we also found that infants' pupillary responses differed between conditions. Specifically, during the Post-phase infants' pupil size was smaller for antisocial versus prosocial events. This effect is unlikely to be attributable to general luminosity differences or measurement errors induced by gaze direction differences (see Brisson et al., 2013), because the luminance of the screen and the room was kept constant across conditions (see Supplemental Material 3); further, no pupil size differences were found during the Outcome phase (during which the climber appeared on opposite sides of the screen between conditions) wherein gaze direction differences should have been maximized. The pupil size differences were also not readily explained by attentional differences between conditions, because pupil size changes during the Post-phase were uncorrelated with infants' looking times at (and fixation shifts between) characters, *p*s $> .13$, and because infants' general attention and looking behaviors showed a different pattern from pupil size changes (see Supplemental Material 8 and Figure 2b). Thus, the pupillary effect observed in the Post-phase seems more likely to reflect infants' arousal responses to the social events occurring in the scene.

But how should this pupillary effect be interpreted? We note that there was a general increase in pupil size from the Pre-phase to the Outcome phase across conditions (see Figure 2); therefore, the differences seemed to be driven by a decrease in pupil size for hindering (vs. helping) during the Post-phase. Indeed, infants' pupil size was significantly reduced from the During-phase to the Post-phase in the hindering condition (Wilcoxon exact test, $V = 535$, $Z = 4.06$, $p < .001$). Although this result may seem contradictory to Hepach and colleagues (2012; which found a decrease in pupil size after 2-year-old children observed a third-party help a needy person), we note that our study design differed from Hepach et al. (2012) in many important ways, including a much younger age group and the use of animations (versus videos of human adults). Crucially, Hepach et al. (2012) assessed children's pupillary responses following the helping event using neutral video stimuli (i.e., colorful bubbles against a colored background), whereas in the current study, infants' pupillary responses were measured as the helping/hindering event unfolded. Considering that the Post-phase of the hill scenario depicted the helper/hinderer's exit of the scene, a pupil size reduction during this phase could (very tentatively) be taken to indicate that infants temporarily disengaged when the hinderer left the scene. Although the exact nature of infants' reduction in pupil size following hindering events is unclear (and see previous studies suggesting that pupil dilation may not reflect the valence or likeability of stimuli; Bradley et al., 2008; Partala & Surakka, 2003; for review, see Hepach & Westermann, 2016; Sirois & Brisson, 2014), it nevertheless represents additional evidence that infants distinguish between helping and hindering events, and is consistent with past research showing that affective processes might be involved in infants' processing of sociomoral scenarios (Steckler et al., 2018). Future studies should incorporate additional psychophysiological measures to further examine the processes supporting infants' sociomoral evaluations.

All in all, the current study represents an early step toward elucidating the nature of infants' responses to sociomoral scenarios through detailed analyses of infants' real-time looking behaviors. We recognize that this study, in itself, does not allow us to rule out all alternative explanations nor can we provide a conclusive interpretation for each individual effect. However, we note that the overall pattern of the data (e.g., longer looking at the helper during the second preference trial, correlations between hilltop looking and helper preferences, longer looking toward the recipient of hindering actions) is consistent with, and parsimoniously explained by, the hypothesis that infants' responses to helping/hindering scenarios are social in nature. Indeed, this hypothesis is supported by behavioral studies showing that infants' helper preferences are influenced by various factors known to be relevant to adults' sociomoral judgments, including the mental states (Hamlin, 2013a, 2015; Hamlin et al., 2013; Woo et al., 2017) and behavioral history (Hamlin, 2014; Hamlin et al., 2011) of the characters involved. To further test this hypothesis and clarify the nature of infants' responses to sociomoral scenes, future research should analyze infants' real-time responses to prosocial and antisocial actions in other contexts using larger samples.

Future research should also examine infants' real-time responses to sociomoral scenarios enacted by other types of stimuli. The current study used simplified agent displays (geometric shapes with eyes) in order to minimize task-irrelevant features (e.g., facial expressions, body movements), and past work suggests that our stimuli contained sufficient information to induce perception of mental states. Indeed, research has shown that adults and children readily attribute intentions and goals to simple moving geometric shapes (Barrett et al., 2005; Castelli et al., 2000; for review, see Scholl & Tremoulet, 2000) and that infants also attribute agency and goals to non-human entities like ours that either look like agents (e.g., have eyes), show self-propelled movement, or both (see, e.g. Biro & Leslie, 2007; Johnson et al., 1998; Luo & Baillargeon, 2005; Premack, 1990). That said, infants' sociomoral evaluations have been examined using other types of stimuli, including live puppet shows, animated human characters, and videos of human adults (Buon et al., 2014; Hamlin & Wynn, 2011; Holvoet et al., 2019; Scola et al., 2015). Future research should explore whether the looking behaviors observed in the current study generalize to other stimulus contexts.

ACKNOWLEDGMENTS

This work was supported in part by NSERC Grant (RGPIN-2016-03775) to J. Kiley Hamlin. The funding source was not involved in study design, analysis of data, interpretation of results, or writing of the report. The authors declare no conflicts of interest with regard to the funding source for this study. We thank Miranda Sitch, Vivian Wong, Rachelle Graham, Alexandra Thatcher, Fibha Khan, Francesca Oldham, Haley Gommerman, and Francis Yuen for their assistance. Special thanks go to Brandon M. Woo for creating the stimuli.

ORCID

Enda Tan https://orcid.org/0000-0002-0512-848X

REFERENCES

Adam, M., & Elsner, B. (2020). The impact of salient action effects on 6-, 7-, and 11-month-olds' goal-predictive gaze shifts for a human grasping action. *PLoS One*, *15*(10), e0240165. https://doi.org/10.1371/journal.pone.0240165

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

Amso, D., & Johnson, S. P. (2006). Learning by selection: Visual search and object perception in young infants. *Developmental Psychology*, *42*(6), 1236–1245. https://doi.org/10.1037/0012-1649.42.6.1236

Aslin, R. N. (2012). Infant eyes: A window on cognitive development. *Infancy*, *17*(1), 126–140. https://doi.org/10.1111/j.1532-7078.2011.00097.x

Barr, R. (2010). Transfer of learning between 2D and 3D sources during infancy: Informing theory and practice. *Developmental Review*, *30*(2), 128–154. https://doi.org/10.1016/j.dr.2010.03.001

Barrett, H. C., Todd, P. M., Miller, G. F., & Blythe, P. W. (2005). Accurate judgments of intention from motion cues alone: A cross-cultural study. *Evolution and Human Behavior*, *26*(4), 313–331. https://doi.org/10.1016/j.evolhumbehav.2004.08.015

Biro, S., & Leslie, A. M. (2007). Infants' perception of goal-directed actions: development through cue-based bootstrapping. *Developmental Science*, *10*(3), 379–398. https://doi.org/10.1111/j.1467-7687.2006.00544.x

Bornstein, M. H., Mash, C., & Arterberry, M. E. (2011). Perception of object–context relations: Eye-movement analyses in infants and adults. *Developmental Psychology*, *47*(2), 364–375. https://doi.org/10.1037/a0021059

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, *45*(4), 602–607. https://doi.org/10.1111/j.1469-8986.2008.00654.x

Brandone, A. C., Horwitz, S. R., Aslin, R. N., & Wellman, H. M. (2014). Infants' goal anticipation during failed and successful reaching actions. *Developmental Science*, *17*(1), 23–34. https://doi.org/10.1111/desc.12095

Brandone, A. C., & Wellman, H. (2009). You can't always get what you want: Infants understand failed goal-directed actions. *Psychological Science*, *20*(1), 85–91. https://doi.org/10.1111/j.1467-9280.2008.02246.x

Brisson, J., Mainville, M., Mailloux, D., Beaulieu, C., Serres, J., & Sirois, S. (2013). Pupil diameter measurement errors as a function of gaze direction in corneal reflection eyetrackers. *Behavior Research Methods*, *45*(4), 1322–1331. https://doi.org/10.3758/s13428-013-0327-0

Brown, D. E. (1991). *Human universals*. McGraw-Hill.

Buon, M., Jacob, P., Margules, S., Brunet, I., Dutat, M., Cabrol, D., & Dupoux, E. (2014). Friend or foe? Early social evaluation of human interactions. *PLoS One*, *9*(2), e88612. https://doi.org/10.1371/journal.pone.0088612

Castelli, F., Happé, F., Frith, U., & Frith, C. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage*, *12*(3), 314–325. https://doi.org/10.1006/nimg.2000.0612

Colombo, J., & Mitchell, D. W. (1990). Individual differences in early visual attention: Fixation time and information processing. In J. Colombo, & J. W. Fagen (Eds.), *Individual differences in infancy: Reliability, stability, prediction* (pp. 193–227). Lawrence Erlbaum Associates.

Cowell, J. M., & Decety, J. (2015). Precursors to morality in development as a complex interplay between neural, socioenvironmental, and behavioral facets. *Proceedings of the National Academy of Sciences*, *112*(41), 12657–12662. https://doi.org/10.1073/pnas.1508832112

Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, *27*(1), 111–133. https://doi.org/10.1207/s15516709cog2701_4

Curry, O. S., Mullins, D. A., & Whitehouse, H. (2019). Is it good to cooperate. *Current Anthropology*, *60*(1), 47–69.

Daum, M. M., Prinz, W., & Aschersleben, G. (2008). Encoding the goal of an object-directed but uncompleted reaching action in 6-and 9-month-old infants. *Developmental Science*, *11*(4), 607–619. https://doi.org/10.1111/j.1467-7687.2008.00705.x

Decety, J., Michalska, K. J., & Kinzler, K. D. (2012). The contribution of emotion and cognition to moral sensitivity: a neurodevelopmental study. *Cerebral Cortex*, *22*(1), 209–220. https://doi.org/10.1093/cercor/bhr111

DesChamps, T. D., Eason, A. E., & Sommerville, J. A. (2015). Infants associate praise and admonishment with fair and unfair individuals. *Infancy*, *21*(4), 478–504. https://doi.org/10.1111/infa.12117

Elsner, B., & Adam, M. (2020). *Infants' Goal Prediction for Simple Action Events: The Role of Experience and Agency Cues*. Topics in Cognitive Science.

Fawcett, C., & Liszkowski, U. (2012). Infants anticipate others' social preferences. *Infant and Child Development*, *21*(3), 239–249. https://doi.org/10.1002/icd.739

Geraci, A., & Surian, L. (2011). The developmental roots of fairness: Infants' reactions to equal and unequal distributions of resources. *Developmental Science*, *14*(5), 1012–1020. https://doi.org/10.1111/j.1467-7687.2011.01048.x

Gredebäck, G., Johnson, S., & von Hofsten, C. (2009). Eye tracking in infancy research. *Developmental Neuropsychology*, *35*(1), 1–19. https://doi.org/10.1080/87565640903325758

Gredebäck, G., Kaduk, K., Bakker, M., Gottwald, J., Ekberg, T., Elsner, C., Reid, V., & Kenward, B. (2015). The neuropsychology of infants' pro-social preferences. *Developmental Cognitive Neuroscience*, *12*, 106–113. https://doi.org/10.1016/j.dcn.2015.01.006

Gredebäck, G., Lindskog, M., Juvrud, J. C., Green, D., & Marciszko, C. (2018). Action prediction allows hypothesis testing via internal forward models at 6 months of age. *Frontiers in Psychology*, *9*, 290. https://doi.org/10.3389/fpsyg.2018.00290

Gulz, A., Londos, L., & Haake, M. (2020). Preschoolers' understanding of a teachable agent-based game in early mathematics as reflected in their gaze behaviors–an experimental study. *International Journal of Artificial Intelligence in Education*, *30*, 38–73. https://doi.org/10.1007/s40593-020-00193-4

Hamlin, J. K. (2013a). Failed attempts to help and harm: Intention versus outcome in preverbal infants' social evaluations. *Cognition*, *128*(3), 451–474. https://doi.org/10.1016/j.cognition.2013.04.004

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*(3), 186–193. https://doi.org/10.1177/0963721412470687

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. (2015). The case for social evaluation in preverbal infants: Gazing toward one's goal drives infants' preferences for Helpers over Hinderers in the hill paradigm. *Frontiers in Psychology*, *5*, 1563.

Hamlin, J. K., Hallinan, E. V., & Woodward, A. L. (2008). Do as I do: 7-month-old infants selectively reproduce others' goals. *Developmental Science*, *11*(4), 487–494. https://doi.org/10.1111/j.1467-7687.2008.00694.x

Hamlin, J. K., Newman, G. E., & Wynn, K. (2009). Eight-month-old infants infer unfulfilled goals. Despite Ambiguous Physical Evidence. *Infancy*, *14*(5), 579–590. https://doi.org/10.1080/15250000903144215

Hamlin, J. K., & Sitch, M. J. (2020). Understanding and evaluating the moral world during infancy. In C. Tamis-Lemonda, & J. Lockman (Eds.), *Handbook of Infant Development* (pp. 777–804). Cambridge University Press.

Hamlin, J. K., Ullman, T., Tenenbaum, J., Goodman, N., & Baker, C. (2013). The mentalistic basis of core social cognition: Experiments in preverbal infants and a computational model. *Developmental Science*, *16*(2), 209–226. https://doi.org/10.1111/desc.12017

Hamlin, J. K., & Wynn, K. (2011). Young infants prefer prosocial to antisocial others. *Cognitive Development*, *26*(1), 30–39. https://doi.org/10.1016/j.cogdev.2010.09.001

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557–559.

Hamlin, J. K., Wynn, K., & Bloom, P. (2010). Three-month-olds show a negativity bias in their social evaluations. *Developmental Science*, *13*(6), 923–929. https://doi.org/10.1111/j.1467-7687.2010.00951.x

Hamlin, J. K., Wynn, K., & Bloom, P. (2012). Reply to Scarf et al: Nuanced social evaluation: Association doesn't compute. *Proceedings of the National Academy of Sciences*, *109*(22), E1427.

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*, *108*(50), 19931–19936. https://doi.org/10.1073/pnas.1110306108

Havy, M., & Zesiger, P. (2017). Learning spoken words via the ears and eyes: Evidence from 30-month-old children. *Frontiers in Psychology*, *8*, 2122. https://doi.org/10.3389/fpsyg.2017.02122

Hepach, R. (2017). Prosocial arousal in children. *Child Development Perspectives*, *11*(1), 50–55. https://doi.org/10.1111/cdep.12209

Hepach, R., Vaish, A., & Tomasello, M. (2012). Young children are intrinsically motivated to see others helped. *Psychological Science*, *23*(9), 967–972. https://doi.org/10.1177/0956797612440571

Hepach, R., & Westermann, G. (2016). Pupillometry in infancy research. *Journal of Cognition and Development*, *17*(3), 359–377. https://doi.org/10.1080/15248372.2015.1135801

Holvoet, C., Scola, C., Arciszewski, T., & Picard, D. (2016). Infants' preference for prosocial behaviors: A literature review. *Infant Behavior and Development*, *45*, 125–139. https://doi.org/10.1016/j.infbeh.2016.10.008

Holvoet, C., Scola, C., Arciszewski, T., & Picard, D. (2019). Infants' social evaluation abilities: testing their preference for prosocial agents at 6, 12 and 18 months with different social scenarios. *Early Child Development and Care*, *189*(6), 1018–1031. https://doi.org/10.1080/03004430.2017.1361415

Johnson, S., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. *Developmental Science*, *1*(2), 233–238. https://doi.org/10.1111/1467-7687.00036

Johnson, S. P., Slemmer, J. A., & Amso, D. (2004). Where infants look determines how they see: Eye movements and object perception performance in 3-month-olds. *Infancy*, *6*(2), 185–201. https://doi.org/10.1207/s15327078in0602_3

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*(2), s41562–41016-40037. https://doi.org/10.1038/s41562-016-0037

Kanakogi, Y., & Itakura, S. (2011). Developmental correspondence between action prediction and motor ability in early infancy. *Nature Communications*, *2*(1), 1–6. https://doi.org/10.1038/ncomms1342

Keemink, J. R., Keshavarzi-Pour, M. J., & Kelly, D. J. (2019). Infants' responses to interactive gaze-contingent faces in a novel and naturalistic eye-tracking paradigm. *Developmental Psychology*, *55*(7), 1362–1371. https://doi.org/10.1037/dev0000736

Kochukhova, O., & Gredebäck, G. (2010). Preverbal infants anticipate that food will be brought to the mouth: An eye tracking study of manual feeding and flying spoons. *Child Development*, *81*(6), 1729–1738. https://doi.org/10.1111/j.1467-8624.2010.01506.x

Korhonen, V., Räty, H., & Kärnä, E. (2018). High support need and minimally verbal children with autism playing a preference based computer game: A pilot eye-tracking study of four individual's attendance to eyes. *International Journal of Special Education*, *33*(1), 212–228.

Krüger, M., Bartels, W., & Krist, H. (2020). Illuminating the dark ages: Pupil dilation as a measure of expectancy violation across the life span. *Child Development*, *91*(6), 2221–2236. https://doi.org/10.1111/cdev.13354

Kuhlmeier, V., Wynn, K., & Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological Science*, *14*(5), 402–408. https://doi.org/10.1111/1467-9280.01454

Lee, W., Kim, E. Y., & Song, H.-J. (2020). Do infants expect others to be helpful? *British Journal of Developmental Psychology*, *38*(3), 478–490. https://doi.org/10.1111/bjdp.12331

LoBue, V., Buss, K. A., Taber-Thomas, B. C., & Pérez-Edgar, K. (2017). Developmental differences in infants' attention to social and nonsocial threats. *Infancy*, *22*(3), 403–415. https://doi.org/10.1111/infa.12167

Lucca, K., Pospisil, J., & Sommerville, J. A. (2018). Fairness informs social decision making in infancy. *PLoS One*, *13*(2), e0192848. https://doi.org/10.1371/journal.pone.0192848

Luo, Y., & Baillargeon, R. (2005). Can a self-propelled box have a goal? *Psychological Science*, *16*(8), 601–608.

Margoni, F., & Surian, L. (2018). Infants' evaluation of prosocial and antisocial agents: A meta-analysis. *Developmental Psychology*, *54*(8), 1445–1455. https://doi.org/10.1037/dev0000538

Marsh, H. L., Stavropoulos, J., Nienhuis, T., & Legerstee, M. (2010). Six- and 9-month-old infants discriminate between goals despite similar action patterns. *Infancy*, *15*(1), 94–106. https://doi.org/10.1111/j.1532-7078.2009.00002.x

Monroy, C., Chen, C. H., Houston, D., & Yu, C. (2020). Action prediction during real-time parent-infant interactions. *Developmental Science*, *24*(3), e13042. https://doi.org/10.1111/desc.13042

Nighbor, T., Kohn, C., Normand, M., & Schlinger, H. (2017). Stability of infants' preference for prosocial others: Implications for research based on single-choice paradigms. *PLoS One*, *12*(6), e0178818. https://doi.org/10.1371/journal.pone.0178818

Oakes, L. M. (2012). Advances in eye tracking in infancy research. *Infancy*, *17*(1), 1–8. https://doi.org/10.1111/j.1532-7078.2011.00101.x

Orquin, J. L., Ashby, N. J., & Clarke, A. D. (2016). Areas of interest as a signal detection problem in behavioral eye-tracking research. *Journal of Behavioral Decision Making*, *29*(2–3), 103–115. https://doi.org/10.1002/bdm.1867

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of human-computer Studies*, *59*(1–2), 185–198. https://doi.org/10.1016/S1071-5819(03)00017-X

Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition*, *36*(1), 1–16. https://doi.org/10.1016/0010-0277(90)90051-K

Rennels, J. L., & Cummings, A. J. (2013). Sex differences in facial scanning: Similarities and dissimilarities between infants and adults. *International Journal of Behavioral Development*, *37*(2), 111–117. https://doi.org/10.1177/0165025412472411

Salvadori, E., Blazsekova, T., Volein, A., Karap, Z., Tatone, D., Mascaro, O., & Csibra, G. (2015). Probing the strength of infants' preference for Helpers over Hinderers: two replication attempts of Hamlin and Wynn (2011). *PLoS One*, *10*(11), e0140570. https://doi.org/10.1371/journal.pone.0140570

Scarf, D., Imuta, K., Colombo, M., & Hayne, H. (2012). Golden Rule or valence matching? Methodological problems in Hamlin et al *Proceedings of the National Academy of Sciences*, *109*(22), E1426. https://doi.org/10.1073/pnas.1204123109

Scarf, D., Imuta, K., Colombo, M., & Hayne, H. (2012b). Social evaluation or simple association? Simple associations may explain moral reasoning in infants. *PLoS One*, *7*(8), e42698. https://doi.org/10.1371/journal.pone.0042698

Schlingloff, L., Csibra, G., & Tatone, D. (2020). Do 15-month-old infants prefer helpers? A replication of Hamlin et al (2007). *Royal Society Open Science*, *7*(4), 191795.

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, *4*(8), 299–309. https://doi.org/10.1016/S1364-6613(00)01506-0

Scola, C., Holvoet, C., Arciszewski, T., & Picard, D. (2015). Further evidence for infants' preference for prosocial over antisocial behaviors. *Infancy*, *20*(6), 684–692. https://doi.org/10.1111/infa.12095

SensoMotoric Instruments (2014). *SMI REDn scientific –Specifications*. SensoMotoric Instruments.

Shimizu, Y., Senzaki, S., & Cowell, J. M. (2021). Cultural similarities and differences in the development of sociomoral judgments: An eye-tracking study. *Cognitive Development*, *57*, 100974. https://doi.org/10.1016/j.cogdev.2020.100974

Sirois, S., & Brisson, J. (2014). Pupillometry. *Wires Cognitive Science*, *5*(6), 679–692. https://doi.org/10.1002/wcs.1323

Steckler, C. M., Liberman, Z., Van de Vondervoort, J. W., Slevinsky, J., Le, D. T., & Hamlin, J. K. (2018). Feeling out a link between feeling and infant sociomoral evaluation. *British Journal of Developmental Psychology*, *36*(3), 482–500. https://doi.org/10.1111/bjdp.12232

Tafreshi, D., Thompson, J. J., & Racine, T. P. (2014). An analysis of the conceptual foundations of the infant preferential looking paradigm. *Human Development*, *57*(4), 222–240. https://doi.org/10.1159/000363487

Tan, E., Mikami, A. Y., & Hamlin, J. K. (2018). Do infant sociomoral evaluation and action studies predict preschool social and behavioral adjustment? *Journal of Experimental Child Psychology*, *176*, 39–54. https://doi.org/10.1016/j.jecp.2018.07.003

Ting, F., Dawkins, M., Stavans, M., & Baillargeon, R. (2020). Principles and concepts in early moral cognition. In J. Decety (Ed.), *The social brain: A developmental perspective* (pp. 41–65). MIT Press.

Woo, B. M., Steckler, C. M., Le, D. T., & Hamlin, J. K. (2017). Social evaluation of intentional, truly accidental, and negligently accidental helpers and harmers by 10-month-old infants. *Cognition*, *168*, 154–163. https://doi.org/10.1016/j.cognition.2017.06.029

Wynn, K., & Bloom, P. (2014). The moral baby. In M. Killen, & J. G. Smetana (Eds.), *Handbook of moral development*, 2nd ed. (pp. 435–453). Psychology Press.

Yu, C., & Smith, L. B. (2011). What you learn is what you see: using eye movements to study infant cross-situational word learning. *Developmental Science*, *14*(2), 165–180. https://doi.org/10.1111/j.1467-7687.2010.00958.x

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Tan, E., & Hamlin, J. K. (2021). Mechanisms of social evaluation in infancy: A preregistered exploration of infants' eye-movement and pupillary responses to prosocial and antisocial events. *Infancy*, 00, 1–22. https://doi.org/10.1111/infa.12447