Social interaction targets enhance 13-month-old's associative learning

Supplementary Information

Table of contents

Data availability

The eye tracking raw data as well as all R scripts for pre-processing and analyzing the data are openly accessible on the Open Science Framework (https://osf.io/4a9b6/). The preregistration can be found on AsPredicted (https://aspredicted.org/zt975.pdf).

Supplementary Information Visual Learning Task

Stimuli

Although seemingly acting in dyads the actors were filmed individually. This ensured an accurate positioning of the dyad partners, consistent timing of actions between actors and trials, and identical levels of motion between social and control stimuli within trials. The control stimuli were created by horizontally mirroring the actions of the individual actors. To clarify that the agents in the control videos were not interacting with another individual outside the visible region of the video, the actors were positioned in such a way that they never "touched" the border of the video. All actors were filmed in front of a green screen to control for color and luminance differences between and within videos. Adobe Premiere Pro was used for cutting and editing the videos. Adobe Premiere's Ultra Key tool was used to isolate the actors from the background and replace it with an even colored, grey background layer which was identical over all videos. All social interactions contained mutual touch between the actors. The actors were visible from the waist up. They were all female, wore white t-shirts, had their hair tied back, and did not wear any glasses or jewelry.

Timing of the Task

The timing of the visual learning task was based on prior studies and our observations during piloting. The final timing is illustrated in Figure 1 in the main document and explained in detail in Table S1 in the end of this document.

Additional Analysis and Results

Saccadic latency and looking time. Following visual inspection of the data, we ran further exploratory analyses in addition to the analyses described in the main document. First, we observed that the latencies in the social interaction condition seemed to rapidly decline during the first half of trials (trials 1-6), before inclining during the second half again (trials 7-

12). In order to explore possible habituation effects to the target videos as one possible explanation of this pattern, we conducted the same GLMMs for looking time at the target videos as we ran for our main analysis of saccadic latency. Looking times were assessed by calculating the total duration of fixations within the social interaction and control AOI for each trial, including fixation data from target video onset until video offset. We found a continuous decrease in looking time at the target videos, indicating a general decrease in sustained attention throughout the experiment (main effect of trial, $\chi^2(1) = 11.12$, $p < .001$, estimate = −187.09, $SE = 52.06$). Following the assumption that this general decrease in attention had caused the continuous increase in saccadic latencies throughout the second half of the experiment, we ran an exploratory GLMM for saccadic latency over the first six trials only. The results of this analysis are reported in the main manuscript.

First look. In complementing the first look analysis in the main document, Table S2 shows the total number of first looks at the two target AOIs following social interaction and control cue as well as the proportional number of first looks used for the first look analysis.

Supplementary Information Preferential-Looking Task

Additional Analyses and Results

As described in the pre-registration, we explored sub-group differences between enhanced and less-enhanced learners further by using a more sophisticated group assignment procedure (see also Mani & Huettig, 2014). For this purpose, we divided the sample based on a median split of a beta-coefficient difference score, calculated for each individual by subtracting the beta-coefficient of their learning function during social interaction trials from the beta-coefficient of their learning function during control trials. This procedure was more sophisticated than using mean latencies, since it focused on latency changes over time. Children with enhanced performance in the learning task looked relatively longer at the social interaction shape ($M = .48$; $SD = .13$) compared to less enhanced learners ($M = .43$; SD = .18, $t(30) = -1.38$; $p = .18$, $d = -.49$).

Supplementary Information Manual Forced-Choice Task

Procedure

The procedure was adapted from previous studies (e.g., Hamlin & Wynn, 2011). The experimenter, sitting opposite of the child, held the choice board upside down with the stimuli facing her body. At the beginning of the task, the experimenter said "Hi" while looking at the child. As soon as the child looked at the experimenter, she flipped around the board out of the child's reach, saying "Look!". Once the child had looked at both shapes, the experimenter pushed the board forward and asked "Which one do you want?". We ended the task if the child had made a choice or if the child did not make any choice for two minutes.

Additional Gaze Following Task (Exploratory)

In addition to our main research question, we included a gaze-following task to explore the possible relation between children's performance in the visual learning task and their gaze following abilities. Infants' seeking of learnable content might not only be reflected in their tendency to learn associations between arbitrary shapes and situations with observational learning opportunities, but also relate their tendency to use others' gaze as social cue guiding them to relevant information. Since we added the gaze-following task as an exploratory measure, it was conducted during an additional eye tracking phase (5 min) at the very end of each testing session. The same eye tracking hardware was used as for the main tasks of the study. We run the task by using Tobii Studio (version 3.4.8.1348).

Stimuli and Design

Each child was presented with six videos during which an actress shifted her gaze to one of two target objects. During half of the trials the actress looked to an object being located to her right side, while she looked to an object located to her left side during the other half of the trials. The order of trials was randomized for each participant. As dependent variable we measured the participant's first gaze shift to either of the two objects. We used video stimuli designed for previous study by Astor & Gredebäck (2019).

Data Analysis and Coding

We defined three areas of interest: one elliptical AOI for the head, and two rectangular areas, one for each target. A gaze following difference score was calculated by subtracting the number of incongruent trials in the gaze following task (child first looked at the not-attended object after actor's head turn) from the number of congruent trials (child first looked at the attended object after the actor's head turn). Trials were only counted as valid when children had looked at the head of the actor before looking at one of the two target objects. On average, each participant contributed 4.8 valid gaze-following trials to the analysis (SD = .71, range = 3-6 trials). To check whether the participants followed gaze at all, we tested the gaze following score against chance level by running a one sample test against zero. To investigate possible relations between visual learning and gaze following abilities, we correlated the latency difference score from the visual learning task with a gaze following difference score by using Pearson's r correlation. We used the MATLAB based open source software TimeStudio (Nyström, Falck-Ytter, & Gredebäck, 2016) for defining AOIs and pre-processing the data (TobiiStudio version 3.19; MATLAB version R2018b).

Results

Children followed gaze in the additional gaze-following task (M = 3; SD = 1.66; t(29) = 9.89, p = .00, d = 1.8). Gaze following abilities were not correlated with visual learning abilities, N = 30 , r(28) = .20, p = .30.

Complementary analyses. In addition to our pre-registered plan, we compared the proportional number of congruent trials between enhanced and less enhanced learners in the visual learning task (see main document section ‘Data Analyses and Coding’ for the procedure of sub-sample creation). The proportional number of trials in which children followed gaze was higher for the sub-sample of enhanced learners (M = .87; SD = .13) compared to less enhanced learners (M = .74; SD = .19, t(25) = 2.05; p = .05).

Description Supplementary Videos

The Videos provided as supplemental material are gaze replays of one exemplary child. The red dots represent fixations, the red lines saccades.

VIDEO 1: Illustrates the entire sequence of the experiment without sound. The video Includes the calibration check, pre and post preferential-looking task, and visual learning task.

VIDEO 2: Shows four exemplary trials in the visual learning task with sound to illustrate the auditory elements in the procedure.

References

Astor, K., & Gredebäck, G. (2019). Gaze following in 4.5- and 6-month-old infants: The impact of proximity on standard gaze following performance tests. *Infancy*, *24*, 79–89. https://doi.org/10.1111/infa.12261

Deligianni, F., Senju, A., Gergely, G., & Csibra, G. (2011). Automated gaze-contingent objects elicit orientation following in 8-month-old infants. *Developmental Psychology*, *47*, 1499–1503. https://doi.org/10.1037/a0025659

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

Mani, N., & Huettig, F. (2014). Word reading skill predicts anticipation of upcoming spoken language input: A study of children developing proficiency in reading. *Journal of Experimental Child Psychology*, *126*, 264–279. https://doi.org/10.1016/j.jecp.2014.05.004

Nyström, P., Falck-Ytter, T., & Gredebäck, G. (2016). The TimeStudio Project: An open source scientific workflow system for the behavioral and brain sciences. *Behavior Research Methods*, *48*, 542–552. https://doi.org/10.3758/s13428-015-0616-x

Reuter, T., Emberson, L., Romberg, A., & Lew-Williams, C. (2018). Individual differences in nonverbal prediction and vocabulary size in infancy. *Cognition*, *176*, 215–219. https://doi.org/10.1016/j.cognition.2018.03.006

Wang, Q., Bolhuis, J., Rothkopf, C. A., Kolling, T., Knopf, M., & Triesch, J. (2012). Infants in Control: Rapid Anticipation of Action Outcomes in a Gaze-Contingent Paradigm. *PLoS ONE*, *7*, e30884. https://doi.org/10.1371/journal.pone.0030884

Table S1

Detailed information regarding the timing of the visual learning task

Phase	Duration	Explanation
Gap	200 ms	We added this gap between trials to clearly separate the trials from one another. The delay created the impression that the next cue "popped up" on the screen after the target video from the previous trial had stopped playing.
Cue	Gaze dependent (max 4000 ms) + 300 ms delay	The cue was gaze dependent to support children's learning by (a) making sure that children had seen the cue, (b) adjusting the task to individual differences as infants could control the trial-to-trial speed with their gaze.
		After the child had fixated the cue for 150 ms (Deligianni, Senju, Gergely, & Csibra, 2011), it remained on screen for another 300 ms before it disappeared. This delay was to ensure that children had processed the cue shape before it disappeared. Initially, we piloted a delay of 1500 ms (Reuter, Emberson, Romberg, & Lew-Williams, 2018), but reduced it to 300 ms as children started to explore the screen during the longer delay. We aimed to (a) clearly point out that cues were associated with following targets, and (b) support the sensation of eliciting the target by looking at the cue.
Gap	600	The delay between cue and target was due to give infants' the opportunity to perform predictive gaze shifts to the cued target region (timing as in Wang et al., 2012).
Target	4000	The duration of the target video was based on our aim to keep it as short as possible to keep the "flow" of the task while keeping it long enough that it contained the relevant phases: (a) facing forward, (b) turning towards/away from one another, (c) establishing the interaction/non-interactive action.

Table S2

Total Number of First Looks in the Visual Learning Task

	Social Interaction AOI	Control AOI	Prop. Number of Looks to Target AOI *(SD)*
Social Interaction cue	261	103	.72 *(.32)*
Control cue	101	266	.72 *(.31)*

Notes. Total numbers over all participants and trials and proportional number of first looks to the correct target region (i.e., number of first looks to the cued target AOI divided by the total number of first looks in both AOIs). Total number of trials over all conditions and participants = 768.