Supplementary Material

Discrepancies between reported and pre-registered analyses

The manuscript reports on a novel paradigm, combining EEG and predictive looking measures. With the exception of the EEG measure, for which prior studies served as a template for planned analysis and expected effect, there was no prior study using either the paradigm, or the combination of the two measures, based on which we could have confidently pre-specify the exact analyses that we will perform. Therefore, the pre-registration includes several variables, which were found to be uninformative upon reflection and were therefore not analyzed or are excluded from the main manuscript for conciseness. Here, we specify how the analyses reported in the manuscript differ from those specified in the preregistration, along with reasons for these changes.

Inference phase
Preregistration specifies that infants' predictive saccades towards the sorting boxes will be analyzed and compared in both Prediction phase (reported in the manuscript) as well as Inference phase. The latter was omitted from analyses as very few saccades were recorded in that period (the average number of saccades was <2 per infant across all trials) and we therefore deemed this measure uninformative.

Number of looks
We had analyzed the number of looks to either AOI during the prediction phase, however we realized upon completion of the study that this measure is likewise not informative. Because the Prediction phase lasted 4 seconds, most infants would make on average one shift from the initially fixated box to the alternative before the end of the trial. In line with other studies (e.g. Southgate et al., 2007, Kaldy et al., 2016) we concluded that First Look and DLS are the measures that are most likely to reflect infants' expectations and predictions, and therefore do not report Number of Looks for conciseness.

Sample size
The data was collected during the first authors' stay at Copenhagen University, following lab closures in the US due to the COVID-19 pandemic. Due to the time pressure and several uncertainties regarding the potential to complete the study, we had diverged from the pre-registered data collection plan. Instead of planned sequential analysis, the data collection was completed without analysis at 0.75% of the planned sample, and resulted in 34 of the tested 57 infants included in the sample for Experiment 1; and 32 (of the 67 tested) included in Experiment 2. The crucial analysis - relating EEG to predictive gaze - is based on data of Experiment 2, for which the sample size matches the preregistration. The rest of inclusion criteria - number of valid trials for each of the measures needed to be included in the sample - was followed precisely.

Exploratory analyses
We did not perform any of the exploratory analysis, as they were provisionally planned for the case in which our main predictions and measures did not produce results.

References
Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. Psychological science, 18(7), 587-592.

Kaldy, Z., Guillory, S. B., & Blaser, E. (2016). Delayed Match Retrieval: A novel anticipation-based visual working memory paradigm. Developmental science, 19(6), 892-900.