Developmental Psychology

Remote Infant Studies of Early Learning (RISE): Scalable Online Replications of Key Findings in Infant Cognitive Development

Elena J. Tenenbaum, Caitlin Stone, My H. Vu, Madeleine Hare, Kristen R. Gilyard, Sudha Arunachalam, Elika Bergelson, Somer L. Bishop, Michael C. Frank, J. Kiley Hamlin, Melissa Kline Struhl, Rebecca J. Landa, Casey Lew-Williams, Melissa E. Libertus, Rhiannon J. Luyster, Julie Markant, Maura Sabatos-DeVito, Stephen J. Sheinkopf, Jennifer B. Wagner, Kayle Park, Anna I. Soderling, Ashleigh K. Waterman, Jordan N. Grapel, Amit Bermano, Yotam Erel, and Shafali Jeste

Online First Publication, November 11, 2024. https://dx.doi.org/10.1037/dev0001849

CITATION

Tenenbaum, E. J., Stone, C., Vu, M. H., Hare, M., Gilyard, K. R., Arunachalam, S., Bergelson, E., Bishop, S. L., Frank, M. C., Hamlin, J. K., Kline Struhl, M., Landa, R. J., Lew-Williams, C., Libertus, M. E., Luyster, R. J., Markant, J., Sabatos-DeVito, M., Sheinkopf, S. J., Wagner, J. B., Park, K., Soderling, A. I., Waterman, A. K., Grapel, J. N., Bermano, A., Erel, Y., & Jeste, S. (2024). Remote Infant Studies of Early Learning (RISE): Scalable online replications of key findings in infant cognitive development.. *Developmental Psychology*. Advance online publication. https://dx.doi.org/10.1037/dev0001849

AMERICAN PSYCHOLOGICAL ASSOCIATION

Developmental Psychology

© 2024 American Psychological Association
ISSN: 0012-1649

https://doi.org/10.1037/dev0001849

Remote Infant Studies of Early Learning (RISE): Scalable Online Replications of Key Findings in Infant Cognitive Development

Elena J. Tenenbaum[1, 2], Caitlin Stone[2], My H. Vu[3], Madeleine Hare[4], Kristen R. Gilyard[5], Sudha Arunachalam[6], Elika Bergelson[5], Somer L. Bishop[7], Michael C. Frank[8], J. Kiley Hamlin[9], Melissa Kline Struhl[10], Rebecca J. Landa[11, 12], Casey Lew-Williams[13], Melissa E. Libertus[14, 15], Rhiannon J. Luyster[16], Julie Markant[17], Maura Sabatos-DeVito[1, 2], Stephen J. Sheinkopf[18], Jennifer B. Wagner[19], Kayle Park[2], Anna I. Soderling[2], Ashleigh K. Waterman[2], Jordan N. Grapel[2], Amit Bermano[20], Yotam Erel[20], and Shafali Jeste[4]

[1] Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine
[2] Duke Center for Autism and Brain Development, Duke University
[3] Biostatistics and Data Management Core, The Saba Research Institute, Children's Hospital Los Angeles, Los Angeles, California, United States
[4] Department of Neurology, Children's Hospital of Los Angeles, Los Angeles, California, United States
[5] Department of Psychology, Harvard University
[6] Department of Communicative Sciences and Disorders, New York University
[7] Department of Psychiatry and Behavioral Sciences, University of California San Francisco
[8] Department of Psychology, Stanford University
[9] Department of Psychology, University of British Columbia
[10] Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology
[11] Center for Autism Services, Science and Innovation, Kennedy Krieger Institute, Baltimore, Maryland, United States
[12] Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine
[13] Department of Psychology, Princeton University
[14] Department of Psychology, University of Pittsburgh
[15] Learning Research and Development Center, University of Pittsburgh
[16] Department of Communication Sciences and Disorders, Emerson College
[17] Department of Psychology, Tulane Brain Institute, Tulane University
[18] Department of Pediatrics, Thompson Center for Autism and Neurodevelopment, University of Missouri
[19] Department of Psychology, City University of New York Graduate Center, City University of New York College of Staten Island
[20] Blavatnik School for Computer Science, Tel-Aviv University

The current article describes the Remote Infant Studies of Early Learning, a battery intended to provide robust looking time measures of cognitive development that can be administered remotely to inform our understanding of individual developmental trajectories in typical and atypical populations, particularly infant siblings of autistic children. This battery was developed to inform our understanding of early cognitive and language development in infants who will later receive a diagnosis of autism. Using tasks that have been successfully implemented in lab-based paradigms, we included assessments of attention, memory, prediction, word recognition, numeracy, multimodal processing, and social evaluation. This study reports results on the feasibility and validity of administration of this task battery in 55 infants who were recruited from the general population at age 6 months (n = 29; 14 female, 15 male) or 12 months (n = 26; 14 female, 12 male; 62% White, 13% Asian, 1% Black, 1% Pacific Islander, 22% more than one race; 6% Hispanic). Infant looking behavior was recorded during at-home administration of the battery on the family's home computer and automatically coded for attention to stimuli using iCatcher+, an open-access software that assesses infant gaze direction. Results indicate that while some tasks replicated lab-based findings (attention, memory, prediction, and numeracy), others did not (word recognition, multimodal processing, and social evaluation). These findings will inform efforts to refine the battery as we continue to develop a robust set of tasks to improve the understanding of early cognitive development at the individual level in general and clinical populations.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Koraly Pérez-Edgar served as action editor.

Elena J. Tenenbaum https://orcid.org/0000-0002-0973-142X

This study's design and hypotheses were preregistered with the Open Science Framework (https://osf.io/yhg95). Where consent was provided, videos are available on Databrary (https://nyu.databrary.org/volume/1619). Code and raw data are available on GitHub (https://github.com/elenatenenbaum/RISE). The authors declare no competing financial interests. Elena J. Tenenbaum is a consultant for Maplight Therapeutics. This work was supported by funding from the Simons Foundation Autism Research Initiative in Human Cognitive and Behavioral Science (Award 874889, awarded to Elena J. Tenenbaum and Shafali Jeste).

Elena J. Tenenbaum played a lead role in conceptualization, funding acquisition, investigation, methodology, project administration, resources, supervision, writing–original draft, and writing–review and editing and an equal role in data curation, formal analysis, software, and visualization.

continued

Public Significance Statement
We present pilot data from the first instantiation of the Remote Infant Studies of Early Learning Battery that was designed to assess infant cognitive development remotely. Findings show results consistent with preregistration for some (attention, memory, prediction, and numeracy) but not all tasks. These results will inform the continued development of a remote battery for the assessment of cognitive development in general and clinical populations.

Keywords: cognitive development, remote, infant

Supplemental materials: https://doi.org/10.1037/dev0001849.supp

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Research on infant cognitive development is often restricted by the need for in-person assessment, limiting the feasibility of large sample sizes. This constraint is particularly relevant when attempting to gather information about atypically developing populations where diagnoses are often not made until well into toddlerhood or preschool age. In 2019, the Simons Foundation Autism Research Initiative held a workshop titled "Next Steps in Infancy Research on Autism," which brought together experts in child development, neurodevelopmental disorders, genetics, and epidemiology. From that workshop emerged a plan to develop a battery of tasks that could assess infant development remotely. Such a battery would allow researchers to establish a scalable approach to assessing cognitive and language development in infants (including those at elevated likelihood for autism), thereby creating opportunities for meaningful progress in our understanding of typical and atypical development. This workshop led to the instantiation of the Remote Infant Studies of Early Learning (RISE) Battery and Consortium. Contemporaneous progress in methods for online data collection via Children Helping Science (previously called LookIt; Scott et al., 2017; Scott & Schulz, 2017) and iCatcher+ (Erel et al., 2023), an open-source software for automated coding of infant looking behavior, improved the feasibility of these efforts. The ultimate goal of the RISE Battery is to test individual differences in infant cognitive and language development through looking-time tasks administered in the infants' homes. The primary goal for this first analysis was to determine whether these specific tasks would replicate from the lab setting to the remote setting. Here, we report these feasibility results with 6- and 12-month-old infants recruited from the general population.

The RISE Battery

Established laboratory measures of cognitive development have identified remarkable perceptual abilities within the first year of life that predict language and cognitive development (Lockman & Tamis-LeMonda, 2020). Recent work suggests that these established lab-based measures can be translated to in-home remote assessment. Studies have shown that we can reliably administer looking time measures using remote methods (Bacon et al., 2021; Bánki et al., 2022; Chuey et al., 2024). Expanding on this work, the RISE Battery includes a set of tasks representative of key developmental domains that have been studied in infancy using looking time measures. Of particular interest is which of these domains may diverge (and which may stay intact) in children who go on to be diagnosed with autism.

Caitlin Stone played a lead role in project administration, software, and validation, a supporting role in conceptualization, investigation, writing–original draft, and writing–review and editing, and an equal role in data curation. My H. Vu played a lead role in formal analysis and visualization, a supporting role in methodology, and an equal role in data curation, writing–original draft, and writing–review and editing. Madeleine Hare played a supporting role in methodology, project administration, writing–original draft, and writing–review and editing and an equal role in validation. Kristen R. Gilyard played a supporting role in writing–review and editing and an equal role in data curation, validation, and writing–original draft. Sudha Arunachalam played an equal role in methodology and writing–review and editing. Elika Bergelson played a lead role in conceptualization and an equal role in methodology and writing–review and editing. Somer L. Bishop played a supporting role in conceptualization and writing–review and editing. Michael C. Frank played an equal role in conceptualization, investigation, methodology, and writing–review and editing. J. Kiley Hamlin played a supporting role in conceptualization and writing–review and editing and an equal role in methodology. Melissa Kline Struhl played a lead role in software, a supporting role in writing–review and editing, and an equal role in conceptualization, investigation, methodology, and project administration. Rebecca J. Landa played an equal role in conceptualization and writing–review and editing. Casey Lew-Williams played an equal role in conceptualization, funding acquisition, investigation, methodology, and writing–review and editing. Melissa E. Libertus played a supporting role in conceptualization and an equal role in methodology, resources, and writing–review and editing. Rhiannon J. Luyster played a supporting role in writing–review and editing and an equal role in methodology. Julie Markant played a supporting role in writing–review and editing and an equal role in methodology. Maura Sabatos-DeVito played a supporting role in conceptualization and writing–review and editing. Stephen J. Sheinkopf played a supporting role in conceptualization, investigation, methodology, and writing–review and editing and an equal role in resources. Jennifer B. Wagner played an equal role in methodology and writing–review and editing. Kayle Park played a supporting role in methodology and writing–review and editing. Anna I. Soderling played a supporting role in validation and writing–review and editing and an equal role in methodology. Ashleigh K. Waterman played a supporting role in validation and writing–review and editing. Jordan N. Grapel played a supporting role in writing–review and editing. Amit Bermano played a lead role in software and a supporting role in conceptualization. Yotam Erel played a lead role in software and a supporting role in writing–review and editing. Shafali Jeste played a lead role in funding acquisition, investigation, and resources, a supporting role in data curation, formal analysis, software, visualization, and writing–original draft, and an equal role in conceptualization, methodology, project administration, supervision, and writing–review and editing.

Correspondence concerning this article should be addressed to Elena J. Tenenbaum, Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Hock Plaza, 2424 Erwin Road, Suite 501, Durham, NC 27705, United States. Email: elena.tenenbaum@duke.edu

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

The domains selected for this pilot study included attention (social and nonsocial), memory (faces and nonface objects), pattern recognition (visual prediction and multisensory processing), language (word comprehension), social cognition (preference for prosocial actions), and numeracy (numerical change detection). This battery of tasks includes developmental paradigms that have been associated with or hypothesized to be linked with autism, language, and cognitive delays (Arunachalam & Luyster, 2016; Bradshaw et al., 2011; Hamlin et al., 2007; Li et al., 2024; Pierce et al., 2016; Reuter et al., 2018; Senju, 2013; Sinha et al., 2014). In the present study, we piloted the battery in a sample of infants from the general population to determine whether the RISE Battery, administered in the remote setting, replicates findings from laboratory-based versions of these tasks. These results will help guide our selection of the battery to be tested as early markers or predictors of autism in a larger cohort of infants at elevated likelihood for receiving an autism diagnosis.

Background

Background for each domain tested is reviewed below, and specific task designs are described in detail in the Method section.

Attention

Attention to visual stimuli is one of the most widely studied areas of infant cognitive development and is gaining interest as an essential domain in studying learning trajectories associated with autism (Colombo et al., 2004; Jones & Klin, 2013; Oakes & Amso, 2018). There is a large body of evidence to suggest that infants later diagnosed with autism show a reduced preference for social stimuli in the presence of nonsocial stimuli when compared to nonautistic controls (Campbell et al., 2019; Chawarska et al., 2016; Pierce et al., 2016). Here, we used a preferential looking task that contrasts dynamic geometric shapes with animated children dancing to explore social preference.

Memory

Memory in infancy is often assessed with familiarization to a stimulus followed by presentation of the familiarized stimulus, coupled with a novel item of the same category (Manns et al., 2000). Using this paradigm, Bradshaw et al. (2011) demonstrated that 6-month-olds later diagnosed with autism show atypical habituation for faces, and toddlers with autism show lack of recognition for familiarized faces but intact recognition for objects. In the RISE Battery, we assessed memory for both face and nonface stimuli using a Visual Paired Comparison (VPC) task.

Pattern Recognition

Two tasks in the RISE Battery can be described as Pattern Recognition tasks. The first assesses prediction. The theory that autism is associated with challenges in prediction suggests that individuals with autism may demonstrate a diminished capacity to predict what comes next (Cannon et al., 2021; Sinha et al., 2014). Though there has been significant work on prediction in infants in the general population, research in this domain in infants with autism is limited (Jeste et al., 2015; Saffran & Kirkham, 2018). In this study, we measured prediction as a function of anticipatory looks to a target visual stimulus following cueing of the target location. This approach has been used in prior work to assess the ability of infants to predict what comes next and to update their predictions and was shown to correlate with language abilities in infants without autism (Reuter et al., 2018).

Another form of pattern recognition assessed in the RISE Battery is the co-occurrence of speech sounds with videos of a speaker talking. Sensitivity to the multimodality of speech arises early in life, and numerous studies have shown that typically developing infants show a preference for congruent stimuli as early as 4 months (Kuhl & Meltzoff, 1982; Patterson & Werker, 1999). Auditory–visual misalignment may be a key feature in autism that contributes to language delays in this population (Stevenson et al., 2016; Venker et al., 2018). Differences in multimodal processing in autism have been reported as early as 9 months, with infants at elevated likelihood for autism showing reduced sensitivity to mismatched auditory/visual signals (Guiraud et al., 2012). Early sensitivity to audiovisual (AV) alignment has been linked to language abilities in autistic and typically developing toddlers (Righi et al., 2018; Tenenbaum et al., 2015). In the RISE Battery, we assess sensitivity to AV Synchrony with the presentation of two identical videos of a woman speaking in highly animated infant-directed speech. One video matches the audio stimuli, and the other precedes the audio track by 666 ms.

Language Comprehension

Though expressive language delays are well documented in children with autism, we know less about early receptive language skills in this population. Evidence from toddlers suggests that the development of receptive language skills in autism may follow similar trajectories to those in typical development but with significant delays (Venker et al., 2013). One study comparing infants with an elevated likelihood for autism to controls found no differences in accuracy or processing speed at 18 or 24 months but lower accuracy in infants with elevated likelihood for autism at 36 months compared to infants with typical likelihood (Chita-Tegmark et al., 2015). Word recognition in infancy has been linked to later language outcomes (Fernald & Marchman, 2012). In the RISE Battery, we assess word recognition using a looking-while-listening paradigm based on stimuli from seminal work suggesting that 6-month-olds do recognize highly familiar words (Bergelson & Swingley, 2012).

Social Cognition

A growing body of empirical work suggests that abilities to evaluate others based on their prosocial and antisocial acts emerge within the first year of life (for review and meta-analysis, see Margoni & Surian, 2018). Performance on tasks assessing preference for prosocial behaviors has been linked to more adaptive social skills in early childhood (Tan et al., 2018). To our knowledge, this has not yet been studied in infants at elevated likelihood for autism. In the RISE Battery, we used a preferential looking task to assess infant preference for prosocial behavior.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Numeracy

Infants are born with the ability to quickly estimate quantities and compare them. For example, newborns prefer to look at images that display the same number of objects as the number of tones that they hear (Izard et al., 2009). Using traditional habituation paradigms, infants have been found to be better at discriminating quantities that have a larger relative difference (e.g., 20 vs. 10 objects) than quantities that are relatively closer together (e.g., 15 vs. 10 objects; Xu & Spelke, 2000). Abilities in these tasks have been linked to later mathematical abilities (Starr et al., 2013). One recent study demonstrated that preschoolers with autism show challenges with nonsymbolic number comparisons relative to typically developing peers (Li et al., 2024). To our knowledge, quantitative skills have not yet been tested in infants at elevated likelihood for autism. In the RISE Battery, we test abilities in numeracy using a preferential looking paradigm.

Each of the seven tasks that comprise the RISE Battery has been used successfully in some form in the laboratory to explore these domains of infant cognitive and language development. Each of the domains tested in early infancy has been associated with developmental outcomes (e.g., language skills, mathematical abilities, social skills, or autism diagnosis). The RISE Battery was created to assess individual differences in the development of infants at elevated likelihood for autism but must first be tested for replication of lab-based findings in the remote setting. The present study provides this test in 55 infants from English-speaking families 6- ($n = 29$) and 12-month-olds ($n = 26$) recruited from the general population in the United States. This project was preregistered with the Open Science Framework at https://osf.io/yhg95.

The current article addresses hypotheses H1 and H2 from the preregistration (copied directly from the preregistration below for reference).

Hypothesis 1: Computer Vision Analysis of Infant Gaze. We predict that the automated coding of looking times accomplished by the iCatcher+ software will be consistent with hand coding of looking behaviors (i.e., left/right/away).

Hypothesis 2a: Social Attention. In the *Geo-Social Attention* task, we predict study participants will look preferentially to the social stimuli.

Hypothesis 2b: Memory. In the *VPC* task, we predict that there will be (a) a stronger novelty preference at 12 months than at 6 months and (b) a stronger novelty preference for faces than objects.

Hypothesis 2c: Pattern Recognition (updating). In the *Prediction* task, we expect that infants will show greater proportion of anticipatory eye movements (AEMs) at 12 months than at 6 months.

Hypothesis 2d: Pattern Recognition (multimodal). In the *AV Synchrony* task, we predict that participants will display preferential looking to the synchronous video, demonstrated by a greater proportion of looking time to the synchronous video relative to overall looking.

Hypothesis 2e: Language Comprehension. In the *Word Comprehension* task, we predict that study participants will show greater difference scores at 12 months than at 6 months.

Hypothesis 2f: Numeracy. In the *Numeracy task*, we expect that infants at both ages will look significantly longer to the numerically changing image stream compared to the numerically nonchanging image stream, regardless of whether the numerically nonchanging image stream contains the smaller or the larger of the two quantities.

Hypothesis 2g: Social Cognition. In the *Helper/Hinderer* task, we predict that infants will look significantly longer to the helper character than to the hinderer character.

Method

Ethical Compliance

This work was completed in compliance with the APA ethical principles regarding research with human participants in the conduct of the research presented in this article. All methods were approved by the Duke Health and Children's Hospital of Los Angeles Institutional Review Boards.

Participants

Participants were recruited from across the United States using the Children Helping Science platform registry and passive recruitment via social media. Inclusion criteria were (a) full-term birth (>37 weeks), (b) no known developmental delays, and (c) no immediate relatives with an autism diagnosis. Exclusion criteria were (a) presence of significant sensory impairments (blindness or deafness) and (b) less than 50% English spoken in the home. A final total of 55 participants included with data were collected at 6 (6:0–6:30; $n = 29$) or 12 months (12:0–12:30; $n = 26$). This sample size was selected to confirm the feasibility and validity of these tasks when administered in the remote setting, where possible, based on power analyses from the studies on which the tasks were based (Bergelson & Swingley, 2012; Bradshaw et al., 2011; Hamlin et al., 2007; Pierce et al., 2016; Reuter et al., 2018; Righi et al., 2018). Participants were representative of the geographic (22 of 50 states) and racial composition of the United States (Scott & Schulz, 2017) but were skewed toward higher education and income (see Table 1). Participants were compensated with $25 Amazon gift cards.

Stimuli

The Geo-Social Attention task assesses infant preference for social stimuli (Pierce et al., 2011). In this task, two videos are displayed side-by-side. In one, a child is dancing, and in the other, a dynamic geometric figure is shown. This is repeated for a total of 10 unique 4-s trials with side of presentation for the social and nonsocial stimuli counterbalanced across trials. The dependent measure is the proportion of looks to the social stimulus relative to social + nonsocial stimuli for the length of stimulus presentation (see Figure 1). The social and abstract stimuli used in this task were based on previous examples in the literature (Imafuku et al., 2017; Pierce et al., 2011; Shi et al., 2015) but have not yet been tested in this specific paradigm remotely or in the lab. Based on previous findings using similar stimuli (Imafuku et al., 2017; Pierce et al., 2011), infants were expected to demonstrate a preference (>.50 of the trial) for the social stimuli relative to the geometric shapes.

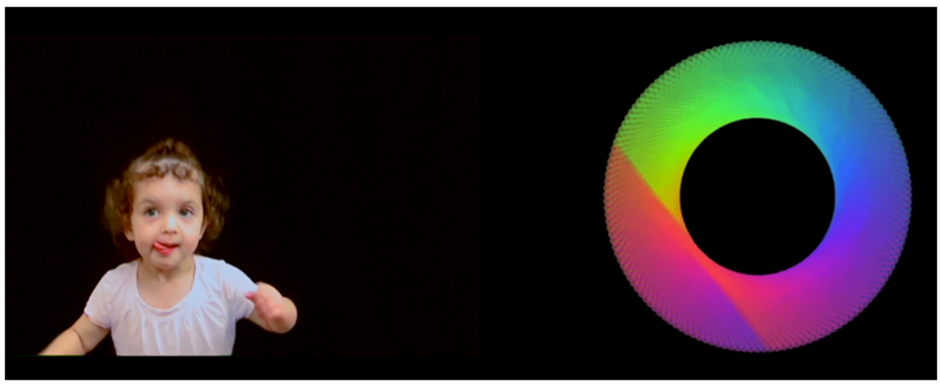
Table 1
Parents' Demographics

Characteristic	Overall (N = 55)[a]	Age 6 (N = 29)[a]	Age 12 (N = 26)[a]
Sex			
Female	28 (50.9)	14 (48.3)	14 (53.8)
Male	27 (49.1)	15 (51.7)	12 (46.2)
Race			
American Indian or Alaska Native	0 (0)	0 (0)	0 (0)
Asian	7 (12.7)	4 (13.8)	3 (11.5)
Black or African American	1 (1.8)	1 (3.4)	0 (0)
Native Hawaiian or other Pacific Islander	1 (1.8)	1 (3.4)	0 (0)
White	34 (61.8)	16 (55.2)	18 (69.2)
More than one race	12 (21.9)	7 (24.2)	5 (19.3)
Ethnicity			
Hispanic or Latino	3 (5.5)	3 (10.3)	0 (0)
Education			
High school	1 (1.8)	0 (0)	1 (3.8)
Associate (some college)	1 (1.8)	1 (3.4)	0 (0)
Bachelor/college	18 (32.8)	12 (41.2)	6 (23.1)
Graduate and professional degree	35 (63.6)	16 (55.2)	19 (73.1)
Annual income[b]	$110,000 ($60,000–$150,000)	$110,000 ($60,000–140,000)	$110,000 ($60,000–$170,000)
Geographic region[b]			
Northeast	13 (24.1)	7 (25.0)	6 (23.1)
Southeast	8 (14.8)	1 (3.6)	7 (26.9)
Midwest	17 (31.5)	12 (42.8)	5 (19.3)
Southwest	1 (1.9)	0 (0)	1 (3.8)
West	15 (27.7)	8 (28.6)	7 (26.9)

Note. IQR = interquartile range.
[a] For categorical variables, value represents n, and values in parenthesis provide %. For continuous variables (only income), value is median and values in parenthesis are interquartile range (25 and 75th percentile). [b] 1–2 participants with missing data (excluded from the denominator).

The VPC task tests memory for faces and nonface objects (e.g., Bradshaw et al., 2011; Pascalis et al., 2002) In this task, infants see a familiarization phase where they view the same stimulus on both sides of the screen for 25 s. They then see the familiar stimulus on one side of the screen and a novel stimulus on the other for 8 s. This is repeated for a second test trial in which the familiar and novel stimulus positions are reversed. Animated attention getters are presented between trials. This series is repeated six times, and the novel and familiar stimuli are counterbalanced between participants. Three trials presented faces (one trial each for White, Black, and Asian faces, paired within trials; Tottenham et al., 2009). Three trials presented complex nonface stimuli ("fribbles") that maintain the complexity of face processing while removing the social nature of the task (Barry et al., 2014). The dependent measure is the proportion of time spent attending to the novel face or object at test (novel/novel + familiar; see Figure 2). While prior studies have used this approach to examine face and object discrimination separately (Pascalis et al., 2002; Pickron et al., 2017; Snyder, 2010), the current task was designed to directly compare these skills in infancy in both 6- and 12-month-olds. Based on previous findings with related stimuli (Manns et al., 2000), we expected infants to focus more attention on the novel faces and objects, and it was hypothesized that

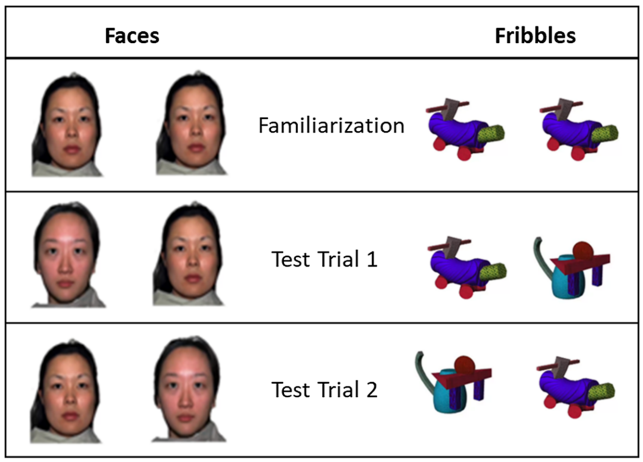
Figure 1
Geo-Social Attention Task



Note. Preferential looking task with video of child dancing contrasted with dynamic geometric image. Abstract geometric stimuli courtesy of Paul Chaikin, https://paulchaikin.net/. Image of child published with permission. See the online article for the color version of this figure.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Figure 2
Visual Paired Comparison Task

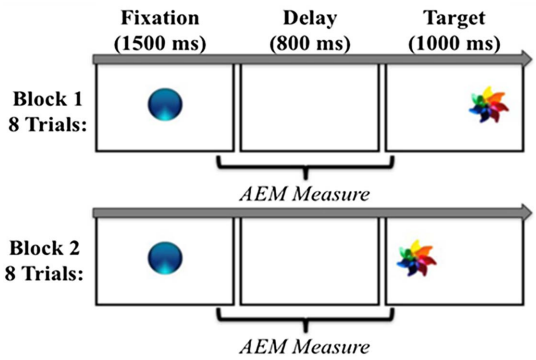


Note. Example trials from the Visual Paired Comparison task. The task showed three face trials (Asian, White, and Black) and three fribbles trials. All faces were female. Stimulus images courtesy of Michael J. Tarr, Carnegie Mellon University, https://www.tarrlab.org/. Images of faces are published with permission. See the online article for the color version of this figure.

this novelty preference would increase between 6 and 12 months (e.g., Wagner et al., 2020).

The Prediction task assesses infants' visual predictions (Reuter et al., 2018). This task consists of eight trials in which a 2-s presentation of an animated central cue is followed by a 2-s presentation of a colorful animated pinwheel (the target stimulus) on one side of the screen (i.e., eight presentations on the left side of the screen or eight presentations on the right side). Then, the target stimulus is presented eight additional times, all on the opposite side of the screen. The full task includes 16 trials. The dependent measure is the proportion of AEMs (conservatively defined as 200 ms before center fixation offset until 200 ms after target onset) to the target image location, analyzed separately for the first and second blocks (i.e., after the presentation side has switched; see Figure 3). Based on previous findings using these specific stimuli with toddlers (Reuter et al., 2018), infants were expected to show reliable AEMs in Block 1 but more variable AEMs in Block 2.

Figure 3
Prediction Task



Note. Prediction task included 16 trials across two blocks. AEM = anticipatory eye movement. From "Individual Differences in Nonverbal Prediction and Vocabulary Size in Infancy," by T. Reuter, L. Emberson, A. Romberg, and C. Lew-Williams, 2018, *Cognition*, *176* (https://doi.org/10.1016/j.cognition.2018.03.006). Copyright 2018 by Elsevier. Reprinted with permission. See the online article for the color version of this figure.

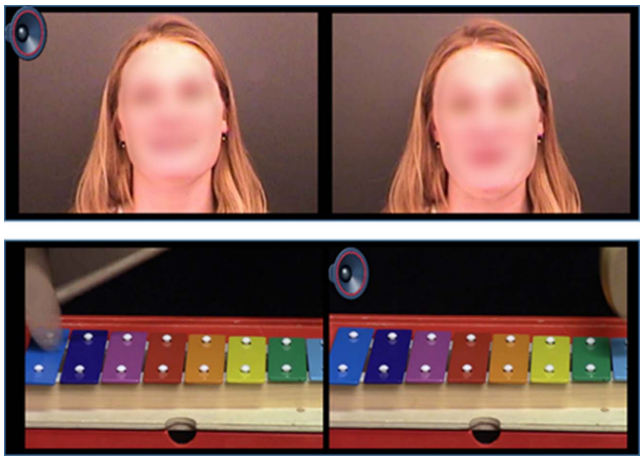
The AV Synchrony task evaluates multimodal processing (Guiraud et al., 2012; Righi et al., 2018). In this task, two videos are shown simultaneously side-by-side, but only one is synchronous with the audio track while the other precedes the audio track by 666 ms. This task includes both social (a woman talking to the camera using infant-directed speech) and nonsocial stimuli (toys being played with in a repetitive manner by adult hands). The task includes a total of 8- and 14-s trials with four social and nonsocial trials. The dependent measure is the proportion of time spent looking at the synchronous video relative to synchronous + asynchronous (see Figure 4). These specific stimuli were used in the 666 ms condition in previous work with toddlers who demonstrated success in discriminating the synchronous videos at this delay (Righi et al., 2018). Based on results from the original study and evidence that neurotypical infants are sensitive to AV Synchrony early in development (Kuhl & Meltzoff, 1982; Patterson & Werker, 1999), infants were expected to show a preference for the synchronous videos.

The Word Comprehension task examines receptive language skills. This specific instantiation was based on classic looking-while-listening tasks (e.g., Bergelson & Swingley, 2012; Fernald et al., 2008), with specific items borrowed from Bergelson and Swingley (2012), but this exact set of images was not used in the original study. In this task, two picture representations with labels that are familiar to young infants are presented side-by-side on the screen while participants are prompted to direct their gaze at one of the images (e.g., "Look at the foot!"). Two blocks containing 10 trials (five pairs of stimuli) are presented. Trials last 8 s and the two blocks are separated by 4 s of dynamic video to recapture attention. In each block, each word was used as both a target and a distracter, such that infants view five sets of images twice in each block, resulting in each picture being presented once as the target and again as a distractor in each block. The order of presentation is randomized between blocks. The dependent measure is the mean proportion of time spent looking at a given object when it is the target in each block (e.g., "Look at the foot!" in Figure 5) minus the mean proportion of time spent looking at that object when it is the distracter in each block (e.g., "Look at the bottle!" in Figure 5). The analysis window was set as 367–4,000 ms posttarget word onset. Items were selected based on words likely to be familiar to young infants and included foot–bottle, cat–book, hand–yogurt, ball–stroller, and diaper–block (Bergelson & Swingley, 2012; Fenson et al., 2007). The design used in the present study differed from that of Bergelson and Swingley (2012) in the following ways: (a) The task did not include attention getters between trials to reduce the length of the task, (b) similarly, there were far fewer trials because this was part of a longer battery of tasks, and (c) the specific items were not identical to the food and body items used in the original study. Based on previous findings, we expected 12-month-olds to perform better in this task than 6-month-olds.

The Numeracy task tests representations of approximate quantities (Libertus & Brannon, 2010). In this task, infants see two simultaneous image streams: one of different images all containing the same number of dots (either 5 or 20) and the other of

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Figure 4
Audiovisual Synchrony Task



Note. Social trials showed two videos of a woman speaking in infant-directed speech with an audio track aligned to one of the videos (Note: faces blurred here due to image permissions. Infants saw unblurred versions). Nonsocial trials showed a noise making toy video with the audio track aligned to one of the videos. See the online article for the color version of this figure.

different images that alternate between the two quantities (i.e., 5 and 20). Each image stream consists of arrays of dots presented for 500 ms followed by 300 ms of blank screen lasting a total of 20 s per trial. A 6-s animated attention grabber is inserted between trials to recapture infants' attention. Each infant sees four trials with the location of the constant and alternating streams counterbalanced across trials. If infants are able to detect the numerical change, they are expected to look longer at the numerically alternating image stream compared to the numerically nonalternating image stream. Though both sides of the screen show sequences of dots that change with respect to location, only the numerically alternating side displays different numbers of dots. To ensure that the two image streams can only be differentiated based on number and not based on other continuous variables of the dot arrays, images are systematically equated for cumulative surface area, cumulative perimeter, or individual element size of the dots as well as density (i.e., the number of dots per cm^2 within each image). Half of the images, that is, every other image across both streams, is identical. One-third of the images across both streams are matched for cumulative surface area, one-third are matched for cumulative perimeter, and one-third are matched for individual element size. Orthogonally, half of the images that differ in number is matched on density. The dependent measure is the proportion of time spent looking at the alternating stream relative to time spent looking at both sides of the screen (alternating/alternating + nonalternating; see Figure 6). The structure of this task is identical to those used in the past (e.g., Decarli et al., 2023; Libertus & Brannon, 2010; Starr et al., 2013), except that the overall trial duration was shorter (20 s compared to 40 or 60 s in previous studies) to ensure feasibility in the context of this much longer battery of tasks. Shorter trials may make it more difficult for infants to detect the numerical differences between stimuli. Hence, we also increased the difference between the numbers of dots (5 and 20) compared to some other studies with infants this age (e.g., 8 vs. 16 in Libertus & Brannon, 2010). Based on previous versions of this task, infants were expected to show a preference for the alternating number stream.

Figure 5
Word Comprehension Task

Note. Visualization of the types of images used in the Word Comprehension task. Note, these images are not the actual stimuli (which are available from the authors but could not be printed due to image restrictions). Trials showed two images of objects familiar to infants while an audio track played infant direct speech, for example, "Look at the foot!" See the online article for the color version of this figure.

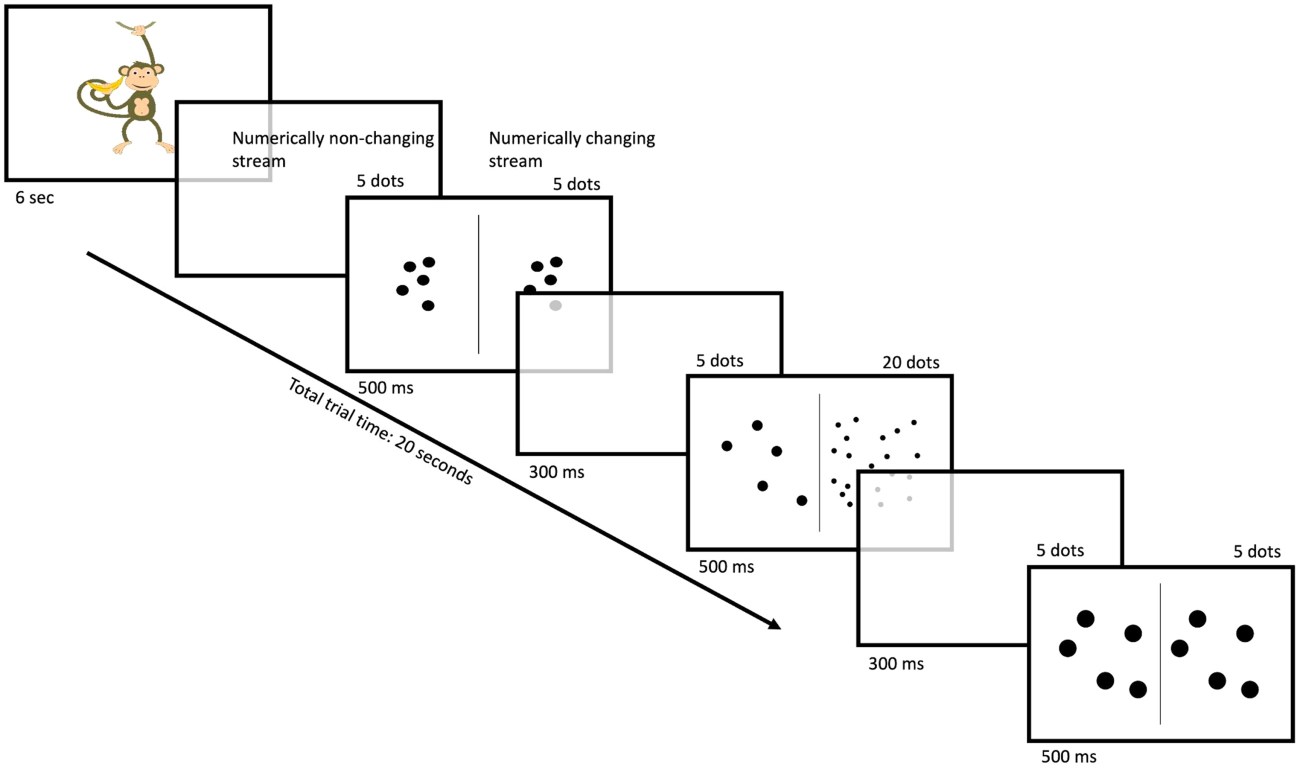
The Helper/Hinderer Task assesses social evaluation. In this task, infants watch videos of a character trying but failing to climb to the top of a steep hill (from Hamlin et al., 2007). On alternating events, a helper character pushes the climber to the top of the hill, facilitating its goal, and a hinderer character pushes the climber to the bottom of the hill, preventing its goal. These events are each shown three times before the characters are shown for a 30-s test trial in which preference for the helper object is measured. The series is then repeated with the same stimuli in the same roles. The Helper/Hinderer stimulus and presentation order is counterbalanced between infants. The dependent measure is the proportion of time spent looking at the helper relative to both objects (helper/helper + hinderer; see Figure 7). Because the RISE Battery is administered remotely and asynchronously, video presentations were not contingent on infant looking. Based on previous work in habituation versions of this task, infants were expected to show a preference for the helper at the test.

Approach

The RISE Battery was presented to infants using the Children Helping Science Platform (formerly LookIt; Scott et al., 2017; Scott & Schulz, 2017), which allows researchers to post studies on a shared website. Families participated via their web browser on their home computer, while webcam video and survey data were sent back to the Children Helping Science server. Study sessions were completed asynchronously, allowing families to participate at their convenience. At the start of every study, the parent recorded a videotaped statement of informed consent. All video consents were reviewed by study personnel prior to accessing the participant videos. The presence of an infant was verified prior to consent approval for all participants. Participant families were able to withdraw their videos prior to releasing them to study personnel. Caregivers were instructed to limit potential distractions during stimulus presentation by avoiding speaking to their child during the tasks and, where possible, removing siblings and pets from the environment. Families had the option to pause the study between tasks and return at any time as long as the browser window remained

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Figure 6
Numeracy Task



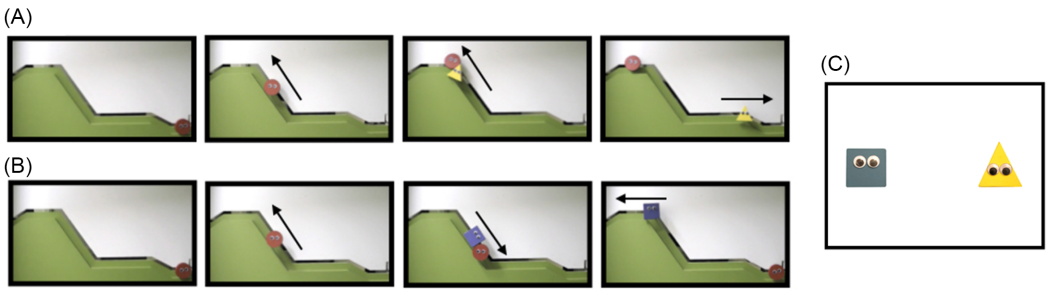
Note. Schematic of numerical change detection paradigm. In this example, the numerically changing image stream alternates between images containing 5 and 20 dots, whereas the numerically nonchanging image stream always contains images with five dots. See the online article for the color version of this figure.

open on their computer. The average time between task administrations ranged from 3 s to 73 min, with a median of 11 s.

Video Coding

iCatcher+ (Erel et al., 2023) is an open-access software that was developed to facilitate automated coding of infant gaze. Using computer vision methods, iCatcher+ performs automatic gaze estimation from low-resolution videos. At the core of this approach is an artificial neural network that classifies gaze direction in real time (see Figure 8). iCatcher+ first selects the face of the infant by capturing the lowest bounding face box identified by the face extractor, thus allowing for detection of infant gaze in the presence of a caregiver (Cao et al., 2021). To confirm the validity of the iCatcher+ detection of infant gaze in the remote setting, 20% of the videos from each task were recoded by human coders.

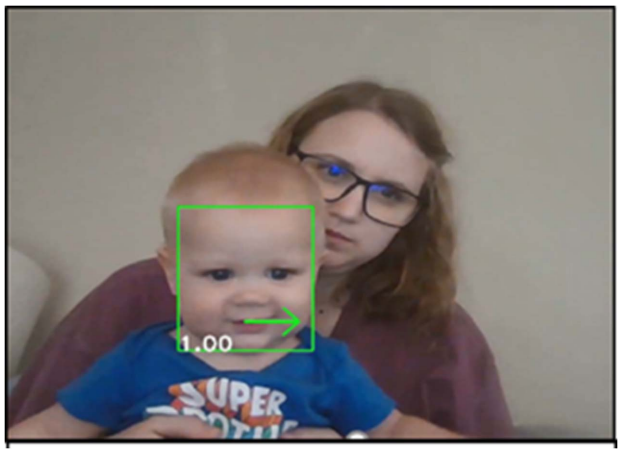
Figure 7
Helper/Hinderer Task



Note. (A) An example of a helper trial where the yellow triangle helps the circle up the hill. (B) An example of a hinderer trial where the blue square knocks the circle back down the hill. (C) A test trial in which preference for the helper was assessed. See the online article for the color version of this figure.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Figure 8
Output From iCatcher



Note. Arrow shows gaze position and number shows estimate of iCatcher confidence in that assessment (0–1.00). Images are published with permission. See the online article for the color version of this figure.

Frame-by-frame human coding was completed using Peyecoder (Olson et al., 2020), an open-source program for gaze-direction coding. For consistency with iCatcher+, coders assessed whether the infant was looking to the left, right, or away (i.e., no "center" code was included). All coders were trained by one master coder. Individual coders then practiced on training videos until interrater reliability was above 0.85. The frames where iCatcher+ reported 85% or more confidence were used for the reliability analysis to maintain consistency with the statistical analyses reported in the Results section. Human-to-human reliability was assessed using Fleiss's κ and was found to be 0.85. Reliability between human coders and iCatcher+ was assessed using Cohen's κ. For the 12-month group, reliability ranged from 0.71 to 0.72; for the 6-month group, reliability was 0.79. While this falls short of our preregistered goal of $\kappa = 0.9$, it is consistent with our preregistered target of $\kappa \geq 0.7$ as acceptable reliability (Hypothesis 1).

Usable Data

Within tasks, a given video frame of infant response was deemed usable if iCatcher+ provided a rating of gaze direction with 85% or more confidence that the participant was looking to the left or the right. This criterion was not preregistered but was determined post hoc based on the significantly negatively skewed data on confidence levels obtained from iCatcher+ (see Table 2). Confidence in iCatcher+ is computed by applying the softmax function over the network predicted outputs and returning the result of the class with the highest value. iCatcher+ is explicitly trained to maximize its confidence in the predicted class using the cross-entropy loss. Applying softmax to the outputs of the network can be interpreted as a valid probability distribution over the classes, and the confidence as the probability of the predicted class is correct.

In order for iCatcher+ to make a confidence assessment, it was necessary for the infant to be looking at the screen and for iCatcher+ to be able to identify the child's eyes. This necessarily excluded frames in which the child was turned away or closing his or her eyes. The proportion of usable frames, or frames with 85% or more confidence, was calculated overall and by trial for each task. As preregistered, for each task, final analysis only included data from trials with more than 33% usable frames. The cut point of 33% usable frames was based on the range of cut points used in the studies on which these tasks were based (e.g., 4% in Righi et al., 2018; 50% in Pierce et al., 2011). Though in the preregistration, we had stated that for the Helper/Hinderer task, the cutoff would be 40% (based on prior work with this task); for consistency with the rest of the battery, the standard cut point of 33% usable frames for at least two helping and hindering familiarization trials and at least five cumulative seconds of usable frames during each of the preferential looking trials was adopted. For the prediction task, consistent with the original methods, infants were required to have an AEM in Block 2 and at least two usable trials in each block for data to be included in the analyses. See Table 2 for a summary of the cause of missing frames.[1]

Statistical Analysis

Demographics and characteristics for parents and participants were summarized using median and interquartile range for continuous variables, and frequency and percentage for categorical variables. The primary outcome was treated as continuous for preferential looking tasks (*Geometric-Social, VPC, AV Synchrony, Numeracy, and Helper/Hinderer*) as each measures the proportion of time a child looks to the target stimulus relative to time spent looking at both sides of the screen. For preferential looking tasks, the expected pattern was defined as a mean proportion greater than 0.5. For the Prediction task, the dependent measure was the proportion of AEMs, which was calculated as the number of usable trials with an AEM divided by the total number of usable trials, and the expected pattern was a mean proportion greater than 0.5. For the Word Comprehension task, the dependent measure was also treated as continuous as it provided a difference score for the proportion of time spent looking at a given item when it was the target minus the proportion when that item was the distracter. For Word Comprehension, the expected pattern was a mean difference score >0.

Graphical representations (e.g., box plots) depict mean outcome values across stimuli for each task. Linear and generalized linear mixed effects models (multilevel) with child and stimulus crossed random effects, as stimuli were not unique to each child, were fit via R packages "lme4" and "lmerTest" to test for these expected patterns in the overall cohort and between the two age groups. We included an interaction term (effect modifier) of stimulus type to test whether the aforementioned patterns were different for social versus nonsocial stimuli for VPC and AV Synchrony tasks. Results are reported as model-based estimates, corresponding 95% confidence limits, and p value of the corresponding hypothesis test. We defined statistical significance using a two-sided $p < .05$. All statistical analyses were conducted in R Studio 4.2.2.

[1] All results were rerun with infants who had at least 50% usable trials in each task (excluding Helper/Hinderer for which there were only two test trials). Results did not change in significance for any task.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 2
Summary of Frames Dropped Due to Lack of Face in Frame, Away Looks, and Low-Confidence From iCatcher

Task (no. of trials)	% of nonface frame		% of away frame		% of frame with iCatcher confidence <.85	
	Age 6	Age 12	Age 6	Age 12	Age 6	Age 12
Geo-Social Attention (10)	3.4	5.8	9.4	2.6	9.7	9.3
The VPC (6)	4.4	11.7	19.7	14.8	10.2	11.4
Prediction (14)	3.0	8.6	8.4	7.2	21.4	19.6
Audiovisual Synchrony (8)	3.9	7.7	11.2	7.1	11.7	10.7
Word Comprehension (20)	5.1	5.8	21.6	11.9	9.0	9.8
Numeracy (4)	4.1	4.3	10.2	7.7	5.8	7.5
Helper/Hinderer (2)	7.8	7.3	12.6	10.9	11.3	12.5

Note. % of nonface frames = iCatcher+ was not able to identify a face in the frame (e.g., child's head was not in view or child was looking back at parent); % of away frames = iCatcher+ detected a face, but gaze was estimated to be off screen; % of frames with iCatcher confidence <.85 (probability that iCatcher+ prediction of class is correct based on Softmax function). VPC = Visual Paired Comparison.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Transparency and Openness

This study's design and hypotheses were preregistered with the Open Science Framework (https://osf.io/yhg95). Where consent was provided, videos are available on Databrary (https://nyu.databrary.org/volume/1619). Code and raw data are available on github: https://github.com/elenatenenbaum/RISE. This work was completed in compliance with the APA ethical principles regarding research with human participants in the conduct of the research presented in this article. All methods were approved by the Duke Health and Children's Hospital of Los Angeles Institutional Review Boards.

Results

Infant Attention and Side Bias

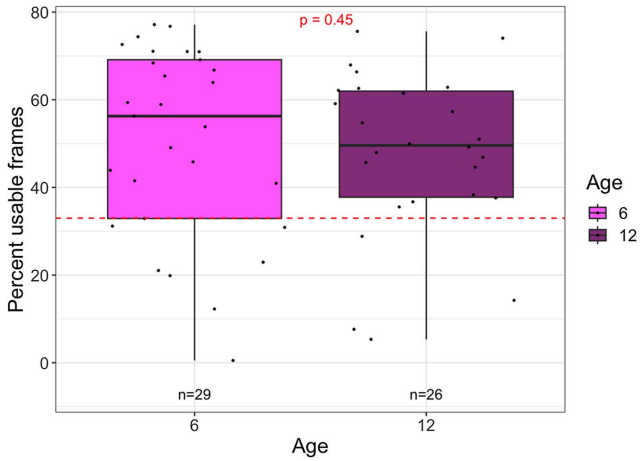
Because this study was intended to test the feasibility and validity of home administration of a battery of looking time tasks, we first explored overall attention to the tasks as a function of order of presentation. As expected, attention to the video and therefore the percent of individuals with usable data decreased with each subsequent task. Additionally, the 12-month-old infants were overall less attentive than the 6-month-old infants (Figure 9). Despite the waning attention, there were usable data for every RISE Battery task (preregistered as >33% usable frames; see Table 2). Participants were randomized to *view Geometric–Social*, *VPC*, *Prediction*, *AV*, and *Word Comprehension* in different orders from 1 to 5. Due to their novelty, the *Numeracy* and *Helper/Hinderer* tasks were initially included as exploratory. They were therefore shown in alternating sequences as the 6th or 7th task in the sequence. There was no difference in the overall percent of usable frames based on the order of presentation (see Supplemental Figures S1 and S2).

Based on preliminary examination of the data, it was observed that infants often fixated on only one side of the screen for the duration of a given trial. To reduce the impact of this behavior on data analysis, tasks for which an infant fixated on only one side of the screen for more than 80% of the usable frames were eliminated from analysis. This side bias criteria was not used in the *Prediction Task* because the window of analysis for an AEM was quite short. This data exclusion was not preregistered. The number of infants for whom this was an issue is shown in Table 3 by task. Though the number of infants excluded from analysis for this reason was large in some tasks (particularly in numeracy), these exclusion rates are not entirely unusual for infant studies, even in the laboratory (Chawarska et al., 2013; Elsabbagh et al., 2009; Frank et al., 2009; Narayan et al., 2010; Peltola et al., 2009).

Geo-Social Attention Task

Our preregistered prediction for the *Geo-Social Attention Task* was that study participants would look preferentially (mean proportion >0.50) to the social stimuli. This was the case among 12-month-olds but not 6-month-olds (Figure 10a; 12-month-olds: 0.59 [0.53, 0.66], $p = .006$; 6-month-olds: 0.51 [0.44, 0.57], $p = .819$; mean difference: 0.09 [0.01, 0.16], $p = .025$). These findings replicate previous work in toddlers (Pierce et al., 2011) and extend those results down to 12 months. The lack of preference for either stimulus type among 6-month-olds may indicate developmental shifts between 6 and 12 months in orientation to social stimuli in the context of visually appealing dynamic geometric figures (see Table 4 and Figure 10 for results for all tasks).

Figure 9
Usable Frames by Age



Note. Percent usable frames between 6-month-olds and 12-month-olds across tasks where each dot represents an individual participant's data. See the online article for the color version of this figure.

Table 3
Summary of Usable Data and Side Bias

Task	Participant with usable data		Participant with side bias		Median (IQR) number of trials with usable data (excluding infants with side bias)	
	Age 6	Age 12	Age 6	Age 12	Age 6	Age 12
Geo-Social Attention	28 (96.6)	26 (100.0)	5 (17.9)	6 (23.1)	6.0 (4.5, 9.5)	8.00 (5.8, 9.0)
The VPC	26 (89.7)	23 (88.5)	1 (3.8)	2 (8.7)	5.0 (3.0, 6.0)	5.0 (4.0, 5.0)
Prediction[a]	28 (96.6)	26 (100.0)			11.0 (10.0, 12.0)	10.0 (8.5, 11.5)
Audiovisual Synchrony	27 (93.1)	25 (96.2)	7 (25.9)	7 (28.0)	7.0 (4.0, 8.0)	8.0 (5.2, 8.0)
Word Comprehension	21 (72.4)	19 (73.1)	3 (14.2)	4 (21.1)	8.0 (4.5, 12.0)	8.0 (4.0, 10.0)
Numeracy	24 (82.8)	23 (88.5)	10 (41.7)	12 (52.2)	4.0 (3.0, 4.0)	4.0 (2.0, 4.0)
Helper/Hinderer	20 (67.0)	19 (73.1)	1 (5.0)	2 (10.5)	2.0 (1.0, 2.0)	2.0 (1.0, 2.0)

Note. Number (and percent) of infants with sufficient usable data (first column, as % of full sample of 55 [6-month-olds: 29, 12-month-olds: 26] infants) and fixation on a single side >80% of the task (second column, as % of infants with usable data). IQR = interquartile range; VPC = Visual Paired Comparison; AEM = anticipatory eye movements.
[a] No side bias was assessed for the prediction task because analysis was restricted to the short AEM window.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

The VPC Task

Our preregistered hypothesis for the VPC task was that 12-month-olds would show a stronger novelty effect than 6-month-olds and that novelty effects for faces would be stronger than for fribbles. With respect to age, we found that both 6- and 12-month-olds showed preference (mean proportion >0.50) for novel faces (12-month-olds: 0.57 [0.54, 0.61], $p < .001$; 6-month-olds: 0.52 [0.49, 0.56], $p = .157$) and nonface objects (fribbles; 12-month-olds: 0.55 [0.51, 0.59], $p = .007$; 6-month-olds: 0.53 [0.49, 0.56], $p = .124$). These effects were only significant in the 12-month-olds. Although 12-month-olds had higher mean proportions of looking times than 6-month-olds, the differences were not statistically significant for either stimulus type ($p = .052$ and $p = .343$). Contrary to our predictions, the effect was not stronger for faces than for fribbles ($p > .05$).

The Prediction Task

Our preregistered hypothesis was that 12-month-old infants would demonstrate more AEM than 6-month-old infants. Indeed, there were larger effects in the 12-month group in both Blocks 1 and 2 than in the 6-month group. The differences between the two age groups, however, were not significant (Block 1: 12-month-olds: 0.57 [0.4, 0.8], $p = .45$; 6-month-olds: 0.45 [0.3, 0.6], $p = .56$; 12-month-olds vs. 6-month-olds: 0.12 [−0.1, 0.4], $p = .34$ and Block 2: 12-month-olds: 0.78 [0.6, 0.9], $p = .01$; 6-month-olds: 0.58 [0.4, 0.8], $p = .35$; 12-month-olds vs. 6-month-olds: 0.2 [−0.1, 0.5], $p = .13$). Both 6-month-old and 12-month-old infants showed an increase in performance between blocks; however, these changes were not significant ($p > .05$).

The AV Synchrony Task

In contrast to our preregistered predictions that infants would show preference for synchronous videos at both 6 and 12 months, we did not find reliable preference for the synchronous stimuli at either age on social (12-month-olds: 0.50 [0.45, 0.55], $p = .956$; 6-month-olds: 0.50 [0.45, 0.55], $p = .946$) or nonsocial trials (12-month-olds: 0.52 [0.47, 0.57], $p = .565$; 6-month-olds: 0.50 [0.45, 0.55], $p = .900$). There was no difference in the effects between social and nonsocial stimuli or between the age groups ($p > .05$). Modifications to this task for future instantiations of the battery are addressed in the Discussion section.

The Word Comprehension Task

As we hypothesized in our preregistration, 12-month-olds showed greater difference scores than 6-month-olds for the Word Comprehension task overall (12-month-olds: 0.05 [−0.04, 0.14], $p = .235$; 6-month-olds: −0.05 [−0.14, 0.03], $p = .191$; mean difference: 0.11 [0.00, 0.22], $p = .044$). However, neither group looked reliably more to the target object when it was the target than they did when it was the distracter (i.e., difference scores did not differ significantly from 0 in either group).

The Numeracy Task

In the Numeracy task, we predicted that infants at 6 and 12 months would show a preference for the alternating streams regardless of whether the numerically nonchanging stream contained smaller (five items) or larger quantities (20 items). While the 12-month-old infants showed reliable preference (mean proportion >0.50) for alternating streams, the same effect was not observed among 6-month-olds when the nonalternating stream contained the smaller quantity (12-month-olds: 0.69 [0.52, 0.86], $p = .027$; 6-month-olds: 0.56 [0.44, 0.67], $p = .354$) nor when the nonalternating stream contained the larger quantity (12-month-olds: 0.60 [0.50, 0.70], $p = .044$; 6-month-olds: 0.54 [0.45, 0.63], $p = .414$). There were no statistically significant differences between age groups ($p > .05$).

The Helper/Hinderer Task

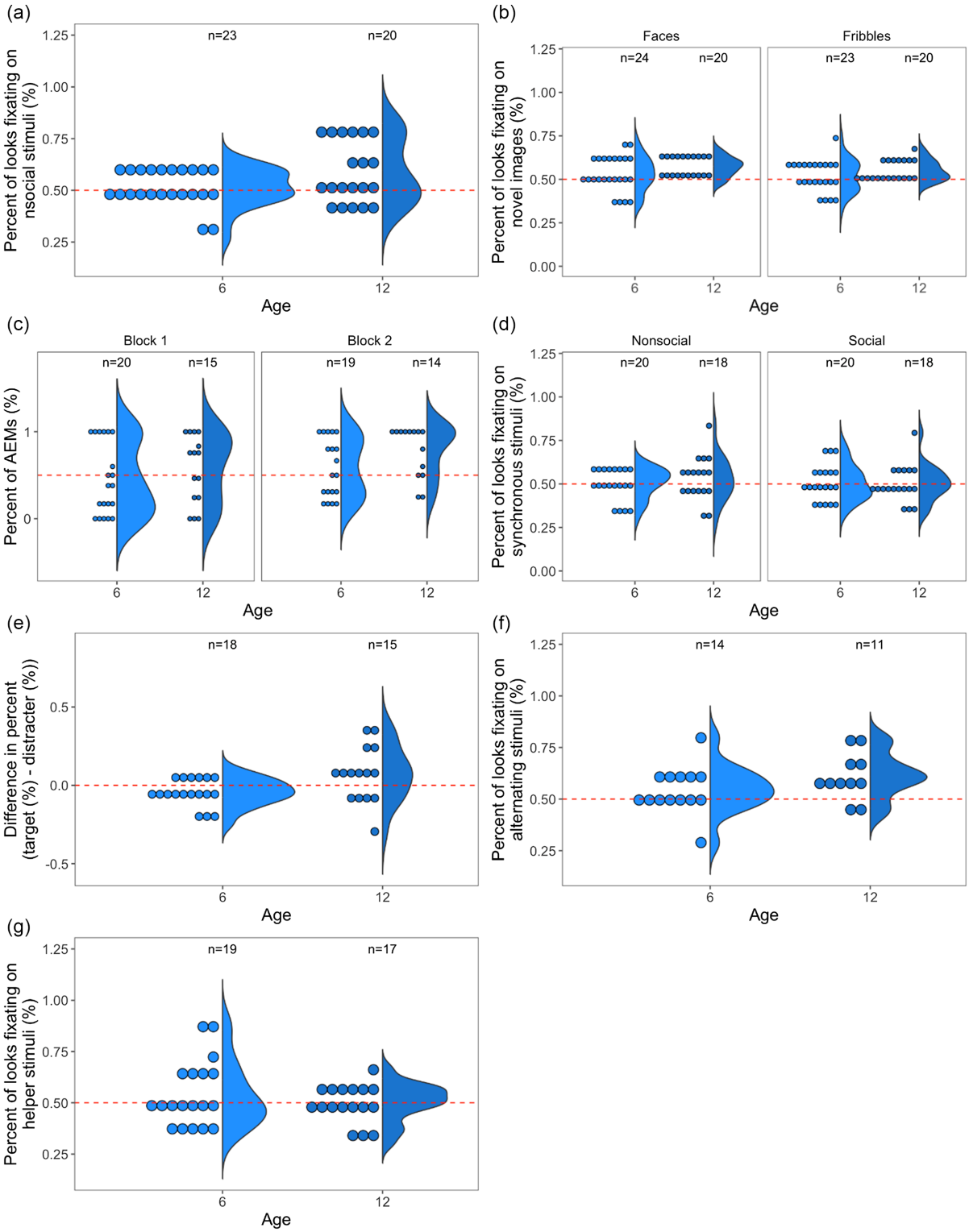
Contrary to our predictions, we did not find reliable preference for the helper at test (12-month-olds: 0.50 [0.44, 0.56], $p = .990$; 6-month-olds: 0.53 [0.47, 0.59], $p = .329$). There were also no statistically significant differences between age groups ($p > .05$). Possible reasons for this are addressed in the Discussion section.

Discussion

This study was intended to test the feasibility of the RISE Battery as a remote and scalable approach to studying infant cognitive and

Figure 10

Results for (a) Geo-Spatial, (b) Visual Paired Comparison, (c) Prediction, (d) Audiovisual Synchrony, (e) Word Comprehension, (f) Numeracy, and (g) Helper/Hinderer



Note. AEM = anticipatory eye movement. See the online article for the color version of this figure.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 4
Summary of Results by Task and Age

Task	Overall		6 months		12 months		Age effect (12 months vs. 6 months)	
	Model estimated mean [95% CI][a]	p[a]	Model estimated mean [95% CI][a]	p[a]	Model estimated mean [95% CI][a]	p[a]	Model estimated mean difference [95% CI][b]	p[b]
Geo-Social Attention	0.55 [.50, .61]	.071	0.51 [0.44, 0.57]	.819	0.59 [0.53, 0.66]	**.006**	0.09 [0.01, 0.16]	**.025**
VPC social (face) trials	0.55 [0.52, 0.58]	**<.001**	0.52 [0.49, 0.56]	.157	0.57 [0.53, 0.61]	**<.001**	0.05 [0.00, 0.10]	.052
VPC nonsocial (fribbles) trials	0.54 [0.51, 0.57]	**.004**	0.53 [0.49, 0.56]	.124	0.55 [0.51, 0.59]	**.007**	0.02 [−0.03, 0.07]	.343
AV Synchrony social trials	0.50 [0.47, 0.54]	.994	0.50 [0.45, 0.55]	.946	0.50 [0.45, 0.55]	.956	−0.00 [−0.07, 0.07]	.931
AV Synchrony nonsocial trials	0.51 [0.47, 0.54]	.617	0.50 [0.45, 0.55]	.900	0.52 [0.47, 0.57]	.565	0.01 [−0.06, 0.08]	.746
Prediction block 1	0.51 [0.39, 0.63]	.86	0.45 [0.29, 0.61]	.55	0.57 [0.39, 0.76]	.45	0.12 [−0.12, 0.36]	.34
Prediction block 2	0.68 [0.55, 0.80]	**.01**	0.58 [0.42, 0.74]	.34	0.78 [0.59, 0.97]	**<.001**	0.20 [−0.05, 0.45]	.12
Numeracy five-item constant stream	0.62 [0.52, 0.73]	**.023**	0.56 [0.44, 0.67]	.354	0.69 [0.52, 0.86]	**.027**	0.134 [−0.06, 0.33]	.183
Numeracy 20-item constant	0.57 [0.50, 0.64]	.062	0.54 [0.45, 0.63]	.414	0.60 [0.50, 0.70]	**.044**	0.06 [−0.06, 0.18]	.328
Helper/Hinderer	0.52 [0.47, 0.56]	.503	0.53 [0.47, 0.59]	.329	0.50 [0.44, 0.56]	.99	−0.03 [−0.12, 0.06]	.492
Word Comprehension	0 [−0.07, 0.07]	.999	−0.05 [−0.14, 0.03]	.191	0.05 [−0.04, 0.14]	.235	0.11 [0.00, 0.22]	**.044**

Note. Bold values indicate $p < .05$. CI = confidence interval; VPC = Visual Paired Comparison; AV = audiovisual.
[a] $H_0 = 50.0$. [b] $H_0 = 0.0$.

language development. We found that a subset of tasks (Geo-Social, VPC, Prediction, and Numeracy) were both feasible and valid in the remote setting using the Children Helping Science platform and automated coding of infant gaze behavior using the open-access software, iCatcher+. Though we did not obtain our anticipated results on all tasks (i.e., AV Synchrony, Word Comprehension, Helper/Hinderer), many of the tasks did replicate in the remote setting (i.e., Geo-Social, VPC, Prediction, and Numeracy) with automated coding of infant behaviors. Here, we summarize task-based findings, age-based findings, limitations of this work, and future directions. For tasks in which we did not obtain the anticipated results, it is possible that this was due to administration in the home setting. However, given growing evidence that remote administration of looking time tasks is comparable to lab results (e.g., Bacon et al., 2021; Bánki et al., 2022; Chuey et al., 2024), we do not believe that this is the cause for these differences. Rather, we provide our interpretations for differences in outcomes by task below.

Task-Based Results

In the Geo-Social task, we observed the predicted preference for social stimuli among 12-month-olds but not 6-month-olds. This task has previously been used in children older than the infants tested here (Pierce et al., 2011), and thus, it may be the case that we would not see these results among 6-month-olds, even in a larger sample of infants (cf. Imafuku et al., 2017). Though preference for faces has been demonstrated in infants as young as newborns (Fantz, 1964; Johnson et al., 1991), the stimuli used here showed children dancing and may not have provided a clear enough view of the faces to appeal to infants as young as 6 months. To address the possibility, we intend to focus more on the children's faces in future versions of this battery.

In the VPC task, we found preference for the novel stimuli at test in both 6- and 12-month-old infants. However, we did not see an age effect nor a stimulus-type effect (fribbles vs. faces), that is, infants were as successful at recognizing the novel fribble as they were at recognizing the novel face. It will be interesting to see if a larger sample leads to an age or stimulus effect.

In the Prediction task, infants were tested on their ability to generate predictions to a visual target before versus after it switched from one side of the screen to the other. While the original study using this task tested 12- and 18-month-olds (Reuter et al., 2018), our results were generally consistent. In the first half of the task, when the target stimulus appeared eight times in the same location, 12- but not 6-month-olds showed reliable AEMs, and the difference between these two age groups was significant. In the second half of the task, when the target appeared eight times in a new location—which required infants to update their prediction—both 6- and 12-month-olds showed reliable AEMs, with no developmental change. These robust results with typically developing infants give us confidence that the task will be feasible for use with infants on atypical developmental trajectories. This would allow us to evaluate whether individual differences documented by Reuter et al. (2018) are related to clinically meaningful developmental outcomes.

In the AV Synchrony task, we did not obtain the anticipated preference for the synchronous stimuli in the social or nonsocial tasks at either age. Though previous work with typically developing toddlers found reliable preference for synchronous videos at both 666 and 1,000 ms (Righi et al., 2018), the remote nature and younger

age range of the RISE Battery may require larger delays in order to be recognized by infants. To address this, we intend to extend the delay in future versions of the battery. It is also possible that challenges with internet speed and reliable connections in the home setting will prevent us from obtaining the anticipated results, even with a longer delay.

Relative to other tasks in the battery, the Word Recognition task may have been less engaging for infants due to the absence of animated videos or faces. It is also the case that the current implementation of this task differed from previous work (e.g., Bergelson & Swingley, 2012) in two critical ways (in addition to those noted in the Method section): (a) Object location was not counterbalanced between blocks (making the task more repetitive) and (b) trial lengths were nearly twice as long. These differences may have contributed to the lack of reliable difference scores, even among the 12-month-olds. To address this, we intend to add more attention grabbing stimuli between trials. Though the looking-while-listening paradigm has been used in many lab-based studies (Bergelson & Swingley, 2012; Fernald et al., 2008; Fernald & Marchman, 2012), when included in a longer battery context that included more dynamic stimuli, increased attention grabbing may be required.

Though we obtained the predicted effects for numeracy among 12-month-old infants, replicating effects previously seen in the lab with younger and older children (Braham & Libertus, 2017; Libertus & Brannon, 2010), the same results were not observed in 6-month-olds. Previous lab-based studies with 6-month-olds (Libertus & Brannon, 2010; Starr et al., 2013) used longer trials (60 s instead of 20 s used here) or presented the image streams on two separate screens, resulting in larger displays that may help younger infants distinguish the sets and individual dots in each set more readily. One other study presented image streams on the same screen, albeit with 8-month-old infants, 10-s trials, and three different ratios between the quantities, but unfortunately did not report whether infants preferred the numerically alternating streams over the nonalternating ones (Schröder et al., 2020). However, the reported means for the 1:4-ratio trials in this study (55% and 61% for two different groups of infants at pretest) are very similar to the means in the present study (56% and 54% for trials with small and large quantities in the nonalternating streams, respectively). Thus, it is possible that our failure to replicate previous lab-based studies in 6-month-olds may be due to the shorter trial duration or presentation on a single screen with smaller stimuli. Future iterations of this task should test whether increasing the trial duration would replicate lab-based results.

The lack of anticipated results in the Helper–Hinderer task may reflect the challenges of transitioning what has historically been administered as an infant-controlled habituation task to asynchronous administration in the remote setting where videos must be of predetermined length. Indeed, the vast majority of past studies demonstrating preferences for helpers over hinderers in infants between 6 and 12 months of age have utilized infant-controlled habituation procedures, in-lab testing settings with live, in-person puppet shows displayed in 3-D, and preferential reaching rather than preferential looking as the dependent variable of interest (see review in Margoni & Surian, 2018). We are aware of one article that demonstrated a preference for helpers over hinderers in 5-month-old infants using familiarization and preferential looking methods similar to what was utilized here (Tan & Hamlin, 2022); however, the stimuli in that study were cartoons in which the events unfolded significantly more slowly than in the current videos; this slowness may have facilitated processing and memory. Finally, whereas the key differences between stimuli in other preferential looking tasks delivered in the RISE Battery were observable during the preferential looking trials themselves (e.g., the state of being social or not, numerical changes or not, a familiar stimulus vs. a novel one, etc.), here infants had to remember what had happened during previous, equally familiar events in order to distinguish the helper from the hinderer. Thus, presumably, processing and memory requirements for this task were quite high. Future pilots should attempt to reduce these processing requirements, for instance, by examining infants' preferential looks to side-by-side displays of helping versus hindering events.

Age Effects

For 12-month-olds, we obtained expected results in the Geo-Social, VPC, Prediction, and Numeracy tasks. Neither age group showed predicted results in the AV Synchrony, Word Recognition, or Helper–Hinderer tasks. Among 6-month-olds, we also obtained unexpected results in the Geo-Social and Numeracy tasks. These unexpected results among 6-month-olds in many of the tasks make it challenging to conclude that this battery is appropriate for use in younger infants. We are currently exploring whether the lack of expected results is due to differences in task design, sample size, or true challenges with this approach or potential differences in performance among 6-month-olds.

Limitations

This pilot study of the RISE Battery had a number of limitations. First, because it was intended to replicate robust effects in lab-based versions of these tasks, the samples were quite small. Replication with larger samples will be important for establishing the validity of these tasks. A second limitation is related to the substantial rates of missing data. In the present study, we tested multiple paradigms to determine which can be implemented effectively in the home setting. This information is valuable as we design future versions of this battery, taking into account the need for animated videos or stimuli that last long enough to be effective even if the infant looks away for large portions of the trial. A third limitation lies in the balance of tasks that are designed for group-level measures being used to identify individual differences. Though our assumption is that infants without known developmental delays should, as a group, perform in a way that conforms to lab-based results, it is possible that variability in performance would actually serve the larger goal of identifying individual differences better than one with consistent results, even in a group such as this. This possibility is addressed further in our future directions discussion. Lack of covariates such as sociodemographic variables was an additional limitation of the present study due to the small sample size of this pilot study and overrepresentation of high-socioeconomic status families in the current sample.

Future Directions

Though not without its limitations, as described, this pilot study provides proof of concept for the potential utility of the RISE Battery and lays the foundation for future investigations including

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

the following: (a) We will explore the predictive value of the RISE Battery on parent-reported developmental outcomes 6 months following participation. This will allow us to examine whether the RISE Battery holds potential value as a marker of infant development at the individual level. (b) We will modify the RISE Battery with lessons gathered in this pilot study so that we may align our results further with the predicted outcomes. Modifications include zooming in on the faces of the children in the Geo-Social Attention task, increasing the lag between synchronous and asynchronous videos in the AV Synchrony task, adding more attention-grabbing displays between stimuli in the Word Comprehension task, and increasing the trial duration for 6-month-olds in the Numeracy task. (c) We will extend data collection to infant siblings of children with autism so that we may explore responses to the battery among infants with typical and potentially atypical trajectories of development. Ultimately, we hope that the RISE Battery will offer a novel way to assess cognitive development that is scalable to a large sample of infants from diverse locations and backgrounds. This approach has the potential to improve our understanding of early developmental trajectories that could contribute to more effective early screening and shed light on specific cognitive processes that underlie early-emerging autism-related difficulties.

Conclusions

This work has demonstrated the feasibility of remote assessment and automated coding of group-level measures of infant abilities in tasks assessing attention (Geo-Social), memory (VPC), learning (Prediction), and approximate number representation (Numeracy). Interestingly, perhaps due to the substantial differences between the tasks in this battery, within each task, the order of presentation did not affect the rates of usable frames, suggesting that infants were able to remain engaged at comparable levels across the 20 min battery (see Supplemental Materials). An alternative description of this result, however, is that infants were just as inattentive at the start of the battery as they were at the end. This was highly task dependent and underscores the need for animated videos and/or trials that allow for significant time looking away from the screen to obtain the anticipated results. As described above, much work remains to be done in establishing the relevance of performance on these tasks for the abilities of infants at the individual level.

References

Arunachalam, S., & Luyster, R. J. (2016). The integrity of lexical acquisition mechanisms in autism spectrum disorders: A research review. *Autism Research*, *9*(8), 810–828. https://doi.org/10.1002/aur.1590

Bacon, D., Weaver, H., & Saffran, J. (2021). A framework for online experimenter-moderated looking-time studies assessing infants' linguistic knowledge. *Frontiers in Psychology*, *12*, Article 703839. https://doi.org/10.3389/fpsyg.2021.703839

Bánki, A., de Eccher, M., Falschlehner, L., Hoehl, S., & Markova, G. (2022). Comparing online webcam- and laboratory-based eye-tracking for the assessment of infants' audio-visual synchrony perception. *Frontiers in Psychology*, *12*, Article 733933. https://doi.org/10.3389/fpsyg.2021.733933

Barry, T. J., Griffith, J. W., De Rossi, S., & Hermans, D. (2014). Meet the Fribbles: Novel stimuli for use within behavioural research. *Frontiers in Psychology*, *5*, Article 103. https://doi.org/10.3389/fpsyg.2014.00103

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(9), 3253–3258. https://doi.org/10.1073/pnas.1113380109

Bradshaw, J., Shic, F., & Chawarska, K. (2011). Brief report: Face-specific recognition deficits in young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *41*(10), 1429–1435. https://doi.org/10.1007/s10803-010-1150-4

Braham, E. J., & Libertus, M. E. (2017). Intergenerational associations in numerical approximation and mathematical abilities. *Developmental Science*, *20*(5), Article e12436. https://doi.org/10.1111/desc.12436

Campbell, K., Carpenter, K. L., Hashemi, J., Espinosa, S., Marsan, S., Borg, J. S., Chang, Z., Qiu, Q., Vermeer, S., Adler, E., Tepper, M., Egger, H. L., Baker, J. P., Sapiro, G., & Dawson, G. (2019). Computer vision analysis captures atypical attention in toddlers with autism. *Autism*, *23*(3), 619–628. https://doi.org/10.1177/1362361318766247

Cannon, J., O'Brien, A. M., Bungert, L., & Sinha, P. (2021). Prediction in autism spectrum disorder: A systematic review of empirical evidence. *Autism Research*, *14*(4), 604–630. https://doi.org/10.1002/aur.2482

Cao, P., Tan, X., Scott, K., & Liu, S. (2021). *iCatcher+: Robust and automatic gaze classification of infant webcam videos*. PsyArXiv.

Chawarska, K., Macari, S., & Shic, F. (2013). Decreased spontaneous attention to social scenes in 6-month-old infants later diagnosed with autism spectrum disorders. *Biological Psychiatry*, *74*(3), 195–203. https://doi.org/10.1016/j.biopsych.2012.11.022

Chawarska, K., Ye, S., Shic, F., & Chen, L. (2016). Multilevel differences in spontaneous social attention in toddlers with autism spectrum disorder. *Child Development*, *87*(2), 543–557. https://doi.org/10.1111/cdev.12473

Chita-Tegmark, M., Arunachalam, S., Nelson, C. A., & Tager-Flusberg, H. (2015). Eye-tracking measurements of language processing: Developmental differences in children at high risk for ASD. *Journal of Autism and Developmental Disorders*, *45*(10), 3327–3338. https://doi.org/10.1007/s10803-015-2495-5

Chuey, A., Boyce, V., Cao, A., & Frank, M. C. (2024). *Conducting developmental research online vs. in-person: A meta-analysis*. https://doi.org/10.31234/osf.io/qc6fw

Colombo, J., Shaddy, D. J., Richman, W. A., Maikranz, J. M., & Blaga, O. M. (2004). The developmental course of habituation in infancy and preschool outcome. *Infancy*, *5*(1), 1–38. https://doi.org/10.1207/s15327078in0501_1

Decarli, G., Piazza, M., & Izard, V. (2023). Are infants' preferences in the number change detection paradigm driven by sequence patterns? *Infancy*, *28*(2), 206–217. https://doi.org/10.1111/infa.12505

Elsabbagh, M., Volein, A., Holmboe, K., Tucker, L., Csibra, G., Baron-Cohen, S., Bolton, P., Charman, T., Baird, G., & Johnson, M. H. (2009). Visual orienting in the early broader autism phenotype: Disengagement and facilitation. *Journal of Child Psychology and Psychiatry*, *50*(5), 637–642. https://doi.org/10.1111/j.1469-7610.2008.02051.x

Erel, Y., Shannon, K. A., Chu, J., Scott, K., Kline Struhl, M., Cao, P., Tan, X., Hart, P., Raz, G., Piccolo, S., Mei, C., Potter, C., Jaffe-Dax, S., Lew-Williams, C., Tenenbaum, J., Fairchild, K., Bermano, A., & Liu, S. (2023). iCatcher+: Robust and automated annotation of infants' and young children's gaze behavior from videos collected in laboratory, field, and online studies. *Advances in Methods and Practices in Psychological Science*, *6*(2). https://doi.org/10.1177/25152459221147250

Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science*, *146*(3644), 668–670. https://doi.org/10.1126/science.146.3644.668

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories—User's guide and technical manual second edition*. Brookes Publishing.

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

and late-talking toddlers. *Child Development*, *83*(1), 203–222. https://doi.org/10.1111/j.1467-8624.2011.01692.x

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 97–135). John Benjamins. https://doi.org/10.1075/lald.44.06fer

Frank, M. C., Vul, E., & Johnson, S. P. (2009). Development of infants' attention to faces during the first year. *Cognition*, *110*(2), 160–170. https://doi.org/10.1016/j.cognition.2008.11.010

Guiraud, J. A., Tomalski, P., Kushnerenko, E., Ribeiro, H., Davies, K., Charman, T., Elsabbagh, M., Johnson, M. H., & the BASIS Team. (2012). Atypical audiovisual speech integration in infants at risk for autism. *PLOS ONE*, *7*(5), Article e36428. https://doi.org/10.1371/journal.pone.0036428

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557–559. https://doi.org/10.1038/nature06288

Imafuku, M., Kawai, M., Niwa, F., Shinya, Y., Inagawa, M., & Myowa-Yamakoshi, M. (2017). Preference for dynamic human images and gaze-following abilities in preterm infants at 6 and 12 months of age: An eye-tracking study. *Infancy*, *22*(2), 223–239. https://doi.org/10.1111/infa.12144

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(25), 10382–10385. https://doi.org/10.1073/pnas.0812142106

Jeste, S. S., Kirkham, N., Senturk, D., Hasenstab, K., Sugar, C., Kupelian, C., Baker, E., Sanders, A. J., Shimizu, C., Norona, A., Paparella, T., Freeman, S. F., & Johnson, S. P. (2015). Electrophysiological evidence of heterogeneity in visual statistical learning in young children with ASD. *Developmental Science*, *18*(1), 90–105. https://doi.org/10.1111/desc.12188

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1–2), 1–19. https://doi.org/10.1016/0010-0277(91)90045-6

Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2–6-month-old infants later diagnosed with autism. *Nature*, *504*(7480), 427–431. https://doi.org/10.1038/nature12715

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, *218*(4577), 1138–1141. https://doi.org/10.1126/science.7146899

Li, X., Li, J., Zhao, S., Liao, Y., Zhu, L., & Mou, Y. (2024). Magnitude representation of preschool children with autism spectrum condition. *Autism: An International Journal of Research and Practise*, *28*(4), 866–880. https://doi.org/10.1177/13623613231185408

Libertus, M. E., & Brannon, E. M. (2010). Stable individual differences in number discrimination in infancy. *Developmental Science*, *13*(6), 900–906. https://doi.org/10.1111/j.1467-7687.2009.00948.x

Lockman, J. J., & Tamis-LeMonda, C. S. (2020). *The Cambridge handbook of infant development: Brain, behavior, and cultural context*. Cambridge University Press. https://doi.org/10.1017/9781108351959

Manns, J. R., Stark, C. E. L., & Squire, L. R. (2000). The visual paired-comparison task as a measure of declarative memory. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(22), 12375–12379. https://doi.org/10.1073/pnas.220398097

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

Narayan, C. R., Werker, J. F., & Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination. *Developmental Science*, *13*(3), 407–420. https://doi.org/10.1111/j.1467-7687.2009.00898.x

Oakes, L., & Amso, D. (2018). Development of visual attention. In S. Ghetti (Ed.), *Stevens' handbook of experimental psychology and cognitive neuroscience* (Vol. 4, pp. 1–33). Wiley.

Olson, R. H., Pomper, R., Potter, C. E., Hay, J. F., Saffran, J. R., Ellis Weismer, S., & Lew-Williams, C. (2020). *Peyecoder: An open-source program for coding eye movements* (v1.0.2-beta) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.3939234

Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science*, *296*(5571), 1321–1323. https://doi.org/10.1126/science.1070223

Patterson, M. L., & Werker, J. F. (1999). Matching phonetic information in lips and voice is robust in 4.5-month-old infants. *Infant Behavior & Development*, *22*(2), 237–247. https://doi.org/10.1016/S0163-6383(99)00003-X

Peltola, M. J., Leppänen, J. M., Mäki, S., & Hietanen, J. K. (2009). Emergence of enhanced attention to fearful faces between 5 and 7 months of age. *Social Cognitive and Affective Neuroscience*, *4*(2), 134–142. https://doi.org/10.1093/scan/nsn046

Pickron, C. B., Fava, E., & Scott, L. S. (2017). Follow my gaze: Face race and sex influence gaze-cued attention in infancy. *Infancy*, *22*(5), 626–644. https://doi.org/10.1111/infa.12180

Pierce, K., Conant, D., Hazin, R., Stoner, R., & Desmond, J. (2011). Preference for geometric patterns early in life as a risk factor for autism. *Archives of General Psychiatry*, *68*(1), 101–109. https://doi.org/10.1001/archgenpsychiatry.2010.113

Pierce, K., Marinero, S., Hazin, R., McKenna, B., Barnes, C. C., & Malige, A. (2016). Eye tracking reveals abnormal visual preference for geometric images as an early biomarker of an autism spectrum disorder subtype associated with increased symptom severity. *Biological Psychiatry*, *79*(8), 657–666. https://doi.org/10.1016/j.biopsych.2015.03.032

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

Righi, G., Tenenbaum, E. J., McCormick, C., Blossom, M., Amso, D., & Sheinkopf, S. J. (2018). Sensitivity to audio-visual synchrony and its relation to language abilities in children with and without ASD. *Autism Research*, *11*(4), 645–653. https://doi.org/10.1002/aur.1918

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology*, *69*(1), 181–203. https://doi.org/10.1146/annurev-psych-122216-011805

Schröder, E., Gredebäck, G., Gunnarsson, J., & Lindskog, M. (2020). Play enhances visual form perception in infancy—An active training study. *Developmental Science*, *23*(3), Article e12923. https://doi.org/10.1111/desc.12923

Scott, K., Chu, J., & Schulz, L. (2017). Lookit (Part 2): Assessing the viability of online developmental research, results from three case studies. *Open Mind: Discoveries in Cognitive Science*, *1*(1), 15–29. https://doi.org/10.1162/OPMI_a_00001

Scott, K., & Schulz, L. (2017). Lookit (Part 1): A new online platform for developmental research. *Open Mind: Discoveries in Cognitive Science*, *1*(1), 4–14. https://doi.org/10.1162/OPMI_a_00002

Senju, A. (2013). Atypical development of spontaneous social cognition in autism spectrum disorders. *Brain & Development*, *35*(2), 96–101. https://doi.org/10.1016/j.braindev.2012.08.002

Shi, L., Zhou, Y., Ou, J., Gong, J., Wang, S., Cui, X., Lyu, H., Zhao, J., & Luo, X. (2015). Different visual preference patterns in response to simple and complex dynamic social stimuli in preschool-aged children with autism spectrum disorders. *PLOS ONE*, *10*(3), Article e0122280. https://doi.org/10.1371/journal.pone.0122280

Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., Diamond, S. P., & Held, R. M. (2014). Autism as a disorder of prediction. *Proceedings of the National Academy of Sciences of the*

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

United States of America, *111*(42), 15220–15225. https://doi.org/10.1073/pnas.1416797111

Snyder, K. A. (2010). Neural correlates of encoding predict infants' memory in the paired-comparison procedure. *Infancy*, *15*(3), 270–299. https://doi.org/10.1111/j.1532-7078.2009.00015.x

Starr, A., Libertus, M. E., & Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(45), 18116–18120. https://doi.org/10.1073/pnas.1302751110

Stevenson, R. A., Segers, M., Ferber, S., Barense, M. D., Camarata, S., & Wallace, M. T. (2016). Keeping time in the brain: Autism spectrum disorder and audiovisual temporal processing. *Autism Research*, *9*(7), 720–738. https://doi.org/10.1002/aur.1566

Tan, E., & Hamlin, J. K. (2022). Infants' neural responses to helping and hindering scenarios. *Developmental Cognitive Neuroscience*, *54*, Article 101095. https://doi.org/10.1016/j.dcn.2022.101095

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

Tenenbaum, E. J., Sobel, D. M., Sheinkopf, S. J., Malle, B. F., & Morgan, J. L. (2015). Attention to the mouth and gaze following in infancy predict language development. *Journal of Child Language*, *42*(6), 1173–1190. https://doi.org/10.1017/S0305000914000725

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B. J., & Nelson, C. A. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. https://doi.org/10.1016/j.psychres.2008.05.006

Venker, C. E., Bean, A., & Kover, S. T. (2018). Auditory-visual misalignment: A theoretical perspective on vocabulary delays in children with ASD. *Autism Research*, *11*(12), 1621–1628. https://doi.org/10.1002/aur.2038

Venker, C. E., Eernisse, E. R., Saffran, J. R., & Ellis Weismer, S. (2013). Individual differences in the real-time comprehension of children with ASD. *Autism Research*, *6*(5), 417–432. https://doi.org/10.1002/aur.1304

Wagner, J. B., Jabès, A., Norwood, A., & Nelson, C. A. (2020). Attentional measures of memory in typically developing and hypoxic-ischemic injured infants. *Brain Sciences*, *10*(11), Article 823. https://doi.org/10.3390/brainsci10110823

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), B1–B11. https://doi.org/10.1016/S0010-0277(99)00066-9

Received February 29, 2024
Revision received June 26, 2024
Accepted August 2, 2024 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.