Title

Examining infants visual paired comparison performance in the US and rural Malawi.

Journal

Developmental Science, 27(5)

Authors

Beckner, Aaron

Arnold, Charles

Bragg, Megan

et al.

Publication Date

2024-09-01

DOI

10.1111/desc.13439

Peer reviewed

HHS Public Access

Author manuscript

Dev Sci. Author manuscript; available in PMC 2024 September 01.

Published in final edited form as:

Dev Sci. 2024 September ; 27(5): e13439. doi:10.1111/desc.13439.

Examining infants' visual paired comparison performance in the US and rural Malawi

Aaron G. Beckner[1], **Charles D. Arnold**[2], **Megan G. Bragg**[2], **Bess L. Caswell**[2], **Zhijun Chen**[3], **Katherine Cox**[2], **Michaela C. DeBolt**[3], **Matthews George**[4], **Kenneth Maleta**[4], **Christine Stewart**[2], **Lisa M. Oakes**[3], **Elizabeth Prado**[2]

[1]College of Human Ecology, Cornell University, Ithaca, New York, USA

[2]Department of Nutrition and Institute for Global Nutrition, University of California, Davis, California, USA

[3]Department of Psychology and Center for Mind and Brain, University of California, Davis, California, USA

[4]School of Global and Public Health, Kamuzu University of Health Sciences, Blantyre, Malawi

Abstract

Measures of attention and memory were evaluated in 6- to 9-month-old infants from two diverse contexts. One sample consisted of African infants residing in rural Malawi (N = 228, 118 girls, 110 boys). The other sample consisted of racially diverse infants residing in suburban California (N=48, 24 girls, 24 boys). Infants were tested in an eye-tracking version of the visual paired comparison procedure and were shown racially familiar faces. The eye tracking data were parsed into individual looks, revealing that both groups of infants showed significant memory performance. However, how a look was operationally defined impacted some—but not other—measures of infant VPC performance.

Keywords

culture-specific attention; eye tracking; infant cognition; visual recognition memory

1 | INTRODUCTION

Decades of research has examined how aspects of infants' looking behavior influences their learning and subsequent memory for objects, faces, and other stimuli. The visual paired comparison (VPC) procedure has been considered a "gold standard" for examining various factors that influence infants' looking and subsequent learning of attended items since the 1970s (Fagan, 1970, 1972, 1974). This is due—at least in part—to the fact that infant looking behavior in the VPC procedure is thought to reflect variation in basic cognitive processes such as information processing speed (for review, see Rose et al., 2004). Infant

Correspondence: Aaron G. Beckner, Martha Van Rensselaer Hall, Ithaca, Cornell University, NY 14850, USA. agb222@cornell.edu.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

VPC studies typically examine look-defined measures of information processing that include the longest individual look infants execute during learning (i.e., peak look duration), how rapidly infants shift attention between multiple objects per second of looking during learning and memory recollection (i.e., shift rate), and infants' visual preference for a novel stimulus paired with a familiar stimulus after an initial learning phase (i.e., novelty preference score). One limitation of the existing literature is that these studies have primarily been administered to a narrow slice of the world's population, making it unclear whether the existing literature provides an appropriate baseline for interpreting information processing in samples of infants reared in culturally distinct contexts.

There are at least three sets of observed findings that have been taken as evidence that look-defined measures in the VPC procedure reflect attention and information processing: (1) the associations between look-defined measures of attention during learning and/or recollection and their impact on subsequent memory performance, (2) the impact of variables such as age, stimulus complexity, and individual differences in attention on infant learning and memory, and (3) prediction of long-term cognitive outcomes and consistent impacts of risk factors and interventions on infant VPC performance. Each of these three patterns is discussed in greater detail below.

The first set of observed findings reflects the associations observed between look-defined measures of attention and memory in infants. Specifically, measures of attention in the VPC procedure predict infants' subsequent memory performance as measured by their visual preference for novel stimuli over familiar stimuli (i.e., novelty preference score) (Courage & Howe, 2001; Rose et al., 2001). Shorter look duration during learning (i.e., peak look duration) and faster shifting between simultaneously presented stimuli during learning and recollection (i.e., shift rate) are associated with stronger novelty preference, supporting the conclusion that these patterns of looking behavior reflect more efficient information processing in infants.

The second set of findings relates to the impact of variables such as age and stimulus complexity on these measures. Infant VPC studies have shown that novelty preference scores vary as a function of age (Colombo et al., 1988; Fagan, 1970, 1972, 1974; Hunter et al., 1983; Rose, 1981, 1983; Rose et al., 2004, 1982) and stimulus complexity (Hunter et al., 1983). Moreover, younger infants, compared to older infants, require more exposure and longer accumulated looking to show a robust novelty preference score (Fagan, 1974; Rose et al., 1982). The effect of age on memory performance as well as age-related differences in the amount of exposure time and accumulated looking necessary to show subsequent memory provide additional evidence that look-defined measures of infant VPC performance reflect variation in information processing. Studies using other procedures have revealed decreases in look duration and increased attentional shifting with increased age (Axia et al., 1999; Colombo & Mitchell, 1990; Courage et al., 2006; Ruff, 1975). Individual differences in these measures have been also reported within specific age groups of infants (Colombo et al., 1991; Freeseman et al., 1993; Jankowski & Rose, 1997). These and other findings are taken as evidence that look-defined measures in the VPC procedure reflect variation in information processing speed.

The third set of observations is that higher novelty preference scores in the VPC procedure are positively associated with longterm outcomes such as IQ (Fagan, 1984; McCall, 1994) and language (Thompson et al., 1991) during later childhood and negatively associated with risk factors such as such as malnutrition (Carter et al., 2010; Nelson et al., 1997), prematurity (Guzzetta et al., 2006; Rose, 1980; Rose et al., 2001, 2004), exposure to toxins (Chiriboga et al., 2007; Emory et al., 2003; Gaultney et al., 2005; Jacobson et al., 2002; Struthers & Hansen, 1992), and developmental conditions such as Down syndrome (Miranda & Fantz, 1974; Nygaard et al., 2001).

Taken together, these three sets of findings are consistent with the traditional view that aspects of looking behavior are an expression of human intelligence (Deary, 2012; Sheppard & Vernon, 2008; Stankov, 1983) and that specific aspects of look-defined behavior in the VPC procedure (1) index information processing and (2) optimally support learning. However, these conclusions are based almost exclusively on studies of infants in Western contexts such as the United States, Europe, and Canada, raising questions about whether the VPC procedure is appropriate for indexing information processing in non-Western samples. The few studies that have examined infant VPC performance in non-Western samples have used patterns observed in Western samples as a benchmark to interpret performance (e.g., Chhaya et al., 2018; Colombo et al., 2014; Drotar et al., 1997; Kennedy et al., 2008; Prado et al., 2020; Rose, 1994; Siegel et al., 2011). Studies have typically assumed that the pattern observed in Western infants reflects “good” performance, often with the goal of examining the impact of clinical interventions (e.g., Colombo et al., 2014; Prado et al., 2020; Siegel et al., 2011) or variables such as health status (Chhaya et al., 2018; Drotar et al., 1997; Kennedy et al., 2008; Rose, 1994) on infant VPC performance and information processing. Studies of this kind have been conducted with Inuit infants in Quebec (Fraser et al., 2012) as well as infants residing in Ethiopia (Kennedy et al., 2008), Uganda (Chhaya et al., 2018; Drotar et al., 1997), India (Rose, 1994), Guadeloupe (Dallaire et al., 2012), Peru (Colombo et al., 2014), and Nepal (Siegel et al., 2011).

A recent eye tracking study of infant attention in rural Malawi by Pyykkö et al. (2019) raises questions about whether it is appropriate to use Western infants as a benchmark or assume that infants in low-resource environments will display poorer performance than infants in higher resourced environments. Pyykkö et al. examined visual search, anticipatory shifting, and attentional biases for faces in infants residing in rural Malawi. They observed age-related changes in attention, but no impact of variables such as gestational age, nutrition status, or maternal measures such as socioeconomic status, psychosocial well-being, and care practices. Although they did not administer a VPC procedure to these infants, the fact that they failed to observe an impact of risk factors on other visual attentional measures in a low-resource environment suggests that it may be problematic to assume that infants in low-resource settings will always display poorer performance.

Our primary objective was to systematically examine VPC performance in infants residing in two culturally distinct contexts: a sample of infants in rural Malawi who were originally recruited as part of a nutritional intervention study (Prado et al., 2020), and sample of infants residing in the US who were originally recruited as part of a pre-registered follow-up study to the research conducted in rural Malawi (https://osf.io/sj67p). The two samples of

infants differed in so many ways that it is difficult to enumerate a comprehensive list of differences—a challenge made more difficult because few studies have examined the lived experiences of infants in rural Malawi. The infants from Malawi lived in the Mangochi District of Southern Malawi, a region predominantly inhabited by the Yao and Chewa tribes. Rural Malawi is considered a low-resource context and there is a high prevalence of poverty. For example, many houses use earth or sand as flooring material and obtain drinking water from community wells rather than piped water in their homes (National Statistical Office, Malawi, 2017). The literacy rate among mothers is lower than what is typically observed in high-income countries (e.g., approximately half of the mothers in our study were able to read, Prado et al., 2020). Family members who work outside the home are typically employed in fishing or farming. Most rural households have either no access or limited access to media such as newspapers, radio, television, smart phones, and internet and families generally spend much of their time outside (National Statistical Office, Malawi, 2017). In contrast, the US sample was recruited from urban and suburban racially and socioeconomically diverse communities in the greater Sacramento Valley (US Census Bureau, 2021), although the families who participated in this study were well educated and middle class. Families with infants in this area generally live in single family homes, own cars, have running water, and are exposed to technology. There are a range of religious and cultural practices that may influence infants' daily lives, but in general the infants came from two-parent middle-class American families.

These profound differences in lived experiences raise questions about whether infants in rural Malawi will display expected associations between age and measures of attention and learning. The VPC task is an ideal procedure for examining these questions because there is robust evidence that look-defined measures in this task (1) reflect information processing (e.g., older infants display higher shift rates, higher novelty preference scores, and shorter peak look durations than younger infants) and (2) display specific patterns of associations with one another (e.g., shorter peak look durations and higher shift rates are typically associated with higher novelty preference scores)—but it's unclear whether these patterns should be evident for infants in rural Malawi. The fact that robust claims have been made infant VPC performance indexing fundamental aspects of human cognition such as intelligence and information processing speed (for review, see Rose et al., 2004) highlights the need for systematic studies of infant VPC performance in understudied contexts.

Several challenges arise when conducting infant VPC research in non-Western contexts such as rural Malawi that include stimulus selection, variation in procedure and apparatus, and operationalization of measures. For example, one important consideration is whether to use stimuli identical to those used in studies of infants in Western samples or to use culturally appropriate stimuli. Presenting culturally disparate samples of infants with identical stimuli introduces potential familiarity confounds because the stimuli are likely more familiar to one group than to the other. This is an important consideration for VPC studies in which researchers infer memory from infants' visual preference for novel versus familiar stimuli. Presenting different stimuli for infants in disparate cultural contexts; however, introduces the possibility that physical differences between stimulus sets will contribute to differences in results. In the present study, we used several trials with different stimuli (thus minimizing any physical differences between any individual stimulus) and carefully matched the image

pairings on each trial; we therefore chose to use different, culturally appropriate stimuli in our two samples.

A second challenge is that infants from different cultural contexts likely have different levels of familiarity with aspects of the testing context. That is, even if the exact same procedure and apparatus are used in culturally distinct samples, infants in one sample may be more familiar with aspects of the methods (i.e., looking at a computer screen) than the infants in another sample. Using different apparatuses and equipment in the two contexts; however, has the potential to introduce different kinds of noise and measurement error in the two samples. In our study, we opted to use identical procedures and apparatuses in our two samples to minimize the noise and measurement error from using different procedures and apparatuses, even though the infants in our different samples may be differentially familiar with the setting.

A third challenge is how eye tracking measures of attention should be operationalized. There are likely differences between infants in different cultural contexts that shape how specific behaviors should be measured and defined. Indeed, there is growing evidence that infants reared in distinct cultural contexts (Forssman et al., 2017; Geangu et al., 2016) or bilingual language environments in the same cultural context (Arredondo et al., 2022; D'Souza et al., 2020; Singh et al., 2015) display different patterns of looking behavior in attention and learning tasks. Eye trackers provide a valuable tool for automating the operationalization of attention measures in culturally disparate samples of infants, but it is unclear whether the algorithms these systems employ are appropriate for infants and adults who reside—or were tested—in different contexts (see Birawo & Kasprowski, 2022, for a review of different eye tracking algorithms). In the present study, we chose to administer an eye tracking VPC procedure to automate stimulus presentation and mitigate experimenter bias.

The use of eye tracking in the present study presented an additional challenge for measure operationalization. Methodological gaps in the existing VPC literature make it difficult to critically examine claims about infant VPC performance using eye tracking methods. In the classic VPC literature, measures of attention are based on human observers' judgments of infants' *looking*. Peak look durations and shift rates derived from looks, therefore, likely reflect cognitive processes that differ from those that researchers typically interpret from eye tracking measures of fixations and saccades in infants (Papageorgiou et al., 2014). This makes it unclear whether measures that have been traditionally calculated from individual looks in the VPC procedure—such as shift rate and peak look duration—can be derived from eye trackers. Our second objective was to develop a computational approach that retains the advantages of modern eye tracking while simultaneously indexing the information processing measures that are conventionally calculated by human observers. Bridging this methodological gap is an important first step towards systematically examining infant VPC performance in the US and rural Malawi.

To address our second objective, we calculated peak look duration, shift rate, and novelty preference scores by computationally filtered the eye tracking data into individual looks. Because it is not immediately obvious how fixations should be combined into individual looks — or whether the same criteria will be appropriate for infants with different lived

experiences — we created looks from our fixation data in different ways by systematically manipulating the minimum fixation duration and maximum gap between fixations required for a bout of attention to be classified as a look. Our third objective was to examine the appropriateness of different operational definitions of a look for our two samples of infants. To address our third objective, we examined whether different operational definitions yielded previously reported associations between (1) VPC measures and age (e.g., higher shift rates, more robust novelty preference scores, and decreased peak look durations in older infants compared to younger infants) and (2) measures of attention and memory performance (e.g., shorter peak look durations and higher shift rates predicting higher novelty preference scores) for infants in the US and rural Malawi.

We are unaware of any studies that have (1) systematically examined information processing measures such as peak look duration and shift rate in non-Western samples, (2) explored the feasibility of calculating look-defined information processing measures from eye tracking data, or (3) evaluated the appropriateness of different operational definitions of behavior for infants from understudied contexts. Therefore, it is unknown whether and how looking patterns are related to information processing and learning in under-represented samples and whether different operational definitions are warranted for culturally distinct groups of infants.

2 | METHOD

2.1 | Participants

We analyzed data from two samples of infants aged 6–9 months of age. The Malawi sample included children who were enrolled in the Mazira Project, a randomized controlled trial carried out in the rural Lungwena and Malindi areas of Mangochi District from February 2018 to January 2019 (Prado et al., 2020). The data analyzed here were collected before the intervention period (i.e., at baseline) and the analyses reported here are original and have not been reported elsewhere. Children aged 6–9 months residing within the catchment areas of the Lungwena Health Center and the St Martin's Rural Hospital in Malindi were recruited into the study through direct recruitment from community health worker listings of age-eligible children and community outreach including village meetings and community football tournaments. Of the 660 infants initially recruited and enrolled in the Mazira Project, 270 infants participated in the version of the VPC task used here during the baseline eye-tracking assessment (the source of the data for the present analyses) before being randomized into intervention groups. We excluded from our analyses the data from 42 infants because they failed to accumulate at least 1-s of looking to the familiarization or test arrays across all operational definitions of a look. Thus, our final sample from Malawi included 228 infants (M_{age} =218.72 days, SD = 38.81, n = 118 girls, n = 110 boys).

The US sample was drawn from a pool of infants who lived within a 30-mile radius of the University of California, Davis. Infants were recruited to participate in this replication and were tested between February 2019 and December 2019. These families predominantly lived in the urban and rural communities of the greater Sacramento Valley. Parents were recruited through mailings about our research in general, and those who expressed interest were notified when their child reached the age range eligible for the current study. Power analysis

revealed that a sample of 36 would allow for sufficient power (80%) to detect a moderate effect size (d=0.50) for one-sample t-tests compared to chance or a paired comparison. Our target sample size was 50 infants between 6 and 9 months of age (this age range corresponds to the age at baseline assessment for the Malawi sample) so that we could implement more complex mixed-effect modeling approaches. The sample size rationale is included on the Open Science Framework as a part of a pre-registration that was created for the original project (https://osf.io/sj67p). To achieve this sample, fifty-three infants were tested, and our final sample consisted of 48 infants (M = 224.59 days, SD = 36.57, n = 24 girls, n = 24 boys) who provided usable data in the eye-tracking assessment; we excluded five infants because they failed to accumulate at least 1-s of looking to the familiarization or test arrays across all operational definitions of a look. Table 1 displays the demographics for both samples. The distribution of participant ages across both samples for both male and female infants is displayed in Figure 1.

2.2 | Apparatus

All infants were tested using a Tobii Pro X2-60 eye-tracker with external processing unit. Sessions were run using a Dell laptop (Dell Latitude 5480 or a Dell Latitude 7280 in Malawi, and a Dell Precision 17 7000 (7710) in the US), an HP EliteDisplay E222 21.5″ monitor (1920 × 1080 resolution) mounted on an adjustable arm, and a webcam attached to the top center of the monitor. The eye-tracker recorded the x and y coordinates of the focal point of the infant's gaze at a rate of 60 Hz (each data point recorded by the eye-tracker corresponded to approximately 16.67 ms).

In Malawi, each eye-tracking system was in a booth in the study center, created by four black curtains that, when closed, blocked out distractions and only the monitor was visible to the mother and child (see Figure 2a). The infant was either placed in a carrier worn by the mother or on the mother's lap, facing the monitor which was positioned approximately 60 cm from the infant's face. The eye-tracking staff requested that the mother look to the side, away from the monitor, to avoid unintentionally directing the child's gaze. The eye-tracking staff monitored the mother and child during the session on the laptop screen via the webcam and reminded the mother to turn away if she started watching the screen. In the US, the eye-tracking system was in a sound-attenuated room with minimal visual distractions. Infants were seated in a highchair or in their parent's lap. Parents were provided felt-covered glasses to prevent them from watching the stimuli and potentially biasing their infants' responses.

2.3 | Stimuli

The experimental stimuli consisted of two sets of 12 color photographs of faces (see Figure 2b). We selected separate face sets to approximate the kinds of faces infants in each context were likely to encounter in their daily lives. The set used in Malawi included African faces from Strohminger et al. (2016) (see Prado et al., 2020 for additional details), and the set used in the US included White and racially ambiguous faces selected from the Child Affective Facial Expression (CAFE) database (LoBue & Thrasher, 2014) and the Face Research Lab London Set (DeBruine & Jones, 2017). We chose to use face stimuli to maximize interest and engagement in the task and to mitigate potential regional differences in familiarity with

non-face stimuli such as objects or abstract shapes. Although using different sets of faces in the two samples may mean that we introduce a confound in terms of differences in the physical features of the stimuli, using faces that were racially unfamiliar to either set of infants—or using non-face stimuli such as objects or abstract geometric patterns—would have introduced a confound, namely, that the stimuli were more familiar to one group than the other. US infants were shown racially heterogeneous faces to reflect the diversity of the local community whereas infants in rural Malawi were shown African faces to reflect the racial homogeneity of rural Malawi. Across the two stimulus sets, pairs were matched on age, perceived gender, and facial expression of the faces (e.g., both versions of face set 1 involved an adult female face paired with an adult male face with neutral expressions). Each face image was approximately 8.97° by 12.72° (9.41 cm by 13.38 cm) at a viewing distance of 60 cm, and the pairs were presented on a gray background (RGB: 136, 136, 136). Each stimulus array was accompanied by classical music and was immediately preceded by a fixation cross that flashed at a rate of 0.65 Hz and alternated with images of colorful toys (Figure 3).

2.4 | Procedure

All protocols were reviewed and approved by the Institutional Review Board (IRB) at the University of California, Davis, and the protocol for the Malawi sample was also reviewed and approved by the Research Ethics Committee at the Research Ethics Committee of the University of Malawi College of Medicine. The same procedure was used in both locations. The session began with the Tobii Lab 5-point calibration procedure; a looming shape was presented in five different locations (in the center and points near each of the four corners). Calibration quality was verified by visually inspecting vertical and horizontal accuracy information that is presented as part of the Tobii validation procedure. Following calibration, infants were presented with four visual paired comparison recognition trials. Between VPC trials, infants were presented with trials from an unrelated attentional cueing task (which will be reported elsewhere), that involved presentation of a cartoon smiley face and images of common household objects. Alternating between the two tasks helped to maintain infants' interest throughout the entire session, and introduced a clear separation between VPC trials, minimizing any carry-over from exposure from one set of images to the next.

Each of the four VPC trials involved the following sequence: (1) a 20-s familiarization phase, in which two identical images were presented side-by-side, (2) a central fixation cross that remained visible until the infant looked, (3) a 10-s test array, in which the familiarization face was paired with a novel face, (4) a central fixation cross presented until the infant looked, and finally, and (5) a second 10-s test array in which the same familiar and novel faces were presented (Figure 3), but left-right position was reversed (see Figure 2b). An experimenter monitored the infant's gaze and pressed a key to initiate the presentation of each stimulus array when the infant's gaze was judged to be directed toward the central fixation stimulus.

2.5 | Data processing

We used Tobii Studio (Tobii, Stockholm, Sweden), SAS version 9.4 (North Carolina, United States), and R version 4.1.1 (Core Team and Others 2013) to process the data. Individual

fixations were identified using the Tobii I-VT Fixation Filter in Tobii Studio. We used the default parameters of (1) fixations as periods of stable gaze position for at least 60 ms for gaze to be defined as a fixation, (2) saccades differentiation from fixations using a threshold of 30 degrees per second and angular velocity across 20 ms time windows, and (3) interpolation for gaps in the eye-tracking data stream of less than 75 ms.

We created separate areas of interest (AOI) for each half of the screen, starting at the edge of the central fixation and ending at the edges of the screen (i.e., one AOI for the left-half of the screen and one AOI for the right-half of the screen). This allowed us to parse the fixation data into individual looks to the left and right side of the screen and mitigate potential variation in calibration accuracy. We defined an individual look as successive series of fixations summing to a minimum duration threshold to one AOI that were interrupted by no more than a specified duration of time. This approach is not unlike how eye tracking algorithms parse data into fixations using a specific velocity or dispersion thresholds across a sliding window—except that our goal was to approximate the temporal scale that human observers could reasonably obtain using behavioral coding. We then systematically manipulated how individual looks were calculated using different look duration thresholds (similar to minimum look criteria in the literature) and interruption duration thresholds (similar to maximum look-away criteria in the literature). All the eye tracking measures analyzed in the present study were calculated based on this method of computationally filtering the eye tracking data into looks. We generated eight different datasets by using two look duration thresholds (500 or 1000 ms) and four interruption duration thresholds (350, 400, 450, and 500 ms). These values were chosen to reflect the range of parameters that researchers might typically employ in behavioral coding studies. Shorter interruption and look duration thresholds provide a closer approximation to the temporal dynamics of fixations whereas longer thresholds reflect the timing that human observers might achieve through button presses in experimenter-controlled procedures.

To be included in our analyses, infants must have accumulated at least 1 s of looking during familiarization and 1 s of looking across the two test arrays presented during the test phase across all datasets (i.e., trials were only included when infants met these criteria for all eight operational definitions). To minimize the effect of extreme scores on our analyses, trials in which infants displayed extreme total look durations during familiarization and test—as defined by total look durations in the upper 97.5 percent and lower 2.5 percent of the distribution—were excluded from the analyses conducted for each operational definition.

For each look and interruption duration threshold dataset, we calculated peak look durations, shift rates, and novelty preference scores for each participant on each trial. *Peak look duration* was calculated by identifying the longest individual look for each infant on each trial during the familiarization phase. *Shift rate* was calculated by first counting each time a look to one AOI was preceded by a look to the other AOI; two successive looks to the same AOI were not counted as shifts. The total number of shifts during the familiarization and test phase were then divided by the total look duration and shift rates were then collapsed across both phases. *Novelty preference score*, a measure of memory during the memory phase of the task, was calculated by summing the total look duration to the novel stimulus across both test arrays presented during the test phase of each trial and dividing by this

number by the total look duration to the novel stimulus and familiar stimulus during this same period. The trial-level scores for all possible combinations of look and interruption duration thresholds were then analyzed using linear mixed-effect models, such that for each of our eight operational definitions, each infant contributed a single score for each trial and each measure for our statistical analyses.

2.6 | Statistical approach

To address our objectives, we fit a series of linear mixed-effect models (LMMs) on the trial-level scores for each outcome measure (i.e., peak look duration, shift rate, and novelty preference score) and each of our eight operational definitions of a look. All models were fit in R using *lme4* (Bates et al., 2014) and *p*-values were calculated using *lmerTest* (Kuznetsova et al., 2017). We fit separate LMMs on each of our outcome measures (peak look, shift rate, and novelty preference) including fixed effects of *child age* (continuous: in days), infant *sex* (categorical: male infants, female), *trial* (continuous: 1–4), *sample* (categorical: Malawi or US), and a random intercept for participant. Calibration accuracy was included as a covariate in all models to control for participant-level differences in eye-tracking data quality and age was scaled (mean = 0) to aid in interpretation. For models examining novelty preference, we included fixed effects for our information processing measures (peak look and shift rate) as well as interactions between each information processing measure and sample. Estimated marginal means were calculated using the *ggeffects* package (Lüdecke, 2018) and visualized to provide estimates for the effect of our variables in the model (e.g., age) after adjusting for the other parameters in the model (e.g., calibration accuracy). All statistical analyses were conducted in R version 4.1.1 and the analysis data, scripts, and full model results including fixed effect estimates for covariates can be found on OSF (https://osf.io/r5gnw/).

3 | RESULTS

3.1 | Descriptive statistics

The mean age of infants (rounded to the nearest day) in the Malawi sample was 218.72 days (SD = 38.81 days) and the mean age of infants in the US sample was 224.57 days (SD = 36.57 days). Differences in the distribution of ages across the two samples was tested using two-sample Kolmogorov–Smirnov test. This analysis revealed that the distribution of ages differed between the two samples, $D = 0.17$, $p < 0.001$, likely reflecting variation in sample size and sampling approach in each context. Analysis of the number of trials that infants completed revealed a significant difference in the mean number of trials infants completed in rural Malawi (M = 2.91, SD = 1.09) and the US (M = 3.58, SD = 0.77), $t(274) = -4.09$, $p < 0.001$. Validation accuracy was analyzed to evaluate differences between samples in the eye tracking data quality, revealing no significant differences in the degree of deviation for infants in the US (M = 1.22, SD = 1.77) and rural Malawi (M = 1.52, SD = 1.54) $t(262) = 1.18$, $p = 0.24$.

3.2 | Examining our variables of interest for different look definitions

Our objectives were to examine infant VPC performance in two culturally distinct samples of infants by parsing eye tracking data into individual looks and manipulating our

operational definition of a look. To address these objectives, we fit a single model for each of our outcome measures: a peak look LMM, a shift rate LMM, and a novelty preference score LMM. Each of these models was fit for each of our eight operational definitions. The models were specified in R as follows:

P*eaklookduration* ~ *infantsex* + *trial* + *childage* + *sample* + *calibrationaccuracy* + (1|participant)

S*hiftrate* ~ *infantsex* + *trial* + *childage* + *sample* + *calibrationaccuracy* + (1|participant)

N*oveltypreference* ~ *infantsex* + *trial* + *childage* + *shiftrate* + *peaklookduration* + *sample* + *sample* * *peaklookduration* + *sample* * *shiftrate* + *calibrationaccuracy* + (1|participant)

The LMMs for peak look duration revealed intercept estimates that ranged from 3.40 to 3.53 s, suggesting that similar peak look durations were observed across variations in our operational definition of a look. These coefficients reflect model estimates for peak look duration on each trial after controlling for the effects of age, trial, sample, infant sex, calibration accuracy, and random participant-level variation and suggest a high level of engagement during the familiarization phase of the procedure (note that these estimates were significantly different from 0 as indicated by the intercept estimates from all models, but we did not statistically compare the model results for each individual operational definition). Estimated marginal means were then calculated from each peak look duration LMM to examine how operationally defining a look impacted associations between these measures and peak look duration. As can be seen in Figure 4, consistent results were observed for the effects of sample, age, and trial on infants' peak look duration.

We obtained a significant fixed effect of age on peak look duration for seven out of eight operational definitions of a look; the last effect was marginal ($p = 0.053$). Thus, the operational definition of a look did not profoundly impact the effect of age on peak look duration; across all datasets, older infants displayed shorter peak look durations than younger infants in both samples. As can be seen in Figure 4, no operational definitions of a look yielded a significant fixed effect for trial or sample, suggesting that infants in the US and rural Malawi displayed similar peak look durations and that infants as a group did not display difference peak look durations across individual trials. The results for all peak look duration models these is displayed in Table 2.

The LMM for shift rate revealed intercept estimates that ranged from 0.19 to 0.43, suggesting variation in shift rate across our different operational definitions after controlling for the effects of other variables in the model. That is, some operational definitions revealed faster shifting between stimuli than other operational definitions. The LMM on shift rate also revealed significant fixed effects of sample and of trial for eight out of eight operational definitions. Across trials, infants in the US displayed faster shift rates than infants in rural Malawi, and for both samples shift rate decreased across increasing trials (Figure 5). Overall, age was not significantly related to shift rate parameter estimates for any operational definitions of a look (Figure 5). The findings from our models examining shift rate are reported in Table 3.

Next, we fit a novelty preference score LMM to examine the associations between age, sample, trial, peak look duration, shift rate, and two-way interactions between sample and information processing measures (e.g., shift rate and peak look duration) on infants' novelty preference scores after controlling for infant sex, calibration accuracy, and random participant-level variation. For the novelty preference LMM, we centered infants' novelty preference score on each trial by subtracting chance (0.50) to determine whether infants showed a statistically significant preference for the previously familiarized stimulus. Centering our outcome measure allowed us to examine whether infants displayed a novelty preference score that was significantly above or below chance ($p < 0.05$) and the intercept estimate indicates the magnitude of the difference. That is, a statistically significant intercept coefficient of 0.05 would indicate that–as a group–infants showed a significant preference for the novel stimulus (0.55) after controlling for variation in the variables included in the model. Intercept estimates for all operational definitions of a look were significantly different from 0, ranging from 0.09 to 0.26 across look definitions. As can be seen in Figure 6, different operational definition of a look impacted the magnitude of infants' novelty preference scores, but all operational definitions revealed novelty preference scores that were significantly above-chance.

The LMM for novelty preference also revealed a significant fixed effect of age across all operational definitions of a look and all parameter estimates were in the same direction. As can be seen in Figure 6, younger infants displayed lower novelty preference scores than older infants across all look definitions. Additionally, the novelty preference LMM also revealed a significant fixed effect of shift rate on novelty preference score estimates for four out of eight operational definitions (see Figure 7). Significant fixed effects of shift rate on novelty preference scores were observed for all look duration thresholds of 1000 ms–independent of the interruption duration threshold–but were never observed for operational definitions that required a 500 ms look duration threshold. Moreover, as can be seen in Figure 7, all fixed effects of shift rate that were statistically significant were negatively associated with infants' novelty preference scores–the opposite pattern that has previously been reported in the literature (Rose et al., 2001,2003). That is, higher novelty preferences were associated with lower shift rates, or fewer shifts per second.

No significant fixed effects were observed for peak look duration, sample, or trial. That is, infants in the US and rural Malawi displayed similar novelty preference scores and novelty preference scores did not vary as a function of trial or for shorter versus longer peak look durations. The effects of sample, age, and trial are reported in Table 4. The novelty preference score LMM also did not reveal significant interactions between shift rate and sample or peak look duration and sample, indicating that different look definitions did not yield different patterns of association between measures of information processing and memory for infants in the US and rural Malawi (Table 5).

4 | DISCUSSION

Our objectives were to examine infant VPC performance in two culturally distinct samples, develop a novel approach for parsing the eye tracking data into look-defined measures of information processing, and examine the appropriateness of different operational definitions

of a look for infants in our two samples. We observed that infants in both samples showed memory for the familiar stimulus as indicated by their novelty preference scores, consistent with previous studies conducted in Western (Rose et al., 2004) and non-Western contexts (Chhaya et al., 2018; Dallaire et al., 2012; Drotar et al., 1997; Fraser et al., 2012; Kennedy et al., 2008; Rose, 1994; Siegel et al., 2011). In general, therefore, this task seemed to elicit a preference for the novel stimulus in infants who had very different everyday experiences as has been observed in other studies (Chhaya et al., 2018). In addition, higher novelty preference scores and shorter peak look durations were observed for older infants compared to younger infants, consistent with previous studies in the VPC procedure (Colombo et al., 1988; Courage et al., 2006; Fagan, 1970,1972,1974; Rose et al., 2004) and other attention paradigms (Axia et al., 1999; Colombo & Mitchell, 1990; Ruff, 1975) for infants in the US. Importantly, we observed no significant differences in novelty preference scores for infants in the US and rural Malawi, confirming that these two samples showed similar memory performance in the procedure.

We observed no expected associations between measures of information processing (i.e., peak look duration and shift rate) and memory performance. In fact, some operational definitions of a look yielded the opposite pattern that has previously been reported for shift rate and novelty preference scores in Western samples for infants in the US and rural Malawi (Rose et al., 2001, 2003). Significant associations between shift rate and memory performance were only observed for operational definitions of a look with a minimum look duration threshold of 1000 ms. It is important to point out that shift rates were lower when imposing longer minimum look duration thresholds (see Figure 5). That is, by requiring that a single look be at least 1000 ms in duration, many shorter sequences of fixations were excluded. As a result, shift rates were shorter and significantly predicted infants' memory performance in the opposite direction we would expect based on previous studies. Thus, although a longer minimum look duration excludes short glances that perhaps do not allow for full processing of the stimuli and ensures that only significant bouts of looking are included, such definitions also exclude brief looks back and forth between the two stimuli. It is important to point out that none of the definitions that we imposed are "correct", rather they each reflect different assumptions about a look (i.e., how long is enough for significant processing, how long can an interruption be within a single bout). What our results demonstrate is that some measures in this task are robust to variations in look definition whereas others vary considerably across different operational definitions.

The fact that we failed to replicate previously observed relations between shift rate or peak look duration and memory performance challenges the robustness of these associations. The many published findings revealing higher novelty preference scores for infants that display faster shift rates and shorter peak look durations may reflect "the file drawer effect" (Rosenthal, 1979)–a term that has been coined to describe publication bias for positive results over negative results. That is, because of publication biases for positive results, it is unknown whether there exist a body of unpublished findings that have also failed to find associations between shift rate or peak look and novelty preference. This makes it difficult to interpret the lack of associations between information processing measures and memory performance observed in the present study. Another possibility is that associations between

these measures were examined in previous studies, but that they were not reported because null effects were observed.

Of course, it is also possible that our findings reflect qualitative differences between look-defined measures from eye tracking data and behavioral coding. That is, even though our definitions of looks were empirically and theoretically derived, our look definitions may not map onto the same attentional processes that are typically captured by human observers. Although this possibility cannot be completely ruled out, the fact that we observed several patterns that replicate the literature suggests that our measures overlap with those that are typically derived from behavioral coding. Across samples, we observed robust novelty preference scores, as has been found in other studies with similar stimuli and temporal parameters (Jones et al., 2011; Rose & Feldman, 1987). In addition, consistent with other findings (Jacobs, 2000; Rose, 1981; Rose & Feldman, 1987), we observed an increase of between 6 and 9 months in novelty preference score. We also found that the duration of the peak look decreased across this age range, as has been observed by others (Colombo & Mitchell, 1990; Courage et al., 2006). Thus, it seems unlikely that our results reflect qualitative differences between human observer and eye tracker recorded data. Nevertheless, if qualitative differences between human- and eye-tracker-defined measures do exist, it underscores the fragility of these effects and suggests that observed associations between measures of attention and learning rely on very narrow and specific ways of characterizing the data.

To our knowledge, this is the first study to parse eye tracking measures into individual looks in the infant VPC procedure. Studies examining visual measures such as peak look duration and shift rate have yielded important insights about infant attention during learning, but deriving these measures depends on human observers. One advantage of parsing eye tracking data into individual looks is that it provides researchers with a tool for examining look-defined measures of information processing without behavioral coding. This method holds promise for bridging the methodological gap between classic behavioral coding studies and eye tracking research. Moreover, human observers may have implicit biases that shape their expectations about infants' performance and impact observed findings–especially in cross-cultural studies. Methods for filtering eye tracking data into look-defined measures may provide a tool for mitigating these kinds of experimenter biases in studies of information processing in the VPC procedure. This is an important issue for the VPC procedure and other paradigms that are used "off the shelf" as clinical assessment tools as well as research on culturally distinct samples of infants. Visual measures such as peak look duration, shift rate, and novelty preference scores have played a critically important role in motivating theories of infant cognition, yet there is a lack of data on how culture and context affect infant visual behavior. The lack of methods for parsing look-defined from eye tracking data presents an additional challenge for conducting studies with culturally distinct samples.

The current study has some limitations. First, our two samples were shown different image sets, raising the possibility that differences in our experimental stimuli may have contributed to our pattern of results. However, both samples showed novelty preference, indicating that they could form memories for these stimuli and discriminate the novel and familiar stimuli. In addition, overall infants in the two samples showed similar levels of engagement and no

differences were observed in the relation among the variables. This suggests that the two sets of stimuli induced the same behaviors in the two groups of infants. Nevertheless, it is possible that if we had used the same stimuli in the two samples that the differences in familiarity may have induced differences in how the two groups of infants processed the stimuli. One way for future studies to address this limitation might be to include both an overlapping set of stimuli that are identical as well as two distinct sets of stimuli that are culturally appropriate.

A second limitation is that the lived experiences of our two samples vary in multiple ways making it impossible to identify precisely why they responded similarly or differently. Differences or similarities between our two samples could reflect cultural practices that influence interactions between infants and other people, economic differences that have consequences both for health and for access to technology, or environmental factors that determine what infants look at each day. It is likely that any differences between our samples reflect a combination of many differences in lived experience. Our goal was not to determine what specific factors contributed to differences in these two groups of infants, but rather what similarities and differences exist in their visual attention. A third limitation is that, even under ideal circumstances, studies of infant attention and learning rely on measures of infant looking behavior that are inherently noisy. Where and how long infants look is stochastic and multiply determined even in highly controlled experimental settings. Studying infants residing in dramatically different cultural contexts likely introduces additional sources of variation.

These limitations reflect challenges that are in many ways endemic to conducting cross-cultural research. One such challenge is that much of the literature on understudied samples focuses on risk factors for development. This is particularly true for participants residing in low-middle income countries (LMICs). The fact that other aspects of infants' lived experiences are often ignored makes it difficult to systematically evaluate the impact of specific experiences on infant development in these samples. This is compounded by the fact that many researchers assume that the lived experiences of infants residing in the US and Europe are typical or optimal, and that their development is universal rather than reflecting specific aspects of their day-to-day experience. As a result, much of the research on the effect of experience on infant cognition in contexts such as rural Malawi has adopted a deficit model. Another challenge arises when researchers make decisions about what experimental stimuli, procedures, and measures to use for cross-cultural studies. Researchers may choose a specific approach because it mitigates potential confounds without recognizing that the same approach introduces other confounds. For example, a researcher might decide to use naturalistic observations instead of screen-based procedures to prevent differences in exposure to technology from confounding their results—but unknowingly introduce experimenter bias because participants are easily identifiable as belonging to a specific group or culture. As these examples illustrate, researchers must navigate complex challenges when conducting cross-cultural research, as was the case for the present study. Tools and methods for indexing information processing from infant looking behavior hold promise for reducing some of the barriers that make studies of this kind challenging.

Studying variation in attention and memory in infants from different cultures is critical for advancing our understanding of infant cognition and the results of the present study indicate that measure operationalization is an important consideration for studies examining basic visual cognitive abilities in infants from different contexts. The real lack of data on how culture and context influence visual behavior raises questions about whether reported results are generalizable and universal, even if they have been replicated in multiple samples, if those samples have all reflected the same narrow slice of the world's population of infants. A deeper understanding of development in general will be gained by conducting principled research in diverse samples of infants.

ACKNOWLEDGMENTS

We thank the staff and faculty at Kamuzu University of Health Sciences as well as the students and staff in the Infant Cognition Laboratory at the University of California, Davis, for their help with data collection. This research and preparation of this manuscript were made possible by a grant awarded to CS from the Bill & Melinda Gates Foundation (OPP1169033) and a grant awarded to LMO from the National Institutes of Health (R01EY030127).

DATA AVAILABILITY STATEMENT

All data and analysis scripts for the results reported here are openly available in the Open Science Foundation at https://osf.io/r5gnw/.

REFERENCES

Arredondo MM, Aslin RN, Zhang M, & Werker JF. (2022). Attentional orienting abilities in bilinguals: Evidence from a large infant sample. Infant Behavior & Development, 66, 101683. [PubMed: 34999429]

Axia G, Bonichini S, & Benini F (1999). Attention and reaction to distress in infancy: A longitudinal study. Developmental Psychology, 35(2), 500–504. 10.1037/0012-1649.35.2.500 [PubMed: 10082020]

Bates D, Mächler M, Bolker B, & Walker S (2014). Fitting linear mixed-effects models using lme4. In arXiv [stat.CO]. arXiv. http://arxiv.org/abs/1406.5823

Birawo B,& Kasprowski P.(2022). Review and evaluation of eye movement event detection algorithms. Sensors (Basel, Switzerland), 22(22), 8810. 10.3390/s22228810 [PubMed: 36433407]

Carter RC, Jacobson JL, Burden MJ, Armony-Sivan R, Dodge NC, Angelilli ML, Lozoff B, & Jacobson SW (2010). Iron deficiency anemia and cognitive function in infancy. Pediatrics, 126(2), e427–e434. 10.1542/peds.2009-2097 [PubMed: 20660551]

Chhaya R, Weiss J, Seffren V, Sikorskii A, Winke PM, Ojuka JC, & Boivin MJ (2018). The feasibility of an automated eye-tracking-modified Fagan test of memory for human faces in younger Ugandan HIV-exposed children. Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence, 24(5), 686–701. 10.1080/09297049.2017.1329412 [PubMed: 28532283]

Chiriboga CA, Kuhn L, & Wasserman GA (2007). Prenatal cocaine exposures and dose-related cocaine effects on infant tone and behavior. Neurotoxicology and Teratology, 29(3), 323–330. 10.1016/j.ntt.2006.12.002 [PubMed: 17234383]

Colombo J, & Mitchell DW (1990). Individual differences in early visual attention: Fixation time and information processing. Individual differences in infancy: Reliability, stability, prediction, 193–227.

Colombo J, Mitchell DW, Coldren JT, & Freeseman LJ (1991). Individual differences in infant visual attention: Are short lookers faster processors or feature processors? Child Development, 62(6), 1247–1257. 10.2307/1130804 [PubMed: 1786713]

Colombo J, Mitchell DW, & Horowitz FD (1988). Infant visual attention in the paired-comparison paradigm: Test-retest and attention-performance relations. Child Development, 59(5), 1198–1210. 10.2307/1130483 [PubMed: 3168636]

Colombo J, Zavaleta N, Kannass KN, Lazarte F, Albornoz C, Kapa LL, & Caulfield LE (2014). Zinc supplementation sustained normative neurodevelopment in a randomized, controlled trial of Peruvian infants aged 6-18 months. The Journal of Nutrition, 144(8), 1298–1305. 10.3945/jn.113.189365 [PubMed: 24850625]

Courage ML, & Howe ML (2001). Long-term retention in 3.5-month-olds: Familiarization time and individual differences in attentional style. Journal of Experimental Child Psychology, 79(3), 271–293. 10.1006/jecp.2000.2606 [PubMed: 11394930]

Courage ML, Reynolds GD, & Richards JE (2006). Infants' attention to patterned stimuli: Developmental change from 3 to 12 months of age. Child Development, 77(3), 680–695. 10.1111/j.1467-8624.2006.00897.x [PubMed: 16686795]

Dallaire R, Muckle G, Rouget F, Kadhel P, Bataille H, Guldner L, Seurin S, Chajès V, Monfort C, Boucher O, Pierre Thomé J, Jacobson SW, Multigner L, & Cordier S (2012). Cognitive, visual, and motor development of 7-month-old Guadeloupean infants exposed to chlordecone. Environmental Research, 118, 79–85. 10.1016/j.envres.2012.07.006 [PubMed: 22910562]

Deary IJ (2012). Intelligence. Annual Review of Psychology, 63(1), 453–482. 10.1146/annurev-psych-120710-100353

DeBruine L, & Jones B (2017). Face Research Lab London Set (Version 3).

Drotar D, Olness K, Wiznitzer M, Guay L, Marum L, Svilar G, Hom D, Fagan JF, Ndugwa C, & Kiziri-Mayengo R (1997). Neurodevelopmental outcomes of Ugandan infants with human immunodeficiency virus type 1 infection. Pediatrics, 100(1), E5. 10.1542/peds.100.1.e5

D'Souza D, Brady D, Haensel JX, & D'Souza H (2020). Is mere exposure enough? The effects of bilingual environments on infant cognitive development. Royal Society Open Science, 7(2), 180191. 10.1098/rsos.180191 [PubMed: 32257297]

Emory E, Ansari Z, Pattillo R, Archibold E, & Chevalier J (2003). Maternal blood lead effects on infant intelligence at age 7 months. American Journal of Obstetrics and Gynecology, 188(4), S26–32. 10.1067/mob.2003.244 [PubMed: 12712133]

Fagan JF (1970). Memory in the infant. Journal of Experimental Child Psychology, 9(2), 217–226. 10.1016/0022-0965(70)90087-1 [PubMed: 5452116]

Fagan JF (1972). Infants' recognition memory for faces. Journal of Experimental Child Psychology, 14(3), 453–476. 10.1016/0022-0965(72)90065-3 [PubMed: 4676459]

Fagan JF (1974). Infant recognition memory: The effects of length of familiarization and type of discrimination task. Child Development, 45(2), 351–356. [PubMed: 4837713]

Fagan JF (1984). The relationship of novelty preferences during infancy to later intelligence and later recognition memory. Intelligence, 8(4), 339–346. 10.1016/0160-2896(84)90016-3

Forssman L, Ashorn P, Ashorn U, Maleta K, Matchado A, Kortekangas E, & Leppänen JM (2017). Eye-tracking-based assessment of cognitive function in low-resource settings. Archives of Disease in Childhood, 102(4), 301–302. 10.1136/archdischild-2016-310525

Fraser SL, Muckle G, Abdous BB, Jacobson JL, & Jacobson SW (2012). Effects of binge drinking on infant growth and development in an Inuit sample. Alcohol, 46(3), 277–283. 10.1016/j.alcohol.2011.09.028 [PubMed: 22444956]

Freeseman LJ, Colombo J, & Coldren JT (1993). Individual differences in infant visual attention: Four-month-olds' discrimination and generalization of global and local stimulus properties. Child Development, 64(4), 1191–1203. 10.2307/1131334 [PubMed: 8404264]

Gaultney JF, Gingras JL, Martin M, & DeBrule D (2005). Prenatal cocaine exposure and infants' preference for novelty and distractibility. The Journal of Genetic Psychology, 166(4), 385–406. 10.3200/GNTP.166.4.385-406 [PubMed: 16463605]

Geangu E, Ichikawa H, Lao J, Kanazawa S, Yamaguchi MK, Caldara R, & Turati C. (2016). Culture shapes 7-month-olds' perceptual strategies in discriminating facial expressions of emotion. Current Biology: CB, 26(14), R663–R664. 10.1016/jxub.2016.05.072 [PubMed: 27458908]

Guzzetta A, Mazzotti S, Tinelli F, Bancale A, Ferretti G, Battini R, Bartalena L, Boldrini A, & Cioni G (2006). Early assessment of visual information processing and neurological outcome in preterm infants. Neuropediatrics, 37(5), 278–285. 10.1055/s-2006-955929 [PubMed: 17236106]

Hunter MA, Ames EW, & Koopman R (1983). Effects of stimulus complexity and familiarization time on infant preferences for novel and familiar stimuli. Developmental Psychology, 19(3), 338–352. 10.1037/0012-1649.19.3.338

Jacobs SE (2000). Visual recognition memory for objects in six- and nine-month-old infants. The Occupational Therapy Journal of Research, 20(4), 271–285. 10.1177/153944920002000404

Jacobson SW, Chiodo LM, Sokol RJ, & Jacobson JL (2002). Validity of maternal report of prenatal alcohol, cocaine, and smoking in relation to neurobehavioral outcome. Pediatrics, 109(5), 815–825. 10.1542/peds.109.5.815 [PubMed: 11986441]

Jankowski JJ, & Rose SA (1997). The distribution of visual attention in infants. Journal of Experimental Child Psychology, 65(2), 127–140. 10.1006/jecp.1996.2363 [PubMed: 9169207]

Jones EJH, Pascalis O, Eacott MJ, & Herbert JS (2011). Visual recognition memory across contexts. Developmental Science, 14(1), 136–147. 10.1111/j.1467-7687.2010.00964.x [PubMed: 21159095]

Kennedy T,Thomas DG, Woltamo T, Abebe Y, Hubbs-Tait L, Sykova V, Stoecker BJ, & Hambidge KM (2008). Growth and visual information processing in infants in Southern Ethiopia. Journal of Applied Developmental Psychology, 29(2), 129–140. 10.1016/j.appdev.2007.12.003 [PubMed: 19684873]

Kuznetsova A, Brockhoff PB, & Christensen RHB (2017). lmerTest package: Tests in linear mixed effects models. Journal of Statistical Software, 82(13), 1–26. https://orbit.dtu.dk/en/publications/id(2e6e11c6-6a44-43df-b809-d96b897d9bf2).html https://doi.org/10.18637/jss.v082.i13

LoBue V,& Thrasher C. (2014). The child affective facial expression (CAFE) set. Databrary, 10, B7301K.

Lüdecke D. (2018). Ggeffects: Tidy data frames of marginal effects from regression models. Journal of Open Source Software, 3(26), 772. 10.21105/joss.00772

McCall RB (1994). What process mediates predictions of childhood IQ from infant habituation and recognition memory? Speculations on the roles of inhibition and rate of information processing. Intelligence, 18(2), 107–125.

Miranda SB, & Fantz RL (1974). Recognition memory in Down's syndrome and normal infants. Child Development, 45(3), 651–660. 10.2307/1127831 [PubMed: 4143821]

National Statistical Office, Malawi. (2017). Malawi demographic and healthy survey.

Nelson KG, Goldenberg RL, Hoffman HJ, & Cliver SP. (1997). Growth and development during the first year in a cohort of low income term-born American children. Acta Obstetricia et Gynecologica Scandinavica. Supplement, 165, 87–92.

Nygaard E, Reichelt KL, & Fagan JF (2001). The relation between the psychological functioning of children with Down syndrome and their urine peptide levels and levels of serum antibodies to food proteins. Down's Syndrome, Research and Practice: The Journal of the Sarah Duffen Centre / University of Portsmouth, 6(3), 139–145. 10.3104/reports.107

Papageorgiou KA, Smith TJ, Wu R, Johnson MH, Kirkham NZ, & Ronald A (2014). Individual differences in infant fixation duration relate to attention and behavioral control in childhood. Psychological Science, 25(7), 1371–1379. 10.1177/0956797614531295 [PubMed: 24815614]

Prado EL, Maleta K, Caswell BL, George M, Oakes LM, DeBolt MC, Bragg MG, Arnold CD, Iannotti LL, Lutter CK, & Stewart CP (2020). Early child development outcomes of a randomized trial providing 1 egg per day to children age 6 to 15 months in Malawi. The Journal of Nutrition, 150(7), 1933–1942. 10.1093/jn/nxaa088 [PubMed: 32286620]

Pyykkö J, Forssman L, Maleta K, Ashorn P, Ashorn U, & LeppÄnen JM (2019). Early development of visual attention in infants in rural Malawi. Developmental Science, 22(5), e12761. 10.1111/desc.12761 [PubMed: 30315673]

Rose SA (1980). Enhancing visual recognition memory in preterm infants. Developmental Psychology, 16(2), 85–92. 10.1037/0012-1649.16.2.85

Rose SA (1981). Developmental changes in infants ' retention of visual stimuli. Child Development, 52(1), 227–233. 10.2307/1129235 [PubMed: 7238147]

Rose SA (1983). Differential rates of visual information processing in full-term and preterm infants. Child Development, 54(5), 1189–1198. 10.2307/1129674 [PubMed: 6354626]

Rose SA (1994). Relation between physical growth and information processing in infants born in India. Child Development, 65(3), 889–902. 10.2307/1131426 [PubMed: 8045175]

Rose SA, & Feldman FJ (1987). Infant visual attention: Stability of individual differences from 6 to 8 months. Developmental Psychology, 23(4), 490–498. 10.1037/0012-1649.23.4.490

Rose SA, Feldman JF, & Jankowski JJ (2001). Attention and recognition memory in the 1st year of life: A longitudinal study of preterm and full-term infants. Developmental Psychology, 37(1), 135–151. 10.1037/0012-1649.37.1.135 [PubMed: 11206428]

Rose SA, Feldman JF, & Jankowski JJ (2004). Infant visual recognition memory. *Developmental Review*: DR, 24(1), 74–100. 10.1016/j.dr.2003.09.004

Rose SA, Feldman JF, & Jankowski JJ (2003). Infant visual recognition memory: independent contributions of speed and attention. Developmental Psychology, 39(3), 538–571.

Rose SA, Gottfried AW, Melloy-Carminar P, Bridger WH, & Mello-Carmina P (1982). Familiarity and novelty preferences in infant recognition memory: Implications for information processing. Developmental Psychology, 18(5), 704–713. 10.1037/0012-1649.18.5704

Rosenthal R. (1979). The file drawer problem and tolerance for null results. Psychological Bulletin, 86(3), 638–641. 10.1037/0033-2909.86.3.638

Ruff HA (1975). The function of shifting fixations in the visual perception of infants. Child Development, 46(4), 857–865. 10.2307/1128390 [PubMed: 1201665]

Sheppard LD, & Vernon PA (2008). Intelligence and speed of information-processing: A review of 50 years of research. Personality and Individual Differences, 44(3), 535–551. 10.1016/j.paid.2007.09.015

Siegel EH, Kordas K, Stoltzfus RJ, Katz J, Khatry SK, LeClerq SC, & Tielsch JM (2011). Inconsistent effects of iron-folic acid and/or zinc supplementation on the cognitive development of infants. Journal of Health, Population, and Nutrition, 29(6), 593–604. [PubMed: 22283033]

Singh L, Fu CSL, Rahman AA, Hameed WB, Sanmugam S, Agarwal P, Jiang B, Chong YS, Meaney MJ, Rifkin-Graboi A, & GUSTO Research Team. (2015). Back to basics: A bilingual advantage in infant visual habituation. Child Development, 86(1), 294–302. 10.1111/cdev.12271 [PubMed: 25074016]

Stankov L. (1983). Attention and intelligence. Journal of Educational Psychology, 75(4), 471–490. 10.1037/0022-0663.75.4.471

Strohminger N, Gray K, Chituc V, Heffner J, Schein C, & Heagins TB (2016). The MR2: A multi-racial, mega-resolution database of facial stimuli. Behavior Research Methods, 48(3), 1197–1204. 10.3758/s13428-015-0641-9 [PubMed: 26311590]

Struthers JM, & Hansen RL (1992). Visual recognition memory in drug-exposed infants. Journal of Developmental and Behavioral Pediatrics: JDBP, 13(2), 108–111. 10.1097/00004703-199204000-00005 [PubMed: 1577955]

Thompson LA, Fagan JF, & Fulker DW (1991). Longitudinal prediction of specific cognitive abilities from infant novelty preference. Child Development, 62(3), 530–538. 10.2307/1131128 [PubMed: 1914623]

Research Highlights

- In both the US and Malawi, 6- to 9-month-old infants showed evidence of memory for faces they had previously viewed during a familiarization period.
- Infant age was associated with peak look duration and memory performance in both contexts.
- Different operational definitions of a look yielded consistent findings for peak look duration and novelty preference scores–but not shift rate.
- Operationalization of look-defined measures is an important consideration for studies of infants in different cultural contexts.

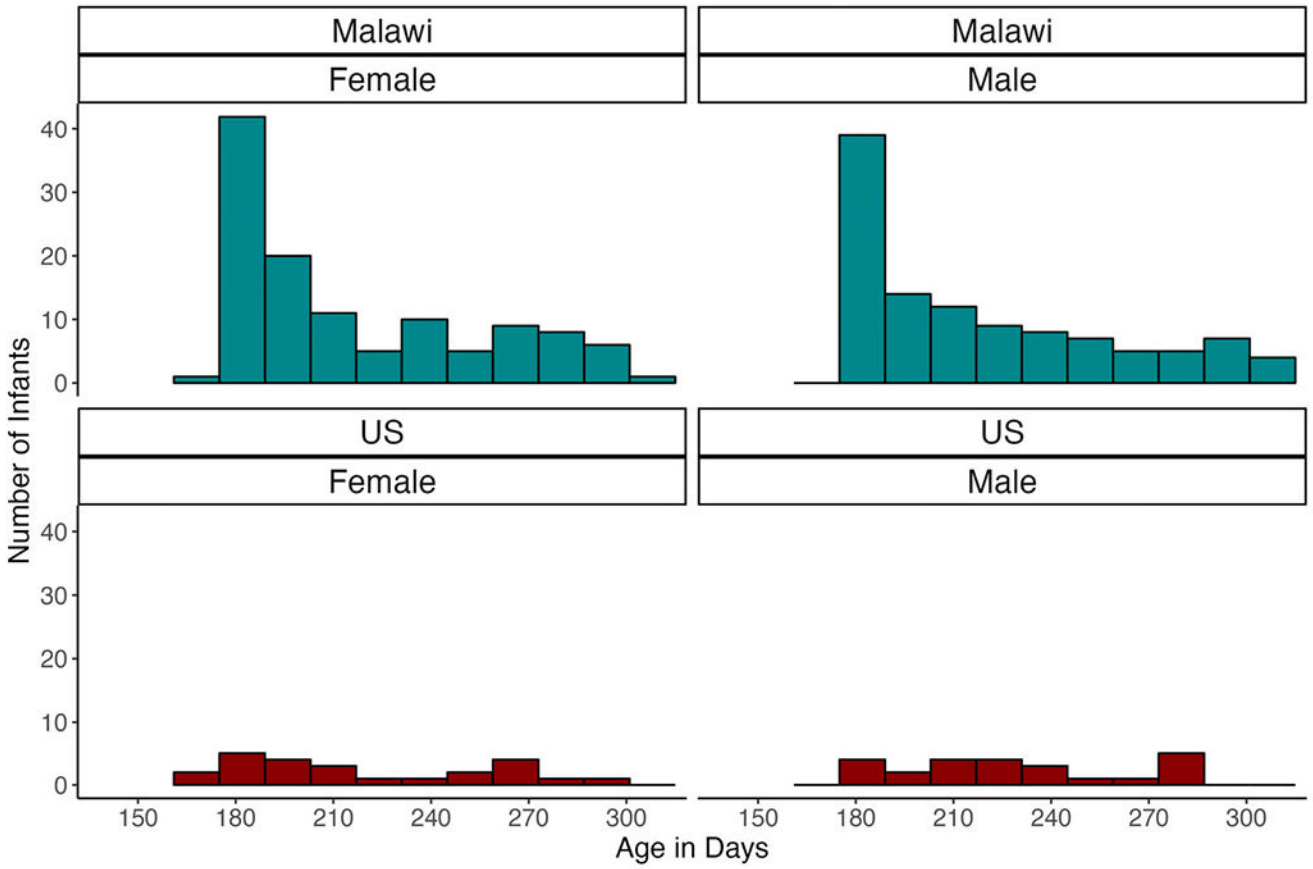

FIGURE 1.
Distribution of ages for female infants (left) and male (right) in Malawi (top) and the US (bottom).

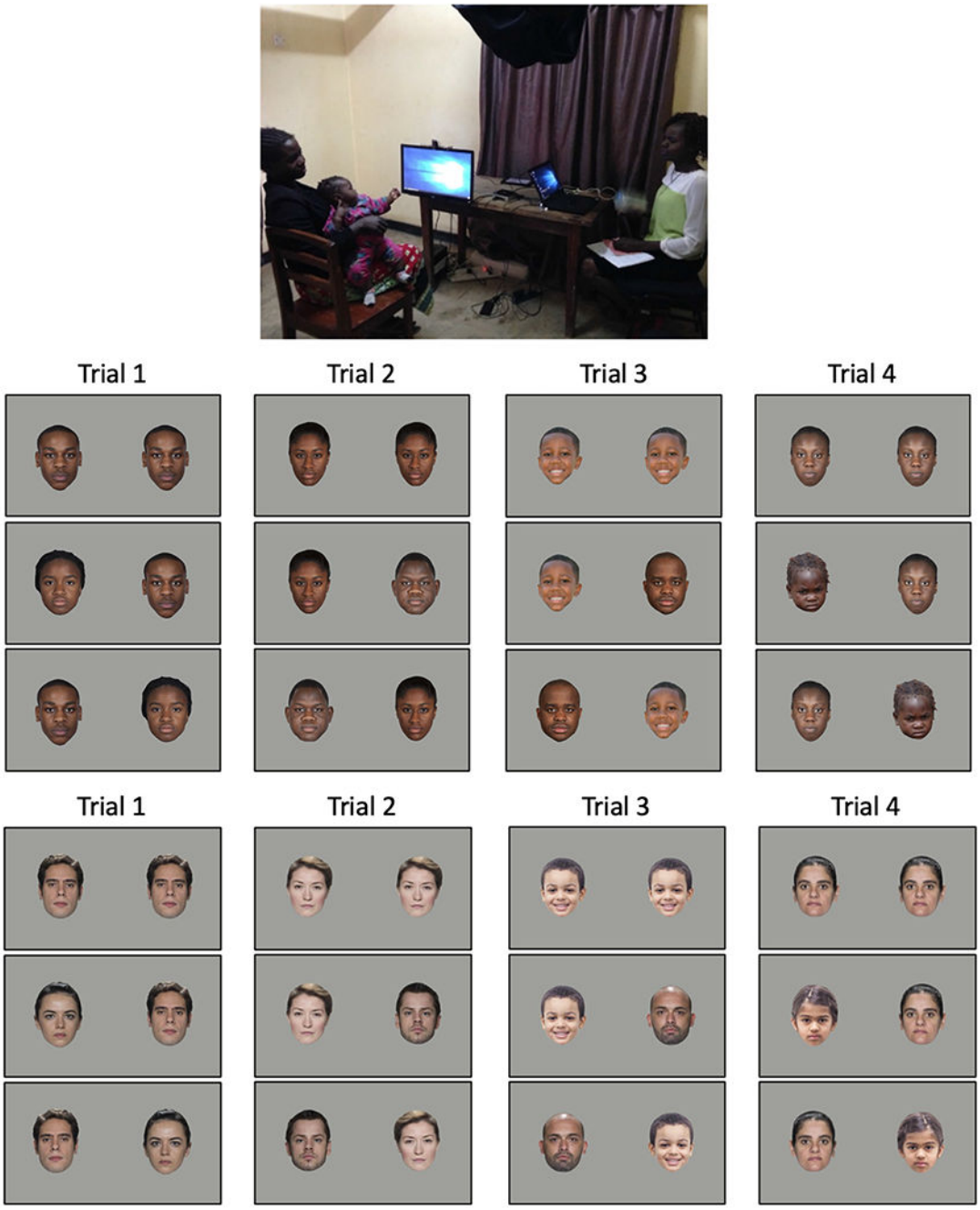

FIGURE 2.

(a) Apparatus for infants in rural Malawi and (b) VPC stimulus set for infants in rural Malawi (top) and the US (bottom). Infants were seated on their parent's lap (rural Malawi and US) or in a highchair (US) during the eye-tracking task and a curtain was used to block out visual distractions for both samples. All aspects of the apparatus and eye-tracking procedure were similar except that infants in the US were tested in a sound-attenuated room. Stimulus pairings were matched in terms of facial expression, perceived gender, and age for infants in rural Malawi and the US.

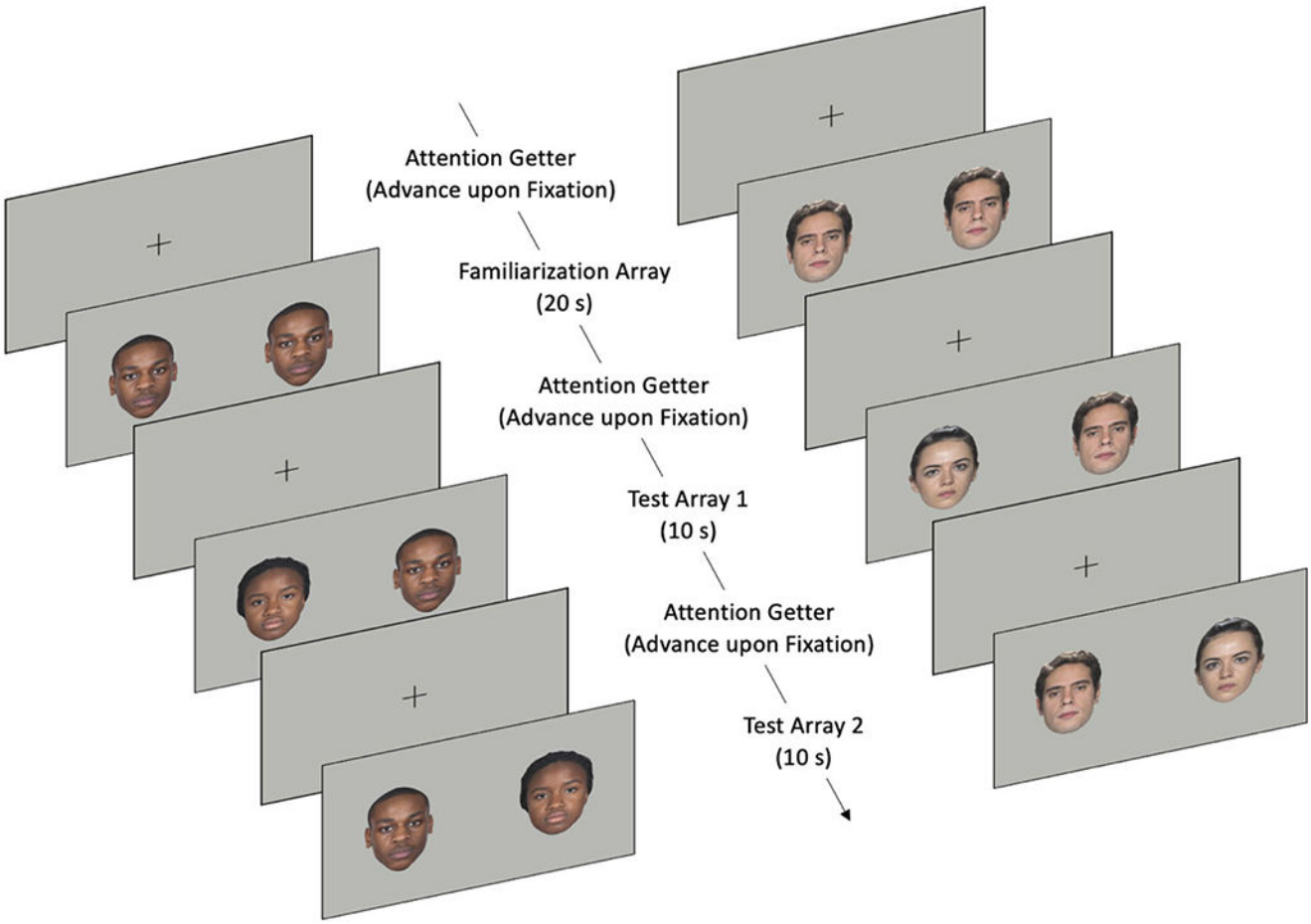

FIGURE 3.

Schematic illustration of a single experimental trial for infants in rural Malawi (left) and the US (right) for the VPC procedure. Infants in both settings were presented with four individual trials that each consisted of an initial familiarization phase and two test arrays. An attention getter was displayed at the beginning of each trial and in-between stimulus arrays; an experimenter initiated a button press when infants' fixation was detected toward the attention getter in the center of the screen.

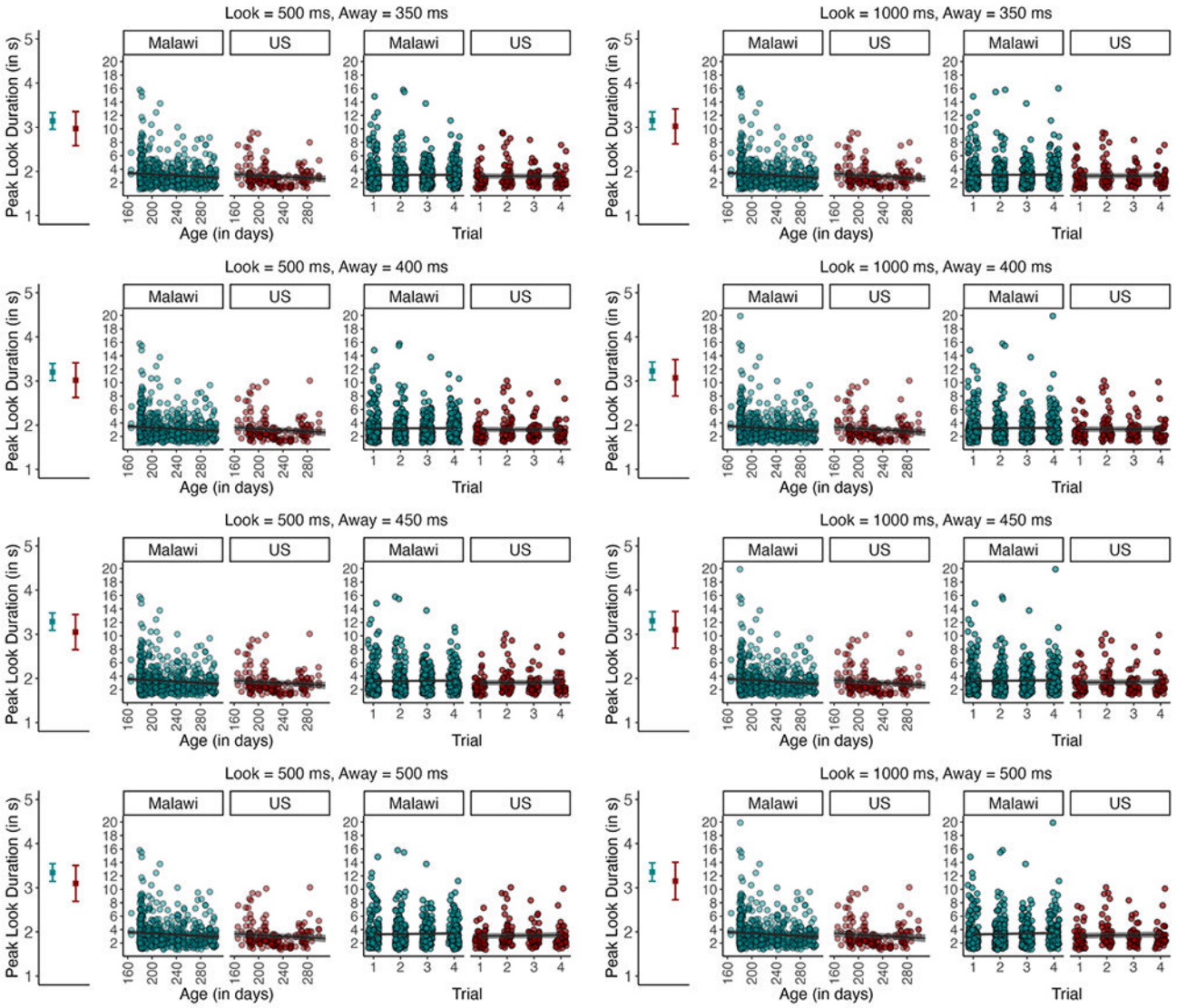

FIGURE 4.

Peak look duration by sample (blue = Malawi, red = US) as a function of operational definitions for a look (i.e., varying the minimum look duration and the maximum look away within a look). Scatter plots contain individual data points representing trial-level observations for infants in each sample and error bars indicate 95% confidence intervals for the estimated marginal means. The estimated marginal means for peak look duration in both samples (left), association between peak look duration and age in days (middle), and association between peak look duration and trial (right) are displayed as squares (sample) or regression lines (age and trial). Each row represents a different interruption duration threshold and each column represents a different look duration threshold.

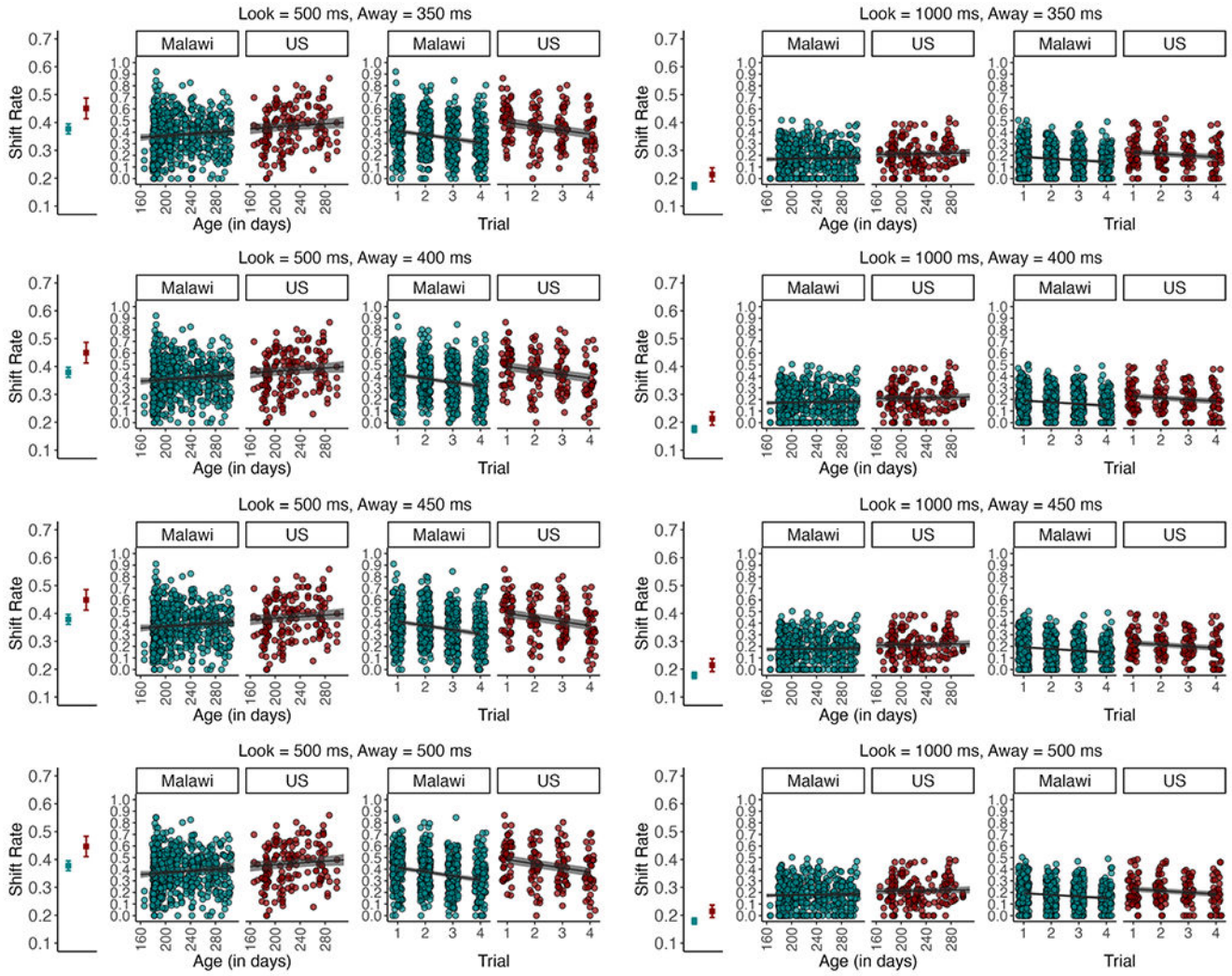

FIGURE 5.

Shift rate by sample (blue = Malawi, red = US) as a function of operational definition. Scatter plots contain individual data points representing trial-level observations for infants in each sample and error bars indicate 95% confidence intervals for the estimated marginal means. The estimated marginal means for shift rates in both samples (left), association between shift rates and age in days (middle), and association between shift rates and trial (right) are displayed for each look and interruption duration thresholds.

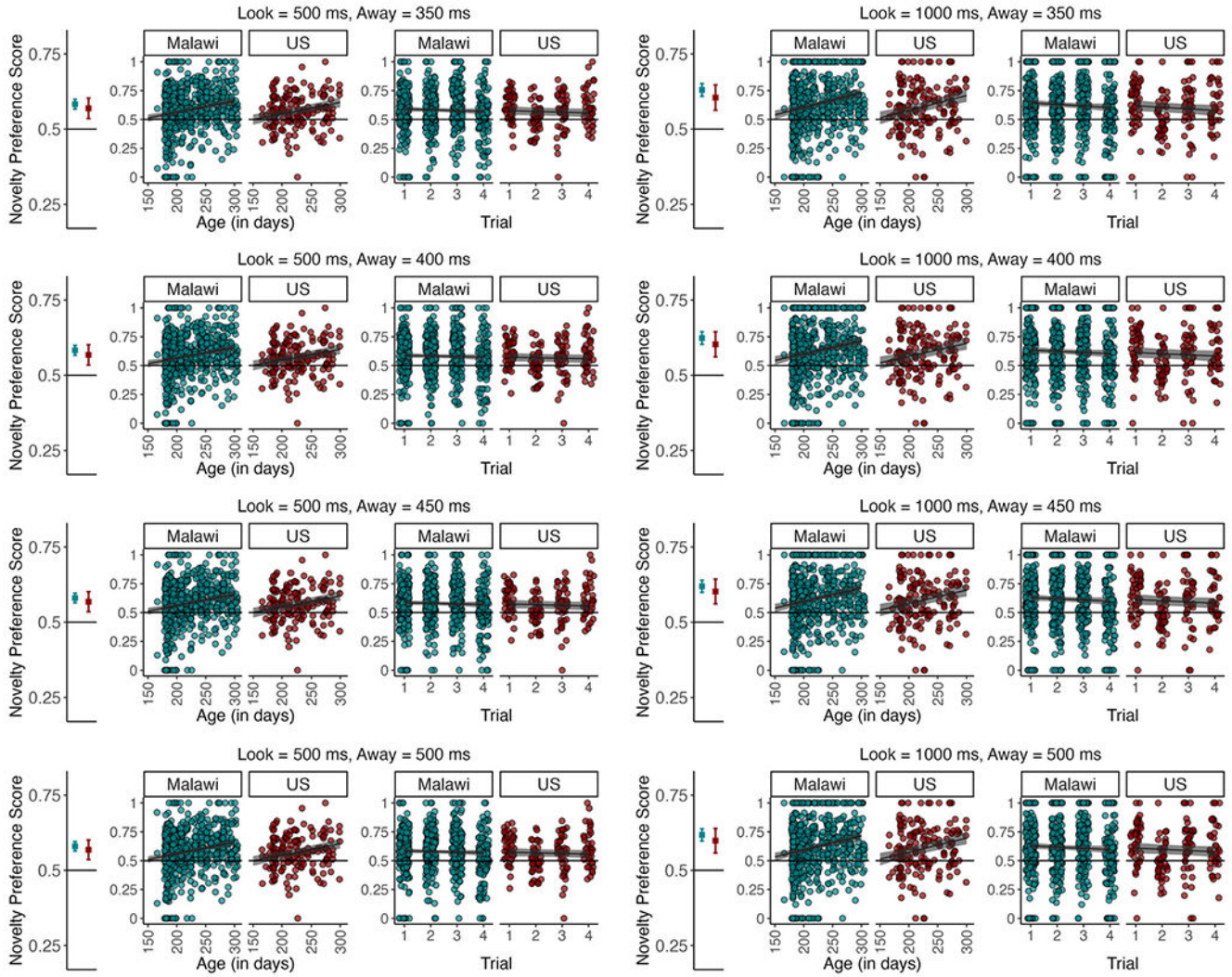

FIGURE 6.

Novelty preference scores by sample (blue = Malawi, red = US) as a function of operational definition. Scatter plots contain individual data points representing trial-level observations for infants in each sample and error bars indicate 95% confidence intervals for the estimated marginal means. The estimated marginal means for novelty preference scores in both samples (left), association between novelty preference scores and age in days (middle), and association between novelty preference scores and trial (right) are displayed for each look and interruption duration threshold. The horizontal line bisecting the vertical axis represents chance (0.50) performance.

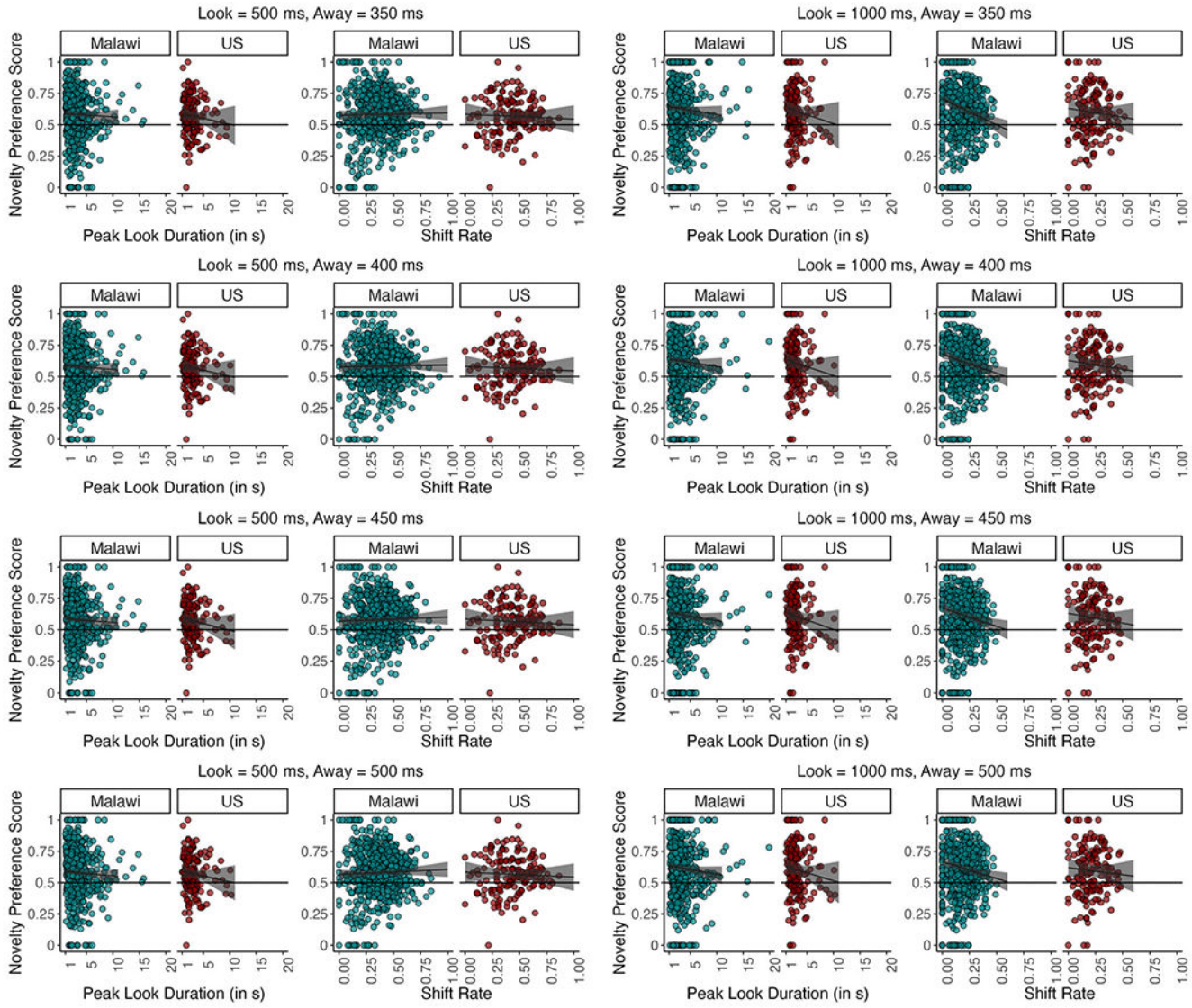

FIGURE 7.

Interaction between information processing measures and sample on infants' memory performance as a function of operational definition. The horizontal line bisecting the vertical axis represents chance (0.50) performance. Individual data points represent trial-level novelty preference scores for infants in the US (red) and Malawi (blue). The lines represent estimated marginal means and the shading around the lines represent 95% confidence intervals of the estimated marginal means.

TABLE 1

Demographic characteristics of each sample of infants.

Sample	Category	Sub-category	Frequency (*N*)	Percent (%)
Malawi	Food insecurity	None	49	21.50
		Mild	8	3.51
		Moderate	19	8.33
		Severe	152	66.7
	Maternal Literacy	Can read	115	50.40
		Cannot read	109	47.80
		Did not report	4	1.75
	Maternal occupation	Work at home	101	44.30
		Service industry	74	32.50
		Fishing or farming	48	21.10
	Tribal affiliation	Yao	190	83.30
		Chewa or other	34	14.90
		Did not report	4	1.75
US	Highest maternal education	High school (no degree)	2	4.16
		High school diploma	1	2.08
		Some college (no degree)	4	8.33
		Associate degree	4	8.33
		4-year degree or higher	37	77.08
	Maternal occupation	Health care	11	22.90
		Office environment	6	12.50
		Stay at home parent	11	22.90
		Education	6	12.50
		Unemployed	2	4.17
		Other	12	25.00
	Race	White	30	62.50
		Asian or Asian American	4	8.33
		Black or African American	1	2.08
		More than one race	12	25.00

Sample	Category	Sub-category	Frequency (*N*)	Percent (%)
	Hispanic	Yes	12	25.00
		No	36	75.00

TABLE 2

Results of peak look LMMs for sample, age, and trial effects across different operational definitions.

Look threshold (in s)	Interruption threshold (in s)	DV	Estimate	*SE*	*t*-statistic	*df*	*p*-value
500	350	Sample	−0.17	0.22	−0.79	202.67	0.43
		Age	**−0.18**	**0.09**	**−2.08**	**227.99**	**0.04**
		Trial	0.00	0.06	0.01	596.89	0.99
	400	Sample	−0.18	0.22	−0.83	203.64	0.41
		Age	**−0.19**	**0.09**	**−2.13**	**227.82**	**0.03**
		Trial	0.00	0.06	0.05	595.24	0.96
	450	Sample	−0.24	0.23	−1.04	207.51	0.30
		Age	**−0.18**	**0.09**	**−2.05**	**231.64**	**0.04**
		Trial	0.02	0.06	0.34	598.07	0.73
	500	Sample	−0.25	0.23	−1.07	208.45	0.29
		Age	**−0.18**	**0.09**	**−1.97**	**232.28**	**0.049**
		Trial	0.04	0.06	0.75	597.86	0.45
1000	350	Sample	−0.13	0.22	−0.58	202.43	0.56
		Age	**−0.19**	**0.09**	**−2.19**	**229.77**	**0.03**
		Trial	0.01	0.06	0.20	599.06	0.84
	400	Sample	−0.15	0.23	−0.66	207.33	0.51
		Age	**−0.20**	**0.09**	**−2.19**	**235.36**	**0.03**
		Trial	0.01	0.06	0.22	599.6	0.83
	450	Sample	−0.20	0.24	−0.86	210.27	0.39
		Age	**−0.19**	**0.09**	**−2.10**	**238.2**	**0.04**
		Trial	0.03	0.06	0.54	600.03	0.59
	500	Sample	−0.20	0.24	−0.85	211.59	0.40
		Age	−0.18	0.09	−1.95	237.89	0.05
		Trial	0.07	0.06	1.08	597.66	0.28

Note: The fixed effect of sample reflects differences in peak look duration during familiarization for infants in the US in relation to infants in rural Malawi and all estimates reflect fixed effects after controlling for covariates (i.e., child sex and calibration accuracy) and participant-level variation.

TABLE 3

Results of shift rate LMMs for sample, age, and trial effects across different operational definitions.

Look threshold (in s)	Interruption threshold (in s)	DV	Estimate	*SE*	*t*-statistic	*df*	*p*-value
500	350	**Sample**	**0.07**	**0.02**	**3.50**	**222.08**	**<.001**
		Age	0.01	0.01	1.71	242.86	0.09
		Trial	**−0.03**	**0.00**	**−7.24**	**585.11**	**<.001**
	400	**Sample**	**0.07**	**0.02**	**3.35**	**223.53**	**<.001**
		Age	0.01	0.01	1.67	242.76	0.10
		Trial	**−0.03**	**0.00**	**−7.36**	**581.96**	**<.001**
	450	**Sample**	**0.07**	**0.02**	**3.37**	**222.28**	**<.001**
		Age	0.01	0.01	1.64	241.43	0.10
		Trial	**−0.04**	**0.00**	**−7.90**	**581.99**	**<.001**
	500	**Sample**	**0.07**	**0.02**	**3.35**	**221.63**	**<.001**
		Age	0.01	0.01	1.79	240.71	0.07
		Trial	**−0.04**	**0.00**	**−7.99**	**581.88**	**<.001**
1000	350	**Sample**	**0.04**	**0.01**	**2.95**	**208.65**	**<.01**
		Age	0.00	0.01	0.72	234.31	0.48
		Trial	**−0.01**	**0.00**	**−4.12**	**594.07**	**<.001**
	400	**Sample**	**0.04**	**0.01**	**2.80**	**206.93**	**<.01**
		Age	0.00	0.01	0.65	234.16	0.52
		Trial	**−0.01**	**0.00**	**−4.18**	**595.96**	**<.001**
	450	**Sample**	**0.04**	**0.01**	**2.82**	**206.19**	**<.01**
		Age	0.00	0.01	0.57	233.9	0.57
		Trial	**−0.01**	**0.00**	**−4.37**	**597.09**	**<.001**
	500	**Sample**	**0.04**	**0.01**	**2.79**	**207.42**	**<.01**
		Age	0.00	0.01	0.64	234.89	0.52
		Trial	**−0.01**	**0.00**	**−4.23**	**601.02**	**<.001**

Note: The fixed effect of sample reflects differences in shift rate during familiarization and test for infants in the US in relation to infants in rural Malawi and all estimates reflect fixed effects after controlling for covariates (i.e., child sex and calibration accuracy) and participant-level variation.

TABLE 4

Results of novelty preference score LMMs for the fixed effects of sample, age, trial, shift rate, and peak look duration across different operational definitions.

Look threshold (in s)	Interruption threshold (in s)	DV	Estimate	*SE*	*t*-statistic	*df*	*p*-value
500	350	Sample	0.03	0.07	0.43	711.62	0.66
		Age	**0.04**	**0.01**	**5.15**	**191.44**	**<.001**
		Trial	−0.01	0.01	−1.08	625.49	0.28
		Shift rate	0.02	0.05	0.40	690.96	0.69
		Peak look	0.00	0.00	−0.78	735.52	0.44
	400	Sample	0.02	0.07	0.35	722.56	0.73
		Age	**0.04**	**0.01**	**4.87**	**189.94**	**<.001**
		Trial	0.00	0.01	−0.77	623.54	0.44
		Shift rate	0.02	0.05	0.35	676.82	0.73
		Peak look	0.00	0.00	−0.99	734.12	0.32
	450	Sample	0.04	0.07	0.63	725.8	0.53
		Age	**0.04**	**0.01**	**5.09**	**190.3**	**<.001**
		Trial	−0.01	0.01	−0.94	628.96	0.35
		Shift rate	0.03	0.05	0.68	681.52	0.5
		Peak look	0.00	0.00	−0.92	733.06	0.36
	500	Sample	0.04	0.07	0.55	719.33	0.58
		Age	**0.04**	**0.01**	**5.35**	**191.86**	**<.001**
		Trial	−0.01	0.01	−0.96	631.86	0.34
		Shift rate	0.04	0.05	0.86	679.93	0.39
		Peak look	0.00	0.00	−1.08	734.39	0.28
1000	350	Sample	−0.05	0.06	−0.86	525.66	0.39
		Age	**0.05**	**0.01**	**5.24**	**188.94**	**<.001**
		Trial	−0.01	0.01	−1.69	613.2	0.09
		Shift rate	**−0.41**	**0.09**	**−4.69**	**667.36**	**<.001**
		Peak look	−0.01	0.00	−1.45	679.87	0.15
	400	Sample	−0.03	0.06	−0.45	548.7	0.65
		Age	**0.04**	**0.01**	**4.61**	**188.23**	**<.001**

Look threshold (in s)	Interruption threshold (in s)	DV	Estimate	*SE*	*t*-statistic	*df*	*p*-value
		Trial	−0.01	0.01	−1.26	611.41	0.21
		Shift rate	**−0.31**	**0.09**	**−3.52**	**685.32**	**<.001**
		Peak look	−0.01	0.00	−1.30	669.66	0.19
	450	Sample	−0.02	0.06	−0.28	563.11	0.78
		Age	**0.04**	**0.01**	**4.79**	**192.32**	**<.001**
		Trial	−0.01	0.01	−1.33	615.28	0.18
		Shift rate	**−0.29**	**0.09**	**−3.34**	**692.63**	**<.001**
		Peak look	−0.01	0.00	−1.51	666.87	0.13
	500	Sample	−0.03	0.06	−0.53	559.4	0.6
		Age	**0.05**	**0.01**	**5.12**	**192.3**	**<.001**
		Trial	−0.01	0.01	−1.37	611.64	0.17
		Shift rate	**−0.27**	**0.09**	**−3.10**	**708.69**	**<.01**
		Peak look	−0.01	0.00	−1.68	676.09	0.09

Note: The fixed effect of sample reflects differences in novelty preference score for infants in the US in relation to infants in rural Malawi and all estimates reflect fixed effects after controlling for covariates (i.e., child sex and calibration accuracy) and participant-level variation.

TABLE 5

Results of novelty preference score LMMs for interaction terms (note that because a single novelty preference LMM was fit for each operational definition, these estimates were calculated from the same models as the results displayed in Table 4).

Look Threshold (in s)	Interruption Threshold (in s)	DV	Estimate	*SE*	*t*-statistic	*df*	*p*-value
500	350	Shift-rate-by-sample	−0.06	0.10	−0.56	654.07	0.58
		Peak-look-by-sample	−0.01	0.01	−0.6	753.88	0.55
	400	Shift-rate-by-sample	−0.05	0.10	−0.5	670.67	0.61
		Peak-look-by-sample	−0.01	0.01	−0.56	751.9	0.57
	450	Shift-rate-by-sample	−0.08	0.10	−0.76	677.68	0.45
		Peak-look-by-sample	−0.01	0.01	−0.71	752.97	0.47
	500	Shift-rate-by-sample	−0.08	0.10	−0.78	680.97	0.43
		Peak-look-by-sample	−0.01	0.01	−0.52	747.82	0.61
1000	350	Shift-rate-by-sample	0.27	0.18	1.47	636.67	0.14
		Peak-look-by-sample	−0.01	0.01	−0.54	663.88	0.59
	400	Shift-rate-by-sample	0.17	0.18	0.92	650.09	0.36
		Peak-look-by-sample	−0.01	0.01	−0.66	664.51	0.51
	450	Shift-rate-by-sample	0.14	0.18	0.75	665.49	0.45
		Peak-look-by-sample	−0.01	0.01	−0.67	662.2	0.50
	500	Shift-rate-by-sample	0.16	0.18	0.87	683.29	0.38
		Peak-look-by-sample	−0.01	0.01	−0.44	640.03	0.66

Note: All estimates reflect fixed effects after controlling for covariates (i.e., child sex and calibration accuracy) and participant-level variation.