A Critical Examination of Information Processing and Memory Performance in the Infant Visual Paired Comparison Procedure

By

AARON GERALD BECKNER

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Psychology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

_____

Lisa M. Oakes, Ph.D., Chair

_____

Steven J. Luck, Ph.D.

_____

Elizabeth L. Prado, Ph.D.

Committee in Charge

2021

Abstract

The visual paired comparison procedure (VPC) is a useful method that allows researchers to derive measures of observable looking behavior to address fundamental questions about learning and memory in preverbal populations. In this dissertation, I discuss the use of the VPC procedure for assessing information processing during infancy. In the first chapter, I provide general theoretical background about the VPC procedure. In the second chapter, I explore whether infants in a non-Western context display the patterns of associations we would expect from decades of research on infant VPC performance in Western samples. In the third chapter, I examine whether infants in rural Malawi show the same patterns longitudinally as infants in Western contexts. In the fourth chapter, I explore how the time available for infants to learn influences their memory in an online context. The overarching goal of this research is to critically examine the generalizability and reliability of infant VPC performance as well as examine how key aspects of infant attention influences their learning and subsequent memory in diverse contexts.

Acknowledgements

I dedicate this dissertation to my mother, Pam Murphy, and my fiancée, Christine Carlisle, who went above and beyond to support my educational aspirations. I'd like to thank my mom for always believing in me, showing me what it means to persevere through adversity, and pushing me to achieve every step of the way. Amidst your own personal hardships, you were there to lift me up during the most difficult times of my life and continue to support me now during some of the best times of my life. Your guidance, support, and strength have helped pave the way for me to reach this milestone. To my fiancée, Christine, you have had an immeasurably positive impact on my life. You have inspired me in so many ways throughout my educational journey through your passion, initiative, support, positivity, warmth, and patience. I am forever grateful for your enthusiasm and willingness to join me on this journey. I'd also like to thank my second mother, Stacey Mounce, for being a role model and profoundly shaping the person I am today; my father, Michael Beckner, for supporting me and expressing so much interest in my research; my step-mother, Lynette Beckner, for teaching me the value of hard-work and discipline; and all of my siblings and step-siblings. I'd also like to thank my brother Adam Beckner in particular, who was one of my biggest allies throughout my childhood and taught me how to think deeply about things through his sense of wonder and fascination with the world.

This dissertation would not have been possible without the collaborators and co-authors involved in designing various studies, collecting much of the data, and providing input on manuscript preparation. I'd like to take a moment to acknowledge several people in particular for their contributions. I'd like to thank Christine Stewart and Elizabeth Prado for leading the Mazira Project, a large-scale nutrition intervention study in rural Malawi that yielded the data reported in two of my dissertation chapters. I'd also like to thank Kenneth Maleta and Matthews George for

their massive data collection efforts in rural Malawi as well as their role in the implementation of the Mazira Project. Thanks also to Cecilia Chen for your involvement in stimulus creation and file management for the Mazira Project and thanks to my other co-authors including Bess Caswell, Megan Bragg, Katherine Cox, and Charles Arnold.

It should come as no surprise that this dissertation was completed in collaboration with my various members of my advising committee. My advisor, Lisa M. Oakes, is one of the most intelligent, hard-working, and encouraging people that I have ever met. I deeply admire your commitment to mentorship, methodological rigor, and intellectual approach to conducting research. You have an uncanny gift for being direct and structured while simultaneously being kind and supportive. You've taught me not only how to be a good scientist, but how to be a good mentor. I feel incredibly fortunate to have had the opportunity to work with you. I would also like to express my thanks to Steven J. Luck for profoundly shaping the direction of my research as a member of my dissertation committee, as a co-author, and as a mentor. You've helped me with professional development and taught me how to do science in an exciting collaborative environment. Thanks also to Elizabeth Prado for your role on my dissertation committee, your mentorship, and your immeasurable contributions as a collaborator.

I'd also like to express my sincere thanks to the undergraduates, staff, and graduate students of the Infant Cognition Laboratory as well as the broader community of researchers at the Center for Mind and Brain (CMB). Thanks for creating a place where researchers can conduct rigorous scientific research in a warm and positive work environment. I'd especially like to thank Michaela DeBolt who feels more like a sister than a coworker (for better and for worse). We've exchanged thousands of text messages, direct slack messages, and phone calls during graduate school and you've always been someone that I could rely as a researcher and as a

friend. I'd also like to thank Annika Voss for helping me stay organized, assisting with project management, and never being afraid to weigh in with recommendations and ideas. Lastly, I'd like to thank Gabi Blanch and Van Pham for help with project management and stimulus creation and Shannon Klotz, Erim Kizildere, Katie Pomaranski, and Christian Nelson for creating such a fun and positive work environment.

I've also been fortunate to form relationships that are independent from my research. My dissertation support animal, Chip, deserves acknowledgement for distracting me when I needed breaks and always showing me unconditional love. I was lucky to be a part of an amazing cohort of graduate students who shared many fun times together. I spent much of my free time playing pickup basketball at the Activities and Resource Center (ARC) and powerlifting and had the opportunity to make many friends there throughout the years. Having that community made for a more positive graduate school experience and gave me an outlet for expending my excess energy. I also want to thank the staff and regulars at "The Root of Happiness" in downtown Davis. At times it was difficult for me to find other graduate students that shared common life experiences. The odd and eccentric regulars at "The Root of Happiness" helped me feel welcomed and connected to a community of people "like me". Lastly, I'd like to thank Ken Jenkins who has been a lifelong friend and helped support me through difficult times.

TABLE OF CONTENTS

Chapter I. Background and Significance

Infants engage in sophisticated cognitive processes and actively learn about their environment in remarkably complex ways. Looking behavior provides researchers with an important tool for studying internal cognitive and psychological processes in infants. Decades of research has leveraged measures of looking behavior to provide insight about how infants form, maintain, and manipulate internal mental representations. Infants looking behavior varies systematically with age (Fagan, 1974; Harris, 1973; Rose, 1981; Ruff, 1975), stimulus complexity (Hunter, Ames, and Koopman, 1983), risk status (Rose, 2004), and physiological measures of attention (Richards, 1985; Colombo et al., 2001). These and other findings suggest that looking behavior reflects fundamental aspects of information processing in infants. Individual differences in infant looking behavior also predict long-term outcomes such as language (Thompson et al., 1991) and IQ (Fagan, 1984; McCall, 1994) during later childhood and thus provide researchers with a tool for assessing fundamental aspects of cognition in infants and young children. Variation in infants' looking behavior during learning is thought to reflect fundamental aspects of cognition (for review, see Rose et al., 2004) and information processing has traditionally been viewed as an expression of human intelligence (Deary, 2012; Stankov, 1983; Vernon, 1987).

The visual paired comparison procedure (VPC) is the "gold standard" for assessing memory and individual differences in information processing in infants (for review, see Rose et al., 2004). This procedure allows us to answer fundamental questions about infant information processing and learning. Studies of infant VPC performance have revealed several key insights about information processing and memory. First, because the infant VPC performance has good discriminative (for review, see Rose et al., 2004) and predictive validity (Colombo et al., 1988a;

Rose, Feldman, & Wallace, 1988b), the VPC procedure has been implemented as a tool for uncovering deficits in at-risk infants and evaluating the impact of clinical interventions (Colombo et al., 2014; Jacobson et al., 2018; Prado et al., 2020; Siegel et al., 2011), but these patterns are inconsistent (Benasich & Bejar, 1992). Moreover, most studies of infant VPC performance have been conducted with samples of infants residing in Western contexts, raising questions about the generalizability of these findings for infants residing in non-Western contexts. Second, the time available for infants to learn about attended items influences their subsequent memory. Studies of infants' VPC performance have revealed that (1) the time available for learning influences infants' memory (Fagan, 1970, 1971, 1972, 1974; Hunter et al., 1983; Rose, 1983) and (2) the strategies infants implement during learning affect their subsequent memory performance (Colombo et al., 1991, 2001; M. L. Courage & Howe, 2001; Rose et al., 2001, 2003) but less work has systematically examined the unique contributions of these variables on infant VPC performance. Some studies have focused on measures of infant behavior during the learning phase (Colombo et al., 1995, 2001; M. L. Courage & Howe, 2001) but others have systematically examined infants' behavior during the learning and recollection phase (Colombo et al., 1988a; Rose et al., 2001, 2003). Moreover, studies often conflate exposure time, or how long a stimulus was presented, with total look duration, or how much looking time infants accumulated to a stimulus.

In this research, I explore several questions about information processing and memory in infants using this procedure. Chapter 2 reports on similarities and differences in information processing and memory performance between infants residing in the US and rural Malawi. Most studies on infant VPC performance have been conducted with infants residing in Western contexts. Moreover, when studies have been conducted in non-Western contexts, it has generally

been assumed that Western infants are "typical" and that any differences in non-Western infants reflects deficits. However, it's not clear whether these assumptions are true for infants in non-Western contexts. The results reported in Chapter 2 suggest that assumptions based on studies of Western infants may not be true for infants residing in rural Malawi. Chapter 3 reports the results of a longitudinal study examining the association of measures of physical health and environmental risk on information processing and memory performance in infants. The results from this study indicate that infants in rural Malawi may follow a distinct developmental trajectory that is qualitatively different from infants in Western contexts. Chapter 4 reports preliminary results from a study of infant VPC performance in an online context. The purpose of this study is to disambiguate the unique contributions of exposure time and total look duration on infants' memory performance. The results from this study are preliminary but suggest that the influence of exposure time and look duration are nuanced and may be particularly difficult to isolate in non-traditional experimental contexts such as online settings.

Chapter II. Comparing infants' Visual Paired Comparison performance in the US and rural Malawi

Note: This is a draft of a manuscript that will be submitted for publication after input from coauthors.

Introduction

Infants' looking has been linked to information processing (Colombo et al., 1991; Fagan, 1974; Richards, 1985; Richards et al., 2010; Rose, 1981; Rose et al., 1982). Specifically, compared to older infants, younger infants exhibit longer, less numerous looks as they visually inspect an image or object (Harris, 1973; Rose, 1981; Ruff, 1975) and require more exposure to learn about an attended object (Fagan, 1974; Rose et al., 1982). Similarly, infants generally look longer at more complex visual stimuli than less complex visual stimuli (Hunter et al., 1983). Shorter individual looks during learning are associated with more robust memory for an attended item (Rose et al., 2001) and predict higher language and IQ scores during early childhood (Fagan, 2000; Fagan & McGrath, 1981; Rose et al., 1988, 1991; Thompson et al., 1991).

Taken together, these findings have led to the conclusion that shorter looking times during learning are the optimal strategy and reflect faster information processing that leads to more robust learning and better memory performance (Colombo, 2001; Frick & Richards, 2001; Richards, 1997; Rose et al., 1982). This conclusions is further supported by the fact that these measures are negatively associated with a wide range of adverse factors including malnutrition (Nelson et al. 1997; Carter et al. 2010), prematurity (Fagan, 1984; Guzzetta et al., 2006; Rose, 1980; Rose et al., 2001), exposure to toxins (Chiriboga et al., 2007; Gaultney et al., 2005;

Jacobson et al., 2002; Singer et al., 2005; Struthers & Hansen, 1992) (Emory et al., 2003), and developmental conditions such as Down syndrome (Miranda & Fantz, 1974; Nygaard et al., 2001). Thus, the literature is consistent with the default assumptions that (1) looking time covaries with information processing efficiency, and (2) faster information processing is optimal.

However, the work supporting these assumptions has been conducted primarily with infants from Western, high-income countries. Studies using these visual attention and memory tasks in other contexts have been interpreted using these default assumptions to identify optimal performance. For example, researchers asked whether a nutritional intervention effectively improved information processing in a sample of infants in Peru (Colombo et al., 2014). Specifically, the question was whether infants who received the intervention processed information faster and showed more robust novelty preference compared to infants who did not receive the intervention.

To date, no studies have systematically examined whether these default assumptions are true for infants in non-Western contexts. This is problematic because differences in looking behavior associated with environmental influences may not always reflect deficits but may also reflect attentional strategies that optimize infants' ability to extract information from their environmental context. Using findings from Western samples as a standard and adopting a focus on deficits in information processing in at-risk infants may cause researchers to ignore adaptive differences in looking behavior. As a result, the field fails to recognize the nuanced ways in which experiential influences shape developmental outcomes. Environmental influences shape biological reactivity, particularly to stress, and adaptive phenotypic plasticity allows for matching between an organism's phenotype and the context in which that phenotype is expressed (Ellis et al., 2006). The development of attentional processes broadly and infant looking behavior

in particular may reflect such processes. That is, observed differences in infants from different environments may reflect "ontogenetic adaptations" (Oppenheim, 1981) that provide an optimal fit between the brain and context.

The purpose of this study was to investigate contextual differences in information processing. We tested two samples of infants living in dramatically different contexts using a visual paired comparison (VPC) procedure. The VPC procedure has been considered the "gold standard" for assessing attention, information processing, learning, and memory in prelinguistic infants since the 1970s (Fagan, 1970, 1972, 1974). This procedure consists of a familiarization phase, during which infants are briefly shown a novel stimulus, followed by a test phase during which infants are then presented with the now-familiar stimulus paired with a different novel stimulus. This procedure takes advantage of infants' proclivity to orient toward novel stimuli over familiar stimuli (Sokolov, 1963) and infants' learning and memory is inferred from their pattern of looking behavior. Specifically, if during the test phase the proportion of total looking time devoted to looking at the novel stimulus (i.e., novelty preference score) is significantly greater than chance (.50), it is determined that infants have remembered the familiar stimulus. A novelty preference score of .50 indicates that infants were equally likely to look at the novel and familiar stimulus and suggests that they did not remember the familiar stimulus. Under some circumstances infants also show a significant preference for the familiar stimulus (Hunter et al., 1983), providing a different kind of evidence that they remember that stimulus.

The VPC procedure also has been used to provide evidence for individual differences in information processing. For example, the duration of looking time to a stimulus during learning and infants' subsequent memory for that stimulus is related to how efficiently infants process information (Colombo & Mitchell, 1990). Some infants are "short lookers" (i.e., spontaneously

exhibit shorter looks during learning), and are thought to be *fast* processors. These infants presumably process more quickly and thus require shorter looks to acquire and encode information. Shorter look durations also provide more opportunities for active comparison between two simultaneously presented stimuli. Infants who are short lookers can more efficiently disengage and rapidly reorient their attention than longer looking infants (Frick et al., 1999).

This work has yielded two measures of developmental and individual differences in information processing: (1) infants' peak look duration during the familiarization phase, and (2) infants' shift rate during both the familiarization and test phases. Peak look duration refers to the duration of infants' longest individual look during familiarization, and it is an indicator of information processing during learning. The assumption is that infants who process information more quickly will have shorter peak looks. In Western samples, peak look is negatively associated with novelty preference scores in the VPC procedure (Colombo et al., 1995, 2001). The rate at which infants shift their gaze between simultaneously presented stimuli is an indicator of information processing during the learning and recollection phases. In Western samples, shift rate is positively associated with infants' novelty preference scores in the VPC procedure (Rose et al., 2001). The assumption is that infants that shift more rapidly between simultaneously presented items are more actively engaged in information processing.

Less is known about the association between measures of information processing and memory performance in infants from non-Western contexts. Often when the VPC procedure has been used in non-Western samples (Dallaire et al., 2012; Fraser et al., 2012; Rose, 1994; Siegel et al., 2011) and with non-White participants (Carter et al., 2010; Chiriboga et al., 2007; Jacobson et al., 2002; Singer et al., 2005), studies have focused on memory as the primary outcome measure, rather than emphasizing how information processing influences memory.

For example, two studies of Ugandan infants examined the influence of HIV exposure on novelty preference and neither study examined the association between novelty preference scores and information processing (Chhaya et al. 2018; Drotar et al. 1997). However, Chhaya et al. found that novelty preference scores on an eye-tracking version of the VPC procedure were associated with infants' performance on the cognitive scale of the Mullen Scale for Early Learning (MSEL), a standard developmental battery. Kennedy et al. (2008) did examine measures of information processing and memory in infants residing in rural Southern Ethiopia, revealing that for the group as a whole novelty preference scores were significantly *less* than chance. This indicates that infants showed preference for the familiar item. In fact, only 17% of infants showed a novelty preference score that was greater than chance (Kennedy et al., 2008). Measures of information processing (e.g., *peak look duration, shift rate, and total look duration*) were related to physical growth outcomes, such as weight, head circumference, and body length, but total look duration was the only information processing measure related to novelty preference scores (Kennedy et al., 2008).

These studies indicate that the VPC procedure can be used to yield insights about memory and information processing in samples of infants residing in rural Sub-Saharan Africa. However, the results of these studies were interpreted with the assumption that shorter looks reflect faster, optimal information processing, derived from studies conducted with infants from Western, high-income countries. Because little work with on infants in non-Western samples has systematically examined the association between information processing measures and learning, it is unknown whether these assumptions are appropriate for infants in these contexts. More importantly, we are not aware of any study that has tested samples of infants from both Western and non-Western countries using the same VPC task, limiting the conclusion that can be drawn

with respect to contextual differences in attentional strategies during learning.

The goal of the present work was to compare performance on the VPC task in two samples of infants living in very different contexts. Specifically, we examined measures from a VPC task in infants from a high-resource, middle-class community in the US and infants from a low-resource, rural, low-income community in Malawi. The data from infants in rural Malawi was originally collected as a part of a nutritional intervention study evaluating the effects of eggs on child growth and development in which infants were randomly assigned to an intervention and control condition (Prado et al., 2020). The data from infants in the US was collected as part of a pre-registered study (https://osf.io/sj67p?view_only=18a5fe4096014b5f8674f266ab372267) designed as a follow-up to the work conducted in Malawi. These two groups of infants differed not only in the economic and educational status of their communities, but in cultural practices, exposure to technology, child rearing approaches, and many other factors that might contribute to early development. For example, infants in both communities live with their biological parents, but in Malawi children are more likely to live with extended family and multiple adults than in the US (Malde et al., 2015). These differences may have a profound influence on the attentional strategies that are optimal for infants in any given context.

We had three objectives. Our first objective was to determine whether infants in our two samples replicated the patterns previously reported in the VPC task literature. Thus, we replicated the analyses reported in Rose et al. (2001, 2003) separately in each sample of infants. Our second objective was to examine whether we observed the expected pattern of associations between our variables of interest (i.e., measures of information processing, memory and child age) in each sample using linear mixed-effect models. Our third objective was to test whether the patterns of associations we identified in our first two objectives differed significantly between

our two samples of infants.

Method

Participants

We analyzed data from two samples of infants aged 6 to 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, Malawi from February 2018 to January 2019 (Prado et al., 2020). It should be noted that the nutrition intervention had a limited impact on growth and development. The analyses reported here are original and were not reported when examining the effect of the nutrition intervention on developmental outcomes. Children aged 6 to 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 community outreach including village meetings and community football tournaments. Children were assessed at baseline before being randomized to the intervention or control group and again 6 months after baseline. There were 270 infants who participated in the baseline eye-tracking assessment before being randomized into intervention groups. Of the 270 infants that participated in the eye-tracking task at baseline, 251 were included in the analyses reported here; the data from 8 infants were excluded because we did not detect a fixation during the familiarization period on any trial, and the data from 11 infants were excluded because they did not contribute any memory problems that met our minimum inclusion criteria (described below). The US sample was drawn from a pool of infants who lived within a 30-mile radius of the University of California, Davis. Thus, these families predominantly lived in the urban and rural communities of the greater Sacramento Valley. Parents were recruited through mailings, and those who expressed interest were notified when their child reached the age range eligible

for the current study. Our target age range was between 6 and 9 months of age, corresponding to the age range for the baseline assessments in the Malawi sample. Fifty-three infants were tested, and our final sample consisted of 48 infants (M = 224.81 days, SD = 36.24, n = 24 girls) who provided usable data in the eye-tracking assessment; we excluded 5 infants because they did not contribute usable data on any memory problems. Table 1 displays the demographics for both samples.

Table 1. Demographic characteristics of each sample of infants.

Sample	Category	Sub-Category	Frequency (N)	Percent (%)
Malawi	Sex	Male	123	49.00
		Female	128	51.00
	Food Insecurity	None	55	21.90
		Mild	9	3.59
		Moderate	22	8.76
		Severe	165	65.70
	Maternal Education	Can Read and Write	118	47.00
		Cannot Read or Write	126	50.20
	Maternal Occupation	Work at Home	110	43.80
		Service Industry	84	33.50
		Fishing or Farming	51	20.30
	Tribal Affiliation	Yao	204	81.30
		Chewa or Other	42	16.70

		Did not Report	5	1.99
United States	Sex	Male	24	50.00
		Female	24	50.00
	Maternal Education	Earned 4-Year Degree	37	77.10
	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	13	27.08
	Ethnicity	White	30	62.50
		Asian American	4	8.33
		African American	1	2.08
		Multiracial	12	25.00
	Hispanic	Yes	12	25.00
		No	36	75.00

Apparatus

All infants were tested using the same make and model eye-tracking system, 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 X 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 placed in a room in the study center fitted with four black curtains hung from custom-built curtain rods (see Figure 2). When closed, the curtains created booth that blocked out distractions and only the monitor was visible to the mother and child, who were seated in a chair facing the monitor. The infant was either placed in a carrier worn by the mother or on the mother's lap. The monitor was positioned so that it was 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 (if they became fussy during calibration, lacked the ability to independently support themselves, or the parent was uncomfortable with the infant sitting in the highchair). To prevent parents from watching the stimuli and potentially biasing their infants' responses, they were provided felt-covered glasses to wear as their infant participated in the study.

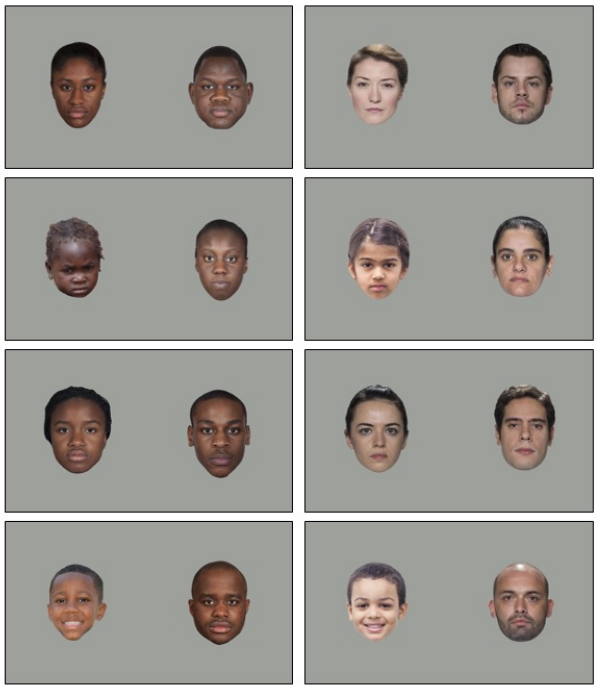

Figure 1. Stimulus pairings for the VPC procedure. The pairings are displayed in the order that they were shown in the VPC problems. The four pairs on the left were shown to infants in Malawi (Strohminger et al., 2016) and the four pairs on the right were shown to infants in the US (DeBruine & Jones, 2017; LoBue & Thrasher, 2014).

Stimuli

For each sample of infants, the experimental stimuli consisted of 8 face images that were selected to approximate the kinds of faces infants were likely to encounter in their daily lives (see Figure 1). Infants in Malawi were shown African faces from Strohminger et al. (2016) stimulus set(see Prado et al., 2020, for additional details). Infants from the US were presented with a combination of 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). Across the sets, the pairs were matched in age and perceived gender (e.g., an African boy's face in the Malawi procedure was matched with a racially ambiguous boy's face in the US procedure). Each face stimulus 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 .65 Hz and alternated with images of colorful toys.

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 University of Malawi College of Medicine. The same procedure was used in both locations. The session began with a standard 5-point calibration procedure (part of the Tobii Lab suite) in which a looming shape is 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 memory problems, interspersed with trials from an unrelated attentional cueing task (which will be reported elsewhere). The attentional cueing task 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 problems, minimizing any carry over from exposure from one set of images to the next.

Each of the four VPC problems involved the following trial sequence: a 20-s

familiarization array, in which two identical images were presented side-by-side, followed by two 10-s test arrays, in which the familiarization face was paired with a novel face (see Figure 2) The two test arrays differed only in the left-right position of the novel and familiar faces. A central fixation cross was presented prior to the onset of each memory problem and in-between stimulus arrays. 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.

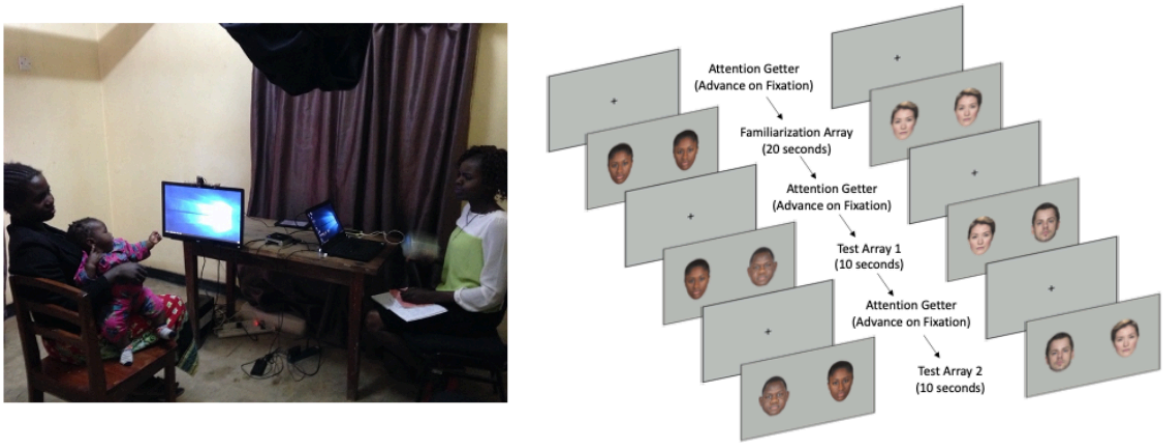

Figure 2. Apparatus and procedure for the VPC procedure. Infants were seated on their parents' lap during the eye-tracking task. Malawian infants were presented African faces and US infants were presented with a mix of White faces and racially ambiguous faces. All aspects of the apparatus and eye-tracking procedure were similar, except that infants in the US were tested in a sound-attenuated room free from visual distractions.

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 & Others, 2013), to process the data. Individual fixations were identified using the Tobii I-VT Fixation Filter in Tobii Studio. Fixations were defined as periods in which gaze position was stable for at least 60 ms. Saccades were differentiated from fixations using a threshold of 30 degrees per second and angular velocity was calculated across 20 millisecond time windows. Periods in which eye-tracking data was missing for less than 75 milliseconds were interpolated.

We created separate areas of interest (AOI) for the each half of the screen, starting at the edge of the central fixation and ending at the edges of the screen. The output from the datastream was the number of 16.67 ms samples comprising each fixation to an individual AOI. We calculated infants' *total fixation duration* on each trial by summing all samples classified as fixations within each AOI and then multiplying this number by the sample rate of 16.67 ms. Total fixation duration was calculated separately for the familiarization and test phase of each memory problem. We calculated infants' *novelty preference score* during test trials by dividing the total fixation duration to the AOI for the novel face by total fixation duration to the AOIs for both faces combined.

To approximate measures used in classic studies of information processing and memory performance (i.e., using human observers rather than eye tracking systems), we parsed the Tobii fixation data into individual *looks*. A fixation is a period of time in which infants maintained a stable point of gaze (POG) towards a location on the screen that was uninterrupted by saccades or fixations to other locations, whereas an individual *look* is a period of time in which an observer judged that an infant maintained *visual attention* to a particular stimulus that was uninterrupted by looks away from that stimulus. Thus, each individual look, as defined by a human observer, typically includes several individual fixations as defined by an eye-tracker. For

the present analyses we defined an individual look as a bout of visual attention to a single AOI that was at least 1 s in duration and contained no interruptions that were more than 1 s in duration. That is a *look* was a series of *fixations* to one AOI that in total were at least 1 s in duration with periods of looking away from that AOI (i.e., to the other AOI or off-screen) that were no longer than 1 s. This operational definition has been shown to yield the best balance between reliability and data quality in behavioral coding studies (Colombo & Horowitz, 1985). Note that because looks can include short interruptions (i.e., less than 1 s of looking away), the total amount of *looking* on a memory problem may be longer than the total amount of *fixation*. To be included in our analyses, infants must have accumulated at least 1 second of *looking* during familiarization and at least 1 second of *looking* during the two test phases.

Our final two measures, *peak look duration* and *shift rate*, were calculated from the *looks*. Peak look duration, a measure of information processing during learning, was calculated by identifying the longest individual look for each infant during the initial familiarization array of each memory problem (so infants had a peak look for each of the memory problems they completed). Shift rate, a measure of information processing during learning and memory recollection, is the rate at which infants shifted their looking from one stimulus to the other, or the number of gaze shifts between the two stimuli per second of looking. We first identified the number of shifts by counting each time a look to one AOI was preceded by a look to the other AOI (although there may have been a look away between the two looks or multiple individual looks within the same stimulus). We calculated shift *rate* by dividing the number of shifts in the familiarization or test phase of each memory problem by the total looking time on that phase of that memory problem.

Statistical Approach

All the analysis code will be made publicly available online (https://osf.io/r5gnw/). All statistical analyses were conducted using R Studio. To address our first objective, we conducted analyses on participant-level averages, replicating statistical approaches from classic studies of visual recognition memory. For each infant, we calculated an *average novelty preference* score by averaging the novelty preference scores for the memory problems completed, an *average shift rate* score by averaging shift rate during both the familiarization and test arrays for each problem, and an *average peak look duration* by averaging peak look during familiarization for each problem. We conducted one-sample *t*-tests comparing infants' average novelty preference score against chance (.50), and we assessed descriptive statistics and intercorrelations among these variables to compare our findings with previous studies (Rose et al., 2001, 2003).

We used linear mixed effects models (LMMs) to address our second and third objectives. All models were fit in R using lme4 (Bates et al., 2014) and *p*-values were calculated using lmerTest (Kuznetsova et al., 2017). Model assumptions were assessed using the performance package (Lüdecke et al., 2019). All results were confirmed using robust linear mixed-effect models (Koller, 2016). We found that models fitting the memory-problem-level novelty preference scores for the US sample resulted in a singular fit (i.e., the model could not estimate the variance components for the given random effect structure). Because dropping the random intercept for each participant and fitting a linear regression model with the same fixed effects did not change the overall pattern of significance and yielded nearly identical fixed effects estimates, we report the original model here for consistency. Interaction terms that were not statistically significant were dropped from the models and the fixed effects were interpreted excluding these nonsignificant interactions.

All models included fixed effects of *child age* (continuous: in days), infant *sex*

(categorical: male, female), and *memory problem* (continuous), and a random intercept for participant. For our second objective, we fit separately for each sample memory-problem-level data for each outcome measure (novelty preference, peak look, and switch rate), and an LMM on novelty preference scores with all the fixed effects in initial models plus peak look and shift rate. For models on novelty preference, we included a fixed effect of total fixation duration during *familiarization*. Note that this was included in the novelty preference models, but not the other models, because total fixation duration during familiarization is a measure of attention during learning. To address our third objective, we fit LMMs on our outcome measures on the combined data from the two samples, including a fixed effect of sample (Malawi vs. US) and interactions between sample and our measures of interest.

Results

Descriptive Statistics

Infants in both samples viewed most of the memory problems (Malawi: $M = 3.04$, $SD = 1.10$, US: $M = 3.69$, $SD = 0.78$). To address our first objective, we examined the descriptive statistics for the average scores (see Table 2). Comparisons of the novelty preference scores to chance indicated above chance performance for both the Malawi sample, $t(250) = 9.62$, $p < .0001$, $d = .61$, and the US sample $t(47) = 5.64$, $p < .0001$, $d = .81$. Thus, both groups of infants formed a memory of the familiarization stimulus.

Table 2. Descriptive statistics for our measures of interest for each sample.

			Familiarization		Test	
Sample	Measure	*N*	Mean	SD	Mean	SD
US	Novelty Preference	48	-	-	0.56	0.07
	Total fixation duration (in s)	48	10.4	3.71	11.4	3.35
	Peak look duration (in s)	48	3.73	1.96	-	-
	Shift Rate (shifts per second of looking)	48	0.17	0.08	0.16	0.08
Malawi	Novelty Preference	251	-	-	0.58	0.14
	Total fixation duration (in s)	251	7.98	3.89	8.26	4.16
	Peak look duration (in s)	251	3.77	1.55	-	-
	Shift Rate (shifts per second of looking)	251	0.11	0.07	0.14	0.10

Although we did not statistically compare the sample means averaged across memory problems, inspection of the means in Table 2 shows that mean peak look duration and mean novelty preference scores were remarkably similar for infants in Malawi and the US. Infants from the US showed slightly higher mean shift rates than infants in Malawi (note that a shift rate of 0.11 indicate 1 shift per each 9.1 seconds) and somewhat longer mean total fixation durations.

The intercorrelations between each child's mean shift rate across memory problems, mean novelty preference score, mean peak look duration, and mean total fixation duration, stratified by sample, are presented in Table 3. Both samples yielded expected significant positive

correlations between total fixation duration and shift rate, and between total fixation duration and peak look duration, confirming the validity of our method for combining fixations into looks to calculate shift rate and peak look duration.

Table 3. Intercorrelations (Pearson's r) among our measures of interest averaged across memory problems for each child and stratified by sample.

Sample	Measure	*N*	Novelty preference	Total fixation duration	Peak look
US	Novelty Preference	48	-	-	-
	Total look duration	48	-0.15	-	-
	Peak look duration	48	-0.21	0.48***	-
	Shift Rate	48	0.37*	0.45***	-0.08
Malawi	Novelty Preference	251	-	-	-
	Total look duration	251	-0.04	-	-
	Peak look duration	251	0.002	0.46***	-
	Shift Rate	251	-0.07	0.57***	0.01

* $p < .05$, ** $p < .01$, *** $p < .001$

The relations between information processing and novelty preference were strikingly different between the two samples. We observed the expected positive correlation between shift

rate and novelty preference score for infants in the US; higher shift rates were associated with stronger novelty preference. Peak look was not correlated with novelty preference. In contrast, the Malawian sample did not show correlations between shift rate or peak look duration with novelty preference.

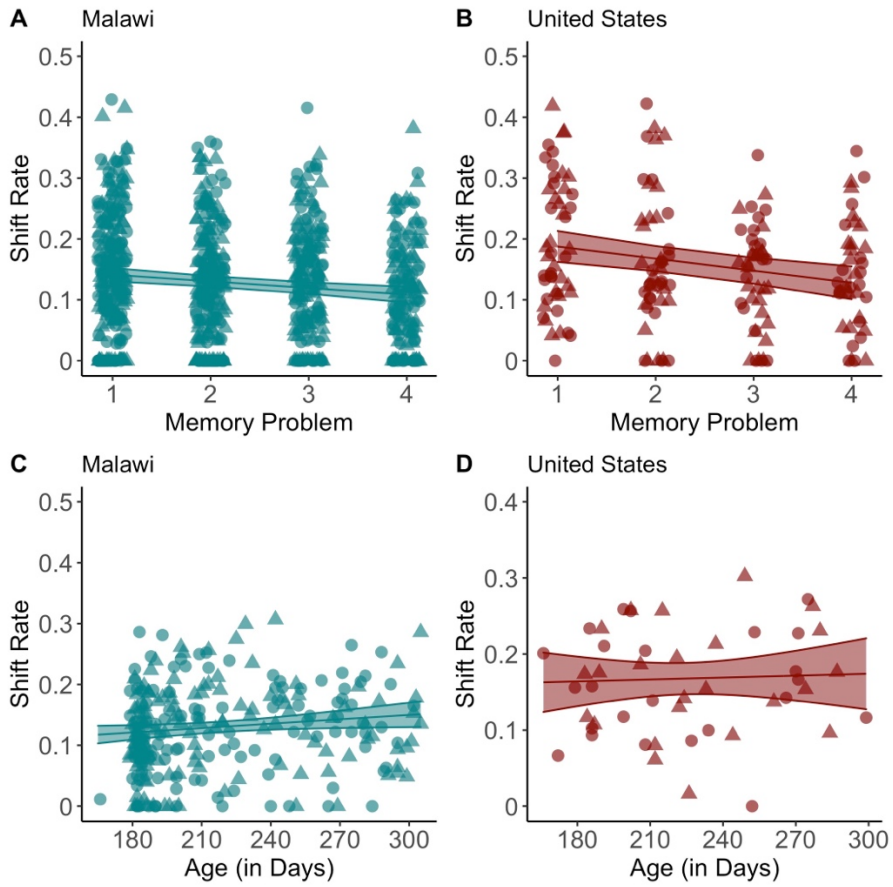

Figure 3. Association between shift rate and memory problem (top) as well as shift rate and age (bottom). Note that these data were generated from models fit separately to the data for both samples of infants. Individual data points represent infants' shift rate (top) and mean shift rate (bottom) across all memory problems. The color of each individual dot indicates sample (red = US, teal = Malawi) and shape reflects infants' biological sex (circle = females, triangle = males). Shading around the lines represent 95% confidence intervals of the estimated marginal means.

Examining the associations among our variables of interest separately for each sample

In this section, we examine the patterns of correlation separately for each sample. The following section will provide comparisons of the samples. The LMMs on the memory-problem-level data for peak look duration as specified above revealed no significant effects for either

sample. The models evaluating shift rate revealed a significant fixed effect of memory problem for both the US sample, β = -0.02, *SE* = 0.01, *t* = -3.51, *df* = 134.65, $p < .001$, and the Malawi sample, β = -.01, *SE* = 0.00, *t* = -3.83, *df* = 601.86, $p = .0001$, due to shift rate decreasing over memory problems for infants in both samples (see Figure 3A and 3B)There was also a significant fixed effect of age for infants in Malawi, β = 0.01 *SE* = 0.004, *t* = 2.126, *df* = 228.56, $p = .03$ (Figure 3D), but not for infants in the US, β = 0.00, *SE* = 0.01, *t* = 0.29, *df* = 45.09, $p = .77$ (Figure 3C). For infants in Malawi, shift rate increased with age. No other fixed effects were significant for either model (the full model results are reported in the supplementary materials at (https://osf.io/r5gnw/).

For novelty preference models, we centered infants' novelty preference score on each memory problem by subtracting chance (.50) to aid in intercept interpretation. Because we centered novelty preference scores, if we observe that our intercept for novelty preference score is significantly different from 0, it will indicate that infants showed a preference for the novel face that is significantly different from chance. The models revealed intercepts that were significantly different from 0 for both samples, US sample, β = 0.06, *SE* = 0.03, *t* = 2.42, *df* = 173.00, $p = 0.02$, Malawi sample, β = 0.09, *SE* = 0.02, *t* = 5.49, *df* = 741.10, $p < .0001$. Thus, consistent with the *t*-tests reported earlier, these effects suggest that, as a whole, both groups of infants showed evidence of remembering the familiarized item after accounting for the effects of total fixation duration during familiarization, age, memory problem, and sex (see Figure 4A).

We also observed a fixed effect of age for the Malawi sample, β = 0.03, *SE* = 0.01, *t* = 4.56, *df* = 194.48, $p < .0001$, but not for the US sample, β = 0.00, *SE* = 0.01, *t* = 0.25, *df* = 173.0, $p = .80$ (see Figure 4). Infants in Malawi showed higher novelty preference scores with increasing age.

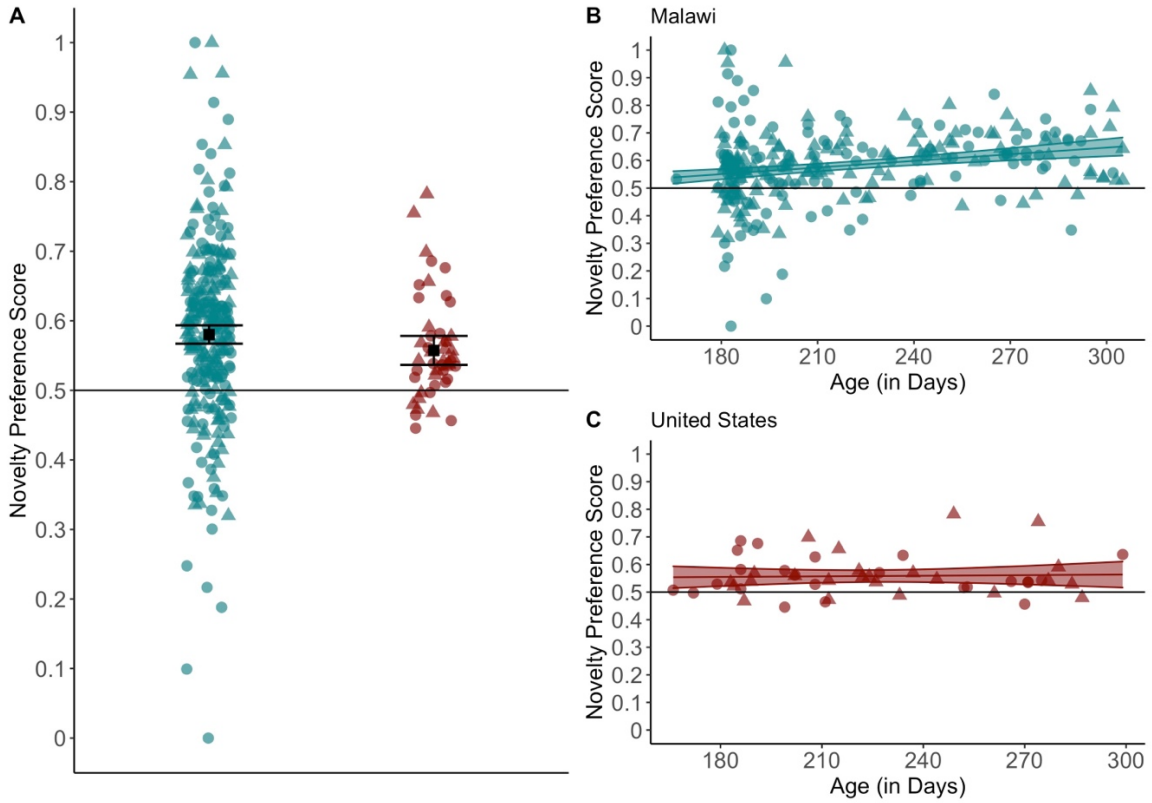

Figure 4. Novelty preference scores (each individual data points represent one infants' mean novelty preference score averaged across all problems) by sample (teal = Malawi, red = US) and biological sex (circle = females, triangle = males). The horizontal line bisecting the vertical axis in each figure represents chance (.50) performance. (A) The black squares display the intercept estimates for novelty preference score from each model and the error bars surrounding these squares represent 95% confidence intervals. (B) The association between novelty preference scores and age in days for infants in Malawi. (C). The association between novelty preference scores and age in days for infants in the US.

Next, we evaluated the association between our measure of memory (novelty preference) and our measures of information processing (peak look, and shift rate). We fit the models

described in the Statistical Approach. In the Malawi sample, novelty preference was significantly associated with age, $\beta = 0.03$, $SE = 0.01$, $t = 4.37$, $df = 194.70$, $p < .0001$, and peak look duration, $\beta = 0.01$, $SE = 0.003$, $t = 3.33$, $df = 723.40$, $p < .001$, but not shift rate. As reported earlier, older children had higher novelty preference scores. However, the effect of peak duration reflected the fact that longer peak look durations were associated with higher novelty preference scores. In the US sample, novelty preference was associated with shift rate and novelty preference scores, $\beta = 0.25$, $SE = 0.12$, $t = 2.06$, $df = 170.00$, $p = .04$, corroborating the correlation reported earlier. In this sample, more shifts were associated with higher novelty preference score.

Examining regional differences in our variables of interest

To address our final objective, and directly compare responding in the two groups of infants, we fit our outcome variables to models that included a fixed effect of *sample* (categorical) and *age-by-sample interaction*. Our model on peak look duration revealed no significant fixed effects or interactions, confirming the general pattern we observed when we analyzed the data separately for each sample (i.e., there were no significant associations of peak look duration with memory problem, age, or infant sex) and further suggest that there were no sample differences in peak look duration or in the association of peak look duration with age.

Because our initial model on shift rate did not reveal a significant interaction between age and sample, this interaction term was dropped from the final model. The final model revealed significant fixed effects of sample, $\beta = -0.03$, $SE = 0.01$, $t = -3.14$, $df = 249.88$, $p = .002$, and age, $\beta = 0.01$, $SE = 0.004$, $t = 2.03$, $df = 278.31$, $p = .04$. The first effect confirms the impression of the means in Table 1 that infants in the US had a higher shift rate than did infants in Malawi. The second effect reflected shift rate increasing with age. This model also revealed a negative association between shift rate and memory problem, $\beta = -0.01$, $SE = 0.003$, $t = -4.99$, $df = 736.49$,

$p < .0001$, due to infants exhibiting a higher shift rate on earlier memory problems than on later memory problems, confirming the effects observed when analyzing the two samples separately.

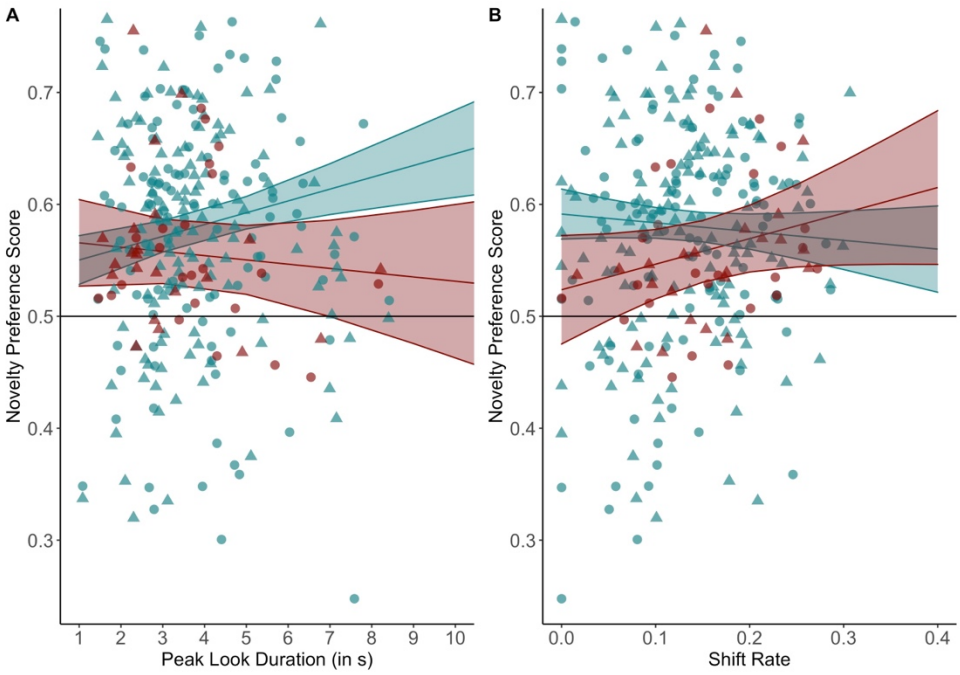

Figure 5. Association between novelty preference score and peak look duration (Figure 5A) and shift rate (Figure 5B). The horizontal line bisecting the vertical axis represents chance (.50) performance. Individual data points represent infants' mean novelty preference scores averaged across all memory problems for infants in the US (red) and Malawi (teal) and shape represents biological sex (circles = girls, triangles = boys). The lines represent estimated marginal means and the shading around the lines represent 95% confidence intervals of the estimated marginal means.

Next, we examined the association between measures of information processing and memory performance on the combined data from both samples. We fit an LMM on novelty preference that included fixed effects of our information processing measures (i.e., peak look and

shift rate), total fixation duration during familiarization, and interactions between our information processing measures and sample (i.e., *peak look-by-sample* and *shift rate-by-sample* interactions). An interaction between age and sample was also included but dropped from the final model because it was not statistically significant. The intercept was significantly different from chance, $\beta = 0.08$, $SE = 0.04$, $t = 2.07$, $df = 596.70$, $p = .04$, indicating that infants' preference for the novel face was significantly greater than chance after accounting for the factors in our model. We also observed a significant fixed effect of age, $\beta = 0.025$, $SE = 0.01$, $t = 4.25$, $df = 233.57$, $p < .0001$, as was revealed when analyzing the Malawi data alone. Importantly, this analysis revealed a significant interaction between sample and peak look duration, $\beta = 0.01$, $SE = 0.006$, $t = 2.51$, $df = 678.24$, $p = .01$ (see Figure 5A). Thus, the association between peak look duration and memory performance differed for infants in our two samples. The US sample showed the expected pattern; shorter peak look durations were associated with higher novelty preference scores. However, the Malawi sample showed the opposite pattern; longer peak look durations were associated with higher novelty preference scores.

We also observed a significant interaction between sample and shift rate, $\beta = -0.31$, $SE = 0.14$, $t = -2.15$, $df = 814.74$, $p = .03$. This pattern of results is displayed in Figure 5B. Again, the US sample showed the expected pattern: higher novelty preference scores were associated with higher shift rates. The Malawi sample showed the opposite pattern: higher novelty preference scores were associated with lower shift rates.

Discussion

In this study, we observed similarities and differences in visual recognition memory in two samples of infants, one in rural Malawi and the other in suburban US. We had three

objectives. First, we sought to determine whether infants in our two samples replicated the patterns previously reported in the literature. We observed that infants in both samples showed similar memory performance as indicated by their novelty preference scores, consistent with previous studies conducted in Western (Rose et al., 2004a) 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). This supports the conclusion that measures of memory performance in this task reflect general cognitive abilities that are generalizable across cultures and contexts.

Our second objective was to examine the relation between variables of interest in each sample. Infants in both samples decreased their shift rate over memory problems, indicating changes in how infants deployed their attention to two simultaneously presented stimuli over time. Peak look and novelty preference did not change over successive memory problems. Furthermore, age was associated with some measures of interest but not others, a finding that is in conflict with results from Rose et al (2001, 2003), who observed that novelty preference, peak look, and shift rate generally changed between 6 and 9 months. This discrepancy likely reflects the fact that the current study was conducted with a narrower and more normally distributed range of ages whereas Rose et al. examined cohorts of infants longitudinally at discrete time periods (e.g., 5 months, 7 months, 9 months).

Finally, our third objective was to test how the patterns of associations were similar or different in our two samples of infants. Thus, although the conclusion from our first objective that infants' memory performance in the VPC procedure reflect cognitive processes that are generalizable across environments and contexts may be true, the particular processes infants use to encode and form those memories seem to vary across environments and contexts. Examination

of the association of information processing measures, specifically shift rate and peak look duration, with novelty preference revealed the classic pattern for our Western sample, but not for our non-Western sample. As has been previously observed in other Western samples, in our US sample, higher shift rates (Rose et al., 2001, 2003) and shorter peak look durations (Colombo et al., 1995, 2001) were associated with stronger novelty preference. For our Malawi sample, in contrast, lower shift rates and longer peak look durations were associated with higher novelty preference scores, the opposite of the pattern that has been observed in Western samples.

Thus, these findings challenge default assumptions about associations between measures of information processing and memory scores, or novelty preference. Specifically, it has been assumed that more efficient information processing, as indicated by shorter, faster individual looks, is the path to stronger memory (Rose et al., 2001, 2003, 2004a). This is the pattern we observed for infants in the US. However, for infants in Malawi, stronger memory (as indicated by higher novelty preference scores) was associated with what has been assumed to reflect less efficient information processing - longer peak looks and slower shift rates. One possible explanation for this finding is that there are multiple paths to forming a strong memory of the familiar stimulus, and that the contexts in which infants develop shape the attentional strategies that are optimal for infant learning. Thus, the developing infant mind is much more flexible than has previously been appreciated, with multiple routes to the same overall memory ability.

This possibility is supported by findings that other developing cognitive abilities are influenced by contextual factors. For example, infants exposed to significant maternal anxiety adopt different attentional strategies, especially attention toward threat, than infants who do not experience maternal anxiety. Specifically, compared to infants with mothers who score low on measure of maternal anxiety, infants with mothers who score high on measures of maternal

anxiety show an attention bias towards threatening stimuli such as angry faces (Morales et al., 2017). Results such as these suggest that rather than there existing a single optimal developmental trajectory for attention to threat, different attentional phenotypes may be adaptive depending on the context. Heightened sensitivity to faces expressing negative affect may be particularly adaptive for infants reared in unsafe environments where maternal anxiety may provide an important signal about the presence of threatening stimuli (LoBue & DeLoache, 2010; Moulson et al., 2009; C. A. Nelson & Dolgin, 1985; Peltola et al., 2008).

The different attentional strategies that we observed may similarly reflect distinct attentional phenotypes that are uniquely adaptive for infants in their particular context. Specifically, rather than a single optimal profile of attention and memory, the present results suggest that different attentional processes or strategies may be adaptive for learning and memory in different contexts. Each of our group of infants may have exhibited a pattern that reflected the strategy that was adaptive and optimally supported learning in their respective contexts. To our knowledge, this is the first study to systematically investigate whether differences in attention in Western and non-Western samples are adaptive for the context in which those infants are reared.

These findings raise important questions about *why* infants in the US and rural Malawi showed a different pattern of results and *how* these attentional processes might be adaptive for infants in their unique contexts. One possible explanation for the distinct patterns we observed across our two samples is that these differences may reflect tradeoffs between exploration and exploitation (Berger-Tal et al., 2014; Kaelbling et al., 1996; Stephens & Krebs, 1986). Humans have evolved to flexibly modify their behavior in response to the specific demands of their environment and infants in particular may be highly sensitive to these kinds of contextual

influences. Efficient visual foraging requires implementing specific strategies based on the unique demands of a particular context. Whether the most adaptive strategy is to explore new information or exploit an existing resource is dramatically influenced by the context in which infants develop (Gopnik, 2020). Exploitation involves taking advantage of an existing resource whereas exploration entails effectively sampling multiple sources of information in the environment. The fact that infants in the US shifted more rapidly and executed shorter looks during learning suggests that they implemented an attentional strategy that was focused more on exploration. In contrast, infants in rural Malawi showed longer looks and lower shift rates, indicating that their attentional strategy may have been driven more by exploitation. Future research will be necessary to explore whether the differences we observed are related to tradeoffs between exploration and exploitation as they relate to visual foraging.

It must be pointed out that although we recruited a sample in the US to match the average age and age distribution of the sample in Malawi, the two samples were not precisely the same in how age was distributed. It is possible, therefore, that the differences between the samples reflected the fact that the Malawi sample had a larger proportion of infants at the young end of the distribution than did the US sample. We addressed this by generating subsamples of the Malawi sample that were age-matched to the US sample. Given the difference in sample size, we were able to generate 1000 such subsamples, and redo all our analyses comparing these subsamples to the US sample. Most of these comparisons yielded the same results as reported here, confirming that our effects did not reflect differences in the distribution of ages in our two samples (the complete results of this resampling approach can be found in our supplemental materials,).

The studies that have been conducted on cognitive abilities in non-Western samples of

infants have predominantly focused on uncovering deficits in at-risk infants (Chhaya et al., 2018; Dallaire et al., 2012; Drotar et al., 1997; Fraser et al., 2012; Jacobson et al., 2018; Rose, 1994; Siegel et al., 2011). The current work ,represents one of the first studies to assess the influence of culture and context on infant cognitive development using a nearly identical version of the same procedure. The fact that our tasks were nearly identical except that they incorporated culturally appropriate stimulus arrays gives us more confidence that the similarities and differences we observe are valid. One challenge in interpreting the differences we observed, however, is that there were so many differences between our two samples that it is difficult to pinpoint the precise experiences that shaped infants' attentional strategies. Our observed findings could reflect cultural practices such as marital status and parenting approach (e.g., monogamous versus polygamous parents, gender norms around primary caregiving, involvement of extended family in parenting), economic differences such as reliable access to technology and resources, and environmental factors such as the distinction between rural agricultural contexts and suburban areas. These differences make it difficult to derive strong conclusions about what specific experiences shaped the differences we observed. However, that was not the goal of the present study; instead, the goal was to determine whether the pattern of correlations previously observed in Western samples reflect universal, domain-general principles of development or simply characteristics that are specific to infants in Western samples. What we can conclude is that the two groups responded similarly in terms of their memory, despite showing different attentional strategies during learning.

Taken together, these findings highlight the need for systematic studies of information processing and memory performance in non-Western samples of infants. Whereas previous research with non-Western samples has focused on identifying risk factors, our findings indicate

that non-Western samples must be studied for the field to truly uncover the basic principles that govern information processing during learning in infants. By analogy, linguists learned many years ago that general principles of linguistics can be understood only by examining a broad range of languages. For the field of infant cognitive development to uncover the *general principles* of development, and not be limited to the principles of development within a narrow set of cultural contexts, future research must go beyond Western samples.

Chapter III. Examining the stability and reliability of information processing and memory in infants residing in rural Malawi

Note: This is a draft of a manuscript that will be submitted for publication after input from coauthors.

Introduction

Infants must shift their gaze between multiple objects, features, and locations to efficiently learn about their environment. Studies of infants residing in Western contexts have revealed that measures of infant looking behavior, such as the duration of individual looks or the amount of active comparison between items, are associated with more robust learning for attended items (Colombo et al., 1995, 2001; Frick & Richards, 2001; Richards, 1997; Rose et al., 2001, 1982), indicating that some infants are better learners than others. In addition, individual differences in infants' looking behavior remain relatively stable across infancy (Colombo et al., 1988b; Rose et al., 2001) and predict long-term cognitive outcomes during early childhood (Fagan, 2000; Fagan & McGrath, 1981; Rose et al., 2005; Rose, Feldman, & Wallace, 1988a; Rose et al., 1991). Taken together, this literature has supported the conclusions that specific aspects of infant looking behavior are (1) associated with better learning and (2) reflect stable individual differences in information processing across infancy.

Two such aspects of infants' looking behavior are (1) how often infants shift their gaze between multiple simultaneously presented objects or images, and (2) the duration of individual looks to objects or images. Infants who exhibit more shifts and who execute shorter individual looks (e.g., "short lookers") have traditionally been assumed to process information more

efficiently than infants that shift more slowly or execute longer individual looks (e.g., "long lookers"). These differences in shift rate and look duration are thought to reflect differences ins processing speed or the development of cortical structures involved in inhibitory control (Amso & Scerif, 2015; Colombo, 2001; Colombo et al., 1991; Freeseman et al., 1993; Johnson et al., 1991; McCall, 1994; Rose et al., 2001). These observations reflect the assumption that individual differences in infant looking behavior reflects variation in basic aspects of human cognition (for review, see Rose et al., 2004).

Evidence that these aspects of visual behaviors reflect variation in cognitive processes derives from findings that infants who display shorter individual looks and shift more rapidly between multiple items show more robust learning in the visual paired comparison (VPC) procedure (Colombo et al., 1995, 2001; Rose et al., 2001). Specifically, these infants have higher novelty preference scores than infants of the same age who have longer individual looks and fewer shifts between items (Rose et al., 2001). In addition, peak look durations, or the longest individual look during learning, decrease whereas shift rates increase with infant age (Axia et al., 1999; Colombo & Mitchell, 1990; Colombo et al., 1988b; Rose et al., 2001; Ruff, 1975), supporting the notion that these measures reflect how effectively and efficiently infants are processing and encoding the information.

VPC information processing measures also show moderate test-retest reliability over periods of weeks and months (Colombo et al., 1988b; Rose, Feldman, & Wallace, 1988a) and predict a broad range of cognitive abilities during later childhood including receptive and expressive language, vocabulary, IQ, and other forms of memory (for review, see Rose et al., 2004). Additionally, infant VPC performance shows high discriminant validity and is known to be negatively related to risk factors including infant Apgar scores (Caron, Caron, & Glass, 1983),

prematurity (Guzzetta et al., 2006; Rose, 1980; Rose et al., 2001; Rose, Feldman, Mccarton, et al., 1988; Rose et al., 1991), genetic abnormalities (Miranda & Fantz, 1974; Nygaard et al., 2001), nutrition (Carter et al., 2010; Nelson et al., 1997; Rose, 1994), and teratogen exposure (Chiriboga et al., 2007; Gaultney et al., 2005; Jacobson et al., 2002; Singer et al., 2005; Struthers & Hansen, 1992).

However, recent findings suggest that at least some of the patterns that have previously been observed in the VPC procedure in Western samples may not be true for infants in non-Western contexts. Beckner et al. (submitted) compared VPC performance in two groups of infants; one group lived in rural Malawi and the other group lived in suburban United States. There were both similarities and differences in VPC performance in the two groups of infants. Infants from both regions showed evidence of learning and memory for attended items; however, this learning was supported by different attentional strategies for infants in the US and rural Malawi. Shorter peak look durations and faster shift rates were associated with better memory performance for infants in the US, as has previously been observed for infants residing in Western contexts. However, longer peak look durations and slower shift rates were associated with better memory performance for infants in rural Malawi, the opposite pattern that has been observed in studies of infants residing Western contexts. Thus, these findings raise questions about the nature of non-Western infants' performance in this task, and how these patterns will change or be stable across development.

Other studies reveal some potential sources of differences in VPC performance between infants in Western and non-Western contexts. Kennedy et al. (2008) found that only 12 of the 69 6- to 8-month-old infants residing in Southern Ethiopia who completed the VPC procedure showed an age-appropriate novelty preference score. More importantly, when controlling for

age, they found that measures of physical health were related to infants' looking behavior. Specifically, they found that lower weight-for-age z-scores were associated with longer look durations, suggesting that less healthy infants had less efficient information processing. In addition, in a sample of 5- to 12-month-old infants residing in India, Rose et al. (1994) found that low-risk infants, based on measures of physical growth, showed the expected relation between age and familiarization duration on infants' memory in the VPC task; high-risk infants did not show the expected pattern. Thus, measures of physical health predicted some aspects of information processing in infants residing in non-Western contexts, and some of the differences between Western and non-Western infants may be due to differences in growth and health.

The purpose of this study was to examine performance on the VPC tasks longitudinally in a sample of infants between the ages of 6 to 15 months residing in rural Malawi. We assessed infants' performance on this task at two time points, approximately 6 months apart. Our first objective was to determine how these infants' VPC performance changes across development, with a particular interest in whether the direction of change is the same as has previously been observed for infants in Western contexts. Specifically, studies in Western contexts have shown dramatic changes from 3 and 12 months of age in how much familiarization time infants need to show visual recognition memory (Colombo et al., 1988b), the rate of shifting gaze between two simultaneously presented items (Rose et al., 2001), and the duration of individual looks (Colombo et al., 2001; Rose et al., 2001). Older infants require less familiarization time to show a novelty preference, they exhibit more switches between stimuli, and they have shorter individual looks. This has been interpreted to suggest that with age infants become faster at processing and learning new information.

In a previous analysis of the data used in the current study, Beckner et al. (submitted) observed different relations between these aspects of visual behavior in 6- to 9-month-old infants raised in rural Malawi. Because Beckner et al. only examined infants' performance at one time point, it is not immediately obvious what kinds of developmental changes we should observe in this task for these infants. One possibility is that with age performance of infants in the non-Western context becomes more like that of infants in Western contexts. Specifically, infants in rural Malawi may show patterns of behavior that look more like infants in the US as they get older (e.g., these infants' shift rates may increase and their peak look durations may decrease with age). An alternative possibility is that the pattern observed by Beckner et al. (submitted) reflects a qualitatively different approach to this task, and that developmental changes will reflect a strengthening of that approach. Specifically, infants' shift rates may decrease, and their peak look durations may increase with age (the opposite pattern that has been observed in Western contexts).

Our second objective was to examine the stability of these aspects of looking behavior by evaluating the test-retest reliability of infant VPC performance in a non-Western context. Reliability is a necessary but insufficient condition for predictive and discriminant validity (Downing, 2003), so establishing the reliabilty of measures of infant VPC performance in rural Malawi is an important first step towards determining whether these variables show the same predictive and discriminative validity that has been observed for infants in Western cultures. It is possible that these behaviors will show reliability, as has been observed in Western samples. This is the expected pattern if this task is a valid assessment of information processing across contexts, as has often been assumed. However, it is also possible that there is little relation

between how infants perform at the two ages, particularly if the strategy infants use to learn about the familiar stimulus shifts over time.

Our third objective is to determine whether the pattern of results previously observed by Beckner et al. (submitted) at age 6-9 months are observed at age 12-15 months. Specifically, we will examine whether longer peak look durations and slower shift rates predict more robust memory performance at the second visit, 6 months after the initial visit, where this pattern was previously observed for this sample . One possibility is that infants will show the same pattern of associations across both visits, providing additional evidence that the strategies that optimally support learning differ as a function of context. Another possibility is that infants will show the pattern of results we would expect from studies of infants in Western contexts on their second visit, but not their first visit, indicating a different developmental trajectory. Importantly, both patterns would suggest that infants in rural Malawi differ from infants in Western contexts, but yield different conclusions about the developmental process. Finally, because previous studies in non-Western contexts suggest that infant VPC performance is associated with infants' health status, we conducted an exploratory goal of examining the impact of physical health on each of our objectives. Thus, we also conducted our analyses including measures of health status (e.g., standardized length-for-age, hemoglobin) and scores on the Home Observation for Measurement of the Environment Inventory (HOME; Caldwell & Bradley, 2003).

We tested these objectives and our exploratory goals by analyzing the data from an eye-tracking VPC procedure in a sample of infants living in rural Malawi. Infants were tested once between 6 to 9 months of age and again between 12 to 15 months of age and measures of physical health were obtained on each visit. This design allowed us to assess whether infants displayed changes in VPC performance across visits (Objective 1), whether infants' VPC

performance showed stability (Objective 2), and whether infants showed the same pattern of results across both visits that Beckner et al. (submitted) previously observed for their first visit (Objective 3). In addition, this approach allowed us to examine how measures of physical health influenced infants' VPC performance and the relations between our measures of interest.

Method

Participants

Infants were recruited as part of a randomized controlled trial carried out in the rural Lungwena and Malindi areas of the Mangochi District, Malawi, from February 2018 to January 2019 (Caswell et al., 2020; Prado et al., 2020). Some of the results from this study have been reported elsewhere (Beckner et al., submitted; Prado et al., 2020). Beckner et al. (submitted) previously reported results from the 251 infants that contributed usable data in the VPC procedure on their first visit, but only examined infants' performance on this first time point. Of these 251 infants that contributed usable eye-tracking data on their first visit, 196 also contributed usable eye-tracking data on their second visit. Thus, the current study includes 196 of the 251 infants reported in Beckner et al. (101 were assigned to the intervention group and received an egg a day during the period between visit 1 and visit 2 and the and the other 95 were assigned to the control condition and did not receive any nutritional supplement). [1] The mean age of infants was 219.09 days on their first visit (*SD* = 38.24) and 400.22 days (*SD* = 38.54) on their second visit with a mean time elapsed of 181.13 days (*SD* = 3.38).

[1] Initial models included an interaction between intervention group and visit, but as was observed by Prado et al. (2020) the intervention had no influence on our variables of interest. Therefore, none of the models reported here include intervention group as a factor.

Table 1. Descriptive statistics for the full sample.

	Category	Sub-Category	Frequency (N)	Percent (%)
Malawi	Sex	Male	95	48.5
		Female	101	51.5
	Food Insecurity	None	45	4.08
		Mild	8	8.67
		Moderate	17	23.0
		Severe	126	64.3
	Maternal Education	Can Read and Write	101	51.5
		Cannot Read or Write	91	46.4
	Maternal Occupation	Work at Home	92	46.9
		Service Industry	64	32.7
		Fishing or Farming	37	18.9
	Tribal Affiliation	Yao	161	82.1
		Chewa or Other	33	16.8
		Did not Report	2	1.02

Apparatus

Eye-tracking data were collected using a Tobii Pro X2-60 with an external processing unit at a rate of 60 Hz and stimuli were presented on an HP EliteDisplay E222 21.5” monitor (1920 x 1080 resolution). The administration of the stimuli and collection of eye-tracking data were controlled by a single laptop. Infants were either tested using a Dell Latitude 5480 or Dell Latitude 7280 laptop computer. Each eye-tracking system was surrounded by four black curtains

hung from custom-build curtain rods. When the curtains were closed prior to the start of the eye-tracking study, only the monitor was visible to the mother and child, blocking out other potential distractions. The infant was placed on their parents' lap and was seated approximately 60 cm from the monitor. Mothers were instructed to look away during the study to mitigate any potential for them to direct their child's gaze to the screen. The child was monitored by the eye-tracking staff using a webcam on the laptop screen and mothers were instructed to turn away if they started watching the screen.

Stimuli

The experimental stimuli consisted of 8 face images (Figure 1) that were selected to approximate the faces that infants were likely to encounter in rural Malawi. Faces were selected from the database reported by Strohminger et al. (2016). Additional details about stimulus selection can be found in Prado et al. (2020). Faces were presented at a viewing distance of 60 cm and each face stimulus was approximately 8.97° by 12.72° (9.41 cm by 13.38 cm). Stimulus arrays were accompanied by classical music and immediately preceded by a flashing fixation cross (.65 Hz) that alternated with images of colorful toys.

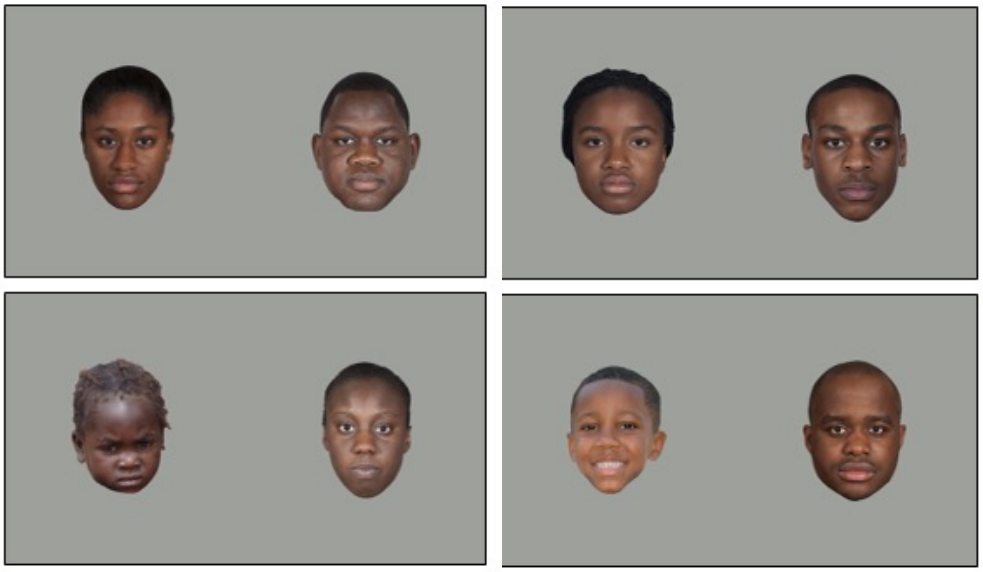

Figure 1. Stimulus pairings for the VPC procedure.

Procedure

All protocols were reviewed and approved by the Research Ethics Committee of the University of Malawi College of Medicine and the Institutional Board (IRB) at the University of California, Davis. All procedures were administered at a local hospital or clinic. Families were brought in for an initial assessment in which they filled out surveys, underwent physical health evaluations, and participated in an eye-tracking procedure. Length-for-age z-scores as well as 5-mL blood samples were collected as a part of this evaluation. Trained anthropometrists calculated child recumbent length to the nearest millimeter using a Holtain length board (Holtain Ltd.) and length-for-age z-scores were calculated using World Health Organization Growth Standards (Caswell et al., 2020). Hemoglobin concentrations were derived from 5-mL blood samples that were collected at the clinic site. Additionally, trained personnel visited participants homes within a week after their first visit and administered brief home environment assessments

to determine the quality and quantity of stimulation and support available to infants. Scores were calculated for infants Home Observation for Measurement of the Environment (HOME) using subscales including parent responsivity, acceptance, organization, learning materials, and involvement (Caswell et al., 2020). These measures assess specific environmental factors rather than socio-economic status and thus may provide a more proximal indicator of developmental risk factors (Caldwell & Bradley, 2003). Systematic analyses of HOME environment scores and infant physical health are reported in Caswell et al.

During the eye-tracking procedure families were brought into a separate room where the infant was seated on their parents' lap. A trained experimenter then initiated the eye-tracking study by presenting a five-point calibration and validation sequence in which infants were shown a smiley face and images of different objects. Calibration quality was assessed by visually inspecting the horizontal and vertical gaze estimates after validation. Calibration was repeated when at least one point was missing or when there was a high amount of noise as indicated by validation. Both the calibration and validation procedures were part of the Tobii Lab suite. After calibration and validation, infants were presented with four VPC problems interspersed with an unrelated attention cueing task. Each of the four memory problems consisted of a 20 s familiarization array and two 10 s test arrays interspersed with an attention getter (Figure 1). The subsequent stimulus array was shown when infants were judged to be fixating the attention getter in the center of the screen and the left-right position of the novel face was counterbalanced such that infants first saw the novel face on the left on half of the memory problems and the right on the other half of the memory problems.

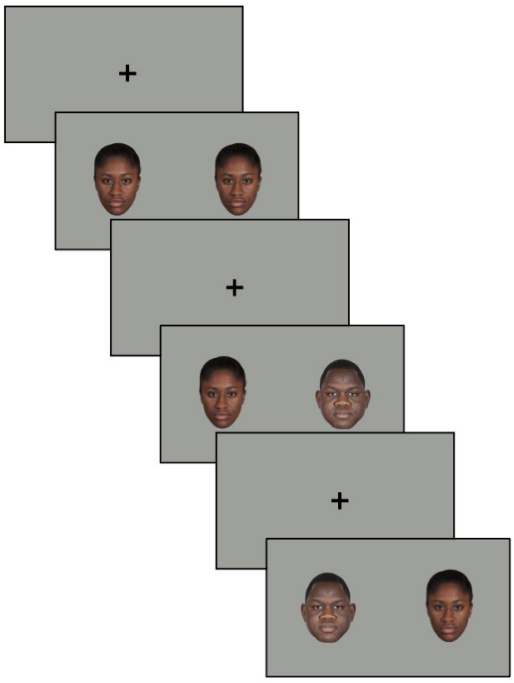

Figure 2. Schematic illustration of a single memory problem. The sequence consisted of a 20 s familiarization array containing two identical images of a face. Then, after a brief delay, infants were shown two 10 s test arrays in which the now-familiar face paired with a novel face.

Data processing

Data processing and statistical analysis was completed in SAS version 9.4 (North Carolina, United States) and R version 4.1.1 (Core Team & Others, 2013). Data were processed as reported in Beckner et al. (submitted). We derived three measures from the eye-tracking data: novelty preference scores, peak look duration, and shift rate. These three measures were used as our primary outcome measures for our statistical models. Novelty preference scores were calculated on each memory problem by dividing total fixation duration to the novel face by total fixation to both faces combined. Novelty preference scores were then averaged across the two test arrays to calculate a single novelty preference score on each individual memory problem.

Peak look duration was calculated by identifying the longest individual bout of *looking* during the familiarization array for each memory problem whereas shift rate was calculated from the number of times infants shifted their looking between the two items on the screen divided by their total look duration. Peak look durations were calculated during familiarization and shift rate was calculated during the familiarization and test arrays. Importantly, look durations were calculated by interpolating bouts of attention from the fixation data to approximate the results from studies of behavioral coding studies. Looks could include short interruptions in fixation that were less than 1 s. Thus, peak look duration and shift rate were calculated using bouts of attention, whereas novelty preference scores were calculated from the total duration of fixations. Descriptive statistics were reported for shift rates during familiarization and test, but analyses were conducted on infants' average shift rate across familiarization and test on each memory problem. This is the standard approach for calculating these measures as has been reported in previous studies (Rose et al., 2001).

Analysis approach

There were three objectives in the current study. The first objective was to assess whether infants' performance significantly improved across visits. To address this objective, we fit linear mixed-effect models on the trial-level data for each outcome measure. These models included a fixed effect of age (continuous), infant sex (categorical), trial (continuous), and visit (categorical). A random intercept for participant and a random slope for visit were also included in the models. We also conducted exploratory analyses including measures of physical health and home environment in models on each dependent variable. For each measure, we fit three additional models: one with a fixed effect of length-for-age z-scores (continuous) at each visit, one with a fixed effect of hemoglobin (continuous) at each visit, and one with a fixed effect of

HOME environment (continuous) on visit 1. Our second objective was to assess whether infants' performance was stable across visits. To address this objective, we performed Pearson correlations between infants' mean performance at visit 1 and visit 2 for each outcome measure. Our third objective was to examine whether infants show the same patterns of associations between information processing and memory performance at both visits. We examined this using linear mixed effect models. Specifically, we fit two linear mixed effect model with novelty preference as the outcome measure. These models included a fixed effect of age (continuous), infant sex (categorical), trial (continuous), and visit (categorical). The first model included a fixed effect of shift rate (continuous) and a shift-rate-by-visit interaction. The second model included a fixed effect of peak look duration (continuous) and a peak-look-by-visit interaction.

Results

Descriptive statistics

Infants contributed an average of 3.09 trials (*SD* = 1.05) at visit 1 and 3.10 (*SD* = 1.05) trials at visit 2. Descriptive statistics for our measures of interest are presented for each visit in Table 2 (note, the results for Visit 1 were previously reported for a larger sample including the 196 infants tested here in Beckner et al, submitted). Compared to Visit 1, infants' mean peak look duration during familiarization, mean shift rate during familiarization and test and total fixation duration during familiarization and test were numerically lower on Visit 2; infants' mean novelty preference scores did not change across visits. Beckner et al. (submitted) reported that the larger sample of infants preferred the novel face on Visit 1; we confirmed that the subsample used in this study also preferred the novel face on Visit 1, $t(195) = 9.148$, $p < .001$, $d = .65$. Comparison of novelty preference scores to chance on Visit 2 revealed that infants also significantly preferred the novel face at the second visit, $t(195) = 10.15$, $p < .001$, $d = .72$.

Table 2. Descriptive statistics for our measures of interest for each sample.

	Familiarization				Test			
	Visit 1		Visit 2		Visit 1		Visit 2	
Measure	M	SD	M	SD	M	SD	M	SD
Novelty preference	-	-	-	-	0.59*	0.13	.59*	.12
Total fixation duration (in seconds)	8.07	3.83	7.49	3.70	8.43	4.02	8.21	3.90
Peak look duration (in seconds)	3.87	1.48	3.59	1.57	-	-	-	-
Shift rate (shifts per second of looking)	0.11	0.07	0.13	0.10	0.14	0.10	0.11	0.08

*Significantly different from chance (.50), $p < .05$. The results from Visit 1 were reported in Beckner et al. (submitted).

Changes in information processing and memory performance across visits

To examine differences in performance between the two visits, we fit LMMs to the trial-level data for each measure of interest. This analysis examined changes across time (i.e., an effect of visit) after accounting for infant sex (categorical: 2 levels) and trial (continuous: 1-4). Each of these models also included a random intercept for participant to account for the nested structure of the data and a random slope of visit for each participant to account for intraindividual differences in the rate of change across visit. Likelihood ratio tests revealed that including a random slope term for visit resulted in a statistically significant improvement in model fit for the models examining shift rate and peak look duration, but not novelty preference. However, the random slope was included in all models to be consistent in how we accounted for variation in change across visits. The models for each outcome measure were identical except that for our model of novelty preference we centered infants' novelty preference scores by

chance (.50) to aid in intercept interpretation.

The model of novelty preference score revealed no significant fixed effects. Thus, even after accounting for the effect of trial and infant sex on memory performance, there was no evidence for changes in novelty preference scores from visit 1 to visit 2 (Figure 3). The intercept was significantly different from 0, $\beta = 0.074$, $SE = 0.01$, $t = 4.566$, $df = 483.60$, $p < .001$, providing corroborating evidence that infants showed a significant preference for the unfamiliar face, even after accounting for the effect of visit, sex, and trial.

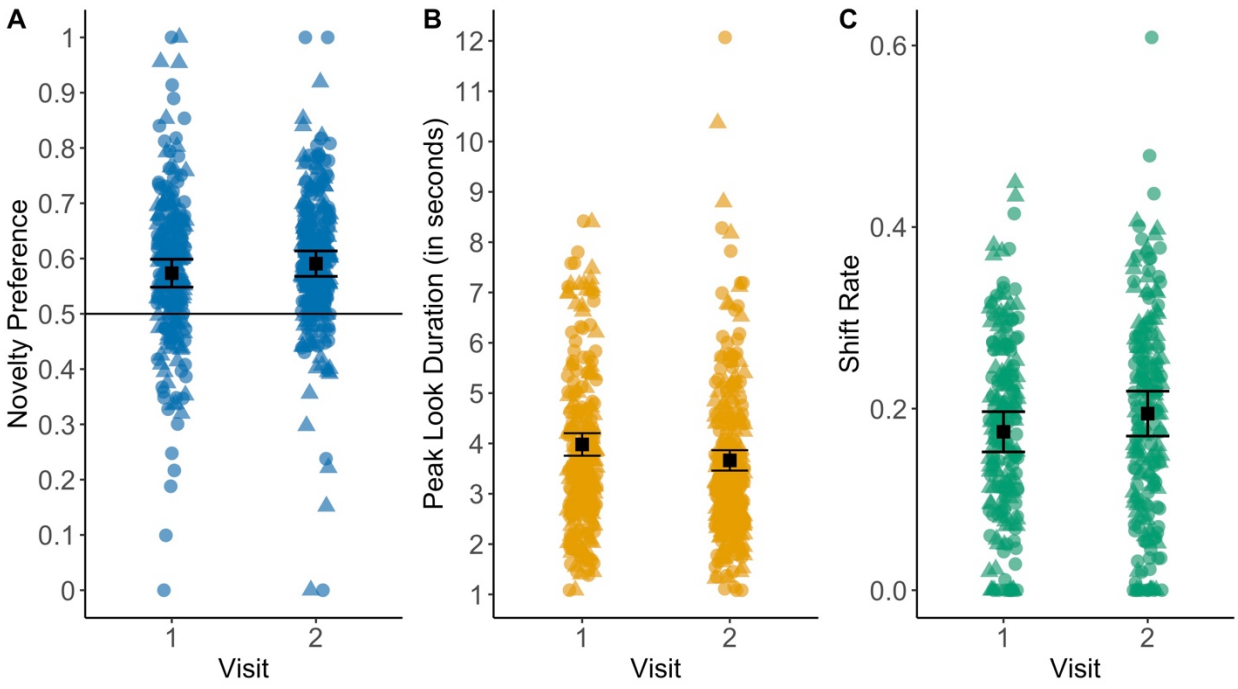

Figure 3. Infants' performance by visit. (A) Novelty preference score, (B), Peak look duration (in s), and (C) Shift rate (i.e., the number of shifts per s of accumulated looking). In each figure, the individual dots represent an individual infant's mean performance averaged across all four memory problems on one visit. The black squares represent the estimated marginal means from each LMM and the error bars surrounding these squares represent 95% confidence intervals for the estimated marginal means. Shape represents infants' biological sex (circle = females, triangle = males).

The model on shift rate revealed significant fixed effect of trial, β = -0.02, *SE* = 0.01, t = -4.27, df = 960.26, $p < .001$; collapsed across visits, infants' shift rate declined across trials. There were no other significant effects. The lack of an effect of visit means that, as was observed for novelty preference, there was no significant change from visit 1 to visit 2 in shift rate.

The LMM evaluating peak look duration, in contrast, did reveal a significant fixed effect of visit, β = -0.32, *SE* = 0.12, t = -2.474, df = 313.31, p = .005, indicating that peak look duration

changed from visit 1 to visit 2. Inspection of the means in Figure 3B shows that infants had a slightly shorter peak look duration in visit 2 than in visit 1, consistent with the assumption that infants' processing becoming more efficient over development. This model did not reveal any other significant effects. Together, these comparisons between the two visits indicate that infants peak look durations, but not their novelty preference or shift rate, changed across visits.

Assessing the stability and discriminative validity of information processing and memory performance across visits

To examine the stability of information processing and memory performance across visits, we conducted Pearson correlations on infants' mean scores for each measure of interest. The correlations between visit 1 and visit 2 were not significant for infants' mean novelty preference scores, $r(194) = -0.04$, $p = .63$, peak look duration, $r(194) = -0.06$, $p = .38$, or mean shift rate, $r(194) = 0.09$, $p = .22$ (Figure 4). Thus, there was little stability in these three measures.

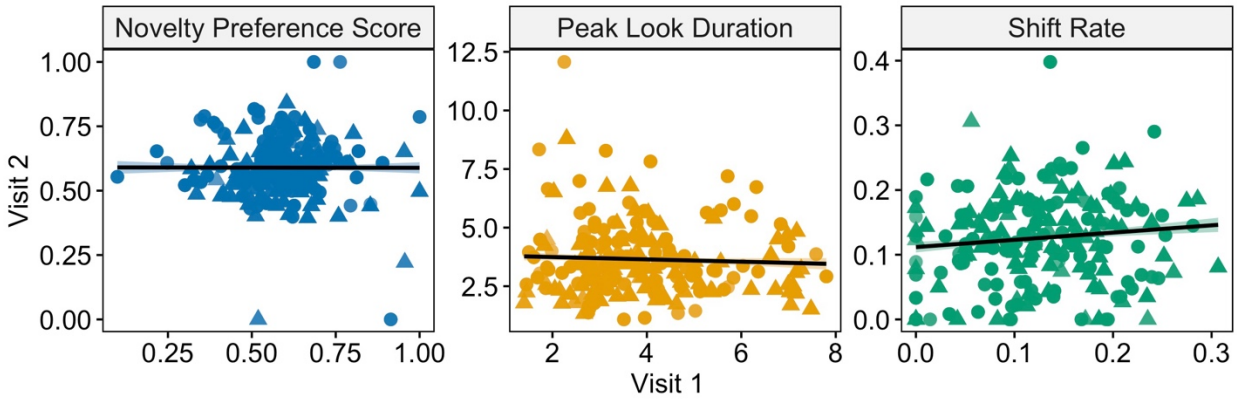

Figure 4. Scatterplots of the data for, visit 1 and visit 2 for Novelty preference (left), Peak look duration (middle), and Shift rate (right) Shift rate. For each figure, the x-axis represents infants' mean performance at visit 1 and the y-axis indicates infants' mean performance at visit 2. Each individual dot represents the score for a single infant across both visits. The regression lines indicate the association between infants' performance at visit 1 and visit 2 and the shading around the line represents 95% confidence intervals.

Similarities and differences in information processing and memory across visits

Next, we examined whether infants showed the same patterns of associations between information processing and memory performance across both visits. First, we examined the intercorrelations between our measures of interest for Visit 2. The intercorrelations between the measures at Visit 1 were reported by Beckner et al. (submitted) for the larger sample; the values provided here in Table 3 are for the subsample of 196 infants who were included in this study. Table 3 also reported the intercorrelations between each child's mean novelty preference score, mean total fixation duration, mean peak look duration, and mean shift rate. At both visits, total fixation duration was positively associated with shift rate and peak look duration.

Table 3. Intercorrelations (Pearson's r) among our measures of interest averaged across memory problems for each child and stratified by sample.

Visit	Measure	*N*	Novelty preference	Total fixation duration	Peak look
Visit 1	Novelty Preference	196	-	-	-
	Total look duration	196	-0.13	-	-
	Peak look duration	196	-0.06	0.39***	-
	Shift Rate	196	-0.12	0.59***	0.004
Visit 2	Novelty Preference	196	-	-	-
	Total look duration	196	0.04	-	-
	Peak look duration	196	-0.08	0.45***	-
	Shift Rate	196	0.00	0.60***	0.07

* $p < .05$, ** $p < .01$, *** $p < .001$

The results of these intercorrelations suggests that infants showed a similar pattern of associations between measures at the 2 visits. We next probed the relation between each information processing measure and novelty preference in separate LMMs; these models had novelty preference as the dependent variable, one of the information processing measures (e.g., peak look duration and shift rate) as a fixed effect, as well as fixed effects of visit, infant sex, and trial. These models also included a random intercept for participant and a random slope for visit.

The model including a fixed effect of shift rate and a shift rate-by-visit interaction did not reveal any significant fixed effects, and only the expected outcome that the intercept was significantly different from 0, $\beta = .11$, $SE = 0.02$, $t = 6.31$, $df = 803.60$, $p < .001$. Table 4 summarizes the results from this model.

Table 4. Results from LMMs evaluating the effect of shift rate and the interaction between shift rate and visit on novelty preference scores.

Parameter	Estimate	standard error	t-value	p-value
Intercept	0.11	0.02	6.30	**<0.001**
Sex	-0.02	0.01	-1.44	0.149
Trial	-0.00	0.00	-0.33	0.743
Visit	-0.00	0.02	-0.15	0.880
Shift Rate	-0.08	0.05	-1.51	0.130
Visit-by-Shift Rate Interaction	0.05	0.07	0.67	0.501
Random Effects				
σ^2	0.03			
τ_{00}	0.00			
τ_{11}	0.01			
ρ_{01}	-0.81			
ICC	0.11			
N	196			
Observations	1213			
Marginal R^2 / Conditional R^2	0.005 / 0.111			

The model including the fixed effect of peak look duration and a peak look duration-by-visit interaction (Table 5) revealed not only that the intercept was significantly different from 0,

β = .04, *SE* = 0.02, *t* = 2.218 *df* = 736.90, $p < .05$, but also revealed significant fixed effects of visit, β = -.06, *SE* = 0.02, *t* = -2.792, *df* = 772.20, $p < .01$, and peak look duration, β = .01, *SE* = 0.003, *t* = 2.704, *df* = 779.90, $p < .01$, as well as a significant interaction between visit and peak look duration, β = 0.01, *SE* = 0.004, *t* = 2.951, *df* = 1161.0, $p < .01$. This interaction indicates that the association between peak look duration and novelty preference scores differed for the two visits (Figure 5). Specifically, as previously reported by Beckner et al. (submitted), more robust memory (i.e., higher novelty preference scores) were associated with longer peak look durations at visit 1. However, at visit 2, more robust memory was associated with shorter peak look durations. Table 5 summarizes the results from this model.

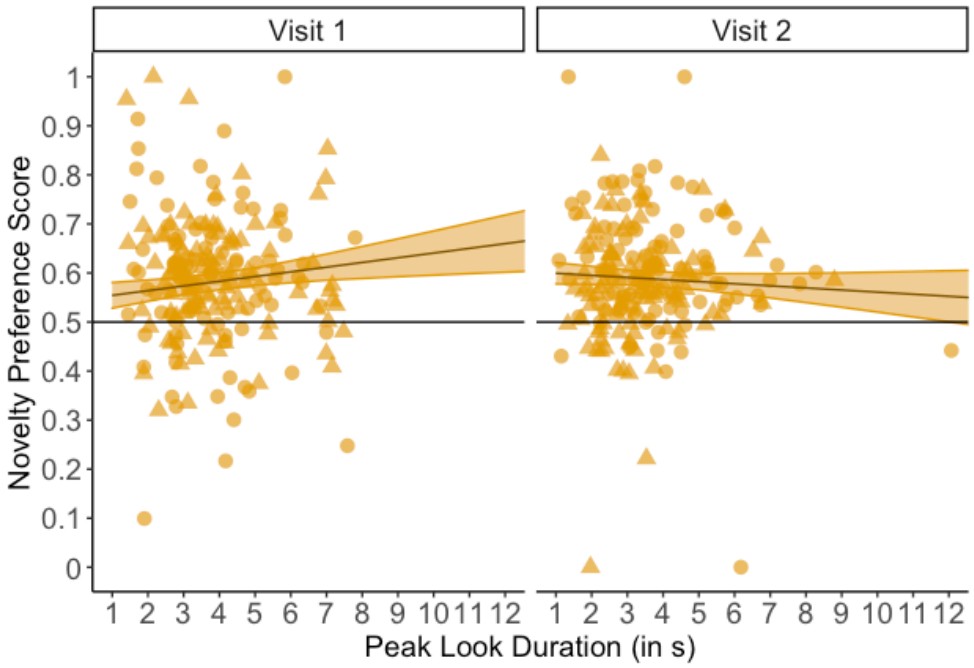

Figure 5. Association between novelty preference score and peak look duration as a function of visit. The horizontal line bisecting the vertical axis represents chance (.50) performance. Individual data points represent infants' mean novelty preference scores averaged across all memory problems for infants at visit 1 (left) and visit 2 (right). The lines represent estimated marginal means and the shading around the lines represent 95% confidence intervals of the estimated marginal means. Shape represents infants' biological sex (circle = females, triangle = males).

Table 5. Results from LMMs evaluating the effect of peak look duration and the interaction between peak look duration and visit on novelty preference scores.

Parameter	Estimate	Standard Error	t-value	p-value
Intercept	0.06	0.02	2.87	**0.004**
Sex	-0.02	0.01	-1.72	0.086
Trial	-0.00	0.00	-0.19	0.853
Visit	0.06	0.02	2.81	**0.005**
Peak Look Duration	0.01	0.00	2.72	**0.007**
Visit-by-Peak-Look Duration Interaction	-0.01	0.00	-2.98	**0.003**
Random Effects				
σ^2	0.03			
τ_{00}	0.00			
τ_{11}	0.01			
ρ_{01}	-0.82			
ICC	0.11			
N	196			
Observations	1213			
Marginal R^2 / Conditional R^2	0.011 / 0.123			

The effect of HOME and health status on infants' VPC performance

Finally, we assessed the influence of HOME environment and physical health on infants' VPC performance. Because previous studies suggested that the pattern typically observed in Western samples was often observed in healthier infants in non-Western samples, we conducted our analyses including a measure of physical health, specifically, length-for-age z-scores and hemoglobin measured at each visit. We used length-for-age because previous studies of infants in non-Western contexts have revealed an effect of this variable on measures of information processing and memory performance in the VPC procedure (Kennedy et al., 2008; Rose, 1994).

Hemoglobin measures on each visit were included as a marker of nutritional status. In addition, environmental risk factors were assessed using infants' HOME environment scores to assess whether environmental risk factors affected infants' VPC performance.

Three separate LMMs were conducted for each dependent variable (novelty preference, peak look, and switch rate). Each model included the same fixed effects of age (continuous), infant sex (categorical), trial (continuous), visit (categorical) as our models reported earlier. All models included a random slope for visit and a random intercept for participant. For each dependent variable, we fit one model that included length-for-age, one model that included hemoglobin, and one model that included HOME scores. None of the models including hemoglobin or HOME revealed any significant fixed effect of hemoglobin or HOME scores on any of our measures of interest (see supplemental materials).

Our models that included length-for-age z-scores, in contrast, yielded a different pattern. For peak look duration, but not novelty preference scores or shift rate, there was a significant fixed effect of length-for-age z-scores on infants' performance, β = -.035, SE = 0.017, t = -2.018, df = 196.10, p = .04, indicating that infants with higher length-for-age z-scores had shorter peak look durations. Figure 6 displays the estimated marginal means for infants' length-for-age z-scores from each LMM and infants' mean scores for each measure.

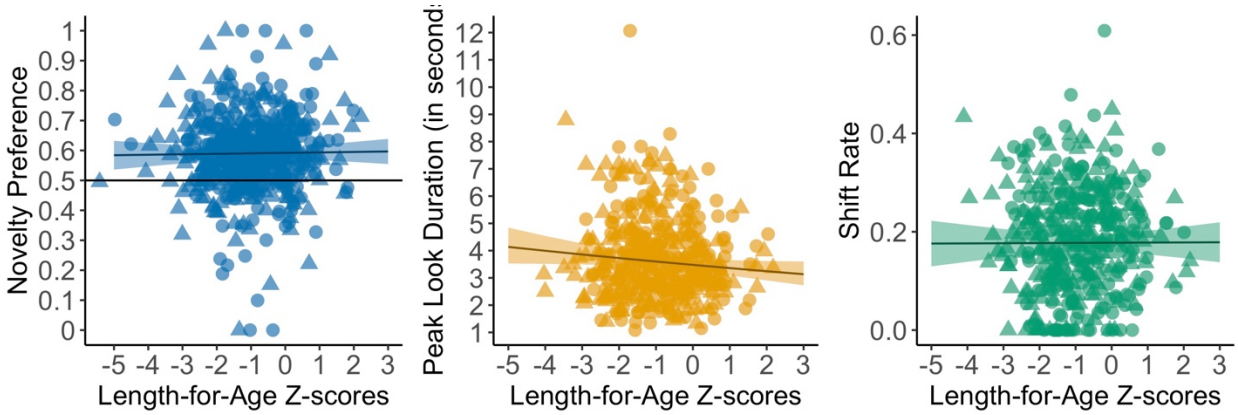

Figure 6. Infants' Novelty preference (left), Peak look duration (middle), and Shift rate (right) Shift rate as a function of their length-for-age z-scores. Individual dots represent infants' mean performance averaged across all memory problems. The lines represent the estimated marginal means from each LMM and the shading surrounding these lines these squares represent 95 % confidence intervals for the estimated marginal means.

Next, we examined whether the association between measures of information processing and memory performance differed as a function of length-for-age. We examined this by fitting two LMMs with novelty preference as our outcome measure, each including a different information processing measure (peak look or shift rate), and the fixed effects, interactions, and random effects described earlier. We also included a fixed effect of length-for-age z-score and an interaction between length-for-age z-score and the information processing measure being tested (peak look duration or shift rate). These models did not reveal a significant interaction between length-for-age z-scores for shift rate, β = -0.01, *SE* = 0.03, *t* = -0.40, *df* = 1155.0, *p* = .507, or peak look duration, β = .002, *SE* = 0.002, *t* = 1.005, *df* = 1161.0, *p* = .315).

Discussion

In this study, we assessed information processing and memory performance longitudinally in a sample of infants residing in rural Malawi and observed similarities and differences in VPC performance across visits. We had three objectives. First, we examined change in performance across visits. We observed only changes in peak look durations across visits, with the duration of infants' peak look becoming shorter, on average, over time. This pattern of decreasing peak look durations with increasing age is consistent with findings from studies of infants residing in both Western (Axia et al., 1999; Colombo & Mitchell, 1990; Colombo et al., 1988; Rose et al., 2001; Ruff, 1975) and non-Western contexts (Rose, 1994). In general, decreasing peak look duration has been interpreted as indicating that infants are becoming faster or more efficient at processing information with age (for review, see Rose et al., 2004).

Our second objective was to examine stability in our measures across visits, as has been observed in Western samples. However, we did not observed stability for any of our measures of interest, reflecting the fact that infants' visit 1 performance did not predict their visit 2 performance for any of our measures. This pattern contrasts with previous studies that have shown moderate stability and test-retest reliability in the VPC procedure across a test-retest interval of 3 months (Colombo et al., 1988b; Rose, Feldman, & Wallace, 1988a). There are at least two explanations for this pattern of results. First, most of the studies establishing the psychometric properties of the VPC procedure have been conducted with infants residing in Western contexts. Thus, it is not clear whether those findings should generalize to infants residing in rural Malawi. Beckner et al. (submitted) found that at their first visit, a larger sample of infants (from which the current sample was drawn) showed the opposite relation between peak

look and novelty preference as has been observed in Western samples. It is possible that this approach was transitory and not stable, reflecting either something about how these young infants learned to visually investigate their world or the general novelty of the experimental setting and procedure.

The second explanation is that the psychometric properties of the VPC task do not lend themselves to stability as assessed here. Even though the VPC procedure has been used for decades to yield insights about information processing and memory performance in infants, there is some controversy about the psychometric properties of the VPC task itself. Much of these criticisms center around issues related to internal consistency, lack of standardization, and inconsistent patterns of predictive validity for the Fagan Test of Infant Intelligence (the commercially available version of the VPC procedure) in particular (Benasich & Bejar, 1992; DiLalla et al., 1990; Tasbihsazan et al., 2003). However, these observations highlight the possibility that the VPC task may not consistently and reliably index aspects of information processing and memory performance in infants. These discrepancies may, in part, reflect the fact that whether infants show evidence of memory in these kinds of procedures is profoundly influenced by item distinctiveness (Oakes et al., 2009), complexity (Cohen et al., 1975), and stimulus categories (e.g., whether faces, abstract patterns, simple objects, or other stimuli are used; Rose et al., 2001). Moreover, some studies have implemented an infant-controlled procedure (Hunter et al., 1983; Rose, 1983) whereas other studies have presented familiarization arrays for fixed durations (Fagan, 1970, 1971, 1972), making it difficult to determine what version of the procedure is best suited for revealing stability in information processing and memory performance. The point is that stability is not always observed for infants VPC performance and there are several potential explanations for why this might be the case.

Our third objective was to determine whether the relation between information processing and novelty preference previously reported by Beckner et al. (submitted) for these infants on their first visit would be observed at their second visit. The results of the current study suggest that the association between peak look duration and memory performance was qualitatively different for infants on their first and second visits. That is, although for these infants longer peak look durations were associated with more robust memory on their first visit, we observed here that on infants' second visit shorter peak look durations were associated with more robust memory. Interestingly, infants' performance on their first visit contrasted with the pattern of associations between peak look duration and memory performance reported in previous studies, whereas their performance on their second visit were consistent with those previous studies (Colombo et al., 2001; Mary L. Courage et al., 2006; Rose et al., 2001).

Alternatively, the shift observed across visits in this task might reflect a distinct trajectory that is qualitatively different from the development that is observed in Western contexts. Specifically, infants in rural Malawi may adopt different attentional approaches to learn about visual stimuli during different periods of development. The developmental trajectory observed in Western samples is generally consistent with the conclusion that infants become increasingly efficient at processing information (see Rose et al., 2004, for review), but the developmental trajectory in this Malawi sample may reflect a shift from *exploration* at the younger age to *exploitation* at the older age or vice versa. When presented with complex or unfamiliar environments, organisms adaptively prioritize novel information (e.g., explore) and spend more time gathering information before moving on to other items (D'Souza & D'Souza, 2021). During their first visit, infants may have been more biased to explore, and as a result longer peak look durations were associated with more robust learning on their first visit. Later, when these infants

were 12 to 15 months, they may have been more biased to an exploitation mode and as a result shorter peak look durations were associated with more robust learning on that visit. In other words, different attentional strategies may have been optimal at different points in development for infants in rural Malawi. This kind of attentional difference has been observed as a function of language experience, with infants exposed to bilingual contexts showing more exploration and infants in monolingual contexts showing more in depth exploitation (D'Souza & D'Souza, 2021). Similarly, children who are behaviorally inhibited seem to exploit more than explore (Pérez-Edgar, 2018). Thus, increasingly there is a recognition that this difference in attentional approaches reflects differences in infants' experiences. Understanding the nature of the developmental trajectory observed in these infants being raised in rural Malawi should be the focus of future research.

Exploratory analyses evaluated the influence of physical health and home environment on infants' VPC performance. These analyses revealed only that infants with higher length-for-age z-scores displayed shorter peak look durations. These findings are generally consistent with other literature showing that in non-Western samples the more typical, Western patterns are observed in healthier infants (Rose, 1994). For example, Rose et al. (1994) observed a decrease in peak look duration in a sample of infants residing in India, but only in relatively healthy infants. Length-for-age z-scores are an indicator of growth faltering and are consistently associated with poor cognitive development (Prado et al., 2020), in part because length-for-age z-scores covary with environmental risk factors that impact both physical and cognitive development. Thus, these findings suggest that healthier infants, as indicated by their length-for-age z-scores, show shorter peak look durations. Importantly, the findings from the current study suggest that peak look duration may be similarly affected by risk factors in Western and non-

Western contexts and thus may index similar cognitive processes across cultures and contexts, but it's difficult to isolate what specific environmental factors are driving this effect. The fact that peak look duration (1) declined with increasing age and length-for-age z-scores and (2) predicted infants' learning and memory suggests that peak look duration may be a more proximal or direct indicator of information processing than measures of shift rate or novelty preference for infants in this non-Western context.

Taken together, the current study indicates that at least one aspect of infants' VPC performance, specifically, peak look as an indicator of more efficient information processing, is consistent across different cultural contexts. These results also extend findings from Beckner et al. (submitted) and indicate that some aspects of information processing may be consistent across cultural contexts, even though no stability was observed for any of our measures of interest over a test-retest interval of six months. The lack of stability that we observed is noteworthy because our sample was large and heterogeneous, providing the necessary preconditions for observing test-retest correlations. However, given that all infants resided in a low-resource environment with low maternal literacy, high food insecurity, and other risk factors, it is also possible that there was not enough variability in these kinds of experiences to yield reliable individual differences. Thus, these findings raise important questions about the psychometric properties of the VPC procedure in understudied cultural contexts. The current study extends work from Beckner et al. (submitted) establishing the "value-added" of conducting research in non-Western contexts and highlights nuances in the interpretation of measures of information processing and memory performance in the VPC procedure.

Chapter IV. Disambiguating the influence of exposure time and total look duration on infant visual recognition memory in an online study

Note: This is a draft of a manuscript that will be submitted for publication after input from coauthors.

Introduction

Decades of research using the visual paired comparison (VPC) procedure has revealed that infants' learning and subsequent memory for attended items are influenced by numerous factors including infant age and stimulus complexity (Colombo et al., 1988b; Fagan, 1970, 1971, 1974; Richards, 1997; Rose, 1983; Rose et al., 1982). In the VPC procedure, infants are first shown one or more images during an initial familiarization array. Then, after a brief delay, a test array is shown containing a now-familiar item paired with a novel item. This procedure takes advantage of infants' differential responsivity to novel and familiar stimuli to reveal insights about a fundamental form of memory called visual recognition memory (for review, see Rose et al., 2004). Infants' visual recognition memory is inferred from their visual preference for the novel and familiar stimuli presented in the test array.

Factors such as stimulus complexity and the time available to learn about attended items influences whether infants will show a novelty preference or a familiarity preference (Hunter et al., 1983). Specifically, infants show a familiarity preference presumably when they have not fully encoded the attended item in memory and a novelty preference presumably when they have formed a robust memory representation (Hunter et al., 1983). This view is supported by the fact that infants require more time to learn about complex stimuli compared to simple stimuli and

younger infants require more time to learn than younger infants (Fagan, 1970, 1971, 1972, 1974; Hunter et al., 1983). Given the same amount of familiarization time with a stimulus, younger infants will show a familiarity preference and older infants will show a novelty preference (Rose et al., 1982), or the same amount of familiarization with more and less complex stimuli will yield a familiarity preference with the more complex stimulus and a novelty preference with the less complex stimulus (Hunter et al., 1983). Findings like this have led to the conclusion that the duration of familiarization time is directly related to how much infants process or encode the stimulus.

However, this literature has left unanswered key questions about how familiarization time is related to learning and memory. In particular, the literature has not clearly delineated the effects of *accumulated looking* and *exposure* times. In some studies, the amount of time infants spend actually looking at the stimulus is manipulated. In these studies, infants show more robust novelty preferences when they have accumulated longer total look durations to the familiarization array (Hunter et al., 1983; Rose, 1983). In other studies, infants' exposure to the stimulus is manipulated, with no effort to account for how much infants look. In these studies, infants show more robust novelty preferences when exposure to the familiarization array (i.e., trials) are longer durations (Fagan, 1970, 1971, 1972, 1974). Thus, this literature does not allow us to disambiguate the contributions of accumulated looking and exposure durations on infant visual recognition memory. Moreover, because studies often conflate total look duration and exposure time, it's unclear whether one of these factors drives the pattern of results observed for the other (e.g., longer exposure times provide infants with opportunities to accumulate longer durations of looking during familiarization).

The time infants spend overtly attending to the familiar stimulus may be critically important in determining what infants encode in memory. Specifically, infants learn by looking. This view derives from the perspective that looking time directly reflects information processing time (Cohen, 1972; Cohen et al., 2002, 1975). Cohen (1998) argued that variation in how long infants look at different items reflects individual differences in how efficiently they process information. Several lines of evidence support this idea. Where infants allocate their gaze, what features they choose to look at, and how long they sustain their attention towards specific items or features profoundly shapes what they learn (Colombo, 2001; Oakes & Rakison, 2019). Infants at risk for deficits in cognitive development tend to have longer look durations than infants not at risk and look durations decrease with age, presumably reflecting infants' ability to process information more quickly with age (for review, see Rose et al., 2004). Infants who exhibit shorter looks appear to be better at rapidly processing and encoding information, as evidenced by the fact that they have more robust memory for attended items than do infants who exhibit longer looks (Colombo et al., 1995; Rose et al., 2001). Thus, how long infants actually look at a stimulus may be the determining factor in whether they show a novelty or familiarity preference.

However, these differences in looking time are confounded with exposure time. Infants who exhibit longer look durations necessarily also have longer exposure times, making it impossible to rule out whether the differences observed are a function of differences in *looking* or differences in *exposure* regardless of the actual amount of time looking at the stimulus. Several aspects of human memory yield more robust memories simply due to the passage of time (Ricker & Cowan, 2014; Squire et al., 2015). Specifically, with the passage of time memories become consolidated and more resistant to interference (Dewar et al., 2009; Dudai, 2004; Nairne, 2002). It is therefore possible that infants who accumulate more looking time during a visual

recognition memory task are benefitting from the opportunity to engage in those processes and that the determining factor in whether infants show a novelty or familiarity preference is the passage of time, regardless of how much infants are engaged in actively looking during the familiarization period.

There is some evidence that the amount of time after initial exposure to a stimulus contributes to infants' storage and consolidation of a memory representation of that stimulus, independent from how long they look at the familiarized item (Catherwood et al., 1996; Oakes & Kovack-Lesh, 2013). For example, Catherwood et al. (1996) familiarized infants with a series of visual stimuli and assessed how the delay between these image sequences impacted their visual recognition memory. Specifically, rather than manipulating how long the stimuli were visible, Catherwood et al. manipulated the time elapsed between stimulus presentations, thus allowing them to examine the effect of the passage of time on memory independent of infants' actually looking at the stimuli. Increasing the time between familiarization trials resulted in more robust memory for difficult stimuli. Thus, even though infants in these conditions did not look more at the stimuli, they showed better memory for those stimuli, indicating that the extra time to consolidate memory played a role. This suggests that processes involved in learning occurred even while infants were not looking at the stimulus during familiarization.

The passage of time may be particularly important when to-be-learned items are highly similar or difficult to differentiate, as this may induce infants to form a memory representation for a particular category or class of stimulus rather than any single item that belongs to a given category. Oakes and Kovack-Lesh (2013) observed that exposure time influenced memory for categorically related items above and beyond the influence of look duration. They familiarized infants to a series 6 trials, each with images of cats or dogs. They then presented test trials in

which one of images presented during familiarization was paired with a novel item from the same category. When test trials were presented immediately after the sixth familiarization trial, infants showed memory for the items presented during the fourth and fifth trials, but not for items presented during the just seen sixth trial. However, when a brief delay was imposed between the sixth familiarization trial and the test trials, infants showed memory for the items presented on the fifth and sixth trials. Thus, the time between familiarization and test was the key variable that determined whether infants showed memory for the familiarized items. These findings indicate that infants may require more exposure time to form and consolidate a memory representation particularly when the memory task is more challenging.

The goal of this study was to systematically examine how exposure time and total look duration are related to infants' memory performance. We expected that familiarization time and exposure time would be related to novelty preference, as has been observed in previous studies. By investigating them in the same study, however, we can determine the relative contribution of each. For example, exposure time may be associated with memory for attended items because longer exposure times provide infants with more time to form and consolidate an internal representation of an item. Alternatively, exposure time may be associated with more robust memory because longer exposures provide infants with more opportunities to accumulate looking towards an attended item. In the current study, we examined the contribution of exposure time and total look duration on infant visual recognition memory using the online platform Lookit (lookit.mit.edu). Our first objective was to determine whether infants would show evidence of memory in an online context as has been observed in more traditional laboratory settings. To date, no published studies have assessed infants' VPC performance using an online context, making the present work methodologically novel. Our second and main objective was to

determine how exposure time versus familiarization time uniquely related to infants' learning and subsequent memory. We predicted that longer exposure times and longer total look durations would be be associated with more robust memory.

Method

Participants

The data presented here are a preliminary analysis conducted on a subsample of the data that will ultimately be collected for this project. The sample used in these analyses was 52 infants (mean age = 240.0 days, range from 180 days to 300 days; 27 girls; 25 boys). The ultimate sample size for this study will be 100 infants. Our final target sample size of 100 infants was determined by gathering effect sizes from eye-tracking data were collected while infants viewed these VPC problems in the laboratory (Beckner et al., under review) and conducting power analyses in R using these effect sizes. Beckner et al. observed effect sizes ranging from .61 to .81 (Cohen's d) for one-sample t-tests against chance (.50). Approximately 45 infants are necessary for sufficient power (80 %) to detect a novelty preference score that is significantly different from chance, but 100 infants are necessary to detect a small to medium effect size for a within-subject manipulation with 80 % power. Because it's difficult to estimate the unique contributions of exposure time, age, and total look duration on infants' novelty preference scores we were unable to conduct systematic power analyses, but one 100 infants provides 80 % power for detecting correlations of .28 between measures such as age or total look duration and novelty preference scores. These results should be considered preliminary until the fully powered study is complete.

To achieve the current sample of 52 infants, and additional 10 infants were tested but excluded from the analyses for failing to provide usable data due to their eyes not being visible

in the webcam recording (n = 1), parent and/or sibling interference (n = 6), fussiness (n = 1), and equipment malfunctions (n = 3). Of the 52 infants included in the current study, 43 were White, 3 were Asian, 5 were more than one race, and 1 parent did not report. All primary caregivers had at least a high school diploma, 49 had at least a bachelor's degree, and 36 had a graduate or professional degree and primary caregivers had a median individual annual income of $160,000.

Families were recruited through multiple sources. First, information about the current study was posted on social media such as Facebook, Instagram, and Twitter as well on outreach websites such as Children Helping Science (https://childrenhelpingscience.com/). Second, parents were recruited through our subject pool at the Center for Mind and Brain. Specifically, we obtained infant names from the California State Office of Vital Records and parents who lived locally were sent flyers with information about our studies and how to participate. Parents who expressed interest in participation were then contacted for this specific study when their infant approached the appropriate age. Third, parents with an existing account on Lookit who expressed interest in study participation were recruited directly through the platform and contacted via email when their infants became eligible to participate in our study.

Stimuli

The experimental stimuli consisted of 8 faces that were either White or racially ambiguous (Figure 1) that were selected from the Child Affective Facial Expression (CAFE) database (LoBue & Thrasher, 2014) and the Face Research Lab London Set (DeBruine & Jones, 2017). Child and adult faces were selected to increase the distinctiveness of each individual face and to approximate the kinds of face stimuli that have previously been used in the VPC procedure (for an example, see Rose et al., 2001). Faces were isolated from their torso and background using Adobe Photoshop (CA, United States). Face pairs were presented on a white

background and accompanied by classical music. Because participants viewed the stimuli on their own computer monitors that varied in size, it is impossible to calculate precise stimulus size or visual angle subtended for each infant. However, each stimulus occupied approximately 20% of the display and the stimulus pairs were separated by a distance that was approximately 30% of the total display.

We also had audiovisual stimuli that were interspersed in-between stimulus arrays. These audiovisual stimuli were animated characters that included Blue from Blue's Clues, Kermit from Sesame Street, Brobee from Yo Gabba Gabba, Tigger from Winnie-the-Pooh, and Elmo from Sesame Street. Elmo was displayed during a calibration sequence and the remaining four audiovisual stimuli were each presented as attention getters for each of the four pairs of faces. Each audiovisual stimulus consisted of an animated image of a given character and a 3-s audio segment of that character speaking. Animations were created for each image using Keynote such that the onset of the image initiated both the animation and the audio segment for a given character. Different animations were used for each character, but the onset of the animations as well as the audio segment occurred as soon as the image was presented, and all audiovisual stimuli were 3 s in duration.

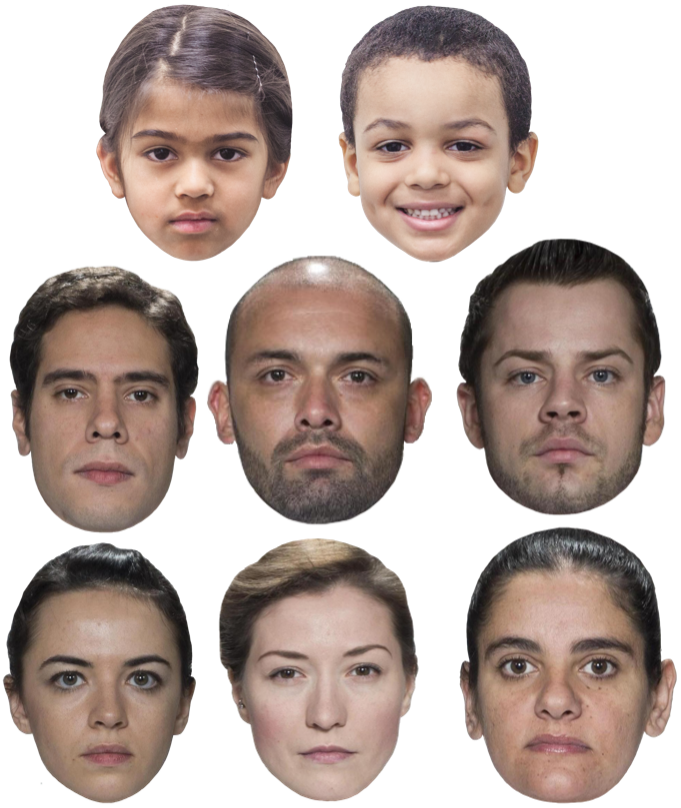

Figure 1. Face images that were used in the VPC procedure.

Procedure

All protocols were reviewed and approved by the Institutional Review Board (IRB) at the University of California, Davis. Sessions were administered online using Lookit (lookit.mit.edu; Scott et al., 2017; Scott & Schulz, 2017). Parents created accounts on the platform and provided basic demographic information. When they were ready to participate, parents logged into their accounts and selected this study. Before the experiment began, parents were asked to confirm that they were currently residing in the United States (due to restrictions from our IRB and funding agency, only infants in the United States were eligible), given instructions in both written and video form about the study and what participation involved, instructed how to configure their webcam and microphone, and asked to provide video consent. Once they

confirmed that they were ready to start the study, they were instructed to initiate the experimental stimuli by pressing a key on their computer and then immediately close their eyes. Once the parent indicated that the infant was ready for the experiment, the four trials began without interruption. Parents could pause the experiment at any time by pressing the spacebar, but we excluded any trial for which the spacebar was pressed during the experimental trial sequence.

The experimental procedure began with a brief 6 s calibration sequence in which a single audiovisual stimulus (e.g., Elmo) was presented first on the left side of the screen for 3 s, then disappeared, and then reappeared on the right side of the screen for 3 s. The purpose of this sequence was to provide a reference point for human observers to reliably code where infants were looking when there was only a single stimulus on the screen.

Immediately after calibration, infants received four memory trials, each involving a different set of faces and attention getter animations. Each trial consisted of an attention getter presented at the center of the screen for 3 s, a familiarization array in which two identical images of a face were presented side-by-side, and two 10-s test arrays in which the now-familiar face was paired with a novel face. The two test arrays differed only in the left-right position of the novel and familiar face to control for side biases. To draw infants' attention to the center of the array an animated cartoon character (accompanied by sounds) presented for 3 s in the center of the screen immediately before each familiarization and test array (see Figure 2).

Familiarization array duration was manipulated (10-s and 20-s) and infants received two trials of each duration. All infants received the same face pairings in the same order, so stimulus set 1 (Figure 2A) was always presented on the first trial, stimulus set 2 (Figure 2B) was always presented on the second trial, and so on. Each face pairing was also accompanied by a specific audiovisual stimulus that was used for the attention getters immediately preceding our stimulus

arrays (see Figure 2). The attention getter was played for 3-seconds before advancing to the next stimulus array.

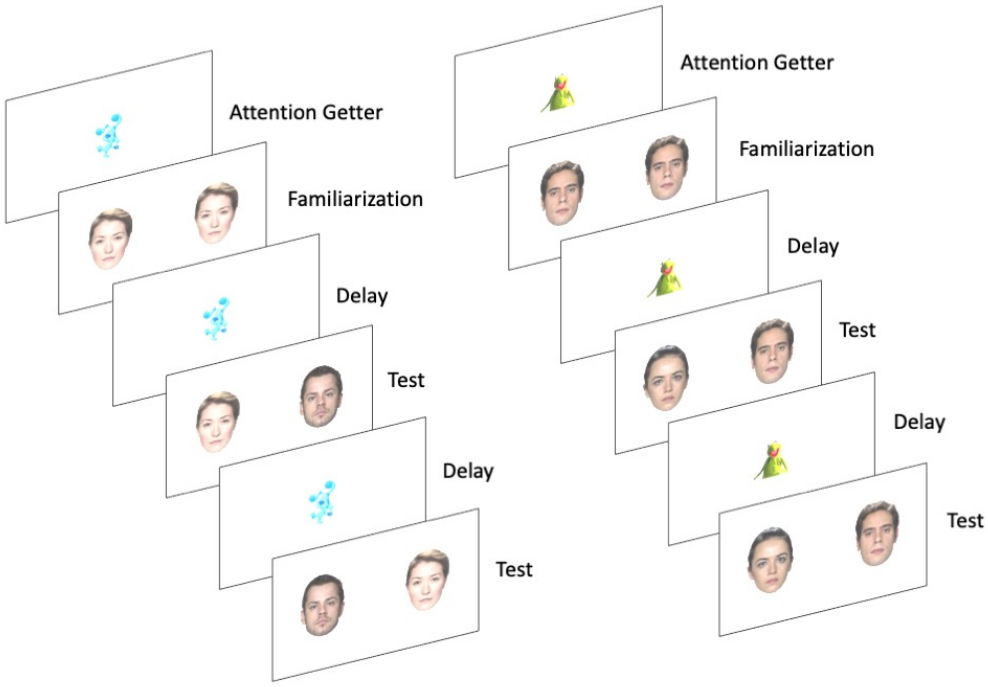

Figure 2. Schematic illustration of the first two experimental trials in the current study (left = trial one, right = trial two).

However, whether the first trial was a 10-s or 20-s familiarization duration varied between infants. The two durations alternated, so infants received trials in one of two orders (10-s, 20-s, 10-s, 20-s, or 20-s 10-s, 20-s, 10-s). The order was randomly determined for each infant by Lookit, and 34 infants received stimulus set 1 as a 10-s familiarization duration on trial 1 and 18 infants received that same stimulus set as a 20-s familiarization duration on trial 1. As a result, although the stimulus pairings and order of stimulus presentation were the same for each infant, the duration of familiarization for any given stimulus pairing varied across infants.

After the last trial, parents were instructed to open their eyes, and they could see a debriefing screen with the full stimulus set and a more detailed explanation of the purpose of the study. Parents were then given the option of withdrawing their data from the study as well as authorizing their data to be uploaded on Databrary (Databrary.org).

Coding

Looking behavior was coded offline using Datavyu (https://datavyu.org/). Across our 52 infants, there were 234 trials to be coded (ranging from 1 to 4 trials per infant). Each trial for each infant was independently coded by two trained coders. Coders first watched the trial at regular speed to get a sense for the infant's behavior. They then used the jog tool to move through the trial from start to finish, identifying periods in which the infant's gaze position remained within a given region of the monitor (i.e., center, left, and right) for at least three consecutive frames. The first of these consecutive frames was identified as the onset of a bout of looking. The offset of each bout of looking was identified as the frame immediately prior to an eye movement away from that region (e.g., left, or right of the monitor). Coders also took notes and recorded any instances of parental and/or sibling interference, equipment malfunction, or other data quality concerns.

Inter-rater reliability was assessed for each phase of each trial (e.g., familiarization and each test phase) by dividing the number of frames in which the two coders agreed (e.g., both indicated the infant was looking to the right, both indicated that the infant was looking to the left, or both indicated that the infant was not looking) by the total number of codable frames in that phase. Only trials with at least 80% agreement between two coders for all phases of the trial (familiarization and test) were included in the analyses. Thus, if any phase of a trial had agreements less than 80 % for any single phase, a third coder was assigned to code that trial.

There were 101 trials that were coded by a third coder. If there was not 80% agreement for all three phases between the third coder and either of the original two coders, it was determined that the trial could not be reliably coded and was excluded from the final data. Twenty-one trials (from 18 individual infants) were excluded for this reason. For the 187 trials included in this analysis, inter-rater reliability was high (M = 94.19 % agreement, SD = 3.18 % agreement).

Data Processing and Analysis

Files downloaded from Lookit provided data about subject demographics (participant age, biological sex, income, education level, etc.), information about the session (e.g., stimulus order, video frame rate), and details about video recordings. We used the frame rate and video recordings information to create coding files in Datavyu. Custom R scripts were created to clean and process the coded data. We created script to integrate data from Lookit about the order the stimuli were presented and the unique timing for each phase on each trial for each individual infant with frame-by-frame data from Datavyu.

As a first step, we identified the start of each phase of the trial using timestamp information provided by Lookit. To account for a known issue regarding a variable delay between when Lookit indicates that a stimulus is presented and when the stimulus actually occurs on the screen, we calculated the duration of time elapsed between the webcam recording and stimulus presentation on each trial for each infant and offset the start of our phases based on this information. Next, we identified the number of frames in which looking was directed towards a specific region of the screen (e.g., left vs. right) for each unique phase (e.g., familiarization and test arrays) on each trial (note that because the recordings were generated using using participants' own webcam and internet, we had to account for differences in frame rate across videos and participants). In R, we calculated look durations to each side of the screen

by multiplying the number of frames in which looking was detected towards a specific screen region by the number of frames per s. We calculated total look duration during familiarization by summing the duration of looking at each item during the familiarization phase of each trial. Novelty preference scores were calculated separately for each 10-s test phase by dividing the total look duration to the novel item by the total looking at the novel and familiar item. The novelty preference for each trial was calculated by averaging the scores for the two phases for that problem.

Statistical analyses were conducted in R (Core Team & Others, 2013). To address our first objective of assessing whether infants would show memory in an online setting, we conducted one-sample *t*-tests on infants' mean novelty preference score averaged across all trials in which infants contributed usable data. Thus, infants' mean novelty preference score reflects how much, on average, they preferred the novel face. To address our second objective of evaluating the unique contributions of exposure time and familiarization time, we performed paired-sample t-tests on infants' mean performance on 10 s exposure trials versus 20 s exposure trials as well as correlations between infants' mean total look duration and mean novelty preference scores across trial types. Lastly, we examined whether infants showed age-related differences in novelty preference scores.

Results

Infants contributed an average of 1.73 ten-s familiarization trials (*SD* = .45) and 1.82 twenty-s familiarization trials (*SD* = .39) to the analyses. Thus, most infants completed both trials of each type. Infants' mean total look duration and mean novelty preference scores for each trial type are shown in Table 1. Unsurprisingly, infants had more looking time during the longer familiarization trials than during the shorter familiarization trials. To address our first objective,

we conducted one-sample t-tests against chance for each trial type. Comparisons of the novelty preference scores to chance indicated above chance performance for both the 10-s familiarization trials, $t(51) = 4.07$, $p < .001$, $d = .56$, and 20-s familiarization trials, $t(50) = 5.57$, $p < .001$, $d = .78$.

Table 1. Descriptive statistics for our measures of interest.

			Familiarization		Test	
Exposure time	Measure	*N*	Mean	SD	Mean	SD
10 s	Novelty Preference	52	-	-	.55*	.10
	Total look duration (in s)	52	6.75	1.77	14.1	2.83
20 s	Novelty Preference	51	-	-	.58*	.10
	Total look duration (in s)	51	12.38	3.32	13.4	3.27

To address our second objective and determine whether exposure time and total look duration influenced infants' memory performance, we conducted paired-samples t-tests between infants' mean novelty preference scores on 10 s and 20 s exposure trials as well as correlations between total look duration and novelty preference scores for each trial type. Comparison of infants' mean novelty preference scores for each trial type revealed that although infants had higher novelty preference scores in the 20-s memory problems, their scores in the two types of problems did not differ significantly, $t(50) = 1.37$, $p = .177$, $d = .19$ (Figure 4).

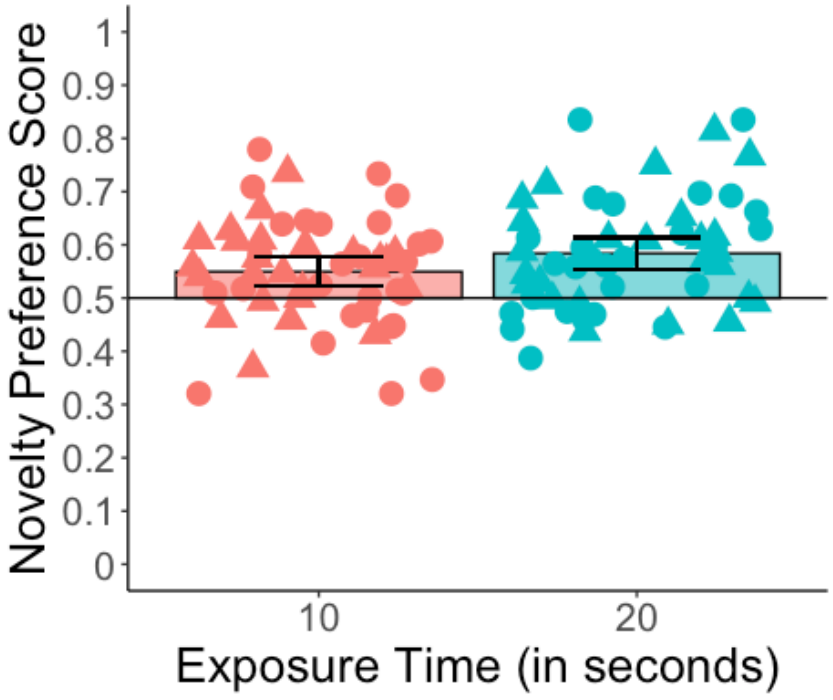

Figure 4. Association between novelty preference score and exposure time. The horizontal line bisecting the vertical axis represents chance (.50) performance. Individual data points represent infants' mean novelty preference scores averaged across all memory problems. The bars represent infants' mean novelty preference score on trials in which exposure time was 10 s (left) and 20 s (right). The horizontal line bisecting the vertical axis in each figure represents chance (.50) performance and error bars represent 95% confidence intervals. Shape represents infants' biological sex (circle = females, triangle = males).

Next, we examined the association between total look duration during familiarization and novelty preference scores collapsed across trial types. Because we expected infants' total look durations to vary on a trial-by-trial basis, we examined the impact of total look duration on memory performance by fitting a linear mixed effect model on the trial-level data. This model included a fixed effect of total look duration during familiarization (continuous; in s) and a

random intercept for participant to account for the nested structure of the data. This model revealed no significant effect of total look duration during familiarization, $\beta = .003$, $SE = 0.002$, $t = 1.118$, $df = 180.0$, $p = .265$, indicating that infants' look duration during learning did not influence their memory performance. Figure 5 displays the effect of total look duration across trial types.

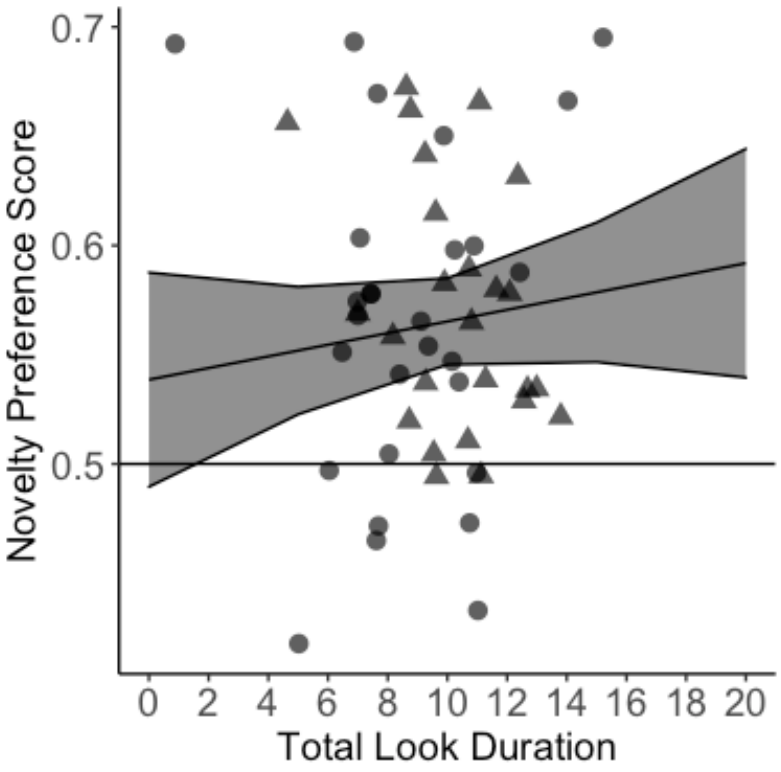

Figure 5. Association between novelty preference score and total look duration during familiarization collapsed across trial type. The horizontal line bisecting the vertical axis represents chance (.50) performance. Individual data points represent infants' mean novelty preference scores averaged across all memory problems. The lines represent estimated marginal means and the shading around the lines represents 95% confidence intervals for the estimated marginal means. Shape represents infants' biological sex (circle = females, triangle = males).

Lastly, we examined whether age influenced infants' memory performance collapsed across trial types. There was a significant correlation between infants' age (in days) and infants' mean novelty preference scores, $r(50) = 0.33$, $p = .018$, indicating that older infants showed higher novelty preference scores than did younger infants (Figure 6).

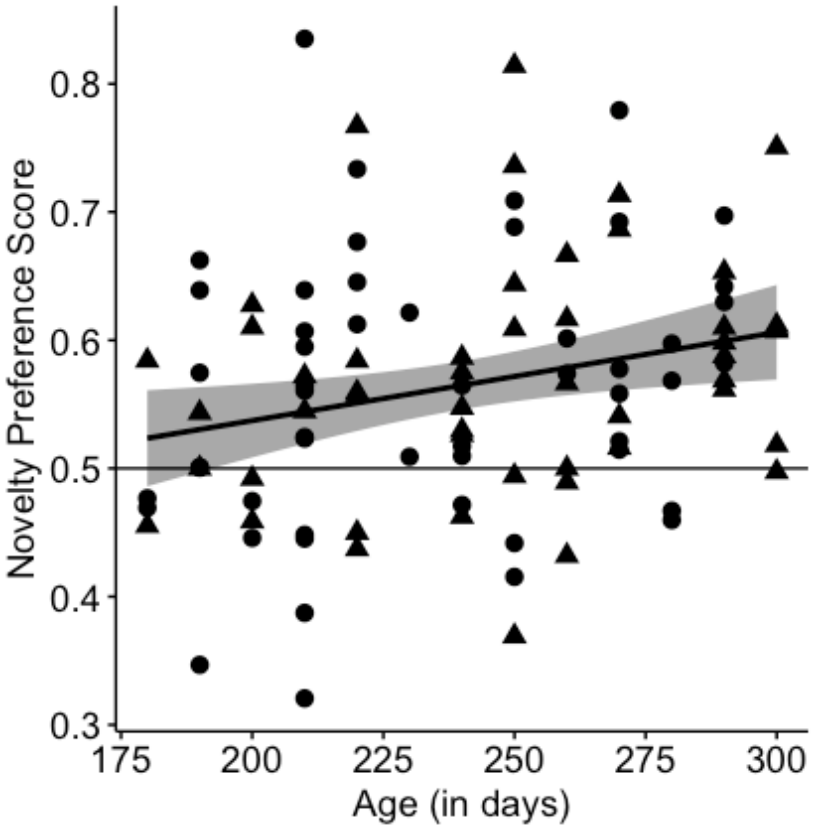

Figure 6. Association between novelty preference score and age in days. The horizontal line bisecting the vertical axis represents chance (.50) performance. Individual data points represent infants' mean novelty preference scores collapsed across all trials. Shading around regression lines represent 95% confidence intervals and shape represents infants' biological sex (circle = females, triangle = males).

Discussion

In this study, we observed that the group of infants, overall, showed evidence of memory in both exposure time conditions, and that infants' memory performance was associated with age but not total look duration. We had two objectives. First, we were interested in determining how assessing infants in an online context influences their memory performance. We observed that infants showed evidence of memory, during test exhibiting a preference for the novel face over

the familiar face. These findings are consistent with studies in more traditional laboratory settings and suggests that assessment context did not prevent infants from encoding, storing, and consolidating memory representations of the attended items during familiarization. Online studies may provide advantages over traditional laboratory assessments such as giving parents more control over the timing of participation, eliminating financial burdens associated with travel and taking time off from work, and increasing the accessibility of developmental science for parents that might not otherwise be able to participate (Scott et al., 2017; Scott & Schulz, 2017). However, participating from home also requires parents to have access to a computer and reliable access to internet and reduces the amount of control that experimenters have over the specific testing conditions. These results clearly demonstrate that the VPC task conducted online, using a platform like Lookit, is sensitive enough to detect whether infants show memory for a familiarized object, at least at the group level.

Our second objective was to disambiguate the unique contributions of exposure time and familiarization time on infants' visual recognition memory. These preliminary analyses were conducted on only about half the projected final sample size. Due to this small sample size, we were not able to systematically examine how these variables uniquely contribute to infants' memory performance. We did observe that infants as a group showed memory for the familiar item in both the 10-s and 20-s memory problems, and that in this sample there was no systematic relation between how long infants looked during familiarization and their subsequent novelty preference. Thus, these initial findings did not replicate the findings in the literature, and we did not observe that infants' memory depended on either exposure time or look duration. Because these analyses are preliminary, and we were unable to conduct our full model, it is unclear

whether the final analyses with the full sample will similarly fail to replicate the patterns observed in the literature.

One significant limitation of the present work was that we did not have sufficient power to test the combined effects of age and exposure or look duration. It is important to point out that we did observe that overall older infants showed higher novelty preference scores than younger infants, just as has been observed in more traditional laboratory settings (Rose & Feldman, 1987; Rose et al., 2001; Rose, Feldman, & Wallace, 1988c; Rose et al., 1982). However, studies in traditional laboratory settings have shown that younger infants must accumulate longer look durations during familiarization than older infants to show equivalent novelty preference scores (Rose, 1981, 1983), indicating that infants in this age range require different amounts of time to encode stimuli. Thus, one possible explanation for the null effects observed in the current study is that we were underpowered for detecting the effects of exposure time and total look duration on infants' memory. Once we have the full data, we will able to determine whether our limited sample size is driving the lack of associations between total look duration and exposure time on infants' memory performance.

An alternative possibility is that exposure and/or looking time do contribute to infants' memory encoding, but in this first study our trials were sufficiently long for infants in this age range to form a memory of the familiar face. Rose et al. (2001) examined visual recognition memory a similar stimulus set and found that infants as young as 5 months showed that a novelty preference after accumulating 10 s of looking. Although Rose et al. used look duration instead of exposure time, infants in that study who accumulated 10 s of looking also necessarily had at least 10 s of exposure time. Thus, in the current study, 10 s of exposure time may have been sufficient for even the youngest infants to learn and subsequently remember the attended items. If this

pattern is upheld in the final sample, one goal for future research is to test even shorter exposure times to see if that disrupts infants' memory formation in this context.

A third possibility is that although online tasks may be sufficient for assessing overall group responding, they do not provide the sensitivity and precision needed to test individual differences in performance. That is, the pattern of results observed in the current study may reflect a signal-to-noise ratio problem. Studies of adult recognition memory for face images typically reveal memory scores that are much higher (at least .90) even when participants are tested in more complex paradigms, store information for longer delays, and are required to retrieve more information about the learned faces (Wang, 2014), but it is important to note that in these kinds of procedures adults are typically prompted to produce an overt verbal or behavioral response. Studies of infant recognition memory depend on our ability to index stochastic cognitive processes associated with looking behavior including infants' natural proclivity to attend to novel over familiar stimuli. Because measures derived from infant looking behavior often have higher measurement error and lower true-score variance than adult studies (DeBolt et al., 2020), it is possible that infants' scores were highly influenced by error variance and that any effect of exposure time and total look duration on infants' memory performance was masked by the noisiness of infant data. The VPC procedure has been criticized for its internal consistency in general (Benasich & Bejar, 1992; DiLalla et al., 1990; Tasbihsazan et al., 2003), but testing infants in an online context may introduce even more measurement error than traditional laboratory assessments. Our full sample may provide us with the statistical power to systematically disambiguate noise in the data from true-score variance. Future research will be necessary to determine whether online studies yield noisier effects than traditional laboratory

assessments, but it should be noted that these conclusions are purely speculative, and the observed patterns could vary dramatically upon completion of data collection.

In conclusion, these findings indicate a similar pattern of results for infant VPC performance in online contexts and more traditional laboratory settings. Despite the potential for increased distractions when families participate from home, infants were able to form, store, and maintain memory representations in visual recognition memory. These findings provide additional evidence that online studies can be feasibly implemented to study aspects of infant cognition. Online studies present researchers with an additional tool for investigating fundamental aspects of infant cognitive development and the results from the current study suggest that these kinds of studies can be implemented to study memory processes as well. Additional data collection will be necessary to derive any robust conclusions about the pattern of results observed in the current study, but the findings presented here are promising and indicate that we can index aspects of infant visual recognition memory in non-traditional experimental contexts.

References

Amso, D., & Scerif, G. (2015). The attentive brain: insights from developmental cognitive neuroscience. *Nature Reviews. Neuroscience*, *16*(10), 606–619.

Axia, G., Bonichini, S., & Benini, F. (1999). Attention and reaction to distress in infancy: a longitudinal study. *Developmental Psychology*, *35*(2), 500–504.

Benasich, A. A., & Bejar, I. I. (1992). The Fagan test of infant intelligence: A critical review. In *Journal of Applied Developmental Psychology* (Vol. 13, Issue 2, pp. 153–171). https://doi.org/10.1016/0193-3973(92)90027-f

Caldwell, B. M., & Bradley, R. H. (2003). *Home observation for measurement of the environment: administration manual*.

Carter, R. C., Jacobson, J. L., Burden, M. J., Armony-Sivan, R., Dodge, N. C., Angelilli, M. L., Lozoff, B., & Jacobson, S. W. (2010). Iron deficiency anemia and cognitive function in infancy. *Pediatrics*, *126*(2), e427-34.

Caswell, B., Arnold, C., Lutter, C., Maleta, K., & Stewart, C. (2020). An egg feeding intervention increased protein quantity and quality among young Malawian children. *Current Developments in Nutrition*, *4*(Supplement_2), 955–955.

Catherwood, D., Skoien, P., Green, V., & Holt, C. (1996). Assessing the primary moments of infant encoding of compound visual stimuli. *Infant Behavior & Development*, *19*(1), 1–11.

Chhaya, R., Weiss, J., Seffren, V., Sikorskii, A., Winke, P. M., Ojuka, J. C., & Boivin, M. J. (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.

Chiriboga, C. A., Kuhn, L., & Wasserman, G. A. (2007). Prenatal cocaine exposures and dose-related cocaine effects on infant tone and behavior. *Neurotoxicology and Teratology*, *29*(3), 323–330.

Cohen, L. B. (1972). Attention-Getting and Attention-Holding Processes of Infant Visual Preferences. *Child Development*, *43*(3), 869–879.

Cohen, L. B., Chaput, H. H., & Cashon, C. H. (2002). A constructivist model of infant cognition. *Cognitive Development*, *17*(3–4), 1323–1343.

Cohen, L. B., Deloache, J. S., & Rissman, M. W. (1975). The Effect of Stimulus Complexity on Infant Visual Attention and Habituation. *Child Development*, *46*, 611–617.

Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, *52*, 337–367.

Colombo, J., Freeseman, L. J., Coldren, J. T., & Frick, J. E. (1995). Individual differences in infant fixation duration: Dominance of global versus local stimulus properties. *Cognitive Development*, *10*(2), 271–285.

Colombo, J., & Mitchell, D. W. (1990). Individual differences in early visual attention: Fixation time and information processing. *Individual Differences in Infancy: Reliability, Stability, Prediction*, 193–227.

Colombo, J., Mitchell, D. W., Coldren, J. T., & Freeseman, L. J. (1991). Individual Differences in Infant Visual Attention: Are Short Lookers Faster Processors or Feature Processors ? Author (s): John Colombo , D . Wayne Mitchell , Jeffrey T . Coldren and Laura J .

Freeseman Published by : Wiley on behalf of the Society for. *Child Development*, *62*(6), 1247–1257.

Colombo, J., Mitchell, D. W., & Horowitz, F. D. (1988a). Infant visual attention in the paired-comparison paradigm: test-retest and attention-performance relations. *Child Development*, *59*(5), 1198–1210.

Colombo, J., Mitchell, D. W., & Horowitz, F. D. (1988b). Infant visual attention in the paired-comparison paradigm: Test-retest and attention-performance relations. *Child Development*, *59*(5), 1198.

Colombo, J., Richman, W. A., Shaddy, D. J., Greenhoot, A. F., & Maikranz, J. M. (2001). Heart rate-defined phases of attention, look duration, and infant performance in the paired-comparison paradigm. *Child Development*, *72*(6), 1605–1616.

Colombo, J., Zavaleta, N., Kannass, K. N., Lazarte, F., Albornoz, C., Kapa, L. L., & Caulfield, L. E. (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.

Core Team, R., & Others. (2013). R: a language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna*.

Courage, M. L., & Howe, M. L. (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.

Courage, Mary L., Reynolds, G. D., & Richards, J. E. (2006). Infants' attention to patterned stimuli: developmental change from 3 to 12 months of age. *Child Development*, *77*(3), 680–695.

DeBolt, M. C., Rhemtulla, M., & Oakes, L. M. (2020). Robust data and power in infant research: A case study of the effect of number of infants and number of trials in visual preference procedures. *Infancy: The Official Journal of the International Society on Infant Studies*, *25*(4), 393–419.

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

Dewar, M., Garcia, Y. F., Cowan, N., & Della Sala, S. (2009). Delaying interference enhances memory consolidation in amnesic patients. *Neuropsychology*, *23*(5), 627–634.

DiLalla, L. F., Thompson, L. A., Plomin, R., Phillips, K., Fagan, J. F., III, Haith, M. M., Cyphers, L. H., & Fulker, D. W. (1990). Infant predictors of preschool and adult IQ: A study of infant twins and their parents. *Developmental Psychology*, *26*(5), 759–769.

Downing, S. M. (2003). Validity: on meaningful interpretation of assessment data. *Medical Education*, *37*(9), 830–837.

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

D'Souza, D., & D'Souza, H. (2021). Bilingual adaptations in early development. *Trends in Cognitive Sciences*, *25*(9), 727–729.

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram? *Annual Review of Psychology*, *55*(1), 51–86.

Fagan, J. F. (1970). Memory in the infant. *Journal of Experimental Child Psychology*, *9*(2), 217–226.

Fagan, J. F. (1971). Infants' recognition memory for a series of visual stimuli. *Journal of Experimental Child Psychology*, *11*(2), 244–250.

Fagan, J. F. (1972). Infants' recognition memory for faces. *Journal of Experimental Child Psychology*, *14*, 453–476.

Fagan, J. F. (1974). Infant Recognition Memory: The Effects of Length of Familiarization and Type of Discrimination Task. *Child Development*, *45*(2), 351–356.

Fagan, J. F. (1984). The relationship of novelty preferences during infancy to later intelligence and later recognition memory. *Intelligence*, *8*(4), 339–346.

Fagan, J. F., Iii. (2000). A theory of intelligence as processing: Implications for society. *Psychology, Public Policy, and Law: An Official Law Review of the University of Arizona College of Law and the University of Miami School of Law*, *6*(1), 168–179.

Fagan, J. F., & McGrath, S. K. (1981). Infant recognition memory and later intelligence. *Intelligence*, *5*(2), 121–130.

Freeseman, L. J., Colombo, J., & Coldren, J. T. (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.

Frick, J. E., & Richards, J. E. (2001). *Individual Differences in Infants' Recognition of Briefly Presented Visual Stimuli*.

Gaultney, J. F., Gingras, J. L., Martin, M., & DeBrule, D. (2005). Prenatal cocaine exposure and infants' preference for novelty and distractibility. *The Journal of Genetic Psychology*, *166*(4), 385–406.

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.

Hunter, M. A., Ames, E. W., & Koopman, R. (1983). Effects of stimulus complexity and familiarization time on infant preferences for novel and familiar stimuli. *Developmental Psychology*, *19*(3), 338–352.

Jacobson, S. W., Carter, R. C., Molteno, C. D., Stanton, M. E., Herbert, J. S., Lindinger, N. M., Lewis, C. E., Dodge, N. C., Hoyme, H. E., Zeisel, S. H., Meintjes, E. M., Duggan, C. P., & Jacobson, J. L. (2018). Efficacy of maternal choline supplementation during pregnancy in mitigating adverse effects of prenatal alcohol exposure on growth and cognitive function: A randomized, double-blind, placebo-controlled clinical trial. *Alcoholism, Clinical and Experimental Research*, *42*(7), 1327–1341.

Jacobson, S. W., Chiodo, L. M., Sokol, R. J., & Jacobson, J. L. (2002). Validity of maternal report of prenatal alcohol, cocaine, and smoking in relation to neurobehavioral outcome. *Pediatrics*, *109*(5), 815–825.

Johnson, M. H., Posner, M. I., & Rothbart, M. K. (1991). Components of visual orienting in early infancy: contingency learning, anticipatory looking, and disengaging. *Journal of Cognitive Neuroscience*, *3*(4), 335–344.

Kennedy, T., Thomas, D. G., Woltamo, T., Abebe, Y., Hubbs-Tait, L., Sykova, V., Stoecker, B. J., & Hambidge, K. M. (2008). Growth and visual information processing in infants in Southern Ethiopia. *Journal of Applied Developmental Psychology*, *29*(2), 129–140.

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

McCall, R. B. (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, S. B., & Fantz, R. L. (1974). Recognition memory in Down's syndrome and normal infants. *Child Development*, *45*(3), 651–660.

Nairne, J. S. (2002). Remembering over the short-term: the case against the standard model. *Annual Review of Psychology*, *53*(1), 53–81.

Nelson, K. G., Goldenberg, R. L., Hoffman, H. J., & Cliver, S. P. (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, K. L., & Fagan, J. F. (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.

Oakes, L. M., & Kovack-Lesh, K. A. (2013). Infants' visual recognition memory for a series of categorically related items. *Journal of Cognition and Development: Official Journal of the Cognitive Development Society*, *4*(1), 63–86.

Oakes, L. M., Kovack-Lesh, K. A., & Horst, J. S. (2009). Two are better than one: comparison influences infants' visual recognition memory. *Journal of Experimental Child Psychology*, *104*(1), 124–131.

Oakes, L. M., & Rakison, D. H. (2019). The "humpty-dumpty" problem in the study of development. In *Developmental Cascades* (pp. 55–74). Oxford University Press.

Pérez-Edgar, K. (2018). Attention mechanisms in behavioral inhibition: Exploring and exploiting the environment. In *Behavioral Inhibition* (pp. 237–261). Springer International Publishing.

Prado, E. L., Maleta, K., Caswell, B. L., George, M., Oakes, L. M., DeBolt, M. C., Bragg, M. G., Arnold, C. D., Iannotti, L. L., Lutter, C. K., & Stewart, C. P. (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.

Richards, J. E. (1997). Effects of attention on infants' preference for briefly exposed visual stimuli in the paired-comparison recognition-memory paradigm. *Developmental Psychology*, *33*(1), 22–31.

Ricker, T. J., & Cowan, N. (2014). Differences between presentation methods in working memory procedures: a matter of working memory consolidation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *40*(2), 417–428.

Rose, S. A. (1980). Enhancing visual recognition memory in preterm infants. *Developmental Psychology*, *16*(2), 85–92.

Rose, S. A. (1981). Developmental Changes in Infants ' Retention of Visual Stimuli. *Child Development*, *52*(1), 227–233.

Rose, S. A. (1983). Differential rates of visual information processing in full-term and preterm infants. *Child Development*, *54*(5), 1189–1198.

Rose, S. A. (1994). Relation between physical growth and information processing in infants born in India. *Child Development*, *65*(3), 889–902.

Rose, S. A., & Feldman, F. J. (1987). Infant visual attention: Stability of individual differences from 6 to 8 months. *Developmental Psychology*, *23*(4), 490–498.

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (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.

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2003). Infant visual recognition memory: independent contributions of speed and attention. *Developmental Psychology*, *39*(3), 563–571.

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2004). Infant visual recognition memory. *Developmental Review: DR*, *24*(1), 74–100.

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2005). Recall memory in the first three years of life: a longitudinal study of preterm and term children. *Developmental Medicine and Child Neurology*, *47*(10), 653–659.

Rose, S. A., Feldman, J. F., Mccarton, C. M., & Wolfson, J. (1988). Information Processing in Seven-Month-Old Infants as a Function of Risk Status. *Child Development*, *59*(3), 589–603.

Rose, S. A., Feldman, J. F., & Wallace, I. F. (1988a). Individual Differences in Infants ' Information Processing: Reliability, Stability, and Prediction. *Child Development*, *59*(5), 1177–1197.

Rose, S. A., Feldman, J. F., & Wallace, I. F. (1988b). Individual Differences in Infants' Information Processing : Reliability, Stability, and Prediction. *Child Development*, *59*(5), 1177–1197.

Rose, S. A., Feldman, J. F., & Wallace, I. F. (1988c). Individual differences in infants' information processing: reliability, stability, and prediction. *Child Development*, *59*(5), 1177–1197.

Rose, S. A., Feldman, J. F., Wallace, I. F., & Cohen, P. (1991). Language: A partial link between infant attention and later intelligence. *Developmental Psychology*, *27*(5), 798.

Rose, S. A., Gottfried, A. W., Melloy-Carminar, P., & Bridger, W. H. (1982). Familiarity and novelty preferences in infant recognition memory: Implications for information processing. *Developmental Psychology*, *18*(5), 704–713.

Ruff, H. A. (1975). The function of shifting fixations in the visual perception of infants. *Child Development*, *46*(4), 857–865.

Scott, K., Chu, J., & Schulz, L. (2017). Lookit (part 2): Assessing the viability of online developmental research, results from three case studies. *Open Mind*, *1*(1), 15–29.

Scott, K., & Schulz, L. (2017). Lookit (Part 1): A New Online Platform for Developmental Research. *Open Mind*, *1*(1), 4–14.

Siegel, E. H., Kordas, K., Stoltzfus, R. J., Katz, J., Khatry, S. K., LeClerq, S. C., & Tielsch, J. M. (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.

Singer, L. T., Eisengart, L. J., Minnes, S., Noland, J., Jey, A., Lane, C., & Min, M. O. (2005). Prenatal cocaine exposure and infant cognition. *Infant Behavior & Development*, *28*(4), 431–444.

Squire, L. R., Genzel, L., Wixted, J. T., & Morris, R. G. (2015). Memory consolidation. *Cold Spring Harbor Perspectives in Biology*, *7*(8), a021766.

Strohminger, N., Gray, K., Chituc, V., Heffner, J., Schein, C., & Heagins, T. B. (2016). The MR2: A multi-racial, mega-resolution database of facial stimuli. *Behavior Research Methods*, *48*(3), 1197–1204.

Struthers, J. M., & Hansen, R. L. (1992). Visual recognition memory in drug-exposed infants. *Journal of Developmental and Behavioral Pediatrics: JDBP*, *13*(2), 108–111.

Tasbihsazan, R., Nettelbeck, T., & Kirby, N. (2003). Predictive validity of the Fagan Test of Infant Intelligence. *The British Journal of Developmental Psychology*, *21*(4), 585–597.

Thompson, L. A., Fagan, J. F., & Fulker, D. W. (1991). Longitudinal prediction of specific cognitive abilities from infant novelty preference. *Child Development*, *62*(3), 530–538.

Wang, B. (2014). Effect of time delay on recognition memory for pictures: the modulatory role of emotion. *PloS One*, *9*(6), e100238.

WHO Multicentre Growth Reference Study Group. (2006). WHO Child Growth Standards based on length/height, weight and age. *Acta Paediatrica (Oslo, Norway: 1992). Supplement*, *450*, 76–85.