Executive functions in infancy: Measurement using a novel tablet task and exploration of longitudinal attentional and cognitive predictors

Emma Macrae[1], Bosiljka Milosavljevic[1,2], Luke Mason[2,3], Marta Perapoch Amadó[4,5], Maria Rozhko[1,5], Laura Katus[6,7], Michelle de Haan[8.9], Clare Elwell[5], Sophie E. Moore[10,11], Sarah Lloyd-Fox[1,2] & The BRIGHT Project Team[12]

[1]Department of Psychology, University of Cambridge, Cambridge, UK.

[2]Centre for Brain and Cognitive Development, Birkbeck, University of London, London, UK.

[3]Department of Forensic and Neurodevelopmental Sciences, King's College London, London, UK.

[4]Department of Psychology, University of East London, London, UK.

[5]Department of Medical Physics and Biomedical Engineering, University College London, London, UK.

[6]School of Human Sciences, University of Greenwich, London, UK.

[7]Centre for Family Research, University of Cambridge, Cambridge, UK.

[8]Great Ormond Street Institute of Child Health, University College London, London, UK.

[9]Great Ormond Street Hospital for Children NHS Foundation Trust
London, London, UK.

[10]Department of Women and Children's Health, King's College London, London, UK.

[11]MRC Unit The Gambia at the London School of Hygiene and Tropical Medicine, Fajara, The Gambia.

[12]The BRIGHT Study team are (in alphabetical order): Maria M. Crespo-Llado, Dominique Taylor & Sophie Yelland

Conflict of interest statement: The authors declare no conflict of interest

Data availability statement: Data supporting this paper will be made available subject to established data sharing agreements.

Acknowledgements: We would like to thank all families participating in this research, without whom this work would not have been possible. This research is funded by the Bill and Melinda Gates Foundation (grants OPP1061089 and OPP1127625). The Nutrition Theme at MRCG is supported by the MRC & the Department for International Development (DFID) under the MRC/DFID Concordat agreement (MRC Programme MC-A760-5QX00). BM is supported by an ESRC Secondary Data Analysis Initiative Grant (ES/V016601/1). SEM is supported by a Wellcome Trust Senior Research Fellowship (220225/Z/20/Z). SLF is supported by a UKRI Future Leaders Fellowship (grant number MR/S018425/1). *This work is supported by the NIHR GOSH BRC. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.*

Research Highlights:

- Assessment of Executive Functioning (EF) skills at 18- and 24-months of age using a novel tablet-based measure.
- The tablet measure was able to detect improvement in executive function scores with age, stability of individual differences, and had good internal consistency.
- Prospective longitudinal design allowed examination of complex relationships between emerging attentional abilities and executive functions.
- Slower attentional disengagement at 8-months, but faster disengagement at 18-months predicted higher EF skills at 24-months.
- No concurrent or longitudinal relationships found between executive functions at 18- or 24-months and general cognitive abilities measured by the Mullen Scales of Early Learning.

Abstract

Executive Functions (EFs) in infancy and childhood are important predictors of later outcomes. The present study used data from a prospective longitudinal study to examine the development and predictors of EF skills among infants during the first 24-months of life. First, we evaluated the use of a tablet-based assessment to measure EFs among infants aged 18- and 24-months. We also examined concurrent and longitudinal associations between attentional disengagement, general cognitive skills and EFs. Participants (*N*=60, 30 female) completed the EF task at 18- and 24-months of age. Attentional disengagement and general cognitive skills were assessed at 5-, 8-, 12-, 18- and 24-months using an eye-tracking measure (the gap-overlap task) and the Mullen Scales of Early Learning (MSEL), respectively. The EF task demonstrated good internal consistency, sensitivity to age-related change in performance and stable individual differences. No associations were found between EF skills and MSEL scores longitudinally or concurrently. The eye-tracking task revealed that *slower* attentional disengagement at 8-months, but *faster* disengagement at 18-months, predicted better EF skills at 24-months. These findings indicate that the tablet-based assessment is a potentially useful tool for measuring emergent EFs early in infancy. The multifaceted relationship between attentional disengagement and EFs suggests that the rapid development of the attentional system in infancy results in distinct attentional skills, at different ages, being relevant for EF development. On the other hand, overt behavioural skills early in infancy do not predict EF skills.

Key words: Infancy; Executive functions; Tablet task; Cognitive development; Prospective longitudinal design

List of Figure Legends

Figure 1. Number of participants at each visit and reasons for withdrawal. Those in bold are the age points used in the current analysis.

Figure 2. Distribution of MSEL ELC scores at the 5-, 8-, 12-, 18- and 24-months visits. The middle line represents the median, upper bound quartile 3 and lower bound quartile 1 of the scores. Violin plots show the distribution of scores. Coloured points represent individual MSEL scores from infants and are coloured by standard deviation from the mean score for the relevant visit.

Figure 3. Distribution of disengagement times as measured by the Gap-Overlap task at the 5-, 8-, 12-, 18- and 24-months visits. Disengagement times presented in box plots where the middle line represents the median, upper bound quartile 3 and lower bound quartile 1 of the scores. Violin plots show the distribution of scores. Points represent individual disengagement times from infants and are coloured by standard deviation from the mean score for the relevant visit.

Figure 4. Distribution of Babyscreen scores at the 18- and 24-months visits. Babyscreen scores at 18 months (left) and 24 months (right) in box plots where the middle line represents the median, upper bound quartile 3 and lower bound quartile 1 of the scores. Violin plots show the distribution of scores. Points represent individual Babyscreen scores from infants and are coloured by standard deviation from the mean score for the relevant visit.

Figure 5: Scatterplot showing the association between gap-overlap disengagement times at 8 months and Babyscreen scores at 24 months.

Figure 6: Scatterplot showing the association between gap-overlap disengagement times at 18 months and Babyscreen scores at 24 months.

Executive functions (EFs) are cognitive processes necessary for achieving goals, flexibly responding to the environment, and regulating behaviours (Carlson, 2005; Thompson & Steinbeis, 2020). These skills are commonly conceptualised as three interrelated components: working memory (WM), inhibitory control (IC) and cognitive flexibility (CF). WM is the ability to retain and manipulate information. IC is the ability to control attention, behaviour and emotions. CF includes the ability to change perspectives and understand rule changes (Diamond, 2013; Garon et al., 2008; Miyake et al., 2000). There is growing interest in the development of EFs in early childhood, as these skills have been found to predict social, emotional, and academic outcomes both cross-sectionally and longitudinally. However, due to methodological constraints, there remains a paucity of research that examines emergent EFs in young, pre-verbal children, and, consequently, potential longitudinal associations between early attentional predictors and EF development remain relatively unexplored.

The present study evaluates the use of a novel, tablet-based assessment of EFs, the Babyscreen (BabyScreen app, Twomey et al., 2021; Twomey et al., 2018), as a viable means of measuring these skills among infants aged 18- and 24-months. Furthermore, we explore the associations between EFs at 18- and 24-months with measures of attentional flexibility (measured by an eye-tracking task of attentional disengagement) and general cognitive skills (measured by a behavioural assessment of language, motor, and perceptual skills), assessed at multiple intervals during the first 24-months of life. The findings from this study will, therefore, provide insight into both the measurement of EFs in infancy and their relationships with other emerging cognitive abilities.

Development of EFs and related cognitive abilities

Garon et al. (2008) proposed a hierarchical model of EF development based on the three-factor structure (WM, IC, CF) observed in adults, with basic cognitive skills developing before 3-years of age, which are later integrated to form EFs. The model suggests that attentional abilities are foundational to all EFs and thus develop first, during the first 6-months of life, and continuously thereafter. WM subsequently develops at around 6-months, IC at 8-months and CF at 12-months.

In line with this proposed developmental timeline, associations between emergent cognitive skills and EFs have been reported both concurrently and longitudinally (Holmboe et al., 2018; Stephens et al. 2019). For example, Holmboe et al. (2018) found concurrent associations between general cognitive skills and IC at 9-months, using both a behavioural assessment of IC (the A-not-B task) and an eye-tracking measure (the Freeze Frame task). Furthermore, cognitive ability at 24-months has been found to be predictive of EFs at 6-years (Stephens et al., 2019). In spite of these promising results, it is often challenging to measure cognitive skills early in infancy using behavioural, examiner-led assessments (Brian et al., 2014; Yaari et al, 2018), which has resulted in a general paucity in research examining EF development during the first months of life.

Furthermore, while attentional mechanisms in early infancy are recognised as important predictors for later EFs, the nature of this relationship is complex and may change during the first 2-years (Hendry et al., 2019). Attentional abilities are described as comprising of alerting, orienting and executive networks (Posner & Rothbart, 2006). The *orienting network*, which promotes fast shifting of attention, begins to develop between 3- to 6-months and is thought to be important for EFs in early infancy. This is supported by Cuevas and Bell (2014), who found that infants who exhibited shorter looking durations (which was posited to reflect faster disengagement) in a behavioural task at 5-months had more advanced EFs at 24-, 36-, and 48-months than those with longer looking times. Using the same behavioural assessment, Devine et al. (2019) reported that shorter looking times at 4-months were a stronger predictor of EFs at 14-months than parent-rated temperament. Taken together, this work supports the idea that attentional flexibility (measured through faster attentional disengagement) in the early part of the first year of life predicts better EF skills later in infancy/childhood.

As infants mature, a proposed a shift occurs, whereby the *executive network* (responsible for sustained attention and resolving conflict) becomes increasingly relevant for cognitive development (Geeraerts et al., 2019; Posner et al., 2012). Kannass et al. (2006) suggest that endogenous control of attention, which is a foundational component of the executive network, starts to develop around 9-months. Thus, beginning at the onset of the second year of life, the ability to sustain attention may become more important than attentional flexibility for emergent EF skills. Indeed, sustained attention at 12-months has been reported to be predictive of EFs (measured using the A-not-B task) at 24-months (Johansson et al., 2015). On the other hand, Nakagawa and Sukigara (2013) found that, at 12-months of age, slower disengagement times predicted less advanced concurrent self-regulation capabilities. However, longitudinally, slower disengagement at 12-months predicted more advanced effortful control at 18- and 24-months of age. This suggests that sustained attention may become relevant for EF skills later than previously thought, between 12- to 18-months. Taken together, prior research provides evidence to support the role of both the orienting network in early infancy and a shift to reliance on the executive network between 9- and 18-months of age in EF development. However, there is a scarcity in longitudinal research, spanning multiple time points during the first 2-years, that examines how the progression of attentional skills impacts EF development. Thus, it is difficult to establish whether reports of both sustained and flexible attention predicting EF skills truly reflect a shift in the attentional networks or if this is an artefact of the diverse experimental methods used to assess both attention and EFs. Furthermore, longitudinal research incorporating assessments at 7- to 11-months, an age largely overlooked in prior research, would be helpful in establishing a clearer frame in the timing of shifts in the attentional skills relevant for EFs.

Measuring EFs in infancy

While theoretical models propose that EFs start to develop during the first 3-years of life (Garon et al., 2008), most research on the early development of EFs focuses on older children (Best & Miller, 2010; Garon et al., 2008). This may be because infants' limited motor and language skills restrict their ability to complete traditional EF tasks (Hendry et al., 2016). Given that neural networks exhibit their highest plasticity during the first 24-months, it is crucial to be able to measure the development of EFs during infancy, as this could support the development of more effective and longer-lasting interventions for delayed EF development (Bornstein, 2014; Fiske & Holmboe, 2019; Wass et al., 2011). Common methods of assessing EFs in infancy include parental report, behavioural and eye-tracking tasks. However, these have limitations including the use of a large battery of tasks, time-consuming manual coding, biased interpretation, and expensive equipment (Frank et al., 2016; Hendry et al., 2016). Therefore, the development of new tasks to measure EFs in infancy is needed. Tasks that use a touchscreen tablet for assessment (hereafter "tablet tasks") have shown promise for reliable measurement of EFs in infancy. They can collect multiple types of variables, including accuracy in item completion, touch patterns and reaction times, and are thought to be engaging and relatively inexpensive so can be used to increase the scale of research (Bhavnani et al., 2019; Frank et al., 2016; Friend & Keplinger, 2003).

Tablet tasks have been used extensively to measure EFs in children over 2-years (Pitchford & Outhwaite, 2016; Semmelmann et al., 2016; Willoughby et al., 2019) but emerging evidence shows promise for their use among younger infants. A tablet task (the Early Childhood Inhibitory Touchscreen Task) developed by Holmboe et al. (2021) has been used in multiple studies and has consistently detected improvements and individual differences in inhibition during infancy. Lo et al. (2021) demonstrated that 18- to 20-month-olds could meaningfully engage with a tablet task measuring reading comprehension. Furthermore, Frank et al. (2016) found that, compared to eye-tracking and storybook paradigms, their tablet task had higher completion rates for 1- and 2-year-olds. However, further research is needed to establish whether tablet tasks can be used to measure global EF abilities in infancy.

The present study

Data used for this study were collected as part of the Brain Imaging for Global Health project (BRIGHT; globalfnirs.org/the-bright-project), a longitudinal study examining infant development from 0- to 24-months of age.

In the current study, our first aim is to evaluate the utility of a novel tablet task, the Babyscreen (Hello Games Ltd, UK), in measuring EFs in infancy. The Babyscreen was developed to measure EFs in children aged 12-36 months and is based on existing EF measures, such as the A-not-B and Object Permanence tasks (Twomey et al., 2018). Initial validation work suggests that the Babyscreen is sensitive to age-related changes in early EFs, for example, children aged 30-36 months completed a greater number of trials and were faster in completing the more complex tasks than those aged 24-29 months (Twomey et al., 2018). Furthermore, Twomey et al. (2021) demonstrated a positive association between performance on the Babyscreen and general cognitive skills, measured by the Bayley Scales of Infant and Toddler Development (BSID), among infants aged 18- to 24-months. This study aims to extend these findings by examining the development of Babyscreen performance with a longitudinal design and examining the association between EFs and both global cognitive skills and attentional disengagement, measured at multiple intervals during the first 2-years of life (at 5-, 8-, 12-, 18- and 24-months). We expect to reproduce and extend Twomey at al.'s (2018) findings, whereby infants will have better performance on the Babyscreen at 24-

months than at 18-months, and that task performance at the two time points will be correlated. Secondly, given that prior work has found an association between global cognitive skills and EFs (e.g., Holmboe et al., 2018; Twomey et al, 2018), we posit that there will be both concurrent and longitudinal positive associations between Babyscreen scores and measures of general cognitive skills. Regarding possible associations between EFs and attention, we expect that, initially, *faster* disengagement at 5-months will predict better EFs at 18- and 24-months. However, coinciding with the shift in salience from the orienting to the executive network in supporting EF development, we also expect that the direction of this association will change around 12-months, when *slower* disengagement thereafter will predict better EF skills. Finally, we do not make a specific hypothesis about the association at 8-months given the scarcity of research examining attention and EFs at 7-11 months of age.

Methods

Participants

This study uses data from the UK cohort within the BRIGHT Project. While the study has been conducted in both the UK and The Gambia, the EF assessments described in the present study were only administered in the UK (see Lloyd-Fox et al, in prep, for further discussion about feasibility work within The Gambian cohort).

Once per week during the recruitment period, all families attending their 32–36-week antenatal visit at the Rosie Hospital, Cambridge University Hospitals were provided with study information. Families were recruited if they provided informed consent and had healthy pregnancies. Infants were only included if they were born between 37- and 42-weeks gestation, were a singleton, had no diagnosis of any major medical or neurological difficulties at birth and had a birth weight of over 2.5kg. Sixty-two infants (50% female) were recruited.

The study was approved by the National Research Ethics Service Committee East of England (REC reference 13/EE/0200).

Participants were invited to 8 scheduled visits from late pregnancy to 24-months post-partum. The visits included eye-tracking and behavioural assessments (for full protocol, see Lloyd-Fox et al., in prep). Figure 1 details the specific ages at each study visit, the number of participants that attended the visit and reasons for participant withdrawal. The current analyses use data from the 5-, 8-, 12-, 18- and 24-month visits. Two participants withdrew before the 5-month visit so the sample examined here comprises 60 participants (50% female).

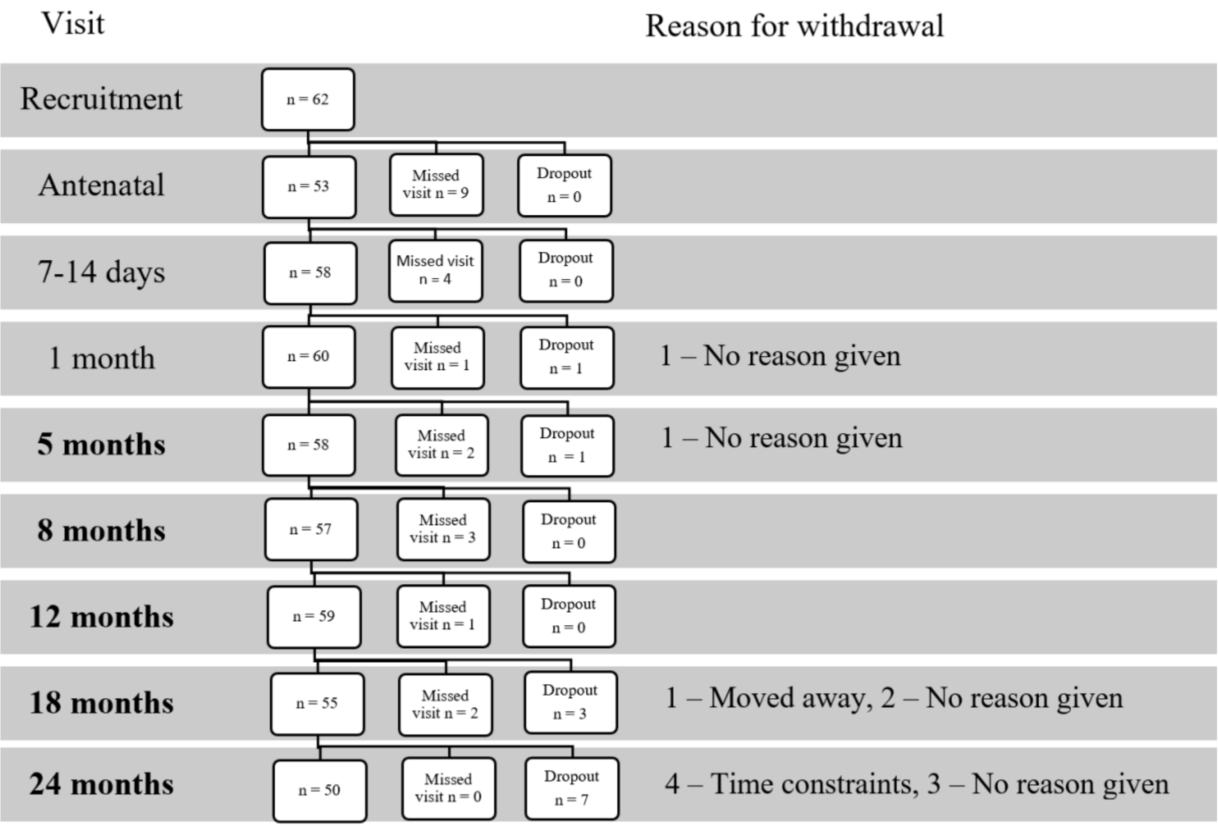

Figure 1. Number of participants at each visit and reasons for withdrawal. Those in bold are the age points used in the current analysis.

Demographic data

Demographic data were collected at the initial antenatal visit, and at 8- and at 18-month post-partum visits by questionnaire. For the current analysis, data from the 18-month visit were used as this was closest in time with the administration of the EF measures. Given prior research, which showed that both maternal education and family income are associated with child EF skills (Hackman et al., 2015; Lawson et al., 2014), information on these demographic characteristics were used in analyses. *Household income* was assessed via a single question asking parents to choose a category that best described their annual household income (<£20,000; £20,000-29,000; £30,000-39,000; £40,000-59,999; £60,000-79,000; £80,000-99,999; £100,000-149,999; >£149,999). They were also given an option not to respond. *Maternal education* was also assessed using a single question asking mothers to indicate their highest level of education (Primary; Secondary; Tertiary graduate; Tertiary postgraduate), also with an option not to respond. Finally, data were collected about *participant racial background* by asking parents to indicate both the mother's and father's ethnicity from a set of 5 options (White, Asian, Black, mixed race and other/don't know). Infant race was ascertained from parents' race and, where parents were from different racial groups, the infant was identified as being biracial or mixed race.

Executive Functions

The Babyscreen software application version 1.5 (Hello Games Ltd, UK) was administered at 18- and 24-months to measure EFs. The task is an 18-item tool that was developed for use with infants aged 12-36 months. It provides a unitary measure of skills but is comprised of items that elicit specific EFs, including working memory and selective attention, and is based on widely used assessments of EFs for older children (Twomey et al., 2018).

Items involve performing a set of problem-solving tasks, which increase in difficulty as the task progresses. The task was presented on an iPad (5th generation, 9.7-inch screen) set to full brightness, 70% of the maximum volume and affixed horizontally to a table. Participants either sat on their parent's lap or stood at the table. Prior to starting the Babyscreen, participants were familiarised with the iPad by playing a game where they could draw on the screen. The Babyscreen task started with three training items, which were followed by the test trials. Participants were given two attempts to solve each trial. They were initially given an opportunity to solve the task independently (first attempt), without any instructions or support. If they did not respond correctly within 20s at 18 months or 30s at 24 months, the experimenter was prompted to give a demonstration. After the demonstration, participants were given another attempt to complete the trial (second attempt). Images of balloons and music were presented as a reward for trial completion. If the trial was not completed correctly on either the first or second attempt, it was skipped. The task was terminated either when infants completed all trials or when they failed to complete three consecutive trials. Experimenters made notes during each trial to indicate if anything affected infant performance (e.g., inattentiveness or fussiness). Parents were also asked to rate their infant's previous touchscreen use on the following scale: never, occasionally, 2-3 times per week, or daily.

The Babyscreen generates two variables for each trial attempt: accuracy (whether the trial was completed successfully) and reaction time (RT; speed of trial completion for successful trials). A feasibility study suggested that the total number of trials completed without demonstration (first attempts) was best able to capture age differences in performance and was a useful measure of overall EFs (Twomey et al., 2021). Therefore, the total number of items completed without demonstration (hereafter "Babyscreen score") was used for primary analyses. The mean RT for trials on the first attempt was also computed and used in analyses.

General cognitive ability

The Mullen Scales of Early Learning (MSEL; Mullen, 1995) are a battery of assessments designed to measure cognitive and gross motor abilities from birth to 68 months. In this study, the MSEL was administered at 5-, 8-, 12-, 18- and 24-months of age. Cognitive abilities are measured by four subscales: fine motor, receptive language, expressive language, and visual reception. The fifth subscale measures gross motor abilities. Each scale is assessed through a series of interactive tasks presented in order of increasing difficulty. Examiners rated whether participants successfully completed each task. Total scores for each subscale were computed and converted to age-normed t-scores based on a US sample (*M* = 50, *SD* = 10; Mullen, 1995). The Early Learning Composite (ELC; *M* = 100, *SD* = 15) was subsequently derived from all cognitive t-scores and was used as a measure of overall cognitive ability. The ELC is used in analyses for the present study.

Attentional disengagement

The gap-overlap task, designed to measure attentional disengagement through testing infants' ability to orient to stimuli in their peripheral vision, was conducted as part of a battery of eye-tracking tasks at 5-, 8-, 12-, 18- and 24-month visits. The procedure of the gap-overlap task was identical to the one described by Glennon et al. (2020) and Jones et al. (2019). Every trial started with the presentation of a central stimulus (CS; image of analogue clock), which was accompanied by an alerting sound. This remained on the screen at 3Hz between 3 and 5cm (2.86°-4.77°) until participant fixated on the CS. Upon fixation, the CS began to rotate at 500° per second for a random interstimulus interval (ISI), which ranged between 500-700ms. After the ISI, the baseline, gap or overlap conditions were presented for 200ms. In the *baseline* and *overlap* conditions, the CS remained on screen, and in the *gap* condition it was removed from the screen. After the 200ms period, the peripheral stimulus (PS, a cartoon cloud) was presented on either the left or right side of screen, 3cm (2.86°) from the edge, accompanied by an alerting sound. It was rotated at 500° per second until participant fixated on it. In the *baseline* condition, the CS disappeared from the screen when PS was presented. In the *overlap condition,* the CS remained on the screen for the duration of the trial. A reward stimulus was presented for 1000ms (cartoon animal accompanied by a sound) when participant successfully fixated on the PS. Trials were presented in blocks of 12, all stimuli were presented at 3cm by 3cm (2.86° by 2.86°). The variable of interest was the saccadic reaction time (SRT) to shift attention from the CS to PS, relative to the onset of PS presentation.

Eye movements were recorded using a Tobii TX300 eye-tracker (Tobii Technology, Stockholm, Sweden) with 300 Hz refresh rate set to a sampling rate of 60Hz. Visual stimuli were presented on a 23-inch monitor. Infants faced the screen while sitting on their parent's lap 60cm from the screen. Once calibrated to infants' eye movements, the task was started, and infants' eye movements were recorded. The session was paused if the infant fussed out and only resumed if possible. Data were subsequently analysed offline.

Participant data were removed if they had fewer than 6 valid trials per condition. Trails were considered valid if they met all of the following criteria: (1) gaze fell on the CS; (2) there were no periods of missing data longer than 200ms during CS period; (3) there was at least one period of gaze on the CS before or after CS onset; (4) there were no periods of missing data longer than 100ms during the PS period; (5) SRTs ranged between 150-1200ms; (6) gaze was not on the opposite side of screen to the PS; (7) gaze was not within the PS area of interest during the period after engagement with the CS but before PS onset. Trials were considered invalid if they violated any of the above criteria. The variable of interest,

attentional disengagement, was computed by subtracting SRTs in the baseline condition from SRTs in the overlap condition.

Statistical analyses

Analyses were conducted in R (R Core Team, 2020). Outlier identification was conducted by the boxplot method using the rstatix package (Kassambara, 2020). Outliers were removed if they were extreme outliers (values that fell beyond Q1 - 3*IQR or Q3 + 3*IQR where Q1 and Q3 are the first and third quartiles and IQR is the interquartile range (Q3-Q1)), skewed the distribution of the data when included or if there were experimenter notes suggesting that the data quality was poor (e.g., participant was upset or highly inattentive during the task). Descriptive statistics (Mean, Standard Deviation) were computed for all variables.

Repeated-measures ANOVAs were conducted to assess age-related change in MSEL and gap-overlap scores between 5- and 24-months. If the ANOVAs showed significant change with age, post-hoc tests using Bonferroni correction were used to identify which age points significantly differed from each other on each task.

To check whether the data met assumptions of repeated-measures ANOVA, Shapiro-Wilk tests were used to assess whether data were normally distributed. Levene's test was used to test the homogeneity of variance assumption. If this assumption was violated, a Brown-Forsythe correction was applied. Mauchly's test was used to determine if the data met the assumption of sphericity.

Evaluation of the Babyscreen

To investigate whether demographic factors influenced Babyscreen scores, one-way between-subjects ANOVAs were conducted. These determined whether there were significant differences between the Babyscreen scores of infants with different levels of each demographic variable (sex, annual household income, maternal education, and previous touchscreen use).

To determine whether the Babyscreen could detect changes in scores between 18- and 24-months, a paired Wilcoxon-signed rank test was conducted. Effect sizes (r) were calculated by Z/√N (Rosenthal, 1991, as cited in Field et al., 2012). To ensure that the change in RT allowance between visits (20s at 18-months and 30s at 24-months) did not affect differences in Babyscreen scores between visits, a general linear model (GLM) was constructed. The GLM was constructed with Babyscreen score as the dependent variable, visit as a fixed effect, and mean RT as a random effect.

Pearson's correlation tests were used to determine whether there was an association between Babyscreen score and mean RT and to determine whether participants' scores were correlated between 18- and 24- months. Cronbach's alpha was used to determine the internal consistency of Babyscreen scores.

Associations between performance on the Babyscreen, cognitive ability and attentional disengagement times

To investigate the concurrent and longitudinal relationships between MSEL ELC scores and gap-overlap disengagement times and Babyscreen scores, multivariate multiple regression models were constructed. This is an extension of multiple regression, in which one can measure the association between multiple dependant variables with a single set of predictors and covariates, accounting for residual correlations (Muñoz-Rocha et al., 2018). Five models were run using data from each study visit separately (5-, 8-, 12-, 18- and 24-months). Babyscreen scores at 18- and 24-months were included as the dependant variables, and MSEL ELC and gap-overlap disengagement scores were included as predictors. For the model with predictors at 24-months, a linear regression was run including only 24-month Babyscreen scores as the dependent variable. Given that there were no significant associations between sex or any of the demographic/family characteristics and Babyscreen performance (see Results for summary), these were not controlled for in the regression models.

Results

Participant demographics

Table 1 summarises participant age and sex ratio at each study visit relevant to present analyses (5-24 months). There were no significant differences in sex distribution at any of the visits.

Table 1*:* Sample size, descriptive statistics for age (days) and sex ratio at each visit.

Visit	*N*	Mean (SD) age in months	Sex ratio (M:F)
5-months	58	5.13 (0.21)	29:29
8-months	57	8.28 (0.33)	28:29
12-months	59	12.35 (0.41)	29:30
18-months	55	18.32 (0.49)	27:28
24-months	50	24.23 (0.52)	23:27

Table 2 summarises participant and family demographic characteristics, measured at the 18-month visit. Of the 60 participants, all families reported an annual household income above £30,000. Furthermore, 78% of the infants' mothers had higher education qualifications, with 47% having postgraduate degrees. Most participants were white (93%) and there were no families where mothers and fathers were from different racial/ethnic backgrounds.

Table 2*:* Infant and family demographic characteristics at 18-months and participant prior touchscreen use

Demographic characteristic	**Frequency of rating (% of sample)**	
Annual household income (£)		
<20,000	0 (0%)	
20,000-29,000	0 (0%)	
30,000-39,999	4 (7%)	
40,000-59,999	13 (22%)	
60,000-79,999	9 (15%)	
80,000-99,999	13 (22%)	
100,000-149,999	6 (10%)	
Do not wish to answer	3 (5%)	
Missing data	12 (20%)	
Parental education level	**Mothers**	
Secondary	2 (3%)	
Tertiary	1 (2%)	
Undergraduate	21 (35%)	
Postgraduate	27 (45%)	
Missing data	9 (15%)	
Parental race	**Mothers**	**Fathers**
White	56 (93%)	56 (93%)
Asian	2 (3%)	2 (3%)
Black	1 (1%)	1 (1%)
Mixed	0 (0%)	0 (0%)
Other/don’t know	2 (3%)	2 (3%)
Previous touchscreen use	**18 months**	**24 months**
Never	6 (10%)	4 (7%)
Occasionally	21 (35%)	15 (25%)
2-3 times per week	3 (5%)	9 (15%)
Daily	8 (13%)	6 (10%)
Missing data	22 (37%)	26 (43%)

Descriptive statistics

Table 3 summarises performance on the experimental tasks (Babyscreen scores, MSEL ELC, gap-overlap disengagement), after removal of extreme outliers for each task.

Table 3*:* Descriptive statistics for each experimental task and number of participants that completed the task at each study visit

Task (Variable (unit))	**Visit**	***N***	***M* (*SD*)**
Babyscreen total score	18-months	38	12.17 (2.63)
	24-months	34	13.64 (2.83)
Babyscreen (Average RT (ms))	18-months	38	13661.99 (2950.78)
	24-months	34	15978.74 (3956.00)
MSEL (ELC score)	5-months	56	93.80 (11.61)
	8-months	57	94.17 (12.70)
	12-months	54	91.12 (14.57)
	18-months	55	101.86 (16.34)
	24-months	49	113.10 (14.21)
Gap-overlap (Average disengagement (ms))	5-months	56	163.42 (77.03)
	8-months	55	129.38 (68.85)
	12-months	49	93.31 (50.85)
	18-months	47	75.50 (52.04)
	24-months	46	59.09 (43.13)

A significant age at visit effect was observed with the MSEL ELC scores, $F(4, 48) = 9.92, p< .001, \eta^2p = .45$. Scores at 24-months were significantly higher than at all other visits, but all other comparisons were non-significant (see Figure 2). Disengagement times during the gap-overlap task decreased across study visits, $F(2,80) = 17.56, p< .001, \eta^2p =.47$, but the only significant difference between consecutive study visits was between 8- and 12-months (see Figure 3).

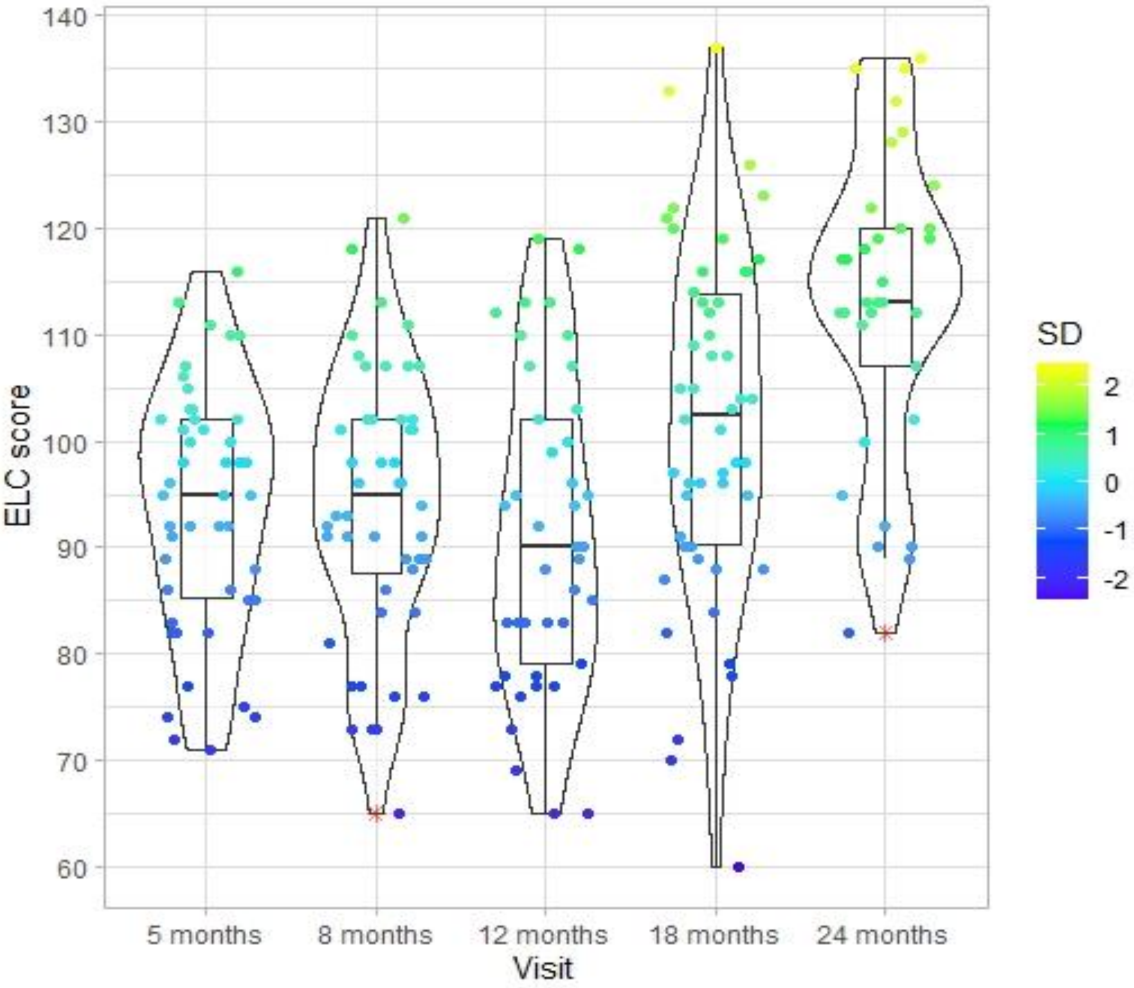

Figure 2. Distribution of MSEL ELC scores at the 5-, 8-, 12-, 18- and 24-months visits. The middle line represents the median, upper bound quartile 3 and lower bound quartile 1 of the scores. Violin plots show the distribution of scores. Coloured points represent individual MSEL scores from infants and are coloured by standard deviation from the mean score for the relevant visit.

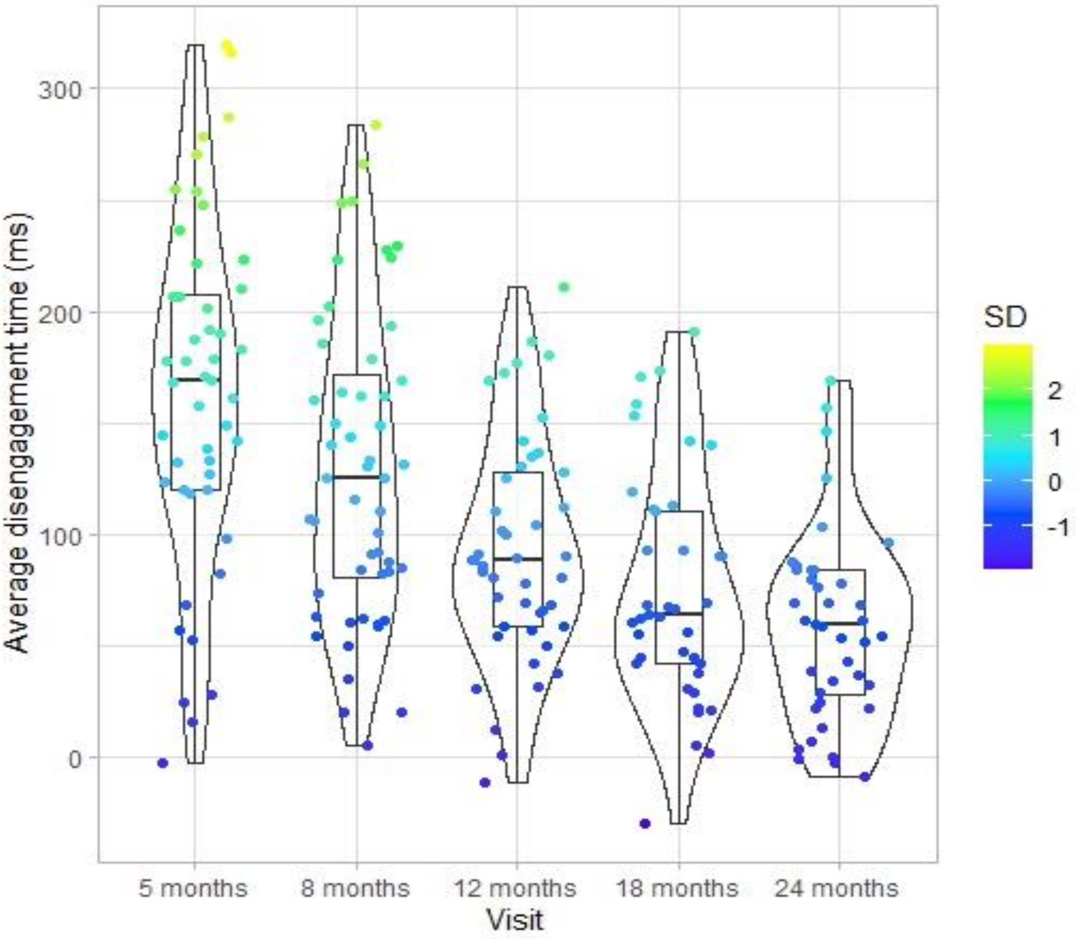

Figure 3. Distribution of disengagement times as measured by the Gap-Overlap task at the 5-, 8-, 12-, 18- and 24-months visits. Disengagement times presented in box plots where the middle line represents the median, upper bound quartile 3 and lower bound quartile 1 of the scores. Violin plots show the distribution of scores. Points represent individual disengagement times from infants and are coloured by standard deviation from the mean score for the relevant visit.

Effect of demographic factors on Babyscreen scores

There was no significant effect of sex on Babyscreen performance at 18-months *F*(1, 34)=.18, *p*=.68, η^2=.01) or 24-months *F*(1, 30)=.02, *p*=.90, η^2=.00). Likewise, there was no effect of annual household income at either 18-months (*F*(5, 25)=1.32, *p*=.29, η^2=.02) or 24-months (*F*(4, 22)=.18, *p*=.94, η^2=.03). There was also no impact of maternal education at either 18-months *F*(2, 28)=2.10, *p*=.14, η^2=.14) or 24-months *F*(3, 26)=2.34, *p*=.10, η^2=.21). Finally prior touch screen use did not impact Babyscreen performance at either 18-months *F*(3, 13)=1.27, *p*=.33, η^2=.07) or 24-months *F*(3, 28)=0.25, *p*=.86 η^2=.03).

Change in Babyscreen scores with age and associations with RT

Babyscreen scores were significantly higher at 24-months than at 18-months, *W*(19)=18.5, *p*=.006, with a large effect size, *r*=.658 (see Figure 4). Additionally, there was a significant correlation between Babyscreen scores at 18- and 24-months, *r*(17)=.50, *p*=.03.

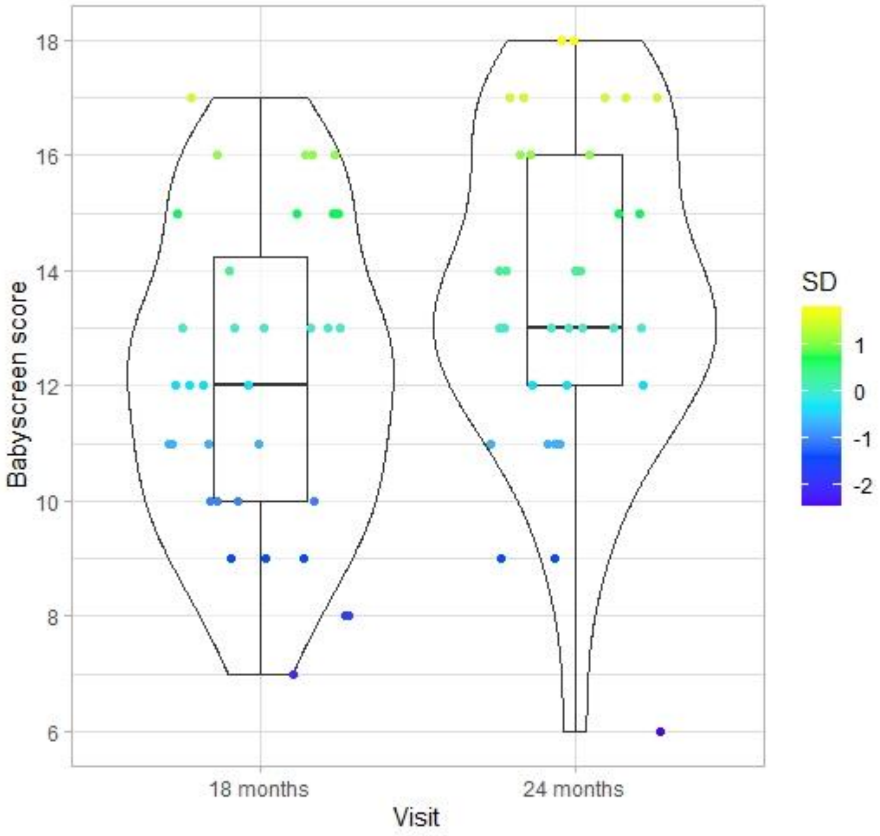

Figure 4. Distribution of Babyscreen scores at the 18- and 24-months visits. Babyscreen scores at 18 months (left) and 24 months (right) in box plots where the middle line represents the median, upper bound quartile 3 and lower bound quartile 1 of the scores. Violin plots show the distribution of scores. Points represent individual Babyscreen scores from infants and are coloured by standard deviation from the mean score for the relevant visit.

There were significant, negative, associations between Babyscreen scores (number of successfully completed items) and RT to complete task at both visits (*r*(34)=-.39, *p*=.02 at 18-months; *r*(31) =-.74, *p*< .01 at 24-months), suggesting that those who scored higher on the Babyscreen also completed trials faster. In spite of this, a general linear model revealed that adding RT as a random effect did not impact on the effect of visit on Babyscreen performance, $F(1,67)=5.00$, $p=.03$. This suggests that differences in RT, and thus the increased time allowance given to complete trials at 24-months, did not account for increases in Babyscreen scores between 18- and 24-months. Finally, the Babyscreen showed good internal consistency at both 18- ($\alpha=.83$) and 24-months ($\alpha=.86$).

Concurrent and longitudinal relationships between Babyscreen scores, cognitive skills and attentional disengagement

Table 4 summarises the multivariate multiple regressions examining associations between Babyscreen performance, MSEL ELC and gap-overlap disengagement scores at each visit. MSEL scores had no longitudinal or concurrent associations with Babyscreen scores at either 18- or 24-months. In contrast, gap-overlap disengagement times at 8-months were positively associated with Babyscreen scores at 24-months, while gap-overlap disengagement times at 18-months were negatively associated with Babyscreen scores at 24-months. This suggests that slower disengagement times at 8-months were associated with higher Babyscreen scores at 24-months, whereas faster disengagement times at 18-months were associated with higher Babyscreen scores at 24-months. No further associations were found between gap-overlap disengagement times and Babyscreen scores. These associations are summarised in Figure 5 and Figure 6.

Table 4: Summary of multivariate multiple regression predicting Babyscreen scores at 18- and 24-months from MSEL ELC scores and gap-overlap disengagement times at 5, 8, 12, 18 and 24 months.

Predictors	**Outcomes**							
	Babyscreen score 18-months				**Babyscreen score 24-months**			
	B	*SE*	*t*	*p*	*B*	*SE*	*t*	*p*
5 months								
MSEL ELC	0.08	0.08	1.09	0.30	0.09	0.08	1.15	0.28
Gap-overlap disengagement	0.01	0.01	0.80	0.44	0.00	0.01	0.45	0.67
8 months								
MSEL ELC	0.07	0.08	0.82	0.43	0.07	0.05	1.50	0.17
Gap-overlap disengagement	0.00	0.01	0.20	0.85	**0.03**	**0.01**	**3.50**	**0.01**
12 months								
MSEL ELC	0.12	0.06	1.98	0.09	0.08	0.05	1.69	0.14
Gap-overlap disengagement	0.01	0.02	0.53	0.61	0.01	0.02	0.83	0.43
18 months								
MSEL ELC	-0.02	0.07	-0.24	0.81	0.07	0.05	1.44	0.19
Gap-overlap disengagement	-0.02	0.02	-0.82	0.44	-0.05	0.02	-2.84	**0.02**
24 months								
MSEL ELC	-	-	-	-	0.02	0.04	0.49	0.63
Gap-overlap disengagement	-	-	-	-	-0.01	0.02	-0.49	0.63

NB Grey cells highlight significant associations

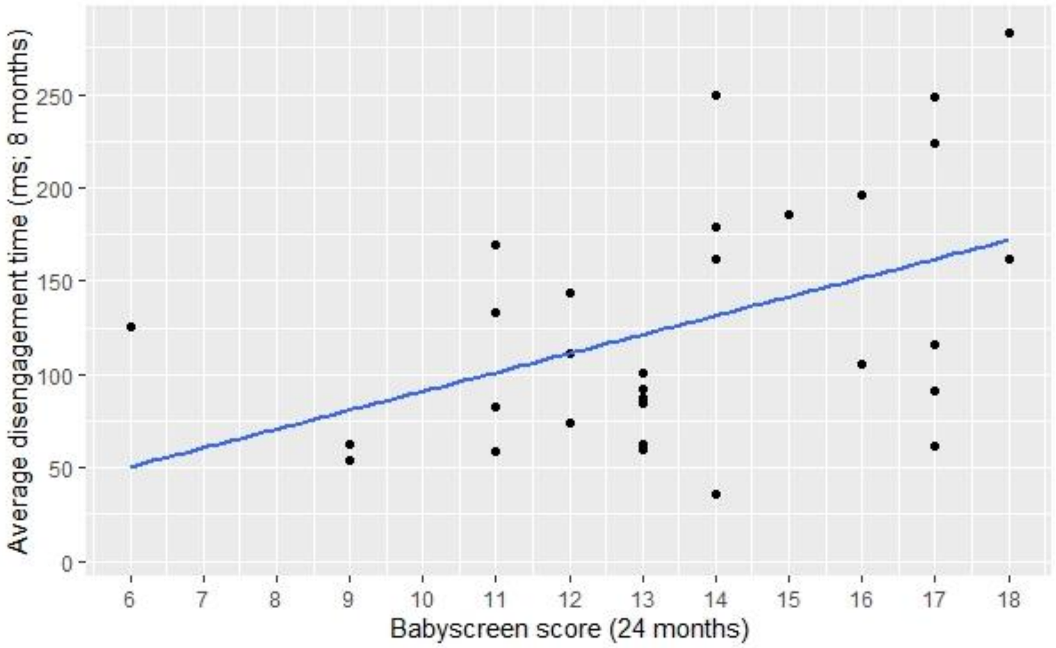

Figure 5: Scatterplot showing the association between gap-overlap disengagement times at 8 months and Babyscreen scores at 24 months.

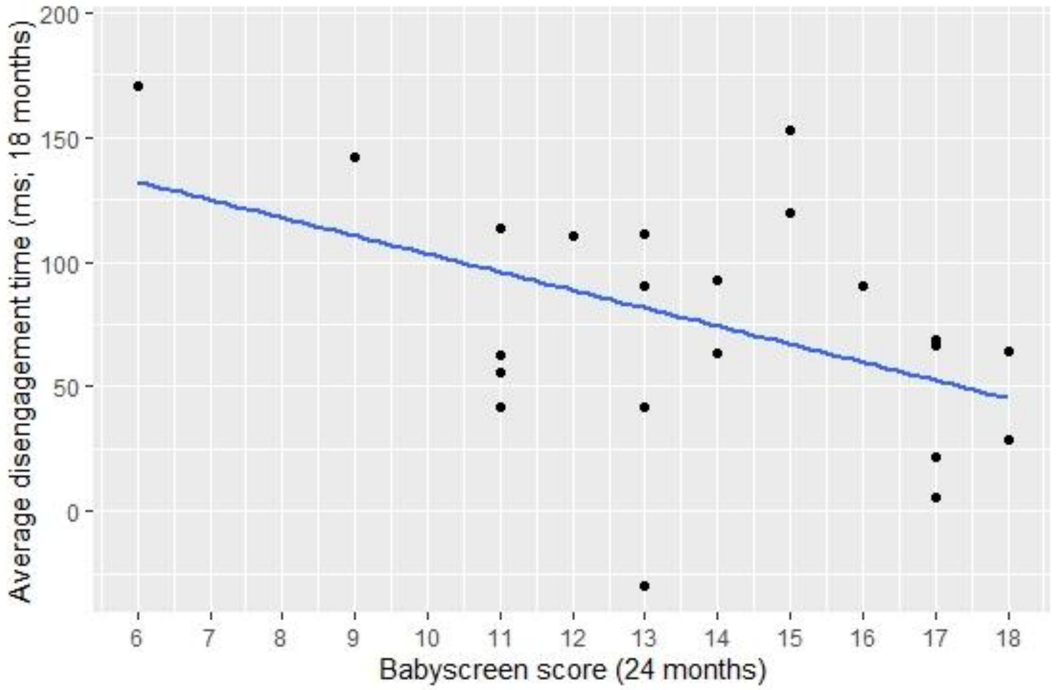

Figure 6: Scatterplot showing the association between gap-overlap disengagement times at 18 months and Babyscreen scores at 24 months.

Discussion

This study evaluated the utility of the Babyscreen task, a novel, tablet-based task in assessing EF skills among infants in the second year of life (aged 18- and 24-months). Longitudinal and concurrent associations between EFs in the second year and cognitive and attentional markers earlier in infancy were also measured. The Babyscreen demonstrated good internal consistency and was sensitive to the development of stable individual differences in EFs between 18- and 24-months. Associations were also found between EFs at 24-months and attentional disengagement at both 8- and 18-months. However, these associations were contrary to expectations – *slower* disengagement times at 8-months predicted better EF performance at 24-months, while *faster* disengagement at 18-months was associated with increased performance at 24-months. There were no further associations between speed of attentional disengagement and EF skills at either time point. Furthermore, there were no significant concurrent or longitudinal associations between global cognitive skills (measured by the MSEL) and EF skills.

Evaluation of the Babyscreen task in assessing EFs in the second year of life

The Babyscreen task demonstrated good performance across several metrics, suggesting that it has promise as a tool to assess EF abilities among infants as young as 18-months of age. Firstly, the task demonstrated good internal consistency at both 18- and 24-months of age. Secondly, consistent with prior research using the task (Twomey et al., 2018; Twomey et al., 2021), Babyscreen scores were higher at 24-months than at 18-months and this age effect remained even after the longer time allowed to complete the task at 24-months was accounted for. Thirdly, there was a significant association between performance at the two time points - infants who had higher scores at 18-months also had higher scores at 24-months. The improvements in Babyscreen scores over a period of 6 months are consistent with demonstrations that infancy is a time of rapid EF development (Garon et al., 2008;

Garon et al., 2014; Hendry et al., 2016). Holmboe et al.'s (2021) tablet task also detected the development of, and stable individual differences in, inhibition between 18- and 24-months. Our findings are therefore consistent with prior work by suggesting that the Babyscreen can discriminate between EF abilities of younger and older infants. The longitudinal design extends previous cross-sectional work with this task and demonstrates that the task can detect both the development of, and stable individual differences in, EFs in infancy.

One limitation of the Babyscreen is that it provides only a unitary, global measure of EFs, rather than examining specific EF domains. Thus, it is difficult to make conclusions about the development of specific EFs. Furthermore, while the psychometric properties are promising, it is also difficult to establish construct validity of the task with other measures, as most other assessments of EFs involve tasks that require verbal instructions and are suitable only for older children.

Demographic factors (sex, household income, maternal education) and previous touchscreen use were not associated with Babyscreen scores. This is inconsistent with research showing that SES influences EF development (Lawson et al., 2018). The null finding here could be due to the homogenous, relatively high-SES sample in which most parents had higher levels of education, meaning there was not enough variation in SES to detect its effect on EFs. On the other hand, the lack of influence of previous touchscreen use on Babyscreen scores is consistent with prior research using the Babyscreen (Twomey et al., 2018; Twomey et al., 2021). However, most of the participants in the sample did have some prior touchscreen exposure, so these findings may have differed if a greater proportion had not used a tablet before.

Associations between EF skills and general cognitive ability

We assessed associations between Babyscreen scores at 18- and 24-months and performance on the MSEL, a behavioural measure of global cognitive skills, both concurrently at these time points and longitudinally (at 5-, 8- and 12-months). Contrary to prior work, there were no associations between the Babyscreen and the MSEL at any age. These findings are surprising given that prior work has demonstrated both concurrent relationships between MSEL scores and inhibitory control at 9-months (Holmboe et al., 2018) and longitudinal associations between MSEL at 2-years and EFs at 6-years (Stephens et al., 2018). It is possible that general cognitive abilities are more relevant for specific types of EFs than others, as demonstrated in the association with inhibitory control (Holmboe et al., 2018), which was not captured by the Babyscreen's global EF score. Likewise, studies with populations with higher prevalence of cognitive delays (e.g., Yaari et al., 2018) report that differences in cognitive skills (measured behaviourally) typically become observable in the second year of life and, thus, we may have been less able to capture meaningful individual differences at the very early time points in this study.

Twomey et al. (2021) found that, among infants referred for neurodevelopmental assessment, those who had cognitive scores consistent with developmental delay on the Bayley Scales of Infant Development (BSID) performed significantly worse on the Babyscreen than infants who had typical development. It is possible that, while the Babyscreen can distinguish between infants with cognitive delay and those with typical development, it is less sensitive to individual differences in cognitive skills among typically developing infants. This is compounded by the fact that the infants in our sample are predominantly from high-SES households and whose parents tended to have high levels of educational attainment. Finally, it is possible that the small sample size in this study did not have sufficient power to detect significant associations between the MSEL and Babyscreen.

Attentional disengagement as a predictor of EF skills

One of the key aims of the present study was to assess whether attentional flexibility, measured through speed of attentional disengagement, in early infancy could predict EF skills at 18- and 24-months. Prior work examining these associations has produced conflicting results, with some research suggesting that faster disengagement in early infancy was important for the development of EFs, while ability to sustain attention became more relevant in later infancy (see Hendry et al., 2019 for review). However, there was substantial variability in prior research in both the associations reported and the specific ages in which they occurred. Therefore, our study was well placed to address some of these inconsistencies and the paucity of research in general at this age point because of its longitudinal design and multiple study visits that were close in time.

We found that *slower* disengagement times at 8-months and *faster* disengagement at 18-months were associated with higher Babyscreen scores at 24-months. However, no significant relationships were found for disengagement times at 5-, 12- or 24-months. The findings are, to a degree, consistent with prior research that showed an association between slower disengagement at 12-months and higher effortful control at 18- and 24-months (Nakagawa & Sukigara, 2013). This prompted the idea that sustained attention, reflecting endogenous control of attention, at the onset of the second year of life, was an important factor in EF development. However, similar work suggested that endogenous control of attention actually emerges earlier, at approximately 9-months of age (Kannass et al., 2006). In line with this work, it is possible that the association between slower disengagement at 8-months and EF skills at 24-months found here reflects the emergence of sustained attention at this age and its importance for later EF development. Considering this alongside prior work, the results could indicate a developmental window, perhaps between 6 to 12 months where slower disengagement is advantageous for later EFs.

The association between faster disengagement at 18-months and higher EF skills at 24-months was contrary to predictions. Sacrey et al. (2013) suggest that, by 12-months of age, typically developing infants start to show more flexible attentional disengagement. They also found that at 12-months of age prolonged disengagement on the gap-overlap task distinguished typically developing infants from those with autism spectrum disorder. Our findings, therefore, support this work because, by 18-months of age, we would expect most typically developing infants in our sample to have fast and flexible attentional disengagement and prolonged disengagement to be associated with difficulties in the cognitive domains.

Finally, the lack of relationships between attentional disengagement measured at 5-, 12- and 24-months and EF skills raises additional questions. Prior studies found no association between attentional disengagement at 4-months and later EF skills (e.g. Holmboe et al., 2018). Therefore, it is possible that 5-months is too early to detect an association between attention and later EF skills. At 24-months, it is possible that attentional disengagement becomes more stable, and participants who showed delayed disengagement at 18-months, caught up. In line with this hypothesis, group differences in attentional disengagement reported between infants with ASD and typically developing controls have been found to be no longer significant by 36-months (Sacrey et al., 2013). It is also important to note that significant associations between attentional disengagement and Babyscreen performance were only found with Babyscreen scores collected at 24-months. This could reflect stabilisation of EF ability at this age, making it a more reliable age to measure EFs than at 18-months.

While future research is required to understand the particular pattern of results that we have found, our work is among the first to examine associations between attention at multiple time points during the first two years of life and emergent EFs. Future longitudinal work would benefit from implementing a similar design with a substantially larger sample size. Furthermore, it would be valuable to include multiple measures of attention, particularly tasks that are specifically designed to measure attentional disengagement and sustained attention.

Strengths, limitations, and implications for future work

This study has several strengths including the multi-method, longitudinal approach. The measurement of the same constructs over 5 time points in the first 2 years of life facilitated intricate investigation of the development of cognitive functions and how they relate to each other during infancy and toddlerhood. This is important as infancy is a time of rapid development of cognitive functions and abilities and relationships are likely to evolve rapidly so examination of multiple time points is needed to find critical points in development (Garon et al., 2008; Hendry et al., 2016). This design is relatively unique within the field with most studies taking measurements at one or two time points or using mixed-age cohorts. This study is also one of the first to measure EFs with a tablet task in children under 2 years.

However, this study is not without limitations. Firstly, the sample size was small (n = 60 overall, with smaller samples for individual tests), limiting power to detect relationships (Button et al., 2013). Secondly, the sample, selected from the city of Cambridge and surrounding rural regions in the UK, was homogenous in terms of race, high-SES and high parental educational attainment. All families reported an annual household income of over £30,000 (cf. UK median of £29,900; ONS, 2021) and 78% of the infants' parents had higher education qualifications (cf. 42% nationally; ONS, 2017). This is likely to explain the lack of variability and relatively high performance in MSEL scores and limits the generalisability of the findings.

Thirdly, there were methodological issues. While the Babyscreen is useful because it measures multiple EFs, understanding of the precise constructs measured is lacking, meaning that analyses were limited to an overall measure of EFs, rather than examining relationships between components of EFs and other abilities. This is problematic because EFs are multidimensional with different developmental trajectories and relationships with other abilities (Hendry et al., 2016).

Conclusions

This study investigated the utility of a new tablet task in measuring EFs in infancy. The Babyscreen was found to be useful for measuring emerging EFs and capturing consistent and improving performance over time with high internal consistency. While the task has been found to discriminate between general cognitive abilities in infants with and without neurodevelopmental delay in other studies, this finding was not replicated in the present, typically developing, sample. The relationship between EFs and attentional disengagement was complex, consistent with the highly varied literature.

Given the limitations of the small, high-SES, typically developing sample used here, it would be useful for future research to repeat the current study with a larger sample. In addition, the inclusion of infants with elevated familial likelihood, or showing signs of developmental neurodivergence would facilitate confirmation of previous results. Furthermore, adding a measure of another facet of attention, such as sustained attention, would provide greater insight into the abilities underlying EF development.

Overall, this study has demonstrated a useful tool for measuring EF development and is one of the first to assess EFs, attention and cognitive skills using a longitudinal, multi-measure design. Use of the Babyscreen could support future research aiming to understand the development of EFs in infancy, to identify those with neurodivergence and could potentially be used in combination with other measures to track the development of EFs throughout life.

References

Austin, G., Bondü, R., & Elsner, B. (2020). Executive function, theory of mind, and conduct-problem symptoms in middle childhood. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.00539

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development, 81*(6), 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Bhavnani, S., Mukherjee, D., Dasgupta, J., Verma, D., Parameshwaran, D., Divan, G., Sharma, K. K., Thiagarajan, T., & Patel, V. (2019). Development, feasibility and acceptability of a gamified cognitive DEvelopmental assessment on an E-Platform (DEEP) in rural Indian pre-schoolers – a pilot study. *Global Health Action*, *12*(1), 1548005. https://doi.org/10.1080/16549716.2018.1548005

Bornstein, M. H. (2014). Human infancy...and the rest of the lifespan. *Annual Review of Psychology*, *65*(1), 121–158. https://doi.org/10.1146/annurev-psych-120710-100359

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), Article 5. https://doi.org/10.1038/nrn3475

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, *28*(2), 595–616. https://doi.org/10.1207/s15326942dn2802_3

Cuevas, K., & Bell, M. A. (2014). Infant attention and early childhood executive function. *Child Development*, *85*(2), 397–404. https://doi.org/10.1111/cdev.12126

Demetriou, E. A., DeMayo, M. M., & Guastella, A. J. (2019). Executive function in autism spectrum disorder: History, theoretical models, empirical findings, and potential as an endophenotype. *Frontiers in Psychiatry*, *10*. https://doi.org/10.3389/fpsyt.2019.00753

Devine, R. T., Ribner, A., & Hughes, C. (2019). Measuring and predicting individual differences in executive functions at 14 months: A longitudinal study. *Child Development*, *90*(5), 618–636. https://doi.org/10.1111/cdev.13217

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Elsabbagh, M., Fernandes, J., Jane Webb, S., Dawson, G., Charman, T., & Johnson, M. H. (2013). Disengagement of visual attention in infancy is associated with emerging autism in toddlerhood. *Biological Psychiatry*, *74*(3), 189–194. https://doi.org/10.1016/j.biopsych.2012.11.030

Field, A. P., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. SAGE Publications Ltd.

Fiske, A., & Holmboe, K. (2019). Neural substrates of early executive function development. *Developmental Review*, *52*, 42–62. https://doi.org/10.1016/j.dr.2019.100866

Frank, M. C., Sugarman, E., Horowitz, A. C., Lewis, M. L., & Yurovsky, D. (2016). Using tablets to collect data from young children. *Journal of Cognition and Development*, *17*(1), 1–17. https://doi.org/10.1080/15248372.2015.1061528

Friend, M., & Keplinger, M. (2003). An infant-based assessment of early lexicon acquisition. *Behavior Research Methods, Instruments, & Computers*, *35*(2), 302–309. https://doi.org/10.3758/BF03202556

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, *134*(1), 31–60. https://doi.org/10.1037/0033-2909.134.1.31

Garon, N., Smith, I. M., & Bryson, S. E. (2014). A novel executive function battery for preschoolers: Sensitivity to age differences. *Child Neuropsychology*, *20*(6), 713–736. https://doi.org/10.1080/09297049.2013.857650

Geeraerts, S. B., Hessels, R. S., Van der Stigchel, S., Huijding, J., Endendijk, J. J., Van den Boomen, C., Kemner, C., & Deković, M. (2019). Individual differences in visual attention and self-regulation: A multimethod longitudinal study from infancy to toddlerhood. *Journal of Experimental Child Psychology*, *180*, 104–112. https://doi.org/10.1016/j.jecp.2018.11.012

Glennon, J. M., D'Souza, H., Mason, L., Karmiloff-Smith, A., & Thomas, M. S. C. (2020). Visuo-attentional correlates of Autism Spectrum Disorder (ASD) in children with Down syndrome: A comparative study with children with idiopathic ASD. *Research in Developmental Disabilities*, *104*, 103678. https://doi.org/10.1016/j.ridd.2020.103678

Hendry, A., Johnson, M. H., & Holmboe, K. (2019). Early development of visual attention: Change, stability, and longitudinal associations. *Annual Review of Developmental Psychology*, *1*(1), 251–275. https://doi.org/10.1146/annurev-devpsych-121318-085114

Hendry, A., Jones, E. J. H., & Charman, T. (2016). Executive function in the first three years of life: Precursors, predictors and patterns. *Developmental Review*, *42*, 1–33. https://doi.org/10.1016/j.dr.2016.06.005

Holmboe, K., Bonneville-Roussy, A., Csibra, G., & Johnson, M. H. (2018). Longitudinal development of attention and inhibitory control during the first year of life. *Developmental Science*, *21*(e12690). https://doi.org/10.1111/desc.12690

Holmboe, K., Larkman, C., de Klerk, C., Simpson, A., Bell, M. A., Patton, L., Christodoulou, C., & Dvergsdal, H. (2021). The early childhood inhibitory touchscreen task: A new measure of response inhibition in toddlerhood and across the lifespan. *PLoS ONE*, *16*(12 December). Scopus. https://doi.org/10.1371/journal.pone.0260695

Hughes, C., & Ensor, R. (2011). Individual differences in growth in executive function across the transition to school predict externalizing and internalizing behaviors and self-perceived academic success at 6 years of age. *Journal of Experimental Child Psychology*, *108*(3), 663–676. https://doi.org/10.1016/j.jecp.2010.06.005

Johansson, M., Marciszko, C., Gredebäck, G., Nyström, P., & Bohlin, G. (2015). Sustained attention in infancy as a longitudinal predictor of self-regulatory functions. *Infant Behavior and Development*, *41*, 1–11. https://doi.org/10.1016/j.infbeh.2015.07.001

Jones, E. J. H., Mason, L., Begum Ali, J., van den Boomen, C., Braukmann, R., Cauvet, E., Demurie, E., Hessels, R. S., Ward, E. K., Hunnius, S., Bolte, S., Tomalski, P., Kemner, C., Warreyn, P., Roeyers, H., Buitelaar, J., Falck-Ytter, T., Charman, T., Johnson, M. H., & Eurosibs Team. (2019). Eurosibs: Towards robust measurement of infant neurocognitive predictors of autism across Europe. *Infant Behavior & Development, 57,* 101316. https://doi.org/10.1016/j.infbeh.2019.03.007

Kannass, K. N., Oakes, L. M., & Shaddy, D. J. (2006). A longitudinal investigation of the development of attention and distractibility. *Journal of Cognition and Development*, *7*(3), 381–409. https://doi.org/10.1207/s15327647jcd0703_8

Lawson, G. M., Hook, C. J., & Farah, M. J. (2018). A meta-analysis of the relationship between socioeconomic status and executive function performance among children. *Developmental Science*, *21*(e12529). https://doi.org/10.1111/desc.12529

Lo, C. H., Rosslund, A., Chai, J. H., Mayor, J., & Kartushina, N. (2021). Tablet assessment of word comprehension reveals coarse word representations in 18–20-month-old toddlers. *Infancy*, *26*(4), 596–616. Scopus. https://doi.org/10.1111/infa.12401

Miyake A., Friedman N.P., Emerson M.J., Witzki A.H., Howerter A., Wager T.D. (2000). The unity and diversity of executive functions and their contributions to complex

"Frontal Lobe" tasks: a latent variable analysis. *Cognitive Psychology*, *41(1)*, 49-100. https://doi.org/10.1006/cogp.1999.0734

Mullen, E. (1995). *Mullen Scales of Early Learning*. American Guidance Service Inc.

Muñoz-Rocha, T. V., Tamayo y Ortiz, M., Romero, M., Pantic, I., Schnaas, L., Bellinger, D., Claus-Henn, B., Wright, R., Wright, R. O., & Téllez-Rojo, M. M. (2018). Prenatal co-exposure to manganese and depression and 24-months neurodevelopment. *NeuroToxicology*, *64*, 134–141. https://doi.org/10.1016/j.neuro.2017.07.007

Nakagawa, A., & Sukigara, M. (2013). Individual differences in disengagement of fixation and temperament: Longitudinal research on toddlers. *Infant Behavior and Development*, *36*(4), 728–735. https://doi.org/10.1016/j.infbeh.2013.08.001

Office for National Statistics. (2017). *Graduates in the UK labour market: 2017*. Office for National Statistics. https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/graduatesintheuklabourmarket/2017#steady-increase-in-the-number-of-graduates-in-the-uk-over-the-past-decade

Office for National Statistics. (2021). *Average household income, UK: financial year 2020*. Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/bulletins/householddisposableincomeandinequality/financialyear2020#analysis-of-average-income

Pitchford, N. J., & Outhwaite, L. A. (2016). Can touch screen tablets be used to assess cognitive and motor skills in early years primary school children? A cross-cultural study. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01666

Posner, M. I., & Rothbart, M. K. (2006). Research on attention networks as a model for the integration of psychological science. *Annual Review of Psychology*, *58*(1), 1–23. https://doi.org/10.1146/annurev.psych.58.110405.085516

Posner, M. I., Rothbart, M. K., Sheese, B. E., & Voelker, P. (2012). Control networks and neuromodulators of early development. *Developmental Psychology*, *48*(3), 827–835. https://doi.org/10.1037/a0025530

Sacrey, L.-A. R., Bryson, S. E., & Zwaigenbaum, L. (2013). Prospective examination of visual attention during play in infants at high-risk for autism spectrum disorder: A longitudinal study from 6 to 36 months of age. *Behavioural Brain Research*, *256*, 441–450. https://doi.org/10.1016/j.bbr.2013.08.028

Semmelmann, K., Nordt, M., Sommer, K., Röhnke, R., Mount, L., Prüfer, H., Terwiel, S., Meissner, T. W., Koldewyn, K., & Weigelt, S. (2016). U can touch this: How tablets can be used to study cognitive development. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01021

Siqueiros Sanchez, M., Pettersson, E., Kennedy, D. P., Bölte, S., Lichtenstein, P., D'Onofrio, B. M., & Falck-Ytter, T. (2020). Visual disengagement: Genetic architecture and relation to autistic traits in the general population. *Journal of Autism and Developmental Disorders*, *50*(6), 2188–2200. https://doi.org/10.1007/s10803-019-03974-6

Stephens, R. L., Langworthy, B., Short, S. J., Goldman, B. D., Girault, J. B., Fine, J. P., Reznick, J. S., & Gilmore, J. H. (2018). Verbal and nonverbal predictors of executive function in early childhood. *Journal of Cognition and Development*, *19*(2), 182–200. https://doi.org/10.1080/15248372.2018.1439493

Thompson, A., & Steinbeis, N. (2020). Sensitive periods in executive function development. *Current Opinion in Behavioral Sciences*, *36*, 98–105. https://doi.org/10.1016/j.cobeha.2020.08.001

Twomey, D. M., Ahearne, C., Hennessy, E., Wrigley, C., Haan, M. D., Marlow, N., & Murray, D. M. (2021). Concurrent validity of a touchscreen application to detect early cognitive delay. *Archives of Disease in Childhood*. https://doi.org/10.1136/archdischild-2019-318262

Twomey, D. M., Wrigley, C., Ahearne, C., Murphy, R., De Haan, M., Marlow, N., & Murray, D. M. (2018). Feasibility of using touch screen technology for early cognitive assessment in children. *Archives of Disease in Childhood*, *103*(9), 853–858. https://doi.org/10.1136/archdischild-2017-314010

Wass, S., Porayska-Pomsta, K., & Johnson, M. H. (2011). Training attentional control in infancy. *Current Biology*, *21*(18), 1543–1547. https://doi.org/10.1016/j.cub.2011.08.004

Willoughby, M. T., Piper, B., Kwayumba, D., & McCune, M. (2019). Measuring executive function skills in young children in Kenya. *Child Neuropsychology*, *25*(4), 425–444. https://doi.org/10.1080/09297049.2018.1486395

Yaari, M., Mankuta, D., Harel- Gadassi, A., Friedlander, E., Bar-Oz, B., Eventov-Friedman, S., Maniv, N., Zucker, D., & Yirmiya, N. (2018). Early developmental trajectories of preterm infants. *Research in Developmental Disabilities*, *81*, 12–23. https://doi.org/10.1016/j.ridd.2017.10.018