Applying gaze-contingent training within community settings to infants from diverse SES backgrounds

Ballieux, H., Wass, S., Tomalski, P., Kushnerenko, E., Karmiloff-Smith, A., Johnson, Mark H. and Moore, Derek G.

NOTICE: this is the authors' version of a work that was accepted for publication in Journal of Applied Developmental Psychology. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Journal of Applied Developmental Psychology, 43, 8-17, 0193-3973.

Journal of Applied Developmental Psychology is available online at:

https://dx.doi.org/10.1016/j.appdev.2015.12.005

Abstract

Even in infancy children from low SES backgrounds differ in frontal cortex functioning and, by the start of pre-school, they frequently show poor performance on executive functions including attention control. These differences may causally mediate later difficulties in academic learning. Here, we present a study to assess the feasibility of using computerized paradigms to train attention control in infants, delivered weekly over five sessions in early intervention centres for low-SES families. Thirty-three 12-month-old infants were recruited, of whom 23 completed the training. Our results showed the feasibility of repeat-visit cognitive training within community settings. Training-related improvements were found, relative to active controls, on tasks assessing visual sustained attention, saccadic reaction time, and rule learning, while trend improvements were found on assessments of short-term memory. No significant improvements were found in task switching. These results warrant further investigation into the potential of this method for targeting 'at-risk' infants in community settings.

Keywords: cognitive training; attention training; early development; socioeconomic status; community settings; infant.

Applying gaze-contingent training within community settings to infants from diverse SES backgrounds

Research suggests that, by the time children from low Socio-Economic Status (SES) backgrounds start school, they can show poor performance on a variety of measures of Executive Functions (EF). These include attention control, which can be defined as 'the capacity to choose what to pay attention to and what to ignore' (Blair & Razza, 2007; Razza, Martin, and Brooks-Gunn, 2010). It has been suggested that attention control can mediate learning and subsequent cognitive development across a range of domains (Cornish, Scerif, and Karmiloff-Smith 2007; Karmiloff-Smith, 1998) - including language acquisition (Rose, Feldman, and Jankowski, 2009), initiating and maintaining social interactions (Mundy, Sullivan, and Mastergeorge, 2009) and learning in academic settings (Welsh, Nix, Blair, Bierman, and Nelson, 2010; see also Scerif, 2010; Wass, 2014). It has even been suggested that EFs may play a protective role in development, such that children with good EFs are better able to compensate for atypicalities in other areas, making them less likely to receive a clinical diagnosis later in development (Johnson, 2012).

Research suggests that, in cases where we want to improve attention control, the earlier the intervention, the greater the potential to effect change (Wass, Scerif, and Johnson, 2012). Neural plasticity is thought to be greatest at very early stages of postnatal development (Huttenlocher, 2002; Spencer-Smith et al., 2011), consistent with how functional patterns of brain activation change with increasing age (Johnson, 2010). In terms of behaviour, Heckman similarly argued that plasticity is greater earlier in development (Heckman, 2006). He maintained that the mastery of skills needed for economic success follows hierarchical rules, with later attainments building on earlier ones (Karmiloff-Smith, 1998; Karmiloff-Smith et al., 2012; Sonuga-Barke, Koerting, Smith, McCann, and Thompson, 2011).

There is evidence that children from low-SES backgrounds, who often experience impoverished early environments or *in utero* exposure to toxic substances (e.g., drugs, alcohol), show reduced sustained attention and poorer attentional control (Hackman & Farah, 2009; Tomalski & Johnson, 2010), and that these difficulties may increase the likelihood of later negative outcomes such as ADHD (Noble, Norman, and Farah, 2005). In particular, being raised in a low-SES context contributes to poorer performance in visual attention and novelty detection tasks, accompanied by reduced prefrontal brain activity (Kishiyama et al., 2009). Disparities in SES have also been linked to differences in selective attention to speech, with children from low-SES families showing reduced ability to filter out irrelevant sound streams as well as a reduced response to attended sounds (D'Angiulli, Herdman, Stapells, and Hertzman, 2008; Kaldy & Blaser, 2013). SES-related differences in frontal gamma power have also been identified in infants as young as six months of age (Tomalski et al., 2013).

These findings suggest that it may be desirable to investigate the effect of early interventions to strengthen the early development of executive functions within low SES populations during infancy. Although a variety of parent- and teacher-mediated interventions are available for children of pre-school and upwards (Thompson et al., 2009), no behavioural techniques have yet been devised for providing training that is directly targeted at infants. In this case, our focus was on computer-mediated interventions, since these have a variety of potential practical advantages over parent- and teacher-mediated interventions. First, if found effective, they can potentially be run with minimal supervision, and in home settings, and therefore can be scaled up at a much lower cost than is possible with clinician-mediated interventions. Second, it is much easier to ensure that paradigms can be administered consistently across multiple sites. Third, more sensitive and rapid criteria can be devised to change task difficulty contingent on performance than is possible with human-mediated interventions. Computerised training techniques also have the long-term potential for integration with human-mediated interventions as part of a multi-component training battery.

A recent review evaluating computerised studies that trained EF across the lifespan reported that very little previous work has attempted to directly target these cognitive functions during infancy – whether in low-SES or other populations; these researchers did note, however, that those studies targeting younger participants tended to report more widespread transfer of training effects, suggesting the possible usefulness of targeting this age range (Wass et al., 2012; see also Diamond, Barnett, Thomas, and Munro, 2007; Wass, 2014). The absence of previous work targeting infant populations is partly due to the methodological challenges involved in recruiting and repeatedly testing children of this age range, and partly due to methodological difficulties in applying training.

Researchers working with infants face the problem of identifying a means by which the individual can interact with a computerised training paradigm, since fine motor skills and action planning are obviously poor at this age (Aslin, 2007). One solution is to use eye-gaze contingent control as the means by which the infant interacts with the training – by using eye-trackers to design training stimuli that change contingent upon where on the screen the infant looks. Using this interface in a lab-based context, Wass and colleagues administered a battery of tasks to typically developing 11-month-olds targeting interference resolution, inhibition, task switching, and working memory for objects embedded in scenes of varying complexity (Wass, Porayska-Pomsta, and Johnson, 2011). An average of 77 minutes of training was administered over four visits spread over 2 weeks, with the effects of training assessed relative to an active control group. Immediately post training, increased cognitive control and sustained attention were observed (Wass et al., 2011); attentional disengagement latencies and saccadic reaction time latencies were reduced following training, and a trend emerged in changes in looking behaviour during free play. No changes were found in working memory.

Thus, it seems that the practical problems of delivering a form of training that can engage infants in contingent training can be overcome by using eye-tracking and has clear potential. However, another problem in effectively applying this approach is that laboratory-based studies tend to recruit less diverse and less representative samples, consisting predominantly of infants from families with higher SES (Henrich, Heine, and Norenzayan, 2010). To surmount this, one potentially fruitful approach is to take eye-tracking equipment out of the lab and into child care centres that enrol primarily low-SES infants. In the UK, early intervention centres are called Sure Start Children's Centres (CCs), which were created for this purpose in 1998 in recognition of the importance of investing in universal early education, particularly for low-SES populations (Guidance, 2013). CCs are mostly found in low-income areas, with high indices of multiple deprivation (Government, 2010; Noble, Mclennan, and Wilkinson, 2010). They are closely linked with their communities, and specifically tasked with helping parents with children under five. The Sure Start programme is comparable to the American Head Start programme and, for example, to the Ontario Early Years Plan approach in Canada, the recently-created Biztos Kezdeta in Hungary, and to approaches recently created in Australia. CC environments, while not as controlled as lab-based testing settings, are likely to be better controlled than the home, with the researcher able to set up in a dedicated room in advance of testing, and to test in a single day a cluster of infants, under similar conditions.

In this study, we assessed whether training paradigms previously employed in lab settings could be successfully administered in CCs, within community settings. We had two goals:

1. To explore whether working in CCs facilitated recruitment of participants from diverse backgrounds, and whether weekly scheduling of training proved manageable for these parents and infants.

2. To test whether training effects observed in the lab could also be demonstrated in a community setting.

The study design was closely based on that used by Wass and colleagues (Wass et al., 2011). The training stimuli used, and three of four pre-post assessments, were identical to those used in this lab-based study. Participating parents, and infants, attended sessions once a week as part of a scheduled weekly drop-in. As with the previous lab-based study, approximately half of the infants in the current study underwent training. The other half was assigned to an active control group, who attended an equal number and duration of sessions, but instead of training watched infant-appropriate animations and TV clips on the eye-tracker monitor. Assessment of training effects was analysed in-task. Transfer of training effects was also assessed using pre-post assessments involving non-trained tasks examining aspects of attention control – namely visual sustained attention, saccadic reaction time latencies, attentional disengagement latencies, anticipatory saccades during rule learning, and short-term memory. Although these tasks differ in task paradigm, their unifying feature is that they all require infants to exercise endogenous (effortful) control over the focus of their visual attention (Rueda, Rothbart, McCandliss, Saccomanno, and Posner, 2005; see also Colombo & Cheatham, 2006). We predicted that, as in the previous lab-based study, training attention control would lead to an improvement in performance on these non-trained attention control tasks from pre- to post-training periods.

Methods

Participants

Infants were recruited by CC staff through phone calls, flyers, and advertisement of our ‘Learn about your baby’ sessions in their quarterly calendars. Parents were either contacted by CC staff, or contacted the Centre or researcher directly, to book an appointment (further details on the set-up and recruitment in CCs are given in Ballieux et al., in press). Parents were made aware that they needed to attend sessions during five consecutive weeks, and were asked whether it was likely that they would be able to attend all sessions. All parents who participated confirmed that they would be able to.

In total, 33 infants were enrolled in the programme. The inclusion criteria were: age range 11 months 0 days to 12 months 30 days, no pre-term infants, no major medical conditions, and no major delivery complications. Of these 33 enrolled infants, 8 dropped out after one session and a further two after two sessions. Reasons for drop-out included equipment failure (2 infants), sickness in the family (1 infant), and lack of parental engagement in the programme (5 infants). Of the infants included in the final study, 6 trained and 5 control infants completed all five planned visits, 2 trained and 7 control infants completed four visits, and 2 trained and 1 control infants completed three visits, making a total of 10 trained and 13 control infants. Gender ratios were 5 male/5 female for the training group, and 5 male/8 female for the control group. Mean ages (with standard deviations in parentheses) for the trained and control groups were 347 (14.3) and 362 (17.6) days, respectively.

Mean gross household income per year was £42478 (Median = £21404; SD = £51193; range from £6000 to £212500; N = 20 because not all parents gave or knew their own or their partner's income, or preferred not to say). Taking the UK government definition of poverty as earning less than 60% of median gross household income (i.e., earning below £13,920 per year), 8 of the 20 households in our sample live below the poverty line (2 of these households were comprised of single mothers), with an additional 4 living below or on the median gross household income level (some of the low-SES professions were cleaner, shop assistant, nursery assistant, postman, security staff, and hotel cloak room staff). The large range of our sample further confirms the interesting demographic make-up in East London, with very high and very low SES living in the same area, using the same facilities (some of the high-end professions were banker, computer programmer, director homeless charity, manager of a train company, accountant, and retail manager). Table 1 shows additional demographic information, with a mean age in days and mean birth weight in grams (standard deviations in parentheses) for the total sample of 352 (14.8) and 3100.5 (573.8), respectively.

Study protocol.

All pre-post, training and control sessions were conducted in quiet rooms that were made available within the CCs (see Ballieux et al., in press). The researcher visited each CC once a week. Prior to their first visit, children were randomly allocated to either trained or control groups. This was performed based on recruitment order, and before the researcher had met the families in person.

Materials and Procedures

Testing equipment consisted of a Tobii T120 eye-tracker with a 17'' monitor, a portable and adjustable Ergotron MX desk mount arm, and a MacBook Pro laptop. The eye-tracker was positioned directly facing the child. The experimenter sat with the laptop, behind a screen or barrier, out of sight of the infant. All testing and training materials were administered via MATLAB and Psychtoolbox. Of note, in this pilot study the same researcher conducted all testing sessions, and therefore was not blinded to group allocation. Note, however, that the paradigms were self-determined, with performance contingent on infant behaviour, and not experimenter responses.

At visit 1, all children underwent the pre-test battery, which lasted for approximately 20 minutes. This was followed immediately by the first training session. At visits 2-4, children participated in either training or control sessions. Each training session lasted until the child no longer engaged with the materials presented. Control sessions were matched, participant-by-participant and visit-by-visit, so that they were the same length as the training session for the yoked participant. At visit 5, all participating children conducted the post-test battery, which was identical to the pre-test battery.

Training.

All four training tasks used were presented at each training session, until the participant no longer engaged with the task. Mean time spent engaged with the training tasks (with standard deviations in parentheses) for Visits 1 through 4 were 10.0 (5.1), 20.5 (5.6), 19.3 (5.5), and 24.6 (11.5) Min, respectively. Of note, these are approximately equivalent to the average training times observed in the previous, lab-based study, where the equivalent mean training times for Visits 1 through 4 were 6.1, 22.9, 25.1, and 22.7 min (Wass et al., 2011). The mean playing time in seconds (with standard deviations in parentheses) for Tasks 1 through 4 (Butterfly, Stars, Windows, Suspects) were 203 (113), 263 (93), 158 (134), and 289 (83), respectively.

Of those infants completing the study, on average 33.7 days elapsed between the first and last training sessions (Training group = 33.8 days, SD = 5.2; Control group = 33.0 days, SD = 7.2). This is in contrast to the previous (lab-based) study, where the same number of training sessions was completed twice per week and the average interval between first and last testing sessions was 15 days. The training battery consisted of four different training tasks (see Figure 1):

Task 1 (Butterfly). A target (a butterfly subtending 6˚) was presented on the screen. When the child fixated the target, the butterfly 'flew' across the screen, and distractors (a house, a tree, clouds, subtending 5-15˚) scrolled in the opposite direction. When the child looked to any of the distractors they disappeared and only the target, now static, remained on screen. On re-fixating the target it re-commenced moving and the distractors re-appeared and continued scrolling. The salience of the distractors changed adaptively, including faster, larger and more densely packed objects. This task rewards a child for maintaining their fixation on one target, and suppressing the prepotent response to look towards moving distractors in the periphery.

Task 2 (Stars). One of five possible targets (each cartoon characters in brightly coloured stars, subtending 6˚) was presented on screen together with eight distractors (smaller stars, planets, clouds, subtending 4-8˚) against a detailed still image as background. If the infant looked to the target within 3000 ms he or she received an animation as a reward. The target changed from trial to trial. The salience of the distractors changed adaptively. At lower difficulty levels, the eight distractors were smaller, static, and identical to each other and dissimilar from the targets. At higher difficulty levels, they were more varied, moving, brightly coloured, and similar to the targets.

Task 3 (Windows). When the infant fixated the target (an animal in a window subtending 7˚), an animation showed the target disappearing into one of several windows that were then covered with curtains. A fixation target (a flower subtending 4.5˚) appeared elsewhere on the screen and rotated when the infant looked at it. After a delay period, the fixation target disappeared. If the infant looked back to the window behind which the target had disappeared, he or she received an animation as a reward. The number of windows, the salience of the distractors, and the length of the delay changed adaptively. This task trained visuospatial working memory and required acting on stored information about objects embedded in complex scenes.

Task 4 (Suspects). One of two possible targets (either an elephant or a chicken subtending 4.5-8°) was presented with one or more distractor items of the same size. When the infant looked at the target within a time limit, he or she received an animation as a reward. The same target was then re-presented with other distractor(s). The number of distractors varied adaptively with performance; at higher performance levels, more distractors were presented. Between blocks of 12 trials, the target changed: where previously the child had received a reward for looking to the elephant, he or she now was rewarded for looking to the chicken. At higher difficulty levels, the target from the previous block was presented concurrently with the target from the current block (a conflict trial); at lower difficulty levels, only novel distractors were presented (non-conflict). This task targets attention shifting and flexible search for changing targets, whilst ignoring distractors.

Control stimuli.

Control sessions were conducted in the same room, with the same experimenters and using the same eye-tracker as the training sessions, and had the same duration and spacing (yoked to a trained participant). Instead of training, control participants viewed a selection of infant-friendly TV clips and still images. These were identical to those used previously (Wass et al., 2011).

Pre-post tests. In order to assess transfer of training effects, the following pre-post tasks were presented at visits 1 and 5, identically to infants in the trained and control groups. Figure 2 shows schematics of these tasks. The tasks were presented interleaved in order, in a battery that lasted approximately 20 minutes in total. In order to maintain engagement during testing, a number of short clips from TV programs were also presented between experimental blocks. The order in which blocks were presented was pseudo-randomised, with the constraint that no two blocks of the same experiment could be presented contiguously.

Sustained attention. Four different still images were presented: two per block in two blocks. Each block contained one of two ‘interesting’ images (i.e., attractive, detailed images of flowers and fish) and one of two ‘boring’ images (i.e., low-detail, monochrome outlines of a diamond and a cross; see Figure 2a). Trials commenced once the subject had fixated a central target. Trials ended when the subject had looked away from the screen for 1 second, as judged by an experimenter, or when 15 seconds had elapsed. At the end of each trial, a fixation target and brief auditory stimulus (< 1s) were presented. If the subject fixated the target, the next trial started immediately; if not, a sequence of different fixation targets and auditory attention getters was repeated. Stimuli were re-presented until: two successive looks were less than 50% of the longest unbroken look so far, eight successive looks had taken place without reaching criterion, or the total presentation length exceeded 120 seconds. One infant (control, post-test) failed to provide usable data for this task.

Gap-Overlap. This task was presented in three blocks. The first two blocks lasted 20 trials; the third continued until either enough valid trials had been collected (12 usable trials per condition) or 80 trials had been presented, or the child became inattentive. After fixating a central target (CT, a cartoon flower subtending 4.5˚), following a variable ISI a lateral target (LT, a cartoon cloud subtending 3˚) was presented to the left or right; when the participant fixated the LT he or she received a brief audiovisual reward. Three conditions were presented: Gap - CT disappears 200 ms before LT appears; Baseline - CT disappears concurrently with LT appearance; Overlap - CT remains onscreen with LT appearance. The order of trials was randomised between conditions. The reaction time (RT) was the time elapsed between LT appearance and the reported position of gaze leaving the central fixation area (a 9˚ box around the CT). Reaction times less than 100 and greater then 2000 ms were excluded. Participants from whom fewer than 10 usable trials per condition were obtained were excluded from further analyses. Average reaction times were calculated by first averaging the reaction times obtained across the three separate conditions, and then combining the log transformed averages to create a final average. Disengagement latencies were calculated as the participant's average reaction time in the overlap condition subtracted from their average reaction time in the baseline condition (following Elsabbagh et al., 2009). A number of infants (3 test group pre-test, 2 test group post-test, 5 control group pre-test, 9 control group post-test) failed to provide 10 usable trials per condition and so were excluded. This drop-out rate is slightly higher than previous studies that have a recorded a drop-out rate of 18% (Wass et al., 2011).

Cognitive Control. This task was presented in two blocks, each lasting 18 trials. After fixating a central target (a cartoon flower subtending 4.5˚), the trial commenced following a 300 ms delay. Two blank rectangles (10.8˚ x 9˚) were presented left and right, concurrently with an auditory stimulus for 2000 ms (the anticipatory window). A visual reward (lasting 4000 ms) then appeared on one side, in either the left or right rectangle, for nine trials in a row (the pre-switch phase) before swapping sides for the next nine trials (the post-switch phase). If the participant correctly anticipated the presentation of the reward, defined as a saccade beginning between 300 and 2300 ms after trial onset and subject to a minimum look duration of 400ms, then the visual reward stimulus appeared immediately. The outcome measure was proportion of correct anticipatory looks. Results for the initial, pre-switch phase measure initial rule learning and for the subsequent, post-switch phase assess task switching. One infant (C, pre-test) failed to provide any usable trials and so was excluded.

Short-term memory (STM). Unlike the three previous tasks, a different task was used to assess short-term memory from that used in the previous study (Wass et al., 2011). This is because analyses revealed that the test-retest reliability of the working memory task used in that study was low (r = .16) in comparison to test-retest reliabilities for the sustained attention, gap-overlap, and cognitive control tasks (r = .75, r = .56, and r = .60, respectively; Wass, 2011). Therefore we sought a more reliable STM assessment (Kaldy & Blaser, 2013), which was presented in two blocks, each consisting of seven trials. In each trial, two targets (each subtending 5˚) were presented for 6500 ms. Two separate occluders then appeared and covered the objects for 2500 ms. The occluders then revealed the objects; one of the objects had changed colour. The two objects were then presented for 7000 ms. The dependent variable was whether the first look was to the side where the colour of the target had changed following the occlusion period, or to the side where the target was the same colour as prior to the occlusion (following Kaldy & Blaser, 2013). The location of the change side varied between trials. Participants were excluded if fewer than 3 usable trials were obtained. Four control participants (2 pre-test, 2 post-test) failed to provide a sufficient number of trials and so were excluded.

Results

The difficulty level changed adaptively during training in response to participants' performance. We first examined our data for training effects. Where observed results are in the same direction as those observed previously, and are therefore consistent with predictions, one-tailed p-values have been used.

Observed Changes During Training

First, we assessed whether changes in performance on the training tasks were observed across the four training sessions administered in the current, CC based study (see Figure 3). Linear regression lines were calculated based on change in performance across the four training sessions. The gradients of these lines were positive for eight out of ten infants, suggesting that they improved across the training sessions. This change was consistent with that predicted, based on previous research (Wass et al., 2011). A t-test analysis suggested that the regression lines differed significantly from chance (t(9) = 2.25, p = .025, one-tailed). This finding establishes a measurable effect of training.

Pre-Post Assessments

Data quality comparison.

First we wished to evaluate the quality of raw eye-tracking data obtained on this trial. To do this we compared data obtained in this study with data obtained in a previous study, run in lab settings, using a Tobii 1750 eye-tracker (which is an older eye-tracker model than the Tobii T120 used in the present trial). The previous study (Wass et al., 2011) used typical 12-month-old infants. The comparison was conducted on raw data collected during the administration of the gap-overlap experiment; the experimental protocols and visual materials for this task were identical across the two studies. Data quality evaluations were calculated using techniques described in detail in Wass and colleagues (Wass, Forssman, and Leppanen, 2014).

Two measures of data quality have been calculated. First (Fig 4a and 4d), the robustness of tracking was quantified by calculating the duration (in seconds) of usable fragments of eye-tracking data obtained during recording. As is universally the case during remote eye-tracking, as described in detail elsewhere (Leppänen, Forssman, Kaatiala, Yrttiaho, and Wass, 2014; Wass et al., 2014), we found that contact with the eye-tracker tended to 'flicker' on and off during recording – most likely due to the fact that, in some samples, certain elements of the information required to calculate the infant's position of gaze (pupil, corneal reflection, and the position of the head in 3D space) were unavailable or just insufficiently robust by the image processing algorithms built into the eye-tracker, leading to null values being returned. This data tends to 'flicker' on and off at short periods (often < 100ms), which confirms our impression from video coding comparisons that this is not due to the infant looking to and from the eye-tracker. In order to calculate this, therefore, the average duration of data fragments was calculated. A low number indicates more 'flickery' (i.e., less robust) data. Shorter usable fragment durations were obtained in this study relative to the comparison study. The Standard Error Mean (S.E.M.) was 2.0 s (SD = .18 s) in the present study, and 3.7 s (SD = .26 s) seconds in the comparison study, which an independent samples t-test confirmed was a significant difference (p < .001). Second (Figures 4b and 4e), the precision of tracking was calculated by quantifying the degree to which reporting of position of gaze is consistent between samples. A higher value indicates that data obtained were less precise. Markedly less precise data were obtained in the present study relative to the previous study. The S.E.M. was 4.0e-03 s (SD = .2e-03 s) in the present study, and 3.2e-03 s (SD = .1e-03 s) in the comparison study (internal units), which an independent samples t-test confirmed was a significant difference (p = .02). Visual inspection of figures 4c and 4f confirm that average eye-tracking data obtained during this task in the community sample was also more widely dispersed over the screen than in the lab sample which, given that the tasks were identical across the two paradigms, suggests that spatial accuracy may also have been lower in the present study (Dimitrov & Rumrill, 2003; Wass et al., 2014).

Pre-post tests.

Table 2 shows the both the raw and marginal (baseline-corrected) means and standard errors for the variables gathered from the four pre-post tasks together with estimates of Cohen's d. As with previous studies, marginal means are considered more accurate estimates of effect sizes since they correct for differences in performance on certain measures that we observed at pre-testing. Therefore values of Cohen's d were calculated from the marginal means. We conducted analyses of covariance (ANCOVA) with the factor group (trained versus control), post-test scores as the dependent variable, and pre-test scores as the covariate. This is equivalent to an ANCOVA on the difference scores with pre-test as a covariate (Dimitrov & Rumrill, 2003). Prior to conducting the ANCOVA, Kolmogorov-Smirnov tests were conducted to ensure that variables being entered into the analysis showed distributional properties that did not differ significantly from normal. In each case this was found to be the case (all K-values < .276, all p-values > .07). Figure 3 shows the change scores for the pre-post assessments, calculated from the marginal means. As a follow-up analysis, individual repeated measures ANOVAs were also conducted on the results obtained from each task for the trained and control groups (following Rueda et al., 2005).

Sustained attention. An ANCOVA revealed a trend increase in peak look duration to 'Interesting' stimuli following training, corresponding to the predicted effect, and that reported in the previous study (F(21) = 2.16, p = .08). Cohen's d was found to be 0.69, indicating a medium-sized effect, albeit with a small sample size. As a follow-up analysis, repeated measures ANOVAs with two-tailed significance levels were conducted independently on the results of the two groups. These identified a significant increase in sustained attention to the interesting stimuli at post-test in the trained group (F(1,9) = 7.81, p = .01), which was not found in the control group (F(1,11) = 1.47, p = .25). No changes as a result of training were found for look duration to 'Boring' stimuli (F(21) = .001, p = .98). A two-way repeated measures ANOVA was also conducted with group (Trained versus Control) and condition (Boring versus Interesting) as independent variables and sustained attention as dependent variable. A significant interaction between group and condition was identified (F(1,21) = 5.21, p = .028, two-tailed). This suggests that training had the effect of increasing look duration in the Interesting condition significantly more than in the Boring condition.

Gap-overlap. An ANCOVA revealed a significant decrease in Average RT in the Trained relative to the Control group at post-test (F(1,8) = 6.67, p = .024, one-tailed). An ANCOVA revealed no significant change in Disengagement Latencies in the Trained relative to the Control group (F(1,7) = .61, p = .23, one-tailed). Follow-up analyses with repeated measures ANOVAs were also not significant for Disengagement Latencies.

Cognitive control. An ANCOVA revealed a significant increase in proportion of correct anticipatory looks in the pre-switch, initial rule learning phase in the Trained relative to the Control group (F(21) = 4.53, p = .024, one-tailed). No effect of training was identified in the subsequent, post-switch phase, which assesses task switching (F(21) = .005, p = .95). Follow-up analyses with repeated measures ANOVAs were also not significant for the post-switch phase.

Short-term memory. An ANCOVA revealed an increase approaching significance in proportion of correct first looks to the change side in the Trained group relative to the Control group at post-test (F(19) = 2.25, p = .076, one-tailed). Follow-up analyses with repeated measures ANOVAs were not significant.

Discussion

The first aim of the study was to explore whether working in CCs allowed us to recruit participants from diverse backgrounds, and whether scheduling of training delivered over a five-week period in CCs proved manageable for parents. The current CC sample was as diverse in terms of ethnic background as the sample of Ballieux and colleagues (Ballieux et al., in press), which was also recruited in CCs. This confirms that working with CCs is beneficial for recruiting a diverse sample of participants. It should be noted, however, that based on the demographic data we collected, not all of our sample could be classified as low-SES – which is a challenge for future work of this sort, based in CCs. A further challenging aspect with the current, multiple visits training programme in the CCs was that completion rates were relatively low, with 23 out of 33 infants completing, compared to 98% (41 out of 42) in the lab-based study of Wass and colleagues (Wass et al., 2011). In devising the study we anticipated that participants would find a weekly training session over five weeks more convenient than a more intensive twice-a-week lab based training period. However, attendance was not as consistent in this study as in the previous one. It is not clear, however, whether more intensive scheduling would have been more effective for this population.

The lower completion rate may have a number of origins. One issue may have been that parents were less motivated to attend. Generally speaking it is harder to recruit and retain parents from more diverse SES backgrounds for lab studies and this would be expected to influence attendance in CCs also. Secondly, as we were restricted in being able to advertise any potential benefits of the training process, we could not use these potential benefits as an additional motivator to attend. In future studies, parents may need to be made more aware before signing up to the training programme that the training may have potential benefits and that it is essential that they attend every session. This could be done with the help of CC staff, who could explain to parents that these sessions are different from the regular optional CC sessions. A third issue concerned the particular time of year in which our testing took place, which may have been a factor in drop-out rates. Over the summer holiday period it appears more difficult for parents to commit to an unbroken five-week period of attendance, whereas over the winter parents tended to be more likely to attend.

Given these limitations, the fact that we still managed to recruit 33 infants from diverse backgrounds, and had as many as 23 complete the training programme, is encouraging for future training studies in community settings. Provided that parents are made aware not only of the importance of the training programme and what benefits might be, but also of the importance of attending every session, training of visual attention in a CC setting is not only possible, but expands our range of assessments of at-risk infants, especially those from low SES backgrounds.

A less common but important reason why some sessions were not completed was because of technical difficulties encountered during eye-tracking in the CCs. The data quality comparison shown in Figure 4 suggests that lower-quality tracking data were obtained in the current study relative to an equivalent, lab-based study that used an older eye-tracker from the same manufacturers. Data were found both to be less robust and less precise, using metrics from Wass et al. (2014), most likely due to the presence of other light sources in the room disrupting tracking. It was not possible to control the lighting within the room as precisely in the CCs as in the lab-based study. Of note, this has affected the pre- and post-test tasks such as the gap-overlap task, in which higher rates of data loss were encountered than previously. Note, however, that this is unlikely to be critical in determining the effectiveness of the training tasks, which are not as sensitive to data quality. Moreover, overall for those infants who did complete the study, training times were approximately equivalent to those obtained in the previous, lab-based study.

The second aim of the study was to see what training effects could be achieved in a CC setting. Most encouragingly, our main analyses showed that training did produce several improvements from pre- to post-test. First, we found that training in CCs led to an increase in sustained attention to 'Interesting' targets, and that no change was found after training in looking time to 'Boring' targets. This suggests that the effect of training was not simply an overall increase in looking time to the screen. Rather, it is consistent with a model suggesting that training attention leads to increased top-down, selective attention control. Second, average reaction time on the gap-overlap task was significantly improved following training. By contrast, no changes were found on disengaging visual attention. However, this may have been to do with the high rates of data drop-out on this task, due to low eye-tracking data quality. Of note, substantially longer reaction times were obtained on this task compared to lab-based versions of the same task, a pattern that is predicted by lower data quality (Leppänen et al., 2014; Wass et al., 2014). Third, improvements emerged on a task assessing anticipatory saccades during a rule-learning task, but during the initial rule learning phase only. See Supplementary Materials for a further discussion of this point. Fourth, trend improvements were found on a task assessing short-term memory. This is in contrast to the previous study, which used a different assessment, in which no improvements were found. In separate investigations (Wass, 2011) we found test-retest reliability of the task that we previously used to be very low, and it may be that memory training effects can be detected when a more sensitive measurement paradigm is used.

In summary we can conclude that setting up a training programme for attention control in CCs is possible and helps with recruiting a more diverse sample of typically developing as well as potentially at-risk infants. Moreover, in terms of task performance, broadly consistent patterns can be observed across identical tasks administered in very different settings. We also noted a number of similar training improvements to those found in the original study. Given the small sample and the small dose of training administered in the current study, and given that training sessions were administered over a 5-week period, as opposed to a 2-week one as in the previous study, we believe that the broadly similar patterns of findings reported here are encouraging and open the prospect of further trials in this area. Future work should investigate the effect of administering similar training paradigms within home settings – a fact that sadly is particularly necessary given that many of the CCs in which the present study was based have since been cut.

References

Aslin, R. N. (2007). What's in a look? *Developmental Science, 10*(1), 48-53.

Ballieux, H., Tomalski, P., Karmiloff-Smith, A., Kushnerenko, E., Johnson, M. H., & Moore, D. G. (in press). Feasibility of undertaking infant eye-tracking assessments in Children’s Centres. *Infant & Child Development*.

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78*(2), 647-663.

Colombo, J., & Cheatham, C. L. (2006). The emergence and basis of endogenous attention in infancy and early childhood. *Advances in Child Development and Behavior, 34*, 283-322.

Cornish, K., Scerif, G., & Karmiloff-Smith, A. (2007). Tracing syndrome-specific trajectories of attention across the lifespan. *Cortex, 43*(6), 672-685.

D'Angiulli, A., Herdman, A., Stapells, D., & Hertzman, C. (2008). Children's event-related potentials of auditory selective attention vary with their socioeconomic status. *Neuropsychology, 22*(3), 293-300.

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). The early years - Preschool program improves cognitive control. *Science, 318*, 1387-1388.

Dimitrov, D. M., & Rumrill, P. D. (2003). Pretest-posttest designs and measurement of change. *Work, 20*(2), 159-165.

Elsabbagh, M., Volein, A., Holmboe, K., Tucker, L., Csibra, G., Baron-Cohen, S., Bolton, P., Charman, T., Baird, G., & Johnson, M. (2009). Visual orienting in the early broader autism phenotype: disengagement and facilitation. *Journal of Child Psychology & Psychiatry, 50*(5), 637-642.

Government, D. f. C. a. L. (2010). *English Indices of Deprivation.*

Guidance, S. S. C. s. C. S. (2013). Department for Education. http://www.gov.uk/government/publications/sure-start-childrens-centres.

Hackman, D. A., & Farah, M. J. (2009). Socioeconomic status and the developing brain. *Trends in Cognitive Sciences*, *13*(2), 65-73.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*(5782), 1900-1902.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*(33), 61-135.

Huttenlocher, P. (2002). *Neural plasticity: The effects of environment on the development of cerebral cortex.* Cambridge, MA: Harvard University Press.

Johnson, M. H. (2010). *Developmental Cognitive Neuroscience, 3rd Ed.* Oxford, UK: Wiley-Blackwell.

Johnson, M. H. (2012). Executive function and developmental disorders: the flip side of the coin. *Trends in Cognitive Sciences, 16*(9), 454-457.

Kaldy, Z., & Blaser, E. (2013). Red to green or fast to slow? Infants' visual working memory for "just salient differences". *Child Development, 84*(6), 1855-1862.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences, 2*(10), 389-398.

Karmiloff-Smith, A., D'Souza, D., Dekker, T. M., Van Herwegen, J., Xu, F., Rodic, M., &

Ansari, D. (2012). Genetic and environmental vulnerabilities in children with neurodevelopmental disorders. *Proceedings of the National Academy of Sciences of the United States of America*, *109,* 17261-17265.

Kishiyama, M. M., Boyce, W. T., Jimenez, A. M., Perry, L. M., & Knight, R. T. (2009). Socioeconomic Disparities Affect Prefrontal Function in Children. *Journal of*

Cognitive Neuroscience, 21(6), 1106-1115.

Leppänen, J., Forssman, L., Kaatiala, J., Yrttiaho, S., & Wass, S. (2015). Widely applicable MATLAB routines for automated analysis of saccadic reaction times in infants. *Behavior Research Methods. Vol 47(2),* 538-548.

Mundy, P., Sullivan, L., & Mastergeorge, A. M. (2009). A parallel and distributed-processing model of joint attention, social cognition and autism. *Autism Research, 2*(1), 2-21.

Noble, K. G., Norman, M. F., & Farah, M. J. (2005). Neurocognitive correlates of socioeconomic status in kindergarten children. *Developmental Science, 8*(1), 74-87.

Noble, M., Mclennan, D., & Wilkinson, K. (2010). The English Indices of Deprivation 2010.

.

Retrieved from http://eprints.ioe.ac.uk/2461/.

Razza, R. A., Martin, A., & Brooks-Gunn, J. (2010). Associations among family snvironment, sustained attention, and school readiness for low-income children. *Developmental Psychology, 46*(6), 1528-1542.

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2009). A cognitive approach to the development of early language. *Child Development, 80*(1), 134-150.

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., & Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences of the United States of America, 102*(41), 14931-14936.

Scerif, G. (2010). Attention trajectories, mechanisms and outcomes: at the interface between developing cognition and environment. *Developmental Science, 13*(6), 805-812.

Sonuga-Barke, E. J. S., Koerting, J., Smith, E., McCann, D. C., & Thompson, M. (2011). Early detection and intervention for attention-deficit/hyperactivity disorder. *Expert*

Review of Neurotherapeutics, 11(4), 557-563.

Spencer-Smith, M., Anderson, P., Jacobs, R., Coleman, L., Long, B., & Anderson, V. (2011). Does Timing of Brain Lesion Have an Impact on Children's Attention? *Developmental Neuropsychology, 36*(3), 353-366.

Thompson, M. J. J., Laver-Bradbury, C., Ayres, M., Le Poidevin, E., Mead, S., Dodds, C., . . . Sonuga-Barke, E. J. S. (2009). A small-scale randomized controlled trial of the revised new forest parenting programme for preschoolers with attention deficit hyperactivity disorder. *European Child & Adolescent Psychiatry, 18*(10), 605-616.

Tomalski, P., & Johnson, M. H. (2010). The effects of early adversity on the adult and developing brain. *Current Opinion in Psychiatry, 23*(3), 233-238.

Tomalski, P., Moore, D. G., Ribeiro, H., Axelsson, E. L., Murphy, E., Karmiloff-Smith, A., . . . Kushnerenko, E. (2013). Socioeconomic status and functional brain development - associations in early infancy. *Developmental Science, 16*(5), 676-687.

Wass, S., Porayska-Pomsta, K., & Johnson, M. H. (2011). Training attentional control in infancy. [Research Support, Non-U.S. Gov't]. *Current biology, 21*(18), 1543-1547.

Wass, S. V. (2011). Gaze-contingent concentration training for infants. *Unpublished PhD thesis, Birkbeck, University of London*.

Wass, S. V. (2014). Applying cognitive training to target executive functions during early development. *Child Neuropsychology, 21(2),* 150-166.

Wass., S. V., Forssman, L., & Leppanen, J. (2014). Robustness and precision: how data quality may influence key dependent variables in infant eye-tracker analyses. *Infancy, 19*(5), 427-460.

Wass., S. V., Scerif, G., & Johnson, M. H. (2012). Training attentional control and working memory: is younger, better? *Developmental Review, 32*(4), 360-387.

Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., & Nelson, K. E. (2010). The development of cognitive skills and gains in academic school readiness for children from low-income families. *Journal of Educational Psychology, 102*(1), 43-53.

Table 1

Demographic Data for the Children's Centre Sample in the Present Study (N = 23)

Gender	Female	Male
	57%	43%
Ethnicity	White	Non-white
	21.7%	78.3%
Parent education level (%)	Mother	Father
Postgraduate	9.1	5.0
Higher Education	36.4	55.0
Further Education	27.3	25.0
High School A-levels	0.0	5.0
GCSE	13.6	10.0
No qualification	13.6	0.0

Table 2

Raw and Marginal Means for Pre-post Assessments (values for Cohen's d have been calculated based on the marginal means).

	Means (SE)								Marginal means (SE)								
	Trained				Control				Trained				Control				Cohen's d
	Pre-test		Post-test		Pre-test		Post-test		Pre-test		Post-test		Pre-test		Post-test		
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	
A Sustained attention: 'interesting' static - peak look duration (s)	29.4	5	56.6	8	26.1	6	36.9	9	27.2	9	55.9	9	27.2	9	37.4	9	0.69
A Sustained attention: 'boring' static - peak look duration (s)	29.3	12	26.3	9	16.3	3	19.0	7	22.1	7	22.2	7	22.1	6	22.3	6	0.00
B Gap-overlap task: Avg RT (ms)	468	19	428	13	443	11	491	31	461	24	426	24	461	41	519	41	1.99
B Gap-overlap task: Disengagement latencies (ms)	269	49	239	49	246	39	304	76	259	52	248	52	259	98	294	98	0.35
C Cognitive control: pre-switch (proportion correct anticipatory looks)	0.54	0.10	0.71	0.09	0.39	0.09	0.43	0.08	0.46	0.10	0.70	0.10	0.46	0.09	0.42	0.09	0.88
C Cognitive control: post-switch (proportion correct anticipatory looks)	0.36	0.09	0.35	0.10	0.24	0.06	0.28	0.07	0.29	0.09	0.32	0.09	0.29	0.08	0.33	0.08	-0.03
D Short-term Memory: proportion of first look to changed target (s)	0.53	0.04	0.57	0.04	0.53	0.05	0.51	0.04	0.53	0.04	0.57	0.04	0.53	0.04	0.49	0.04	0.58

Figures

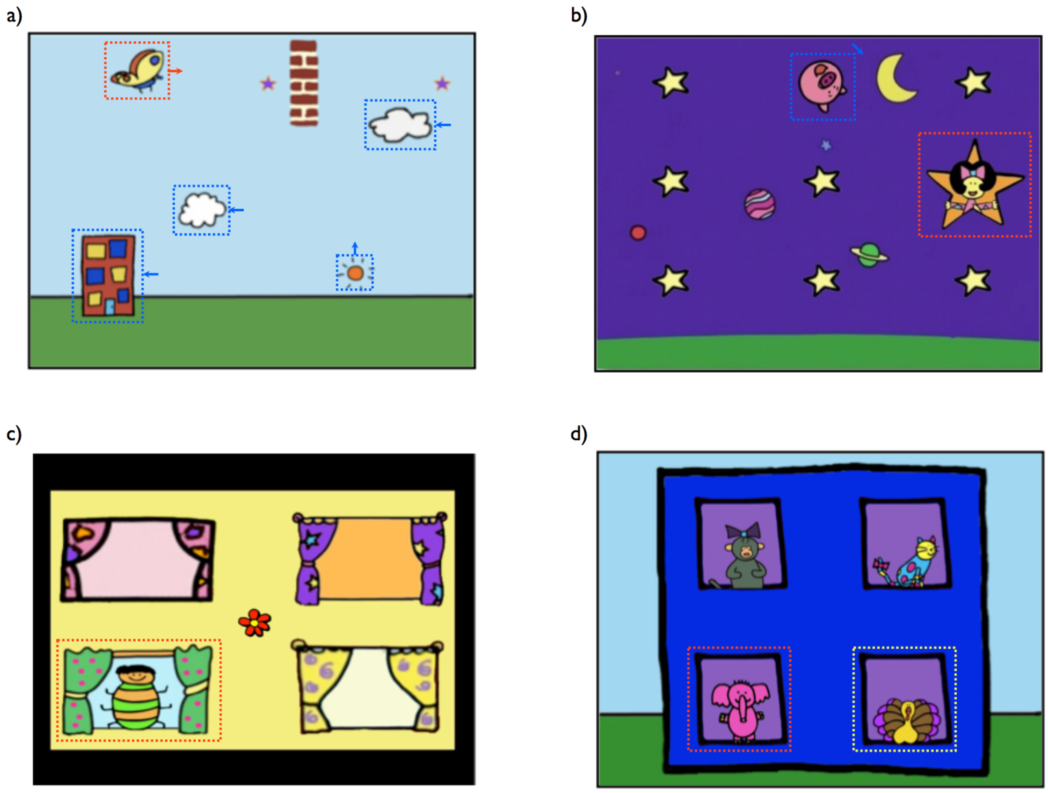

Figure 1: Schematics of the four training tasks administered. Dashed rectangles indicate active areas and arrows indicate objects that were moving on-screen (both were not visible in the original materials); a) Task 1 (Butterfly): the butterfly (indicated in red) scrolled from left to right as long as the child looked directly at it, with static and moving (indicated in blue) distractors presented in the child’s peripheral visual field. If the child looked to any of the distractors, they disappeared and the scrolling stopped; b) Task 2 (Stars): a target (indicated red) was presented on-screen along with a number of static and moving (indicated blue) distractors. If the child looked to the target within a time window, he or she received a reward. Both target and distractors changed between trials; c) Task 3 (Windows): a target (indicated red) was presented in one location on screen. All four windows then closed and fixation target (the red flower) appeared for a variable inter-stimulus interval. After the fixation target disappeared, a look back to the cued window triggered a reward; d) Task 4 (Suspects): a target (indicated red) was presented along with a range of distractors. If the child looked to the target within a time window, he or she received a reward. Once per block of 12 trials the target changed. Targets from the previous block (indicated yellow) were presented concurrently with the current target, as distractors.

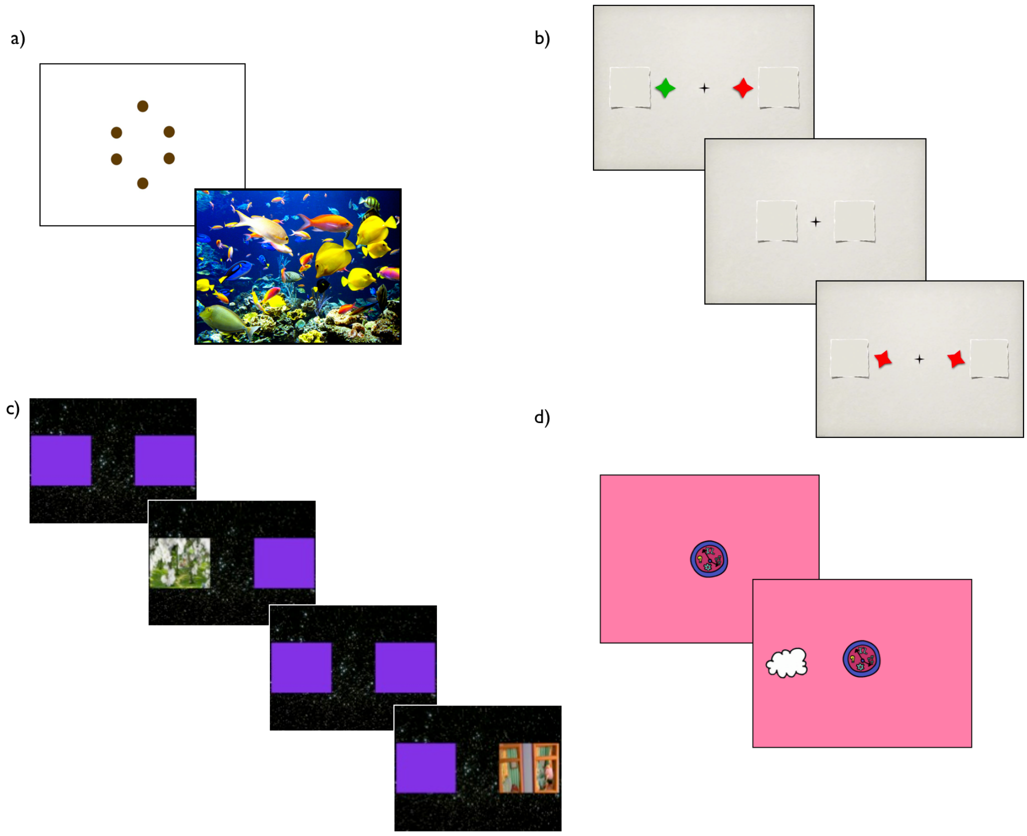

Figure 2: Schematics showing the pre-post tests that were administered: a) Examples of the 'Boring' (top) and 'Interesting' (bottom) stimuli used in the Sustained Attention task; b) Illustration of the screen layout for a trial in the Short Term Memory task; c) Illustration of the screen layout for the Cognitive Control task; d) Illustration of screen layout for the overlap condition Gap-Overlap task (in the baseline condition, the central target disappeared as the lateral target was presented).

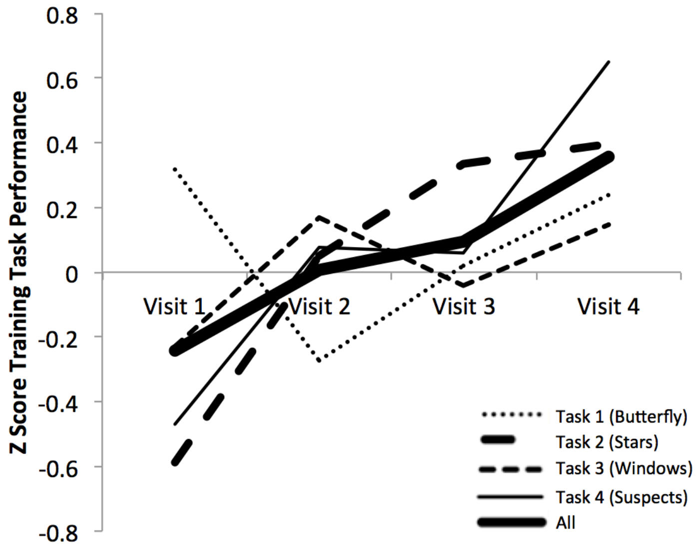

Figure 3: Line graph showing how training task performance changed across the four visits in the current study. Z-scores are presented, as described in the main text.

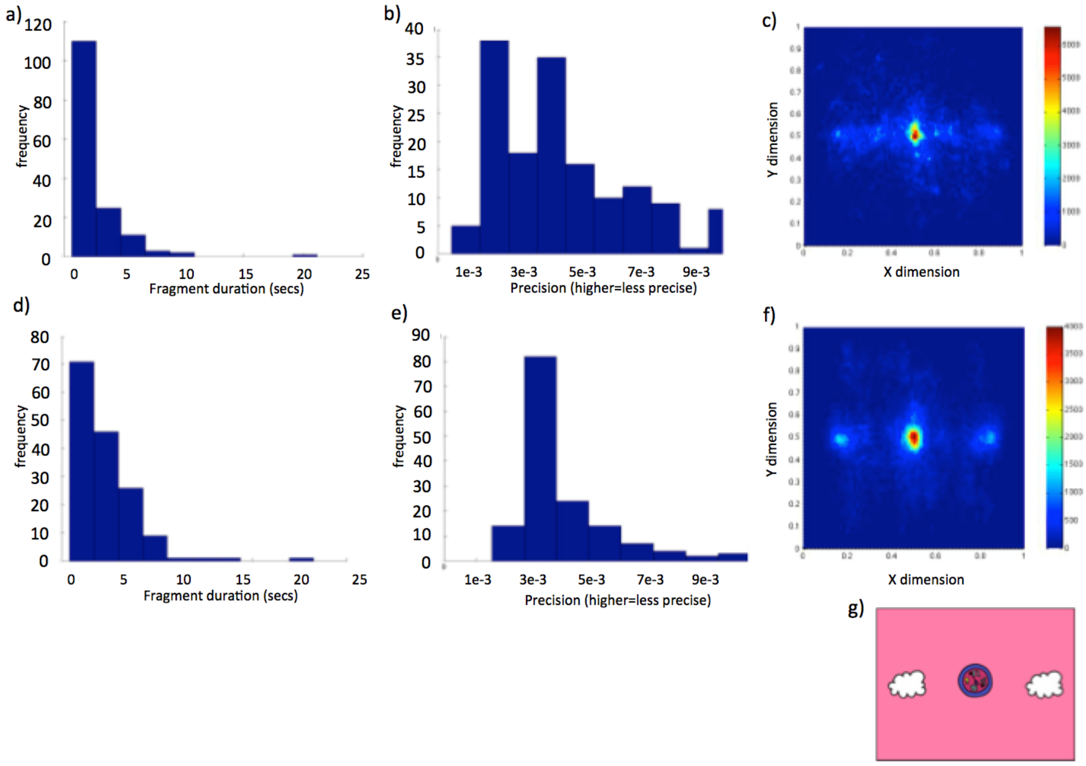

Figure 4: Data quality comparison based on data from the gap-overlap study: a) – c) show data from the present study, and d) – f) show data from a comparison study that used identical procedures, in lab settings, with typical infants; a) and d) show histograms showing the duration of usable fragment durations that were present in our data (calculated on a block-by-block basis). Shorter usable fragment durations were obtained in the present study a) relative to the comparison study d); b) and e) show histograms showing the precision of our data (calculated on block-by-block basis). Less precise data were obtained in the present b) relative to the comparison study e); c) and f) show gaze maps of usable gaze data obtained during the trial, and g) shows a schematic of how images were distributed on the screen during the trials.

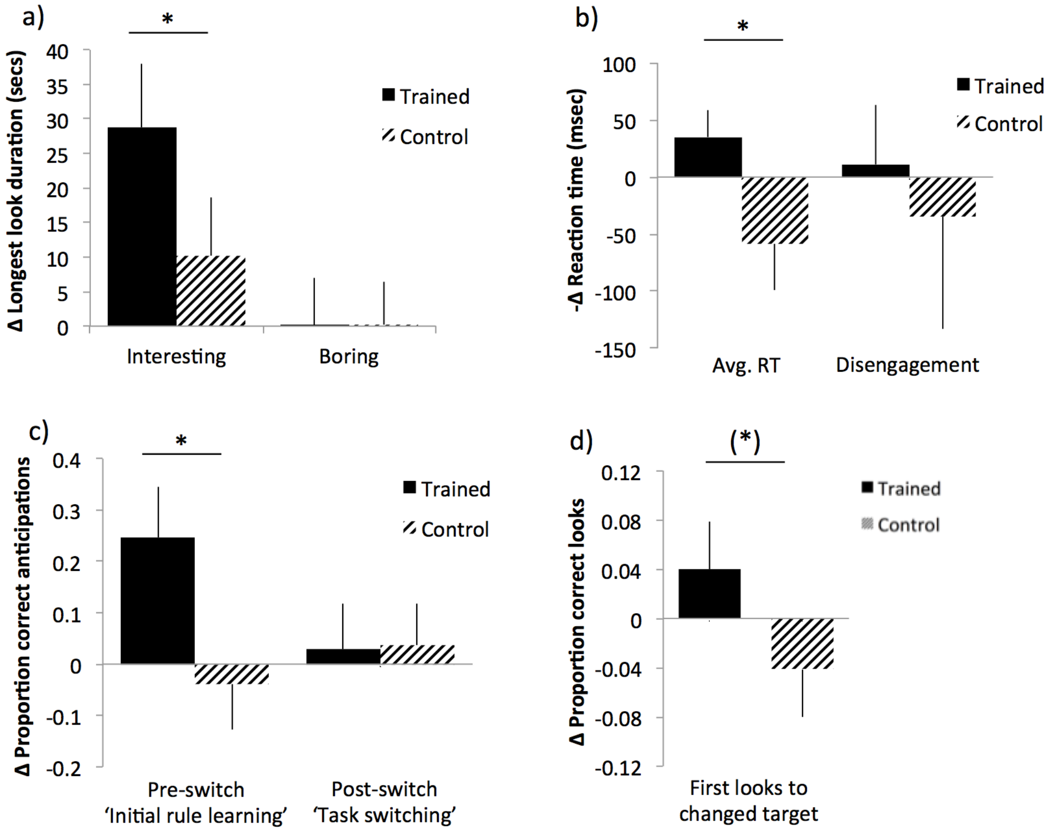

Figure 5. Results of pre-post assessments. Bar charts show change (Δ) scores on pre-post assessments, calculated from the marginal means; a) Sustained attention; b) Gap-overlap task. Because the valence of the predicted and observed change was negative, -Δ scores are presented for ease of comparison; c) Cognitive control task, * - p < .05; d) Short-term memory task. Stars indicate the significance of the analyses presented in the main text, * – p < .10.