Feasibility of undertaking off-site infant eye-tracking assessments of neuro-cognitive functioning in early-intervention centres

Ballieux, Haiko, Tomalski, P., Kushnerenko, E., Johnson, M.H., Karmiloff-Smith, A. and Moore, D.G.

Feasibility of undertaking off-site infant eye-tracking assessments of neuro-cognitive functioning in early-intervention centres

Haiko Ballieux[1, 2], Przemyslaw Tomalski[3], Elena Kushnerenko[2], Mark H. Johnson[4], Annette Karmiloff-Smith[4], and Derek G. Moore[2]*

[1] Department of Psychology, Faculty of Science & Technology, University of Westminster, London, UK

[2] Institute for Research in Child Development, School of Psychology, University of East London, London, UK

[3] Neurocognitive Development Lab, Faculty of Psychology, University of Warsaw, Warsaw, Poland

[4] Centre for Brain and Cognitive Development, Department of Psychological Science, Birkbeck, University of London, London, UK

* Corresponding author: Derek Moore, Institute for Research in Child Development, School of Psychology, University of East London, London, UK; e-mail: D.G.Moore@uel.ac.uk. tel. +44 (0)20 8223 4433.

(IN PRESS) Infant and Child Development

Acknowledgments

This study was supported by the Nuffield Foundation (PI: DM). The Nuffield Foundation is an endowed charitable trust that aims to improve social well-being in the widest sense. It funds research and innovation in education and social policy and also works to build capacity in education, science and social science research. The Nuffield Foundation has funded this project, but the views expressed are those of the authors and not necessarily those of the Foundation. More information is available at www.nuffieldfoundation.org.

We would like to thank all participating families for their vital contribution, as well as management and staff in Children's Centres in Tower Hamlets and Newham (London, UK). We would especially like to thank Sally Parkinson, Head of Commissioning in Newham, for helping with setting up partnerships with CCs, and Monica Forty and her team for the ongoing support in Children's services in Tower Hamlets. We acknowledge additional support of the Wellcome Trust (098330/Z/12/Z, AKS), the UK Medical Research Council (G0701484, MJ), the EU (PCIG10-GA-2011-304255, PT) and additional support from the UEL School of Psychology.

Abstract

Recent work suggests that differences in functional brain development are already identifiable in 6- to 9-month-old infants from low socio-economic status (SES) backgrounds. Investigation of early SES-related differences in neuro-cognitive functioning requires the recruitment of large and diverse samples of infants, yet it is often difficult to persuade low-SES parents to come to a university setting. One solution is to recruit infants through early intervention children's centres (CCs). These are often located in areas of high relative deprivation to support young children. Given the increasing portability of eye-tracking equipment, assessment of large clusters of infants could be undertaken in centres by suitably trained early intervention staff. Here we report on a study involving 174 infants and their parents, carried out in partnership with CCs, exploring the feasibility of this approach, We report the processes of setting up the project and participant recruitment. We report the diversity of sample obtained on the engagement of CC staff in training and the process of assessment itself. We report the quality of the data obtained, and the levels of engagement of parents, and infants. We conclude that this approach has great potential for recruiting large and diverse samples worldwide, provides sufficiently reliable data, and is engaging to staff, parents and infants.

Introduction

Recent work suggests that socio-economic status (SES) disparities are associated with specific profiles of neuro-cognitive differences in childhood (Noble, McCandliss & Farah, 2007; Noble, Norman & Farah, 2005; D'Angiulli, Herdman, Stapells & Hertzman, 2008; Kishiyama, Boyce, Jimenez, Perry & Knight, 2009; for a review see e.g. Hackman, Farrah & Meaney, 2010; Lipina & Posner, 2012). Further, there is emerging evidence of differences in functional, frontal brain development that are already identifiable in 6- to 9-month-old infants from low socio-economic status (SES) backgrounds (Tomalski et al, 2013), and that SES-associated factors may specifically affect cognitive flexibility and attention in infancy Clearfield & Niman, 2012; Clearfield & Jedd, 2013). It is of concern that these early differences may already set some infants on a developmental pathway that leads to poorer educational outcomes (eg. McEwan, 2003; Fernald et al., 2011).

Given this concern, it is important for researchers to begin to work towards developing effective ways of identifying neuro-cognitive differences as early as possible, and to actively recruit more diverse samples. Reliable and specific measures that could reveal early individual differences in components of neuro-cognitive functioning may subsequently be used to inform early interventions. These interventions might target specific areas of difficulty in early infancy and prevent children from entering school already behind in social, attention or language abilities, with all the cascading effects this may have for the individual and society (see Allen & Duncan 2008; Allen, 2010, 2011).

One obstacle to developing measures of SES-associated individual differences in neuro-cognitive functioning in infants is that most experimental studies of infants take place in university babylabs. The reason for restricting testing to laboratory facilities has been the need to employ expensive and bulky technical equipment as well as concerns about maintaining a controlled environment. However, one problem with this approach is that babylab-based studies tend to recruit relatively homogenous samples of infants who have more affluent and better-educated parents. Even with specific targeting strategies, such studies rarely attract large numbers of low-SES or ethnically diverse families. As outlined by Diemer and colleagues (2013), psychologists often pay little heed to social class when formulating theoretical models, conceptualizing studies, recruiting participants, selecting measurement tools, and analyzing data. Consequently, this raises questions about the generalizability to the general population, and to low-SES groups in particular, of much experimental work with infants.

One way to address the issue of non-representative recruitment and non-generalizability is to capitalize on recent goverenment investment in early years services and to locate studies within early intervention centres. In the United Kingdom, these centres are known as Children's Centres (CCs). The creation of "Sure-Start" CCs in 1998 was a UK government response to an increasing worldwide recognition of the importance of investing in universal early education, and the recognition of the benefits this may have particularly for low-SES populations. CCs provide a range of community health services, parenting and family support, integrated early education and childcare, and link to training and employment opportunities for families with children under the age of five (CC Statutory Guidance, 2010). CCs are often found in low-income areas, with high indices of multiple-deprivation (Nobile, McLennan & Wilkinson, 2010). They are closely linked with their communities, and specifically tasked with helping parents with children under five-years-of-age. The Sure-Start programme was modelled on the Head-Start programme developed in the US, and this approach has also been adapted in other countries. For example, Australian's Head Start (early learning centres), Canada's Ontario Early Years Plan, or the recently set up Biztos Kezdeta in Hungary. Consequently, we consider this report to be of interest to researchers wishing to adopt a similar approach in different countries.

While many researchers are engaged in off-site testing in the home or nursery, assessments are often limited to standardized table-top behavioural measures, because of the diversity of environments encountered in the home. Furthermore, testing in the home requires considerable staff time in travel, and is not a cost effective way of assessing large clusters of infants. Furthermore, while it may be possible to set up assessments in other professional settings such as family doctor clinics, CCs specifically target pre-school children in low income areas, so working in CCs offers a far better opportunity to recruit more diverse samples of infants and in large numbers. This approach reduces researcher travel compared to testing in homes, and may also facilitate recruitment, as parents will already be attending CCs. Additionally, CC environments, while not as controlled as babylabs, are likely to be better controlled than the home, with a researcher able to set up in a dedicated room in advance of testing, and test in one day a cluster of infants in each CC under similar conditions. Furthermore, if researchers can engage existing CC staff in delivering assessments, then there is the potential, not only for effective recruitment, but also for the efficient assessment of sizeable samples.

Thus, working in CCs presents an opportunity to recruit and assess more diverse samples. However, it also presents a challenge to neuroscientists to adapt lab-based experimental measures that are normally used in cross-sectional studies of group effects in homogeneous samples, into portable, effective and reliable measures that take account of infant diversity and individual differences. These measures need to be able to be used in the room settings they are likely to encounter in CCs, and be designed to be appropriate for diverse populations, for example by using face stimuli that reflect the full diversity of ethnicities. To maximize recruitment, as well as being interesting to infants, measures also need to be easily understandable and relevant to a diverse range of parents.

While brain-imaging equipment such as EEG and fNIRS is becoming more portable, and offers increasing potential for field-based assessments of infant brain functioning (see Lloyd-Fox et al., 2014), brain-imaging techniques often require many trials to show an effect, are still far from simple to set up, and require staff with specialist knowledge, not only to create tasks and analyse data, but also to administer assessments. On the other hand, eye-tracking equipment is increasingly portable, is far easier to calibrate than in the past, and is much simpler to use for data collection than brain imaging techniques. Thus, while still complex to program and analyse data, the collection of reliable eye-tracking data by CC staff may be possible using pre-programmed eye-tracking paradigms written for integrated monitor systems.

Eye-tracking equipment uses the corneal reflection of an infrared light source, relative to the location of the pupil, to record the direction and duration of looking, and these measurements are very precise, in the order of millimeters and milliseconds (Oakes, 2011; Aslin, 2011), allowing the detailed assessment of a range of attentional and cognitive processes in infants. Unlike earlier equipment, modern eye-trackers are much more tolerant of infant head movement than in the past. Furthermore, newer eye-tracking software allows instant playback showing infant tracking and fixation patterns overlaid on stimuli. This gives a potential added advantage of presenting immediate and engaging feedback to parents about their infant's behaviour.

Previous studies of at-risk groups of infants have already utilized a range of eye-tracking measures in babylabs to define and validate potential early markers of developmental difficulties including in the BASIS study (see http://www.basisnetwork.org/), which explored the early emergence of Autism Spectrum Disorder (ASD; see e.g. Guiraud et al., 2012; for a review see Jones & Klin, 2013; Jones et al., 2014). The current paper reports on a first UK study to take these assessments into CCs to measure individual differences in young infants from highly diverse populations.

In the following sections we present the protocol on the setting up of the study, we outline the contexts in which UK CCs operate and report on how we set up partnerships with the bodies who oversee CCs (local authorities and other providers). We also report on the level of engagement of CC staff in undertaking training in how to deliver eye-tracking assessments. We outline the diversity of the sample we were able to recruit with this approach, report on the diversity of the testing environments encountered, and on the quality of the eye-tracking sampling data obtained in these contexts versus a laboratory, and finally report on the level of engagement of parents with the tasks and procedures.

Candidate eye-tracking tasks selected

For the current study, we used adapted versions of tasks, previous versions of which had also been used in the BASIS study of early autism. We selected five 'candidate' tasks to see whether they might be useful for identifying potential individual differences in early social attention and communication in non-lab settings. SES has been associated with differences in the quality of early mother-infant interactions, which may also be associated with differences in attention to faces, and in the emergence of differences in joint attention behaviours, particularly gaze monitoring, as well as differences in audio-visual speech integration and auditory discrimination. The five tasks were:

Face pop-out task. In this task, infants are presented with arrays of images of different classes of objects arranged in a circle. In half the trials one of the images is a face. The aim is to assess the extent to which infants show a preference for looking at faces over other classes of objects (birds, cars, shoes, etc.). Measures of individual differences in face 'pop-out' include the number of looks, duration of first look, and total looking to the face compared to objects over trials. Group differences have been found in measures of pop-out between infants with and without autism (Gliga et al., 2009; see also Frank, Amso & Johnson, 2014, and Frank, Vul & Johnson, 2009). In our version, we adapted the task for use with a diverse population by including a wider variety of ethnicities of faces (Ballieux et al., 2013).

Gaze following task. From the age of 6 months, infants increasingly use the direction of a person's gaze as a cue for looking towards objects of attention (Senju & Csibra, 2008), and treating gaze references as cues for learning new words (e.g. Gliga & Csibra, 2009; see also Gillespie-Lynch et al., 2013). Differences in infant sensitivity to dynamic eye gaze have been associated with later emerging autism (Elsabbagh et al., 2012). In this task, we measured the frequency of orienting responses towards the location of an object congruent with the direction of gaze, either to the left or right, shown by a face on a video screen. Again, we adapted this task to be appropriate for more diverse samples by including a range of ethnicities of faces.

Audio-visual Speech Integration (AVSI) task. This task assessed infants' expectations of the relationship between pronounced speech sounds and expected lip movements. We used an eye-tracking version (Tomalski et al., 2012) of an audio-visual speech integration task developed by Kushnerenko et al. (2008). The task had already been adapted as part of an earlier Babylab-based ELAS study (Early Language and Attention study) where we deliberately tried to recruit a more diverse sample of participants in order to assess individual differences and begin to examine SES effects (See Kushnerenko et al, 2013; Tomalski et al 2013). Again, we adapted this task for more diverse samples by including a range of ethnicities of faces (Moore et al, 2014). Infants were presented with four types of videos: 2 congruent videos where the auditory track matches the seen articulatory lip movements (using sounds /ba/ and /ga/) and 2 incongruent videos (visual /ba/ dubbed onto auditory /ga/ and vice versa) along with a silent face control condition. Looking times to the eyes and mouth at 6-9 months of age in this task have been shown to be predictors of receptive language development in toddlers and found to be associated with distinctive patterns of brain activity during AVSI (Kushnerenko et al., 2013a,b;

also see Lewkowicz & Hansen-Tift, 2012). In addition, infants at risk for developing autism show atypical patterns of face scanning during audiovisual integration (Guiraud al., 2012).

Vowel discrimination task. A preferential listening procedure (modeled after Polka et al., 2008; Mattock et al., 2008) was used to test the discrimination of two vowels embedded in a word ('dog' vs. 'dug' and 'bet' vs. 'bat'), while infants saw a picture of a mobile phone on the screen. Infants were familiarized with one word presented repeatedly for 30s and then received two test trials – one with the familiarized word only, and another with the novel word presented between instances of the familiarized one. Preference (longer looking times) for the novel vowel/word was treated as indication of vowel discrimination. The reason for including this task was to assess whether delays in the emergence of vowel discrimination may predict later language difficulties.

Free viewing task. Short (30s) video clips showing talking and interacting people were presented to measure orienting to social cues in naturalistic settings as well as tracking of interaction partners. The analysis of fixation distributions allows us to study the allocation of attention to the eyes and mouth and other sources of information in a display (Võ et al., 2012). Similarly, individual differences in allocation of attention to these naturalistic scenes may predict later social and cognitive difficulties.

Parallel papers will report data from each of these tasks and report on the relationship between individual differences, SES and ethnicity. As already outlined, the purpose of the current paper is specifically to report on the practicalities of undertaking this form of research, including the level of engagement of CC staff in undertaking training in how to deliver eye-tracking assessments.

Setting up partnerships with CCs

This study was undertaken in partnership with Children's Services with assessment taking place in six CCs supported by Children's Services and local Health Services in the London boroughs of Newham and Tower Hamlets. These boroughs are in the top five for social and economic deprivation in England (DCLG, 2010, GLA London perspective on EID 2010), and have around 42% of children living below the poverty line (Aldridge et al., 2013). Local authorities in the UK follow different models of CC management, and the CCs in our study reflect this diversity.

In Newham, CCs are semi-independent and often set up by existing schools and nurseries responsible for their management and budget. Managers and staff in three CCs in Newham were approached separately, all responding very positively to the idea of participation in the project. They indicated that they valued the project's goals and could see how it could potentially benefit families in their centres.

Commitment to the project in the borough of Tower Hamlets was also good. The management of CCs in this borough was more centralised, with a clear management structure overseeing all governance, research, and external collaborations. At the early stages of the project we approached and received formal support from the Head of Early Years in the borough and delivered formal presentations at their CC managers meeting and to each centre individually. When applying for funds for the project we received a commitment in kind in staff time to allow staff to take part in training and assessments.

In total, six CCs were selected in the two boroughs where assessment would take place, with an additional partner CC helping with recruitment. The study received clearance from the University ethics committee, and additional clearance was obtained from the Research Governance Directorate of Tower Hamlets.

Recruitment process

All participating CCs advertised our 'Learn About Your Baby' sessions as a potential learning experience for parents, who could come to discover and see for themselves how their infants attended to various stimuli. Sessions were scheduled and advertised in CC quarterly activities calendars for parents alongside other baby-targeted activities (e.g. baby yoga, baby club, parent and toddler group). We also distributed flyers and posters advertising the sessions (see Figure 1). The advantage of being part of the scheduled activities was that CCs were able to actively recruit on our behalf. We provided the CCs with our required age-range and exclusion criteria: age range, 6 months 0 days to 7 months 30 days; no pre-term infants; no major medical condition;, and no major delivery complications. Then, the CCs accessed their own database and sent the flyer and a study information sheet (available on request) to all parents with infants fitting these criteria. In addition, flyers and posters were distributed in the CC reception areas. The information materials were written in English except for the 'calling all babies' phrase on the flyer and poster[1] (many CC staff members were able to speak other languages). Parents who wished to take part contacted the CC, or researcher directly, to book an appointment. Since the parents often already knew the staff members working at the CCs, this may have made them more inclined to join the sessions.

[Figure 1 here]

The CC managers estimated that around 50% of the total number of the parents on their databases whom they contacted actually took part in the study (this estimate varied from 33% to 65% across the centres). All parents were briefed prior to taking part in the study that this was a research project and the results could not be used in diagnosing any difficulties of individual babies before these methods had been validated. One disadvantage of being part of

[1] Given the diversity of languages in East London it was not possible to create materials translated for all. However, in future studies it would be possible to target specific language groups with translated materials if this were the focus of the study.

the timetable was that this restricted us to the same slot each week when we could test participants in a given centre. If this once-weekly slot was not convenient for parents, then they could not always be tested. In a few cases, we assessed infants in another CC if the timing of the session and location were more convenient.

Groups of sessions were timetabled either for a morning or afternoon, or in some cases all day. As we were using one set of equipment, and one team of researchers, it was essential to carefully coordinate the timetabling of sessions throughout the week across CCs. Overall, this allowed us to comfortably assess on average three infants a week in each CC, approximately 20 infants per week across the six CCs.

The testing session itself consisted of an introduction by the experimenter, administration of the five eye-tracking tasks, each lasting 5 minutes, a scripted playback of videos of the infant performance, the completion of parent questionnaires and a session evaluation. Parents were given a certificate of participation for their baby, a £10 shopping voucher, and a children's book. In addition, with permission, we took a picture of the baby and sent an A4 printout on photo paper to the parent's home.

Characteristics of sample

Despite constraints of times of testing in each CC, we managed to recruit a total of 195 infants (of a target of 200) over seven months. Of these, twelve took part in initial piloting, and nine were assessed but later excluded from analysis, as they did not meet age and/or health criteria. This left a total sample of 174 infants who did a full assessment session, and from whom we collected eye-tracking data. All participants included in the sample were born full-term (36-42 weeks gestational age). A comparable proportion of participants came from each of the two boroughs (Newham 54.5%; Tower Hamlets 45.5%).

Income and education: As anticipated, participants were diverse, varying in levels of education and in the income of parents. The mean family income of participants was £49,497 and the median was £30,000 with 40% of the sample having a family income of £20,000 or less. There was a wide range including some families with no income (families new to the UK and not eligible for benefits), and also a handful with incomes greater than £200,000, reflecting higher earning levels in London.

Ethnicity and language diversity: In our previous Early Language and Attention Study (ELAS) it required large investment of effort to recruit a large diverse sample for assessment in the Babylab (see Tomalski et al., 2013a, b); Kushnerenko et al., 2013a, b). By contrast, in a relatively short seven-month recruitment period, and with constraints on timetabling, CCs recruited a large and diverse sample for the current study.

In the UK racial diversity is classified as 'ethnicity' rather than 'race'. Ethnic categories are classified in UK surveys according to the guidelines of the UK government office of national statistics[2] based on population prevalence and self-labeling surveys. Note that some categories are used in the US or other non-UK populations that do not feature in UK classifications and visa versa. For example, the US category 'Hispanic' is not a category recognized in the UK, and Asian is used in the UK as a super-ordinate category with subcategories of ethnicities across the Indian sub continent and South East Asia.

[Table 1 here]

Table 1 shows the profile of our CC sample and that of the earlier lab-based ELAS study. Both studies targeted a diverse sample but, as mentioned, through CCs we recruited far more quickly, and we attracted a significantly greater proportion of non-white participants (Chi-squared = 18.89, p <.01), as well as a larger proportion of people living in rented

[2] http://www.ons.gov.uk/ons/guide-method/measuring-equality/equality/ethnic-nat-identity-religion/ethnic-group/index.html - 1

accommodation (Chi-squared = 17.48, $p < .01$). Only around 58.6% of the families recruited through CCs reported English as their first language, with 71.3% reporting more than one language being spoken at home. A large number of families were bilingual or multilingual, with 42 different languages being spoken at home as the first language.

[Figure 2 here]

Figure 2 shows the distribution of the main ethnicities of the infants assessed across the six CCs. Note that all CCs attracted a highly diverse sample, but as each CC was located within a different community, they recruited sub-samples with very different profiles of ethnicities. This reflects the many different population 'pockets' of ethnicity found across these two boroughs in East London, and indicates that the study managed to reflect well the local diversity of each CC.

Engagement of Children's Centre staff in training sessions

One major aim of this project was to establish the feasibility of engaging CC staff with eye tracking technology to ascertain whether they would be comfortable with undertaking assessments. This would inform researchers on the possibility of using this approach for the assessment of large samples by non-specialists, and inform policy makers of the feasibility in future of developing larger scale screening programmes. To facilitate this process, we worked closely with eye-tracking technology specialists (Acuity Ltd.) to develop a short training programme, and created study materials to educate and engage the staff. The training lasted 2-3 hours, and usually took place in the CC. It comprised a seminar on infant cognitive research to contextualize the work, followed by basic training of practical skills on loading and running the eye-tracker paradigms, including how to set up and undertake eye-tracking recording and how to replay to parents[3] video of the infant eye-tracking trace.

Staff members in all but one CC (due to time restrictions) participated in the training sessions at the beginning of the project. In total, we conducted five training sessions attended by 16 CC staff members. We asked all participants to give feedback by answering questions on the quality of the training, its ability to hold attention and its usefulness. Each question was coded 1 to 5 (poor to excellent). Fifteen of the sixteen staff reported the level of content of training good, very good or excellent; all 16 found the ability of the training to hold their attention good, very good or excellent; and all 16 found the usefulness of the training for their work to be very good or excellent.

Staff involvement in assessment sessions

On the whole, engagement of staff members and their managers was very high. Most managers were very interested in the project and were keen to let staff members take time to conduct the sessions. The participation of staff members who took part in training in assessments was good, with 75% (n= 12) sitting in on and/or partly running at least one session. In one CC, staff were very engaged and set up a rota for staff members to join sessions in a given week. In two other CCs, we had one staff member participate. Staff members included nursery and teaching staff, with ages ranging from 20 to 47 years. However, in three CCs staff members did not participate in the sessions and so the whole session was delivered by the experimenter. This was not due to lack of interest, but rather reflected increased workload in Tower Hamlets in 2010-11 as a result of re-organisation caused by national budget cuts.

Set up and range of testing environments encountered

[3] Slides and course materials from this course are available on request.

Administering sessions required a mobile eye-tracking kit that could be easily moved around and set-up within 20 minutes by a single person (see Figure 3). This kit consisted of a 17" eye-tracking integrated monitor (Tobii T120 model), and a portable Ergotron MX desk mount arm (45-214-026) that could be clamped onto a table and adjusted so we had consistency in the height of the screen relative to the position of the infant. We used a HP EliteBook 8440p laptop to control the eye-tracker using Tobii Studio version 2.0. The eye-tracker kit fitted into a purpose-built, wheeled hard case, supplied by Acuity Ltd, and a standard laptop backpack was used to carry testing materials, the laptop, cables, etc. We used partition screens available in the centres to hide the experimenter, who sat to the side of the infant (see Figure 3). The five eye-tracking tasks took around half an hour maximum to administer, with the rest of the session taken up with questionnaires, evaluation and video playback to parents.

[Figure 3 here]

One of the possible challenges of testing in CC's is the potential variability of available facilities. Indeed, we worked in rooms varying in size, background noise, lighting, temperature, and visual distracters. These were normally employed for group activities (minimum 4x5 meters), or as consulting rooms by health visitors or midwifes (on average 3x3 meters). The background noise in most rooms was low, apart from one centre located next to a busy road. We measured ambient sound levels in 7% of sessions at a distance and height equivalent to the distance and height of the infant's head to the screen (distance approximately 60cm, height approximately 1.3m). Staff members understood the importance of noise levels and were very cooperative in keeping them to a minimum in the corridor during testing sessions. The average overall sound level in the testing rooms was 49.5dB, ranging from 40.2 to 55.1dB. Any obvious visual distractions (colourful posters etc.) were moved out of sight. In three CCs, the amount of daylight entering the room was not easy to control. Lighting conditions were therefore slightly different in each centre.

Quality of eye-tracking data collected compared to laboratory studies

To evaluate the quality of eye-tracking data obtained, we compared the data on the AVSI task on which we had also collected data in the Babylab as part of the ELAS study. The same equipment and experimental paradigm were employed in both studies (see Tomalski et al., 2013b).

[Table 2 here]

As can be seen in Table 2, comparing performance on the same audio-visual speech integration (AVSI) task in two kinds of settings, Babylab versus CC, did not lead to a significant increase in participant dropout, nor to a reduction in the proportion of valid trials per participant. Both datasets were comparable in this regard. However, one difference was that testing in CCs led to a lower proportion of time points at which the eye-tracker recorded valid gaze data (the TobiiT120 eye-tracker was sampling gaze position at 120Hz; proportion of valid samples M=60.5% versus ELAS study M=76.5%), and a corresponding increased variability within the sample (the ELAS study SD=16.39, range 26-97%; current study SD=21.61 and range 0-99%).

We also compared data from our face pop-out task to a slightly different unadjusted version used by the Birkbeck Babylab in the BASIS study (Gliga et al., 2009; Elsabbagh et al., 2013). The percentage of looks to a face in a display of 6 different objects was very similar across the two studies. We conclude that testing in a CC setting does not significantly alter infant engagement with the eye-tracking tasks nor completion rates, but that variable conditions (e.g. light levels, noise, general distractions) and greater sample variability may somewhat reduce data integrity.

Recognising that there may be differences in level of performance across sites on the basis of small differences in room settings, it would seem sensible to undertake an analysis of differences between centres. This is certainly our intention. However, any meaningful examination of differences in levels of data quality between centres cannot be undertaken without considering not only any differences in environment, but also taking into account the unique profiles of each sample recruited at each different site (see sample characteristics). We will be reporting these multilevel analyses in subsequent papers, examining SES and ethnicity effects alongside site-specific effects, having taken into account the unique profiles of the samples at each site.

Level of parental engagement

A benefit of modern eye-tracking is its ability to give immediate (positive) feedback to parents about the behaviour of their infant, by playing back video of their behaviour. Parents were played videos of their infant's performance accompanied by a scripted generic and non-evaluative commentary[4]. The videos presented the eye-gaze scanning trace overlaid on the stimuli, revealing their infant's patterns of visual exploration. Our intention was to show parents the complexity of visual behavior and the rapid nature of attention shifts already apparent in 6- to 7-month-olds. Parents found this particularly engaging and interesting, and often indicated how surprised they were at the level of visual control displayed by their infant. During the remaining time parents answered a set of questionnaires on family demographics, infant social environment, sleep and feeding, and were then asked to give feedback on the session, answering four questions (see Table 3).

[Table 3 here]

The vast majority of parents reported enjoying the eye-tracking session and considered it interesting or very interesting. A large number of parents (84%) indicated that they felt the session had somewhat changed their understanding of their infant, while a smaller group

[4] A scripted commentary was used that did not give interpretations of the behaviours displayed, so we did not cause concern or give false impressions. This script is available on request.

indicated they were now thinking at least a little differently about their baby (52%). These results were consistent with our expectations: we did not expect parents to radically change how they thought about their infant merely following a 1/2-hour assessment session, but we did expect them to change to some extent their understanding of what kinds of things their infant was interested in and capable of doing, even at such a young age (e.g. attention shifting, deciding what to observe). Noteworthy is the fact that parents of lower SES considered the session more important and influential on their knowledge of their baby. Unemployed mothers or those with manual jobs more often reported that it very much changed their understanding compared to employed mothers with more skilled jobs (Chi-squared = 26.89, $p = .001$). Similar differences emerged for fathers with routine jobs or long-term unemployed compared to working fathers (Chi-squared = 18.3, $p = .019$).

Discussion

In this paper we have examined the practicalities of taking eye-tracking technology and methods out of a university research laboratory context into UK Sure Start Children's Centres. The experience of working in these early intervention settings has been very positive. Children's services and CC management in two local authorities demonstrated a clear interest and willingness to collaborate in this type of study. It should be noted that managerial support for the project was crucial for a successful collaboration in each CC in terms of securing long-term room allocation, allowing staff time for testing and allocating recruitment and session scheduling duties. We have demonstrated that by working together with CCs it is possible to recruit a culturally and economically diverse population, and that this approach may be far more successful in recruitment than typically achieved in Babyab-based studies. The sample recruited was more likely to be non-white and more likely to live in rented accommodation than a sample recruited via other routes, even compared to those Babylab-based studies, including our own, where we had specifically targetted low SES populations.

While this study was UK based, the model of children's centres, first started with Head Start in the US, is now being used across many countries worldwide, including Australia (Head Start early learning centres) Canada (Ontario Early Years Plan), Chile (Un Buen Comienzo) and recently Hungary (Biztos Kezdeta). As early intervention centres are often established in areas of most apparent need, they tend to be in low-income areas. They therefore represent a particularly important network of potential recruitment and assessment hubs worldwide for studies aimed at gathering large samples of infants from low SES backgrounds. This study supports the rationale for countries planning to develop the Head start/Sure start CC model, not only for intervention, but also as a potential place to undertake detailed neuro-cognitive research, which could in turn inform early intervention programmes. We have demonstrated that it would be possible to use centres as a base for large-scale studies of early neuro-cognitive functioning using eye-tracking paradigms that may not be possible to do in the home. This study shows that early intervention centres are good contexts in which to recruit more diverse samples and to produce usable eye-tracking data.

Furthermore, we have demonstrated that it is possible to engage and train CC staff to deliver these measures. Staff members, who did not have any prior specialist expertise, found the training and approach informative and were willing to take part in the assessment sessions and were pleased to include these as part of the scheduled programmes. There are however some ongoing concerns about the extent to which CCs are in a position to dedicate staff resources to facilitate this work in the UK. Due to changes in the national budget and the resulting re-organization, staffing levels underwent significant change between 2010 and 2011 when the study took place. Despite this, staff members were prepared and willing to facilitate the study. The positive feedback we obtained even against this backdrop suggests that, under more stable budgetary/staffing conditions, engagement in training and delivery of sessions would be even greater. Given the fact that the cost of CC staff time was covered solely by the centres themselves, they did an admirable job in continuing to help in booking and scheduling families for sessions, incorporating them into their calendar, and in providing testing rooms. We are optimistic that this level of engagement would be greater still should we demonstrate that these techniques serve a useful function in determining which children need particular help. Note that these sessions also allowed CCs to increase their provision of useful and targeted sessions for parent and infants, i.e. working together with CCs can have mutual benefits.

CCs were generous in dedicating a room for testing and we encountered some variation in the settings for assessments across CCs. However, these differences were generally manageable and noise levels, room layout, and lighting were within acceptable ranges. Infants engaged well with the tasks and completion rates were comparable to previous studies in our Babylabs at UEL and Birkbeck, comparing performance on our AVSI task from a lab based study with a smaller but diverse sample from the ELAS study, and with the Pop-out task from the larger scale BASIS study of infants at risk for autism. Nonetheless, on some measures the quality of sampling using the eye-tracker may be somewhat lower. We plan in future papers to undertake more extensive analyses of cross-centre variance for each task, taking account of the variability in settings and variability in the SES and ethnic profiles across sites.

Perhaps most importantly for the future development of a large-scale studies and the use of these measures for wider screening or training, we have found that almost all parents report the assessment process as engaging, informative, and interesting. Feedback received immediately after the session was highly promising, with the vast majority of parents finding the session and the generic information on their infant's visual attention interesting and enjoyable. What is even more promising is the fact that it was parents with lower socio-economic status in particular who found the session influential on their understanding of their child. This suggests that relatively simple measures, that provide video visualisations of infant's gaze data with a short commentary, may prove very useful as an intervention in its own right, and is effective in engaging families from impoverished and deprived areas in discussion about their infants. Another indicator of parental engagement is the return rate for a follow-up session 18 months later. More than half of the parents we approached returned for follow-up. Considering the fact that by then many mothers have gone back to work, and/or have moved house (sometimes to a different borough or even a city), this return rate was most encouraging.

Of course it is possible that there was something particularly engaging to parents about the set of tasks we employed, and that other tasks may not produce the same level of engagement. Our experience suggested, however, that the type of task used was secondary to the impact achieved simply by showing parents the overlaid scanning trace post testing. For all infants and all tasks it was possible to use the video playback to demonstrate to parents in an immediately engaging way their infant's abilities to control their attention, which many parents did not realise they had. We predict with reasonable confidence that any eye-tracking paradigm, that by necessity uses a stimulus that engages infants to attend, and which uses software that can immediately produce a video of infant scanning, would be equally useful in engaging and informing parents.

In conclusion, the approach of taking eye-tracking into early intervention centres offers considerable promise for recruiting and assessing large samples of infants from diverse SES and ethnicity groups not normally easy to recruit to university Babylabs, and this is likely to be applicable in comparable centres and programmes overseas. With good partnerships, recruitment of diverse populations can be greatly facilitated, and the settings encountered are adequate to allow assessments of large samples with good rates of completion. Eye-tracking assessments were successfully incorporated into CC schedules of activities, were engaging to staff, are adaptable to variable testing conditions and can be used to convey a positive message to parents.

There is considerable interest from many bodies including the American Psychological Association office on socio-economic status in ensuring that research takes fuller account of diversity and socio-economic status, and that awareness of SES issues is increased in theory and research (Diemer et al, 2013). Working in early intervention centres such as CCs promises both to facilitate the recruitment and assessment of more diverse participants and also to bring this work to the attention of these staff working with families and young children. In the long-term, partnerships between Babylabs and CCs may lead to the development of theoretically-driven, engaging, and easy-to-apply screening programmes for attention and language difficulties that can be implemented by early years professionals and facilitate the delivery of targeted early interventions. Once we have established which 'candidate' tasks are the best predictors of specific outcomes, we will be in a position to refine further the assessment process, to make it shorter and more targeted, offering the promise of short and cost-effective universal screening that can inform early intervention.

References

Aldridge, H., Bushe, S., Kenway, P., MacInnes, T. & Tinson. A. (2013). London's Poverty Profile. Published by Trust for London & New Policy Institute. ISBN: 978-1-901373-75-4. www.londonspovertyprofile.org.uk/LPP_2013_Report_Web.pdf.

Allen, G. (2010). Early Intervention: the next steps. Early Intervention Commission. UK Government: London.

Allen, G. (2011). Early Intervention: Smart Investment, Massive Savings. Early Intervention Commission. UK Government: London.

Allen, G. & Duncan Smith, I. (2008). Early Intervention: Good Parents, Great Kids, Better Citizens. Centre for Social Justice. UK Government: London.

Aslin, R.N. (2012). Infant Eyes: A Window on Cognitive Development. *Infancy, 17*(1), 126-140.

Ballieux, H., Tomalski, P., Kushnerenko, E., Karmiloff-Smith, A., Johnson, M.H., & Moore, D.G. (2013). Eye-tracking in the Community: Relations Between Infant Attention to Social Cues and their Socio-economic, Ethnic, and Language Background. Poster presented at the *Biennial meeting of the Society for Research in Child Development, Seattle, US.*

Clearfield, M.W., & Niman, L.C. (2012). SES affects infant cognitive flexibility. *Infant Behavioral Development, 35*(1), 29-35. doi: 10.1016/j.infbeh.2011.09.007.

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.

Diemer, M. A., Mistry, R. S., Wadsworth, M. E., López, I. and Reimers, F. (2013), Best Practices in Conceptualizing and Measuring Social Class in Psychological Research. *Analyses of Social Issues and Public Policy*, 13, 77–113. doi: 10.1111/asap.12001

Elsabbagh, M., Gliga, T., Pickles, A., Hurry, K., Charman, T., Johnson, M.H. & the BASIS Team (2013). The development of face orienting mechanisms in infants at-risk for autism. *Behavioural Brain Research, 251*, 147-154.

Elsabbagh, M., Mercure, E., Hudry, K., Charman. T., Pickles, A., Baron-Cohen, S., Bolton, P., Johnson, M.H., & The BASIS Team (2012). Infant neural sensitivity to eye gaze predicts emerging autism. *Current Biology, 22*(4), 338-342.

Department for Communities and Local Government (2010). English Indices of Deprivation. ISBN: 978-1-4098-2924-9. www.gov.uk/government/publications/english-indices-of-deprivation-2010.

Fernald, L. C. H., Weber, A., Galasso, E., & Ratsifandrihamanana, L. (2011). Socioeconomic gradients and child development in a very low income population: evidence from Madagascar. *Developmental Science*, *14*(4), 832–47. doi:10.1111/j.1467-7687.2010.01032.x

Frank, M.C., Vul, E., & Johnson, S.P. (2009). Development of infants' attention to faces during the first year. *Cognition, 110*, 160-170.

Frank, M.C., Amso, D., & Johnson, S.P. (2014). Visual search and attention to faces in early infancy, *Journal of Experimental Child Psychology, 118*, 13-26.

Gillespie-Lynch, K., Elias, R., Escudero, P., Hutman, T., & Johnson, S.P. (2013). Atypical gaze following in autism: A comparison of three potential mechanisms. *Journal of Autism and Developmental Disorders, 43*, 2779-2792.

Gliga, T. & Csibra, G. (2009). One-year-old infants appreciate the referential cues of deictic gestures and words. *Psychological Science, 20*(3), 347-353.

Gliga, T., Elsabbagh, M., Andravizou, A. & Johnson, M.H. (2009). Faces attract infant's attention in complex displays. *Infancy, 14*(5), 550-562.

Guiraud, J., Kushnerenko, E., Tomalski, P., Charman, T., Johnson, M.H., & The BASIS team (2011). Differential habituation to repeated sounds in infants at risk for autism. *Neuroreport, 22*, 845-849.

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

Hackman, D. A., Farah, M. J., & Meaney, M. J. (2010). Socioeconomic status and the brain: mechanistic insights from human and animal research. *Nature Reviews Neuroscience*, *11*(9), 651–659.

Jones, E. J. H., Gliga, T., Bedford, R., Charman, T., & Johnson, M. H. (2014). Developmental pathways to autism: A review of prospective studies of infants at risk. *Neuroscience and Biobehavioral Reviews, 39*, 1-33.

Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature*, *504*(7480), 427–31. doi:10.1038/nature12715.

Kishiyama, M. M., W. T. Boyce, A. M. Jimenez, L. M. Perry and R. T. Knight (2009). Socioeconomic disparities affect prefrontal function in children. *Journal of Cognitive Neuroscience*, *21*(6), 1106-1115.

Kushnerenko, E., Teinonen, T., Volein, A., & Csibra, G. (2008). Electrophysiological evidence of illusory audiovisual speech percept in human infants. *Proceedings of the National Academy of Science USA, 105*(32), 11442–11445.

Kushnerenko E., Tomalski P, Ballieux H, Ribeiro H, Potton A, Axelsson EL, Murphy E, Moore DG. (2013a). Brain responses to audiovisual speech mismatch in infants are associated with individual differences in looking behaviour. *European Journal of Neuroscience, 38*(9), 3363-3369. doi: 10.1111/ejn.12317.

Kushnerenko, E., Tomalski, P., Ballieux, H., Potton, A., Birtles, D., Frostick, C. & Moore D.G. (2013b). Brain responses and looking behavior during audiovisual speech integration in infants predict auditory speech comprehension in the second year of life. *Frontiers in Psychology, 16*(4), 1-8. doi: 10.3389/fpsyg.2013.00432.

Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceeding of the National Academy of Sciences of the United States of America, 109*(5), 1431-1436. doi:10.1073/pnas.1114783109.

Lipina, S. J., & Posner, M. I. (2012). The impact of poverty on the development of brain networks. *Frontiers in Human Neuroscience, 6*, 238. doi:10.3389/fnhum.2012.00238

Lloyd-Fox,S., Papademetriou, M., Darboe, M.K., Everdell, N.L., Wegmuller, R., Prentice, A.M., Moore, S.E. & Elwell, C.E. (2014). Functional near infrared spectroscopy (fNIRS) to assess cognitive function in infants in rural Africa. *Scientific Reports, 4*, art number 4740. doi:10.1038/srep04740.

Mattock, K., Molnar, M., Polka, L. & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition, 106*, 1367-1381.

McEwen, B. S. (2003). Early life influences on life-long patterns of behavior and health. *Mental Retardation and Developmental Disabilities Research Reviews, 9*(3), 149–54. doi:10.1002/mrdd.10074

Moore, D.G. Tomalski, P. Ballieux, H. Birtles, D. Johnson, M.H. Karmiloff-Smith,A. & Kushnerenko, E. (2014) Audiovisual Speech Integration at Six Months as a Predictor of Language Outcomes at Age Two in Infants from Families Varying in Income, Occupation, Ethnicity and Language. Paper presented at: *International Conference on Infant Studies, Berlin.*

Noble, K.G., McCandliss, B.D. & Farah, M.J. (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Developmental Science, 10*(4), 464-480.

Noble, M., Mclennan, D., & Wilkinson, K. (2010). *The English Indices of Deprivation 2010*. Retrieved from http://eprints.ioe.ac.uk/2461/.

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

Oakes, L. M. (2010). Infancy Guidelines for Publishing Eye-Tracking Data. *Infancy, 15*(1), 1–5.

Oakes, L. M. (2011). Advances in eye tracking in infancy research. *Infancy, 17*(1), 1-8.

Polka, L., Rvachew, S. & Molnar., M. (2008). Speech perception by 6- to 8-month-olds in the presence of distracting sounds. *Infancy, 13*(5), 421-439.

Senju, A. & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology, 18*(9), 668-671.

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

Tomalski, P., Ribeiro, H., Ballieux, H., Axelsson, E., Murphy, E., Moore, D.G. & Kushnerenko, E. (2013b). Exploring early developmental changes in face scanning patterns during the perception of audiovisual mismatch of speech cues. *European Journal of Developmental Psychology*, *10*(5), 611-624.

Wass, S. V, Smith, T. J., & Johnson, M. H. (2012). Parsing eye-tracking data of variable quality to provide accurate fixation duration estimates in infants and adults. *Behavior Research Methods*, *45*(1), 229-250.

Võ, M.L.-H., Smith, T.J., Mital, P.K. & Henderson, L.M. (2012). Do the eyes really have it? Dynamic allocation of attention when viewing moving faces. *Journal of Vision, 12*(13), 1-14.

Figure 1: Poster used for recruitment in Tower Hamlets local authority centres

Figure 2: Distribution of infant ethnicity over the six Children's Centres

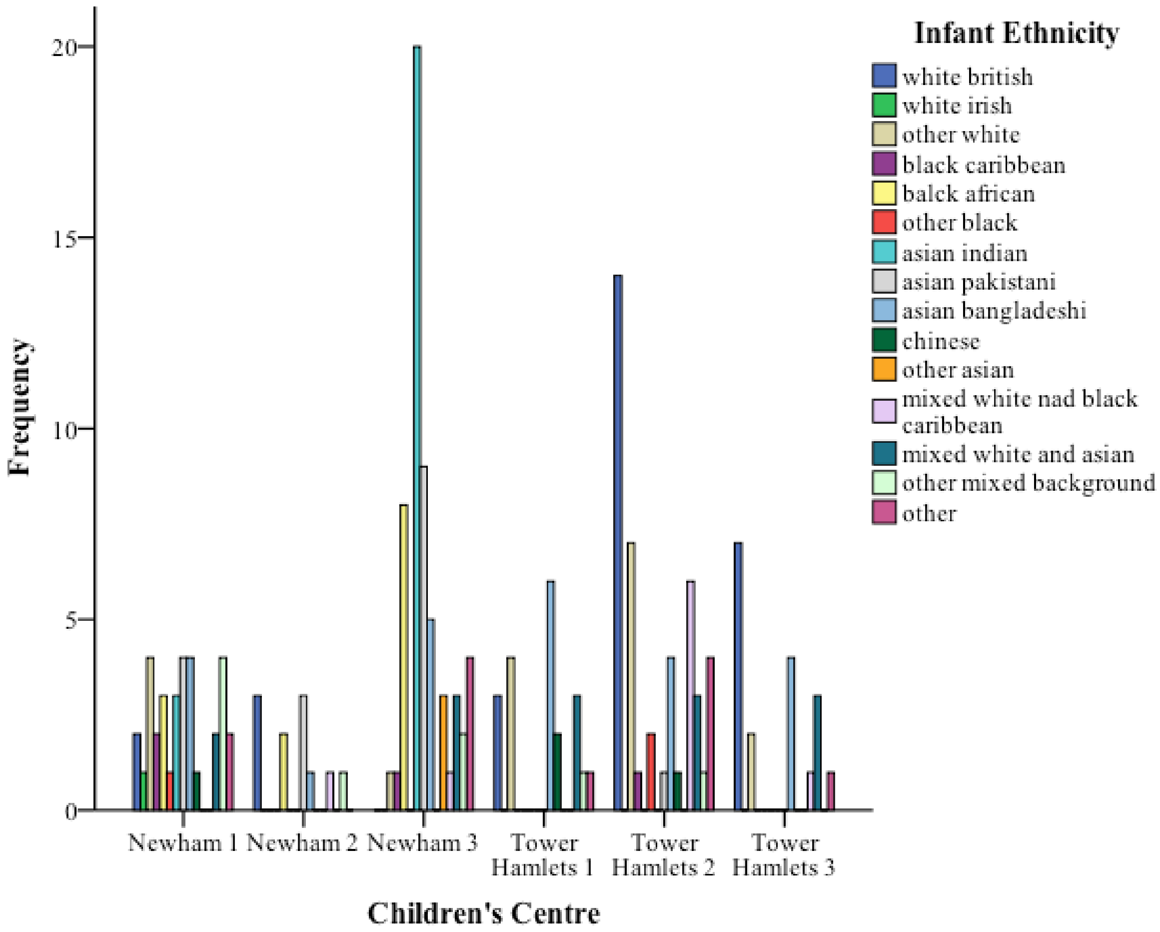

Figure 3: Photographs of differing set-ups in four CCs

Table 1: Diversity of sample recruited via CCs compared to ELAS Babylab study (Tomalski et al., 2013a)

Measure			ELAS study (n = 45)	CC sample (n=174)
Mean family income in £			53,238	49,487
(sd)			(44,712)	(65,456)
Median family income in £			46,000	30,000
Age in days			226.4	209.3
(sd)			(44.3)	(19.7)
Gender (%)	Female		68.9	39.1
	Male		31.1	60.9
Infant ethnicity (%)	White		60.0	25.9
	Non-white		40.0	74.1
Gestational age in weeks			39.6	39.5
(sd)			(1.9)	(1.5)
Birth weight in grams			3374.0	3229.1
(sd)			(566.8)	(501.5)
Mother's age at birth years			31.7	30.1
(sd)			(5.9)	(4.9)
Type of residence (%)	Owned house/flat		62.3	28.7
	Rented house/flat		33.3	65.6
	Rented room		2.2	2.3
	Other		2.2	3.4
Mother's occupational		1	42.2	32.8
level [a] (SEC; %)		2	20.0	9.2
		3	37.7	58.0
Father 's occupational		1	57.9	47.2
level [a] (SEC; %)		2	15.8	13.0
		3	26.3	39.8
Mother's education level [b]		1	44.4	20.1
(%)		2	26.6	35.1
		3	6.6	14.4
		4	13.3	10.3
		5	4.5	14.4
		6	4.5	5.7
Father's education level [b]		1	40.0	19.0
(%)		2	13.3	31.6
		3	11.1	11.5
		4	15.6	12.1
		5	13.3	13.8
		6	6.7	4.6

a: Parental SEC classification: (1) – higher managerial and professional occupations; (2) – intermediate occupations; (3) – routine and manual occupations or long-term unemployed.
b: Parental education: (1) – post-graduate; (2) – higher education degree; (3) – further education; (4) - high school A-levels; (5) – GCSE; (6) no qualification.

Table 2: Quality of eye-tracking data for the AVSI task recorded in Children's Centres versus UEL Babylab

	ELAS laboratory study	CC study
Proportion of participants completing task	84.2%	82.8%
Proportion of valid trials per participant (SD)	94.4% (9.39)	88.8% (16.43)
Proportion of time points at which the eye-tracker collected valid gaze data (SD)	76.5% (16.39)	60.5% (21.61)

Table 3: Parental feedback on the sessions by maternal occupational status (SEC)

		% of parents responding				
	Maternal SEC	Not at all interesting	Not so interesting	Neutral	Interesting	Very interesting
Q1. How interesting did you find this session?	Overall	0	0.6	2.3	26.6	70.5
	1 & 2	0	0	1.4	31.5	67.1
	3	0	1.0	3.0	23.0	73.0
		Not at all	Very little	A bit	Quite a lot	Very much
Q2. How much did you enjoy the session?	Overall	0	0.6	7.5	37.6	54.3
	1 & 2	0	0	9.6	38.4	52.0
	3	0	1.0	6.0	37.0	56.0
Q3. How much has this session changed the way you understand your baby?	Overall	3.5	11.0	29.5	43.4	12.7
	1 & 2	5.4	6.8	37.0	48.0	2.8
	3	2.0	14.0	24.0	40.0	20.0
Q4. How much has this session changed the way you think about your baby?	Overall	36.4	11.0	23.7	21.4	7.7
	1 & 2	39.6	9.6	27.4	22.0	1.4
	3	34.0	12.0	21.0	21.0	12.0

SEC classifications (1) – higher managerial and professional occupations; (2) – intermediate occupations; (3) – routine and manual occupations or long-term unemployed.