Early identification of children at risk of communication disorders:
Introducing a novel battery of Dynamic Assessments for infants

Helen Spicer-Cain, Bernard Camilleri, Natalie Hasson, Nicola Botting

Centre for Language and Communication Science Research, City University of London

Corresponding author contact: Nicola.botting.1@city.ac.uk

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

Funding: The original study was supported by doctoral funding to the first author from City University of London. No other funding was received.

ACCEPTED INTO AJSLP 8th November 2022

Abstract

Purpose

Many children with communication disorders (CD) experience lengthy gaps between parental reporting of concerns and formal identification by professionals. This means that children with CD are denied access to early interventions that may help to support the development of communication skills and prevent possible negative sequelae associated with long-term outcomes. This may be due, in part, to the lack of assessment instruments available for children younger than three years of age. This study therefore reports on promising preliminary data from a novel set of valid dynamic assessment measures designed for infants.

Methods

We recruited 53 low-risk children and two groups of children considered to be at high risk for CD (n=17 social-high-risk and n=22 language high-risk) due to family members with language and social communication difficulties. Children were between 1 and 2 years of age and were assessed using a battery of five dynamic assessment (DA) tasks related to receptive vocabulary, motor imitation, response to joint attention, turn taking and social requesting. A set of standardised measures was also used.

Results

The DA tasks showed high levels of inter-rater reliability and relationships with age across a cross-sectional sample of children from the low-risk group. Three tasks showed moderate to strong correlations with standardised measures taken at the same age, with particularly strong correlations between the DA of receptive vocabulary and other receptive language measures. The DA of receptive vocabulary was also the only task to discriminate between the three risk groups, with the social-high-risk group scoring lower.

Conclusions

These results provide preliminary information about early DA tasks, forming the basis for further research into their utility. DA tasks might eventually facilitate the development of new methods for detecting CD in very young children, allowing earlier intervention and support.

Introduction

Many children experience communication difficulties that require intervention during development. Autism and Developmental Language Disorder (DLD) represent two of the most prevalent disorders of childhood. Roughly 2% (Roman-Urrestarazu et al, 2021) and 8% (Norbury et al, 2016) of all children experience these disorders respectively, and there is compelling evidence that there are lifelong sequelae including employment issues (Autism: Harmuth et al, 2018; DLD: Dubois et al, 2020) as well as for mental health (Autism: Hollocks et al, 2019; DLD: Botting et al, 2016). Yet, for DLD especially, there is relatively low awareness (Thordardottir et al, 2021) and a paucity of research compared to other developmental disorders (Bishop, 2010; McGregor, 2020). There is a view that early intervention is optimum for these children, as language difficulties associate with wider long-term difficulties such as memory impairment (Henry & Botting, 2017), poorer educational attainment and employment prospects (Conti-Ramsden et al, 2018) and increased mental health issues (Botting et al, 2016). However, very early diagnosis and associated intervention services are not yet recommended in many countries including the UK, (e.g. Lindsay et al, 2008; Boyle, 2011; Wallace et al, 2015; Reilly et al, 2015; Bishop et al, 2017; Law et al, 2020; Jullien et al, 2021), in part because there are limited reliable assessments which can accurately identify infants with communication difficulties before the age of 3. In this paper we present preliminary data from a set of novel assessment tasks as a first step towards developing tools for identifying very early social and communication difficulties. We have focussed this 'proof of concept' study on groups of children at risk of Autism and DLD because of their combined prevalence and also because these are groups where we expect communication difficulties to show early signs; however, tools that are applicable to communication difficulties in other risk groups would be a wider long-term aim.

Autism

Autism is a lifelong pervasive developmental disorder which is diagnosed on the basis of impairments of social communication and social interaction, alongside restricted and repetitive behaviours (American Psychiatric Association, 2013). Recent prevalence estimates indicate that approximately one in every 68 children aged four in the USA has an ASD (Christensen, et al., 2016) and that this figure is approximately 4.5 times as high for boys than for girls (1 in 42 as opposed to 1 in 189 respectively; Christensen, et al., 2016). A similar estimate of prevalence was derived for the UK by Baron-Cohen et al (2009), although only 60% of these cases were formally diagnosed before study participation. Because ASD is a spectrum condition, it is vastly heterogeneous in its presentation. Language abilities can range from minimal use of, or comprehension of spoken language, to intact structural language skills in the context of difficulties with pragmatic skills, or language use. ASD can occur both with and without learning disability, and recent estimates suggest that 44% of children with ASD have average or above average intellectual ability (Christensen, et al., 2016). Regression of communication and adaptive skills, usually in the second year of life, is also reported in a subset of cases (Meilleur & Fombonne, 2009).

Developmental Language Disorder (DLD)

DLD is the preferred label for language difficulties of unknown aetiology in children, including conditions that were previously referred to as Specific Language Impairment or Developmental Dysphasia (Bishop et al, 2017). DLD affects approximately 7-8% of children at school-starting age (Tomblin, et al., 1997; Norbury, et al., 2016) and typically occurs in the absence of hearing loss, neurological impairment or severe environmental deprivation which would explain difficulties with language learning. Previous criteria for Specific Language Impairment required normal non-verbal intellectual ability and/or a discrepancy between language skills and IQ, but it has recently been established that there are few significant differences between children with language difficulties in the presence of typical IQ and children who have low abilities in both language and cognition, and IQ criteria are therefore no longer used to define DLD (Bishop et al, 2017; Norbury, et al., 2016). The presentation of DLD is also heterogeneous and may involve difficulties at any level of language processing, including phonology, morphology, syntax, semantics or pragmatics. Individual children may be affected across one or multiple levels of language and across either receptive or expressive modalities or both. The relationship of DLD to delayed language acquisition (or “late talking”) in early childhood is complex, (e.g., Dale et al., 2003; Reilly et al., 2010; Zambrana et al, 2014; Duff et al., 2015;

Rudolph & Leonard, 2016), and early language delays are not always predictive of later language impairment on an individual level. However, at least some children who have DLD are known to have difficulties with spoken language throughout the lifespan (Botting, 2020).

Age of Identification

The age at which children with communication disorders are identified has important implications for early intervention and support. The potential consequences of not identifying and providing input for children with communication disorders are large (Hus & Segal, 2021) and may include poorer employment prospects (Chen et al., 2017; Conti-Ramsden et al., 2017); increased mental health difficulties (Botting et al., 2016); and economic costs for society (Rogge & Janssen, 2019). For ASD, Zwaigenbaum et al (2009) found that most parents of children later diagnosed with ASD identified concerns about their children's development between 12 and 18 months of age, including concerns regarding delayed language development, limited play skills, decreased social responsiveness, extreme behavioural reactions to external stimuli and difficulties with sleep or feeding. Some parents report even earlier concerns starting before 12 months of age (Zwaigenbaum et al., 2005; De Giacomo and Fombonne, 1998; Filipek et al., 1999). However, as noted above, early identification is not straightforward, and reported age of diagnosis in ASD tends to be much older than reported age of first concern. Mean age of diagnosis varies between studies and across countries, with a review by Daniels and Mandell (2014) reporting means ranging from 32 to 120 months of age across different studies, and a recent meta-analysis reporting a mean age of autism diagnosis of 60.5 months based on studies from 40 countries (van't Hof et al, 2021). Other recent studies suggest similar age of diagnosis: a mean of 46 months of age in Australia (Bent et al, 2020), 58 months in France (Rattaz et al, 2022) and 54 months in the USA (Hanley et al, 2021). Indeed, recent prevalence data from the USA suggests that, of 4681 children with an autism diagnosis at the age of 8, only 47% received a diagnosis before the age of 3 (Maenner et al, 2021). Importantly, in the UK context, Crane et al (2016) report an average delay of approximately 3.5 years between first contact with health professionals and confirmed diagnosis in their survey of parents of children with ASD, and only 11% of children in their sample were diagnosed before the age of three. Cohort studies also emphasise the rate of later diagnosis in ASD, with the number of diagnosed children in the Early Language in Victoria Study (ELVS) more than doubling between the ages of four and seven (Veness et al 2014), and percentages of children diagnosed with ASD rising from 0.9% at age five to 1.7% at age seven and 3.5% at age 11 within the Millennium Cohort Study (Dillenburger et al, 2015). Although some children do receive intervention services before they have a diagnosis, these do not appear to start significantly earlier (mean age for first receipt of services = 4.1 years; Hanley et al, 2021).

There is less information on diagnosis/recognition of difficulties among children with DLD, but Rannard et al (2004) reported that a quarter of parents in their sample noticed difficulties in their children's language and communication between 12 and 18 months of age, and a further quarter between 18 and 24 months of age, making around 50% showing concern by the time their child reached their second birthday. Absent or unusual babbling, poor intelligibility and late language onset were the main areas of concern noticed by parents. However, despite roughly half of parents being concerned about their child's communication skills by the age of two, only one third of the children in this study received any input from speech and language therapy services before the age of three, and 21% had no support until they started school (Rannard et al, 2004). Similarly, Tomblin et al (1997) found that only 29% of children who showed evidence of language impairment in Kindergarten had ever been referred to speech and language therapy services. In a more recent study in the UK, Norbury et al (2016) found that, of 9.92% of children who had language difficulties at school entry, only 39% had ever been referred to SLT services, and only 40% received any additional support at school. This figure may be higher in other countries, but a recent study in the USA also found that not all children who have speech and language difficulties are receiving SLT support, with only around 75% ever having received services for these difficulties during their lifetime (Davidson et al, 2022). Thus, there is a need to work towards more sensitive early measures of language and social communication, tapping into pre-verbal behaviours, such as joint attention and turn-taking, which form the foundations of language development (Curtin et al, 2021).

UK Referral figures for children of different ages confirm these results from both disorders, with Broomfield and Dodd (2011) finding that only 6% of referrals made to one speech and language therapy service under study were for children under two, 67% between the ages of two and five, and 27% for children of five or older. The Bercow Report (Bercow, 2008) also highlighted that UK parents continue to have difficulties accessing speech and language therapy services and 28% of those who responded felt they had had to fight for their child to receive a diagnosis and associated services. As noted earlier, inefficient diagnostic pathways may lead to poorer outcomes across a variety of areas in later life for children with communication impairments (Hus & Segal, 2021).

Current assessment issues

There may be many reasons for delayed identification, including limitations in the training of professionals in early assessment, resource issues and service eligibility criteria (Huerta & Lord, 2012). However, certainly in the UK and US, there is generally a lack of valid assessments appropriate for infants. Where they are used, the likelihood of social, cultural and linguistic bias is high (Dockrell & Marshall, 2015), untested adaptations are sometimes made (Cycyk et al, 2021), and the arbitrary cut-off scores are problematic (Spaulding, Plante & Farinella, 2006). Furthermore, the appropriateness of a given test is often not considered properly in practice (Friberg, 2010; Betz et al, 2013). Thus, the addition of appropriate, culturally and linguistically sensitive infant assessment tools, is one area where there is need for urgent improvement to avoid the consequences of late- or missed-diagnoses (Hus & Segal, 2021). In particular, the current model relies heavily on impairment focused assessments that use formal, static approaches – that is measures which tap into performance at once time point, without considering process (Spaulding et al, 2012; Roulstone et al, 2015; Dockrell & Marshall, 2015). These assessments are primarily designed to identify children scoring below a particular threshold rather than predicting risk or assessing change over time. (Hasson & Botting, 2010). When considering very young children, especially those at risk of communication difficulties, these tests may not serve the purpose of assessing possible difficulties because they are not feasible with infants, and because infants tend to show a wide range of ability at a given age (Law & Roy, 2008).

Speech and language therapists, teachers, psychologists and others who assess children using static tests have long known that there are some groups of children who are not well-served by traditional formal assessment methods (e.g. Spaulding et al, 2012). There are many reasons why a child may fail to perform well under static testing conditions, including cultural and linguistic diversity, shyness, difficulties with attention regulation, difficulties with social interaction and lack of familiarity with the formal testing process, as well as difficulties with the specific knowledge and skills being assessed (Chiat & Roy, 2007; Camilleri & Law, 2007; Hasson & Joffe, 2007). Because static testing usually seeks to remove the effects of the individual examiner by making the testing process exactly the same for each child, without environmental support or examiner feedback, it tells us only how the child performs on a specific measure under those conditions on a specific day. What it does not tell us is how the same child performs in more natural situations, or where they are engaged with the examiner in a collective effort to generate correct responses (Peña et al, 2007). This causes an issue with validity whereby the static assessment only captures a one-point estimate of the construct, rather than the construct itself (Messick, 1998; Hasson & Joffe, 2007; Camilleri & Law, 2007; Spaulding et al, 2012) and yet at the same time fails to eliminate all tester input effects (Muskett, Body & Perkins, 2012). Thus, a different approach is needed. Practitioners often take the approach of very informal observation or reliance on parent report (for example, by health visitors or doctors; Law et al, 2020) to counter the lack of formal assessment, but using a 'Dynamic Assessment' to measure emerging skills and learning potential offers a middle ground providing flexible yet objective measurement (Bamford et al, 2022).

Dynamic Assessment (DA)

In contrast to the static formal testing usually used by speech and language therapists, DA is more focused on the process of learning, and what a child's potential level of performance is, when supported by an adult who can provide prompting, cueing or teaching to help them improve their performance on the task . DA arose originally out of the socio-cultural theory of Vygotsky (1978), who described the "zone of proximal development” (ZPD) of the child’s skills in any area of learning. This describes the gap between the child’s habitual unaided performance and the level that they are able to reach when supported by an adult or more experienced peer. That is, children’s learning potential can be measured by observing what they can achieve in a scaffolded paradigm, rather than just their performance in an unaided scenario (Hasson & Joffe, 2007). This difference in static and dynamic methods has been noted and built on for school age children and now has widespread awareness (Deutsch & Reynolds, 2000) and some practice among some school psychologists (Hussein & Woods, 2019). However, to our knowledge no work has been done exploring the use of DA in preverbal infants.

Sternberg and Grigorenko (2002) described two main formats into which DA methodologies can be organised: the “sandwich” and the “cake” (see a recent description by Bamford et al, 2022). The use of test-teach-retest dynamic assessment procedures may be referred to as a “sandwich”, in which children are tested using static assessments before and after a brief intervention, to reveal the amount of change, or ‘gain’ that has taken place. The teaching phase typically involves a metacognitive element, which enables the child to learn which elements/strategies are required for successful completion of the task in question. Ratings of the child’s responsiveness during the ‘teach’ phase, together with the gains achieved between the test and the retest provide an indication of the child’s potential to learn. Within the field of speech and language assessment, this methodology has been adopted for the diagnostic purpose of distinguishing between low language ability and typically developing preschool children from specific culturally and linguistically diverse groups in the United States of America (Kapantzoglou et al., 2012; Peña et al, 2014). Static, standardised assessments can be biased against these children, leading to low scores for both typically developing children as well as children with language disorders. Dynamic assessments were found to reduce this bias, when assessing a range of areas including vocabulary (Kapantzoglou et al., 2012; Pena, Iglesias & Lidz, 2001), categorization (Ukrainetz, Harpell, Walsh and Coyle, 2000) and narrative (Pena et al., 2006).

The “cake” format, which sometimes forms the centre part of the ‘sandwich’ (see below for hybrid methods), is perhaps more suitable when assessing very young children, below the age of four. This method usually involves the integration of graduated prompts or feedback into the assessment session, as described by Campione and Brown (1987) and Carlson and Wiedl (1978) and used more recently by researchers such as Patterson et al. (2020) in preschool children. The examiner provides support to the child as they are completing the assessment, typically using a pre-determined cueing hierarchy that provides the child with increasingly explicit support to reach the correct answer or complete the task. What is measured here and interpreted as the size of the ZPD is the number of cues given to the child to enable them to complete the task, with more favourable scores being achieved by children who require less cueing to achieve success (Campione & Brown, 1987). This methodology was previously adopted with young preschool children (aged 30 to 36 months), who had a specific difficulty with expressive language (Bain & Olswang, 1995; Olswang & Bain, 1996). This DA targeted the immediate potential for children performing at the one-word stage of expressive language development, to produce two-term utterances, by using a series of graduated prompts which facilitated production of the two-term utterance. These prompts included elicitation questions, sentence completion and direct/indirect modelling. The key findings were that children’s scores on the DA were highly predictive of change over a nine-week period, both with (Bain & Olswang, 1995) and without intervention (Olswang & Bain, 1996).

Some DA research with young children in the United Kingdom and Europe has adopted a hybrid approach, incorporating both graduated prompts and an element of metacognitive intervention (Hasson et al., 2013; Camilleri, Hasson & Dodd, 2014). This has included research looking at bilingual and multilingual children in their first year of schooling (MacLeod & Glaspy, 2022). One big difference between this research and that from the USA is that children in the UK/Europe derive from a wide range of bilingual backgrounds, whereas the children in the studies cited above from the USA were recruited from specific linguistic backgrounds (e.g., Hispanic or Native American). The UK studies compared typically developing bilingual children and bilingual children with developmental language delays (on the Speech and Language Therapy caseload). The findings were that caseload children required greater assistance and made fewer gains in both vocabulary and sentence production (Hasson et al., 2013; Camilleri et al., 2013), further extending the evidence base that DA can be used to distinguish between these two groups. All but one of the children with developmental language delays were found to experience difficulties with components of the DA assessment (Camilleri et al., 2013).

Of the different approaches mentioned above, the graduated prompt approach is particularly suited when working with very young children, as it does not require the explicit metacognitive element that is crucial to the 'sandwich' or test-teach-retest approach. Although recent reviews by Hunt et al, (2019) and Orellana et al (2019) indicate that test-teach-retest methods are mostly chosen, the meta-analysis by Orellana et al (2019) concluded that modifiability ratings (similar to those used in the graduated prompts approach) showed more promise as an indicator of typical development vs. language impairment at least in bilingual children. This is therefore the approach which has been selected for the current study.

Present study

As discussed above, there is currently an emerging evidence base for the use of DA in speech and language therapy. Although the evidence to date is mainly from small-scale studies (see Joffe & Hasson, 2007; Orellana et al, 2019), there is a growing awareness in the field suggesting that DA can be used successfully with preschool children to determine the presence or absence of language and communication impairment, and in order to suggest strategies that may be used to support children in their communication development, or predict how children will respond to intervention (Hunt et al, 2019). This paper aims to address some gaps identified in the literature, including exploring the application of DA to children under two years of age, and the use of DA with high-risk children as a predictor of later language and communication skills. The areas of focus for these new DA tasks encompass 5 key communicative gestures and behaviours that have been reported in the literature as predictors of later language or as delayed in children with later social communication difficulties (Law et al, 2017; Rohlfing, 2019; Ramos-Cabo et al, 2019). Namely these areas comprise: Early receptive vocabulary (Markus et al, 2000); Response to joint attention (Salo et al, 2018); Motor imitation (Hanika & Boyer, 2019); Turn taking (Hendenbro et al, 2014); and non-verbal requesting behaviour (Ramos-Cabo et al, 2019). We acknowledge that the full development of a new DA tool for clinical practice will take many iterations. Thus, the objective here is to present work to establish initial 'proof of concept' and feasibility of an early infant measure.

The aims of the study were threefold:

1. To investigate whether reliable normative scores can be gained from a novel battery of very early DA procedures for use with infants under two years of age who have no first-degree relatives with communication or literacy difficulties.
2. To assess performance on these measures in relation to age, sex and standardised tests of communication in a low-risk group of children (normative sample).
3. To explore whether there are early indicators of (known groups) validity, using preliminary comparisons of infants at high-risk of communication disorders (siblings or parents with ASD, DLD or Dyslexia) with low-risk infants (siblings or parents with no known difficulties).

Method

Recruitment

Recruitment took place via social media, where contacts of the researchers were encouraged to share the project website on their own feeds. Parents of children in the correct age range could then visit the project website, view the project information sheet, and contact the research team if they agreed to take part. Children with bilingual exposure were not excluded from this sample, as long as they were exposed to English as one of the main languages of the home and could be assessed in English[1] Informed consent was taken from parents of all infants at the start of the research visit. The infants who participated were too young to give formal assent, but willingness to interact with the researcher and participate in activities was

[1] For the purposes of developing this task, English was the only language assessed. However, we acknowledge that in clinical practice, it is preferable to assess all home languages.

taken to indicate assent. The study was granted ethical approval from City University of London, Language and Communication Science Research Ethics Committee.

Participants

Two groups of children participated in the study: those at low-risk (n=51) and high-risk (n=41) of communication difficulties based on family history. The novel tasks were assessed for feasibility, reliability age relations and preliminary validity using the low-risk group only, to establish how the test performs for a normative sample (Pena, Spaulding & Plante, 2006). The high-risk groups were then used to compare scores to explore clinical usefulness and preliminary known-group validity.

Low-risk children (with typical siblings and parents)

Participants in the first part of this study were 51 low-risk children (25 female and 26 male) and had a mean age of 12.2m (*SD* = 3.0) at the time of assessment (see Table 1 for demographics). The majority of this group were white British (n = 37) with 4 who had mixed ethnicity, 2 who were Asian and 2 reported as being of 'other' ethnicity (6 children had no ethnicity data recorded). The inclusion criteria for this group were that children had no known developmental, physical or sensory difficulties at the time of recruitment, and their parents and elder siblings showed no evidence of language, communication or literacy difficulties. This sample included five children who were exposed to other European languages within their home in addition to English (Swedish (n=1), Finnish (n=1), German (n=2) and Italian (n=1), with exposure to their additional language varying between 20 to 40 hours per week (*M* = 24.8; *SD* = 8.6) as reported by parents on the UKCDI demographic questionnaire (Alcock et al, 2020). Children were largely recruited from Greater London (72.5%) although some were from other parts of England.

In total 31 of the children were first-born, and had parents with no reported history of difficulties with language, social communication or literacy development. Twenty children had older siblings (n=26 siblings), who were reported by parents to be developing typically. Parents completed the Children's Communication

Checklist (CCC-2) (Bishop, 2003) for 18 elder siblings of children in the sample who were aged 4;0 and above. All 18 elder siblings scored within the average range for the General Communication Composite score on the CCC-2, indicating no communication impairments, and none scored within the range of clinical concern on the Social Interaction Deviance Composite score. Of the remaining 8 siblings, 4 were older than 4;0 but did not have a CCC-2 completed by parents, and 4 were younger than 4;0 and therefore the CCC-2 could not be completed. However, in all cases, parents reported no concerns about their development. Additionally, 3 out of 4 siblings under the age of 4;0 were present during the assessment of the infant in the study, and were judged by the first researcher, who is an experienced speech and language therapist, to have language and communication skills within the typical range for their age. All elder siblings were therefore assumed to be typically developing. In addition, none of these 21 infants had parents who reported a history of difficulties with language, literacy or social communication.

Demographic data showed that 79.0% of mothers and 81.8% of fathers of these infants were aged 31 or older. The sample had high levels of parental education, with 95.5% of mothers and 86.4% of fathers reporting an undergraduate or postgraduate degree, and no parents reporting no formal educational qualifications. Overall, 76.7% of the sample reported family annual income of £42,000 or more. It is therefore acknowledged in the data that follow that these infants may not be representative across a broader range of socioeconomic status. See statistical group comparisons below.

High-risk children

For the final research question, a further 41 children were recruited who were considered at high risk of communication difficulty on account of their siblings or parents having existing developmental disorders. These children fell into 2 groups with the following inclusion criteria: i) those with siblings who had a diagnosis of autism or social communication disorder or were being assessed for this diagnosis, or whose siblings fell below the clinical threshold for the Social Interaction Deviance Composite score on the Children's Communication Checklist, 2nd Edition (CCC-2; Bishop, 2003). We refer to this group as the <u>Social-High-Risk</u>

(SHR) group; ii) infants with siblings and/or parents who had a diagnosis of Developmental Language Disorder, Dyslexia or other Speech Language and Communication Needs, or who were late to speak (defined as fewer than 50 single words at the age of two). We refer to these children as the Language-High-Risk (LHR) group. Children were not excluded on the basis of siblings with other genetic syndromes but infants with genetic syndromes, physical disabilities or sensory impairments were excluded. High-risk children were recruited from across England and are detailed below.

In the SHR group there were 18 children, 10 female and 8 male, with a mean age of 15.4 months (*SD* = 3.9). Of these 11 were white British and 1 was of mixed ethnicity (6 children with missing ethnicity data). Ten of these children had elder siblings or half-siblings with a confirmed ASD diagnosis. For the other 8 elder siblings, concerns were raised by parents about their social interaction skills. Where the elder sibling was aged 4;0 or above, parents completed the CCC-2, and a conservative Social Interaction Deviance Composite (SIDC) score of -10 or less was taken to indicate the presence of a social communication impairment in elder siblings who did not have an ASD diagnosis. Where the elder sibling was aged less than 4;0, the younger child was considered to fall into the SHR group if the elder sibling was under assessment for an ASD diagnosis. Two of the eight elder siblings without a formal autism diagnosis had been given a diagnosis of Social Communication Difficulties by a Speech and Language Therapist; a further child had previously been assessed for ASD and not given the diagnosis (although traits that could be consistent with mild ASD were identified), and two were currently undergoing ASD assessment. For three undiagnosed elder siblings, parents did not refer specifically to ASD when describing their elder child but made reference to difficulties interacting with others. All of these children showed low SIDC scores on the CCC-2 whilst no sibling of any child in the low-risk or LHR group had SIDC scores that would indicate significant social impairment. Two elder siblings in this group had additional diagnoses: one of Attention Deficit Disorder and one of Cri-Du-Chat Syndrome. One SHR child was exposed to another language (French) for 24 hours a week. In total 16.7% of SHR children were from Greater London.

In the LHR group there were 23 children, 15 female and 8 male, with a mean age of 13.2 months (*SD* = 3.0). Sixteen were white British and 5 were of mixed ethnicity (2 children had missing ethnicity data). Eight children in this group had elder siblings or half-siblings with concurrent speech and language difficulties, and three had parents, elder siblings or half-siblings with a history of late language emergence (no single words before the age of two). The remaining 12 children had parents and/or siblings/half-siblings with a diagnosis of dyslexia. In two out of eight cases of concurrent LI, the elder siblings also had learning difficulties and global developmental delays. In cases where elder siblings were 4;0 or older and speaking in sentences, the CCC-2 was used to confirm that they had impairments of language but did not have social communication difficulties. Where the elder sibling was younger than 4;0, their parents reported in all cases that they were receiving support from Speech and Language Therapy services for language or speech and that no concerns had been raised about social communication. In total, 47.8% of LHR children were from the Greater London area and one child was exposed to Spanish for 15 hours a week.

Demographic group comparisons

The parental age profile of the high-risk group was similar to that of the low-risk group, with 91.7% of mothers and 83.4% of fathers in the SHR group, and 86.3% of mothers and 91.0% of fathers in the LHR group aged 31 or older. Chi squared analysis using three age categories (30 or younger, 31 to 35 and 36 or older) showed a similar pattern of maternal (χ^2 (4) = 0.893, p=0.926) and paternal age (χ^2 (4) = 0.911, p=0.923) across groups.

Parental education levels were lower for both mothers and fathers in the SHR group, and for fathers in the LHR group, than the low-risk group (66.7% of mothers and 58.3% of fathers in the SHR group and 95.5% of mothers and 68.2% of fathers in the LHR group reported having an undergraduate or postgraduate degree). A significant group difference was found for maternal education level (with categories up to and including Level 3 qualifications collapsed[2] ; (χ^2 (4) = 12.376, p=0.015), such that the SHR group contained more

[2] Level 3 qualifications are end of high-school qualifications such as A-Levels usually taken at 18 years of age

mothers educated to Level 3 or lower, and fewer mothers educated at degree or postgraduate level than the other groups. Analysis of paternal education level using the same categories revealed that group differences were not statistically significant (χ^2 (4) = 8.049, p=0.090). Family income across three categories (£24,000 or less, £24,001 to £42,000, and £42,000 or more) also did not differ significantly across groups (χ^2 (4) = 3.055, p=0.549). 50% of SHR families and 71.4% of LHR families reported family annual income of £42,000 or more.

Groups showed a significant difference in birth order distribution (χ^2 (4) = 20.557, p<0.001), with the low-risk group containing higher numbers of first-born children, and the SHR and LHR groups containing more third or fourth children. In part, this was a function of how the groups were defined, as first-borns could not occur in the SHR group, but could be classified as control/LHR depending on parental dyslexia status.

The children from bilingual households do not appear different in terms of SES status or background from the main sample, but the number of children was too small to statistically analyse.

[Table 1 about here]

Measures

Demographic data were collected via questionnaire to parents for all groups, and which were completed for 45/51 low-risk infants, 20/23 LHR infants, 12/18 SHR infants. This questionnaire was the one included on the front of the published UK-CDI measure (please see Alcock et al, 2020 for more details), and asked about parental age, education and family income, as well as hours of exposure to additional languages. These variables were not used for categorisation into groups, which was done solely on family risk factors as described for each group below. An additional question about family history was also included. This question asked parents to say whether any family member had any of the following difficulties: Hearing Impairment; visual impairment; physical disability; autism spectrum disorder; Asperger Syndrome; speech and language difficulties; dyslexia or other problems with reading and spelling; learning difficulties; other developmental difficulties. Family history was also discussed with all parents by the first author who is a qualified and experienced SLT.

A set of **standardised measures** was administered for validation of novel assessment tasks. These included:

- The UK Communicative Development Inventory (UK-CDI) (Alcock et al, 2020). This is a parent-report measure, adapted from the MacArthur Bates Communicative Development Inventory (Fenson et al, 2007). Parents are given a list of 395 words across 19 categories, and asked to indicate whether their child understands and/or says these words. There is also a checklist of 63 gestures and pretend play actions, which parents are asked to indicate whether their child ever performs. The UK-CDI was normed on 1210 children from the UK, who were selected to match the demographic composition of the UK population, and may therefore represent children with a broader range of parental education levels than those included in this sample. However, the UKCDI demonstrates high internal validity for all scales (receptive vocabulary: α=0.99; expressive vocabulary α=0.99; gesture scale α=0.99). Strong correlations were also observed in the standardisation sample with scores on standardised measures of language and an object selection task that measured comprehension directly (Alcock et al, 2020). Parents were asked to complete the CDI for their children in English, as this was one of the main languages for all families.
- The Infant Toddler Checklist (ITC) from the Communication and Symbolic Behaviour Scales – Developmental Profile (CSBS-DP) (Wetherby & Prizant, 2002). This is a 24-item questionnaire, completed by parents, which generates three subscale scores for Social, Speech and Symbolic aspects of communication. The ITC was initially standardised on more than 2000 children in the USA, many of whom were recruited from the same geographic area. However, the standardisation sample matches that included in the present study in terms of having high levels of infants whose parents have completed degree-level or postgraduate education. The ITC shows good levels of internal consistency (α=0.93) and test-retest reliability (r=0.88), as well as strong correlation in the standardisation sample with other aspects of the CSBS-DP that involve more detailed parent questionnaires and examiner assessment (Wetherby & Prizant, 2002). Additionally, a large cohort study in Australia found the ITC to be a valid clinical tool for measuring early communication skills (Eadie et al, 2010).

- The Modified Checklist for Autism in Toddlers (M-CHAT) (Robins et al, 2001). This is a 23-item checklist, where parents are asked to answer “yes” or “no” to each item, based on their child’s typical behaviour. A subset of 6 items of this questionnaire (the “Core 6 items”) is considered to be particularly indicative of risk for a later diagnosis of ASD (Robins et al, 2001). The M-CHAT shows a high level of internal consistency (α=0.85) and also has high levels of sensitivity (0.97) and specificity (0.95) (Robins et al, 2001). Although the M-CHAT is designed for use from 18 months of age, it was included in this study due to its clear format and its potential for indicating emergent difficulties that are linked to ASD. Scoring for this checklist is according to the number of items failed, and higher scores therefore indicate more symptoms related to ASD.
- The Pre-school Language Scales, 4th Edition (PLS-4) (Zimmerman, Steiner and Pond, 2002). This is a standardised language assessment, providing scores for receptive and expressive language for children aged from birth to 6 years 11 months. In infancy, scores are mainly given from observation of infant communication during natural interaction, although some older children in the sample were administered receptive language items using toys or picture material. The PLS-4 was originally standardised in the USA, on a sample of 2400 children selected to match the demographic characteristics of the US population. The assessment then received additional UK standardisation with a sample of 800 children matching the UK demographic profile, who were similar in ethnicity to the children in this study, but had a broader range of parental education levels. Test-retest reliability (r=0.82-0.95) and internal consistency of this measure were high (α=0.72-0.95) in the standardisation sample, and standardisation of the measure showed a good ability to distinguish typically developing children from those with language disorders (Zimmerman, Steiner & Pond, 2002).

A set of novel **dynamic assessment** measures, designed and piloted by the authors for use in this study, was also administered to the children. These measured skills in five areas found in previous studies, to be associated with early communication skills including:

- Receptive vocabulary
- Motor imitation
- Response to joint attention
- Turn taking
- Requesting

These areas of development were chosen as representing core elements of early communication derived from a number of sources including existing reviews (e.g., Ramos-Cabo et al, 2019) and a review of the early communication literature (Spicer-Cain, 2019). These tasks were then tested in a feasibility phase involving 8 children aged 9-17 months (all monolingual; 6 white and 2 mixed ethnicity), and were judged to be engaging for the children, that children were able to complete the assessment and that parents found them acceptable. We concluded that the tasks formed an appropriate assessment for this age range and were likely to be predictors of later language (Spicer-Cain, 2019). This feasibility pilot also helped to guide scoring and number of trials on each task. Note that although early expressive language may be an important predictor, because of the very young target age of the children (12 months), a dynamic cueing hierarchy for this skill was not considered feasible. We therefore acknowledge that this set of DA tasks is preliminary and serves as a ‘proof of concept’ battery to determine whether initial reliability and feasibility can be achieved.

Based on the principles of DA, graded cueing hierarchies were devised to support children to achieve each of the tasks (Orellana et al, 2019). These are detailed for each task in more detail below, but overall were designed to provide three prompts if the child could not achieve the task independently. Generally, the first of these prompts was a repetition of the instruction, designed to draw the child’s attention to the task and give the child more processing time. The second prompt was more specific, and aimed to reduce the difficulty of the target task. The third prompt provided full support for the child to achieve the task. Administration of all DA tasks was videorecorded for reliability checking.

DA of receptive vocabulary

For the dynamic receptive vocabulary task, children were shown a series of five common items, which were taken out of a bag and placed in front of the child, without naming them. The items (cup, car, duck, ball and spoon) were chosen to represent words a child would typically acquire as part of their early vocabulary. For each of the five items, the child's attention was drawn using their name, and pointing to the array of items. The child was then asked to give one of the items to the researcher, accompanied by an open-hand gestural prompt. The cueing hierarchy in Appendix 1 was then used for each item. Items were returned to the array after each had been tested, so that the child was always looking at a choice of five items.

DA of motor imitation

Motor imitation was tested via imitation of actions on objects using a toy cup and spoon. The list of gestures included in the Actions and Gestures section of the UCKDI (Alcock et al, 2020) was reviewed, and used to choose these objects for use in the motor imitation task, considering previous research showing that young children are more likely to imitation actions involving objects (Kim et al, 2015). Actions were then chosen that could be performed with these objects, but which were mostly unrelated to their typical use, to enable the experimenter to be sure whether the infant was truly imitating the action, as opposed to just showing understanding of object function. A cup and spoon were given to the child at the start of the activity, and the experimenter then demonstrated the action using their own set of objects, and encouraged the child to copy using the phrases "X do it" and "your turn". Animated sound effects were also used by the experimenter to maintain the child's attention, although the child was not required to copy the sound, and most did not attempt to do so. The actions used were:

- Pretending to eat from the cup using the spoon
- Banging the spoon on the bottom of the cup
- Touching the spoon to the experimenter's nose
- Placing the cup upside down on the experimenter's head
- Stroking the spoon on the experimenter's arm

Allowances were made for the children's level of motor development, and any clear attempt to perform the target action was considered as correct, with no requirement for completely correct execution. The child was also credited for using either their own set of items or those of the experimenter, or for performing the actions on their own body or the experimenter's. For each of the five items, the cueing hierarchy in Appendix 1 was used.

DA of response to joint attention (point following)

Response to joint attention (RJA) was assessed based on the child's ability to follow adult pointing, during a picture-book reading task. A first words picture book containing large colourful photographs of everyday objects was used, with several objects pictured on each page. Unlike the other subtests, ten trials were run for this task, because the pilot study suggested both that infants at this age were more difficult to score on this item; and that increased items on this task were better tolerated than for other DA items (ideally all elements would have run with ten trials). For each RJA trial, the experimenter pointed at an item on the page, saying "Look! A (name of item)". To aid the scoring of the task, the items used for each child were chosen so that the child would have to make an obvious gaze shift from where they were currently looking to look at the item to which the adult was pointing. The sequence of cueing in Appendix 1 was used. If the child pointed to items in the book, the experimenter named these, and the child was allowed to look at each page until they lost interest, although only one trial was made on each page.

DA of turn-taking

Turn-taking skills were assessed using a ball-run toy designed for infants, where a ball is put into a hole and then runs down a spiral track. The experimenter first demonstrated the toy for the child by taking a turn, and then encouraged the child to take a turn using the phrase "X's turn"/ "you do it". Once the child was engaged with the toy, the experimenter initiated a turn-taking sequence by taking a turn themselves (see Appendix 1). Five turn sequences were then scored according to the procedure in Appendix 1. The experimenter and child then continued to play with the toy until the child lost interest, although only the first five turns were scored.

DA of social requesting

Requesting was measured using a disco ball, which spun and displayed colourful lights when it was switched on. The child was shown the toy, and once they were engaged with it, the toy was then switched off. Appendix 1 shows the cueing sequence which was then used to support the child to make a request to have the toy turned back on. Requests did not have to be verbal, and could be made using gesture, touch or vocalisation, as long as this was considered to be socially referenced (accompanied by eye contact to the experimenter or parent). Five trials were scored, and then the experimenter and the child continued to play with the toy until the child lost interest in it.

Procedure

Children were assessed by the first author who is a qualified SLT, in their home with a parent present. For the first fifteen minutes of the session, parents and children were video recorded playing with a standard set of toys. During this time, aspects of the PLS-4 which could be rated from observation were completed. The remainder of the appropriate items from the PLS-4 were administered, depending on the age and abilities of the child. The dynamic assessment measures were then administered and scored live during task completion. However, all tasks were videorecorded for later reliability checking. The total duration of the session was around 60 minutes for each child. This included DA administration and scoring of between 10 to 25 minutes. Parents were then given a set of questionnaires to complete and return to the research team, including the three standardised questionnaire measures listed above.

Analysis

Results were analysed using SPSS version 23.

For research question 1, intraclass correlations were used to assess reliability. Cronbach's alpha was used to report internal consistency.

For research question 2, due to the non-normal distribution of some variables, Spearman correlations were used to investigate the relationship between age and scores on each of the dynamic assessment measures. As age was significantly related to most scores, partial correlations were used to establish relationships between dynamic assessment scores and scores on other measures taken concurrently. Mann-Whitney U tests were used to compare scores across biological sex.

For research question 3, ANCOVAs were used to compare all 3 groups on the DA tasks, controlling for age. Assumptions of ANCOVA were checked: Homoscedasticity was verified via scatterplots of predicted against standardised residuals, and there were no outliers for any task. However, Shapiro Wilks tests showed that standardised residuals were significantly non-normally distributed in at least one group for all tasks. Transformation of data did not normalise the distributions. No difference in the pattern of results was observed when combining both high-risk groups and comparing to low-risk groups, thus this analysis is not reported. There was not enough variability in maternal education scores to consider this as a covariate.

No results changed substantively on any analysis when children from bilingual families (n=7) were removed, therefore all children are retained in the analyses that follow.

Results

Descriptive Statistics

Descriptive statistics for the five DA measures for the children in the low-risk control group only are reported in Table 2. This group of children act as a normative sample and thus are the basis for the development of the tool. Adjusted means from all groups are reported later when we compare scores across known groups. Scores on all measures were not normally distributed, with floor effects present on the measures of receptive language, motor imitation and turn taking, and ceiling effects present on the measures of response to joint attention and social requesting. Interquartile ranges are reasonably wide across all tasks.

[Table 2 about here]

Aim1: Reliability of the DA tasks

Inter-rater Reliability

In order to ensure the reliability of the above scores, a random subset of 27 videos, selected using stratified sampling to represent 25% of each risk group, was scored by a second rater to investigate inter-rater reliability for the dynamic assessment measures. The independent rater was given the DA scoring hierarchies, training in the scoring methods used, and the basic information about the project was explained, but they were not told any other information about the children, and so were blind to group status or other scores. Intra-class correlation coefficients for the dynamic assessment measures represent 'good' (>0.75) or 'excellent' (>0.90) agreement for all dynamic assessment measures except Turn Taking which was moderate at 0.70 (Koo & Li, 2016) when the whole sample was considered. All values were good or excellent for our normative (low-risk) and SHR samples. All values except turn-taking were in this range for the LHR group (see Table 3).

[Table 3 about here]

Inter-task correlations and internal consistency

With the effect of age statistically controlled, the five DA measures did not show significant partial correlations with one another, suggesting that they should not be combined into a single scale (see Table 4). Unsurprisingly, internal consistency of the battery was therefore low at $\alpha = 0.594$.

[Table 4 about here]

Aim 2: Relationship of DA tasks with age and standardised measures of communication

Relationship with Age

There were significant positive relationships between age and all but one of the DA tasks. For Receptive Language (r = .553, p < .001), Motor Imitation (r = .640, p < .001) and Turn-taking (r = .777, p < .001) these associations were all strong, whilst for Response to Joint Attention the relationship was moderate (r = .495, p < .001). There was no age relationship with the DA of Requesting (r = -.072, p = .620), with high variability of scores present at all ages. Fig 1 illustrates the findings.

[Fig 1 about here]

Relationship with Sex

Due to the non-normal distribution of scores on the DA measures, Mann-Whitney U tests were used to compare the scores of boys and girls from the low-risk control group on the five tasks. None of the comparisons showed significant differences, although there was a marginal difference on the motor imitation task in favour of girls (see Table 5).

[Table 5 about here]

Relationships with Other Measures

Scores on the DA tasks were compared with scores on other parent-reported and experimenter-administered standardised measures of communication ability. Three DA tasks showed moderate to large associations with at least one other measure taken concurrently. For the DA receptive language task, significant correlations were found with parent-reported receptive vocabulary on the UKCDI, and with receptive and expressive language scores on the PLS-4. The ITC Symbolic and Social subscale scores showed a significant association with the DA turn taking task, which was also significantly correlated with Total Gestures scores on the UKCDI. For the DA social requesting task, significant correlations were found with parent-reported expressive vocabulary on the UKCDI, and the ITC Social subscale (see Table 6).

After correcting for multiple comparisons using the Bonferroni method, only the association between the DA measure of receptive language and the PLS-4 Auditory Comprehension score remained significant.

[Table 6 about here]

Aim 3: Comparison of DA tasks across low-risk and high-risk groups One-way between-groups ANCOVAs with age entered as a covariate were run to evaluate group differences on the DA measures. Adjusted mean scores for each group on all tasks can be seen in Table 7.

[Table 7 about here]

Receptive Language

Age was significantly related to scores on the DA of receptive language within the ANCOVA model (F (1, 88) = 73.669, p < .001). Significant group differences were found on the dynamic receptive language task (F (2, 88)

= 5.218, p = .007, η^2_p = 0.106)[3]. Scores in the SHR group were significantly lower than those of the low-risk group (p = .002) and the LHR group (p = .016). Scores in the LHR group did not differ from the low-risk group (p = .558). See Table 7 for adjusted means and SEs.

Motor Imitation

The association between age and motor imitation scores was significant within the ANCOVA model (F (1, 88) = 53.021, p < .001). Scores on the DA of motor imitation did not differ significantly between groups (F (2, 88) = 1.212, p = .302, η^2_p= 0.027), although the mean score of the SHR group was lower than those of the other two groups.

Response to Joint Attention

The covariate age was significantly related to scores on the DA of response to joint attention within the ANCOVA model (F (1, 88) = 15.997 p < .001). The ANCOVA did not indicate significant differences between risk groups for response to joint attention scores (F (2, 88) = 0.511, p = .602, η^2_p = 0.011), although the mean score of the SHR group was lower than for the low-risk and LHR groups.

Turn-taking

Age was significantly related to score on the DA of turn taking within the ANCOVA model (F (1, 87) = 42.582, p < .001). Turn taking scores showed no significant differences among risk groups (F (2, 87) = 0.461, p = .633, η^2_p = 0.010), although the mean scores in both high-risk groups were lower than for the low-risk group.

[3] Notably the UK-CDI receptive vocabulary scale (Alcock et al, 2017) was not sensitive enough to detect differences across these groups (F (2, 74) = 2.114, p = .128, η^2_p = 0.054).

Social requesting

Age did not show significant relationship to scores on the DA of social requesting within the ANCOVA model (F (1, 86) = 2.879, p = .093) and neither did the ANCOVA show significant group differences in scores on the dynamic requesting task (F (2, 86) = 1.028, p = .362, η^2_p = 0.023).

This pattern of results was identical when combining the two high risk groups in comparison to the low-risk infants.

Discussion

The current study aimed to present 'proof of concept' findings from a set of new dynamic tasks of early communication for infants, with the long term aim of developing a tool that can be used reliably and easily within family homes. This study is unique in using DA to investigate the skills of infants at high risk of communication difficulties. The tasks presented here are a first step towards more reliable communication assessment of children at very early ages before most static measures are appropriate.

Task characteristics

While the five tasks showed only weak inter-correlations and therefore appear to be measuring different constructs, each measure had good inter-rater reliability (across all groups). In addition, four of the five DA tasks showed a significant correlation with age, indicating sensitivity to developing abilities in children within the age range studied here. However, the fact that there were some floor and ceiling effects, indicates that the tasks may need to be refined to capture a fuller range of language and communicative potential in both clinical and typically developing populations. This would eventually enable the creation of norms so that individual child scores on each task can be interpreted appropriately on a clinical basis according to age. No sex differences were observed in our normative sample, suggesting that DA might in the future serve as a useful tool for equality of diagnosis across boys and girls.

The tasks also showed significant correlations with standardised measures for our normative, low-risk sample, suggesting that they are valid and are tapping important constructs relevant to communication. Our DA tasks were also sensitive enough to detect some early differences between risk groups, especially for receptive language, which was the only task out of the five to show statistically significant differences. The SHR group also received lower mean scores than low-risk infants on the motor imitation, point following and turn taking tasks, but these did not constitute significant differences. No differences were found between LHR children and the children in the low-risk group. These findings are now discussed in more detail below.

Receptive language and other group comparisons

Receptive language was the only task to show significant differences across groups. Nevertheless, at this early stage of test development, the trend towards lower scores for children in the SHR group for turn taking, RJA and motor imitation (which all relate to the development of receptive language), is also worth noting and taking forward to the next iteration. Receptive language was also the only DA task to correlate with standardised tasks. This result in itself does not entirely limit the usefulness of the other DA tasks, since it may be that they are more sensitive than standardised tests at this age, or that they are measuring slightly different aspects of communication. However, taken together, our results highlight receptive language as the most promising early assessment domain, especially for children at risk of autism. This is particularly interesting given that receptive language has been found to be a strong predictor of language outcome both for children with autism and for late talkers (Brignall et al, 2019; Fisher, 2017).

We were somewhat surprised that the LHR group showed no differences compared to the low-risk group. This may be due to the group criteria including family members with dyslexia, which is diagnosed later, or because the pathway of difficulties for those with language disorders is more gradual and less identifiable in infancy. Notably receptive language was not different for the LHR group whereas this was already showing signs of impact for the SHR group. Several research studies have demonstrated the instability between early language delay and later language impairment (e.g. Dale et al, 2003; Reilly et al, 2010; Zambrana et al, 2014;

Duff et al, 2015; Rudolph et al, 2016). It may be the case that group differences based on *family history* were not evident here, but that individual children who later receive diagnoses of communication difficulties will show differences on the DA tasks as infants, indicating their predictive validity. The crucial aspect in validating these DA tasks is whether they can be used to identify children who require support early on. Work is ongoing to follow up the current cohort at school age to investigate this very question.

Strengths and limitations of the present study

The present study addressed a number of key gaps in the literature. Firstly, studies of language high-risk children are few, while numerous studies of social-high-risk children exist. Although few differences were evident between LHR and low-risk children, it may be that these will manifest later in childhood, particularly in terms of literacy outcomes (e.g., Zambrana et al. 2014). Secondly, this study is one of only a few studies to use DA methods to assess infants, particularly infants at high risk of communication disorders.

Age of the children

In order to recruit a sufficient sample, the age range of the children was wider than ideal. Our key aim for developing the tool was to keep the age range of the *low-risk children* reasonably tight and this was achieved, with only 3 of this group older than 16 months. However, the age range of the social high-risk group was wider. As with most clinical measures, we anticipate that a tool of this kind may be useful for identifying older children who are at risk of language and communication difficulties, and who are functioning at a lower level than expected for age. Attempts were made to control for age effects in statistical analyses, but it is acknowledged that results would be clearer and easier to interpret in a cohort that had a narrower age range at each assessment time point. There is also a suggestion in existing research that the profiles of children with language difficulties change with age, such that social communication difficulties and features relevant to ASD diagnosis become more prominent over time in children whose language difficulties appeared more specific earlier in childhood (e.g., Conti-Ramsden et al, 2006 Chiat & Roy, 2013; Charman et al, 2015). Replication of similar results in a sample with higher proportions of high-risk children, and with follow-up of the sample at later ages, would lend weight to the conclusions of this study, and would allow analysis of some trends that did not reach significance in the current sample, but appeared to have large effect size.

Sample diversity

In addition, the self-selecting nature of the sample means that results are not necessarily generalisable to the total population of young children. The parents who responded to recruitment advertising were typically educated at degree level or higher, were older than 30, and had higher levels of family income than the population as a whole. Their children may therefore not accurately represent children from a broader range of socio-economic status. It would therefore be useful to recruit a sample of participants more diverse in socio-economic factors such as family income and parental education, to assess whether this affects the pattern of results. Although 7 children came from bilingual families, removing them from analyses did not affect results. However, it was not possible to statistically compare children exposed to other languages to those in monolingual households due to small numbers. Increasing the diversity of the sample would be an important step for future research as this would help to establish the utility of culturally-sensitive assessments. For the purposes of this study the only language tested was English, but we acknowledge that in clinical practice, all home languages should be assessed. One of the potential strengths of DA is that tests may be more easily adapted to work across several languages.

It is also the case that parents who already had concerns about the development of their child may have been more likely to enrol them to participate in this research project, so that their communication development could be evaluated. It is therefore possible that the groups of high-risk children who participated in the study contain higher numbers of children with developmental communication difficulties than would be the case in an unselected sample, although our results suggest this is not the case. Indeed, the opposite may also be true, in that parents with concerns about their second child might have avoided a study where issues could have been revealed.

Measurement issues

Another potential limitation was that some of the siblings were too young to complete the CCC-2 (although we introduced alternative criteria for these few) and some parents did not return sibling questionnaires. Furthermore, there was also no direct assessment of sibling/parent probands with communication or literacy difficulties and classification of children was dependent on parental reports. Together, these limitations reveal that establishing sibling status is not a straightforward process. In future research, more objective classification using direct measures of parent and elder sibling language, social communication and literacy skills could be of benefit.

The range of measures taken in this study allowed evaluation across many areas of development, and also allowed the comparison of novel tasks with established measures for a normative sample. However, it is important to note that the assessments were not blinded, as the first author and assessor had knowledge of the children's risk status. Nevertheless, good inter-rater reliability was achieved for novel tasks after coding by researchers who were blind to the children's group status. We did not include a DA of expressive vocabulary. This is because we felt it would be difficult to define a cuing hierarchy with children who were mostly non-verbal, but assessment of utterances or vocalisations could be explored in further studies.

Finally, the fact that there was only one task per skill may also serve to limit the assessment battery. It should be noted however, that infants are very restricted in terms of the time and attention they can apply to formal tasks. Indeed, we originally intended all DA items to have 10 trials, but our feasibility pilot suggested that only RJA was tolerated sufficiently for this number. The DA tasks presented here are being developed with a view to offering a quick yet reliable addition to tools used by health visitors and other infant-care professionals.

Future directions and Dynamic Assessment in practice

Some adaptations could be useful in future versions of our tasks. In particular, it may be beneficial to combine some of the tasks into one composite subtask tapping into early non-verbal communication skills. The receptive vocabulary task could also be modified to incorporate novel word learning opportunities; an expressive task could be included for older children; and the scoring of tasks could be standardised across areas. Other aspects of communication which are emerging, and may be more easily measurable than vocabulary (such as vocalizations), might also be useful target behaviours to include as part of a DA of early communication. Thus, further development of the DA tasks should be a focus of future research, including data from large diverse samples, trials for additional items and more detailed investigations into reliability and validity of the measures especially at an individual level. Ultimately, if key tasks differentiating children at high risk of later language delay and/or social communication difficulties could be identified through further studies, it would be possible to trial intervention programmes for children who show early signs of these difficulties and evaluate the effect on outcomes. It may be that the children in our high-risk groups do not go on to develop communication disorders, despite their family histories and conversely, some children in the low-risk group may develop a communication disorder. Therefore, the utility and inter-rater reliability established for the DA tasks in the current study warrant future work on developing and formalising these tasks in order to improve the prediction of future difficulties.

In spite of the need for further development of these procedures, we concur with Hasson & Joffe (2007), in believing that dynamic approaches are a promising way forward for providing practitioners with a reliable user-friendly screening tool for identifying infants at risk. DA tasks have the advantage of being a quick to administer direct assessment that is infant-friendly and ecologically valid in contrast to existing standardised tools that are often used inappropriately (Dockrell & Marshall, 2015; Cycyk et al, 2021; Spaulding, Plante & Farinella, 2006; Betz et al, 2013). We have shown that they can also be carried out with reliability. However, further work is needed to establish clinical discrimination between individual children who will need language and social support, and those who do not (Szatmari et al. 2016).

Lastly, for DA to be used in practice, the issue of training would need developing and evaluating at an individual case level. In the present study, all assessments were completed by the same Speech and Language Therapist (SLT; first author), but in practice any infant assessments would ideally be available to a wider group of professionals such as health visitors following full test development. Continuing work suggests that this training would be much easier and quicker than for most DAs, but rigorous further development of the tool is needed before any clinical implementation.

Conclusions and implications

This study suggests that Dynamic Assessment for infants may be feasible and useful, especially in the domain of receptive language. Many children with communication disorders are still being identified too late to access the critical early intervention they need, and appropriate standardised tests are not currently available. Although we have focussed here on children at risk of Autism and DLD, our hope for the long term is that tools can be developed to identify communication issues in a wide range of children.
We acknowledge that there are numerous barriers to early identification, especially in the UK. However, if parents and professionals were to have a reliable screening tool to identify the key markers of communication disorders in early life, this may increase the number of children identified before the age of two or three. This would in turn allow more children to access intervention designed to improve outcomes. This work represents just the first steps towards such a tool. Further longitudinal work will play a key part in determining which skills in infancy are predictive of communication problems not just at preschool age, but into the school years and adolescence. Only then can children receive earlier intervention and thereby attain more positive academic and psychosocial outcomes later in life.

Acknowledgements

The authors would like to thank all the families who took part in this study, City University of London for doctoral funding to the first author, and the Heather van der Lely Trust for supporting follow up work resulting from this project.

Data Availability Statement

The data produced by this study is preliminary and therefore is not hosted on an open access platform. However, data requests can be made to the first or corresponding author.

References

Alcock, K., Meints, K., & Rowland, C. (2020). *The UK Communicative Development Inventories: Words and Gestures*. Havant: J&R Press.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition.* Washington DC: American Psychiatric Association.

Bain, B. A. & Olswang, L. B. (1995). Examining readiness for learning two-word utterances by children with specific expressive language impairment: Dynamic assessment validation. *American Journal of Speech-Language Pathology, 4*(1)*,* 81-91

Baron-Cohen, S., Scott, F., Allison, C., Williams, J., Bolton, P., Matthews, F., & Brayne, C. (2009). Prevalence of autism spectrum conditions: UK school-based population study. *The British Journal of Psychiatry, 194*, 500-509.

Bent, C.A., Barbaro, J. & Dissanayke, C. (2020). Parents' experiences of the service pathway to an autism diagnosis for their child: what predicts an early diagnosis in Australia? *Research in Developmental Disabilities, 103*, epub 103689.

Bercow, J. (2008). *The Bercow Report: A Review of Services for Children and Young People (0-19) with Speech, Language and Communication Needs.* Nottingham: DCSF.

Betz, S.K., Eickhoff, J.R., & Sullivan, S.F. (2013). Factors influencing the selection of standardized tests for the diagnosis of specific language impairment. *Language Speech and Hearing Services in Schools*,44, 133–146.

Bishop, D.V.M. (2003). *The Children's Communication Checklist, 2nd Edition.* London: Pearson.

Bishop D.V.M. (2010). Which neurodevelopmental disorders get researched and why? *PLoS ONE* 5(11): e15112.

Bishop, D., Snowling, M., Thompson, P., Greenhalgh, T., & Consortium (2017). Phase 2 of CATALISE: a multi-national and multi-disciplinary Delphi consensus study of problems with language development: terminology. *Journal of Child Psychology and Psychiatry, 58*(10), 1068-1080.

Botting, N. (2020). Language, literacy and cognitive skills of young adults with developmental language disorder (DLD). *International Journal of Language & Communication Disorders*, *55*(2), 255-265.

Botting N., Durkin K., Toseeb U., Pickles A. & Conti-Ramsden G. (2016). Emotional health, support, and self-efficacy in young adults with a history of language impairment, *British Journal of Developmental Psychology,* 34, 538–554.

Boyle, J. (2011). Speech and language delays in preschool children. *British Medical Journal*, 343, d5181.

Brignell, A., May, T., Morgan, A. T., & Williams, K. (2019). Predictors and growth in receptive vocabulary from 4 to 8 years in children with and without autism spectrum disorder: A population-based study. *Autism,* 23(5), 1322-1334.

Broomfield, J., & Dodd, B. (2011). Children with speech and language disability: caseload characteristics. *International Journal of Language and Communication Disorders, 39*(3), 303-324.

Camilleri, B. & Botting, N. (2013). Beyond static assessment of children's receptive vocabulary: the dynamic assessment of word learning (DAWL). *International Journal of Language and Communication Disorders, 48(5),* 565-581.

Camilleri, B., Hasson N., & Dodd, B. (2014). Dynamic assessment of bilingual children's language at the point of referral. *Educational & Child Psychology, 31(2)* 58-73.

Camilleri, B., & Law, J. (2007). Assessing children referred to speech and language therapy: Static and dynamic assessment of receptive vocabulary. *Advances in Speech Language Pathology*, 9(4), 312-322.

Campione, J. C., & Brown, A. L. (1987). *Linking dynamic assessment with school achievement.* In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential*. New York: Guilford.

Charman, T., Ricketts, J., Dockrell, J., Lindsay, G., & Palikara, O. (2015). Emotional and behavioural problems in children with language impairments and children with autism spectrum disorders. *International Journal of Language and Communication Disorders, 50*(1), 84-93.

Chen, J. L., Leader, G., Sung, C., & Leahy, M. (2015). Trends in employment for individuals with autism spectrum disorder: A review of the research literature. *Review Journal of Autism and Developmental Disorders*, *2*(2), 115-127.

Chiat, S. & Roy, P. (2007). The preschool repetition test: an evaluation of performance in typically developing and clinically referred children. *Journal of Speech, Language and Hearing Research,* 50, 429-443.

Chiat, S., & Roy, P. (2013). Early predictors of language and social communication impairments at aged 9-11 years: a follow-up study of early-referred children. *Journal of Speech, Language and Hearing Research, 56*(6), 1824-1836.

Christensen, D., Bilder, D., Zahorodny, W., Pettygrove, S., Durkin, M., Fitzgerald, R., Rice, C., Kurzius-Spencer, M., Baio, J. & Yeargin-Allsop, M. (2016). Prevalence and characteristics of autism spectrum disorder among 4 year old children in the Autism and Developmental Disabilities Monitoring Network. *Journal of Developmental Behavioural Paediatrics, 37*(1), 1-8.

Conti-Ramsden, G., Simkin, Z., & Botting, N. (2006). The prevalence of autism spectrum disorders in adolescents with a history of specific language impairment (SLI). *Journal of Child Psychology and Psychiatry, 47*(6), 621-628.

Conti-Ramsden, G., Durkin, K., Toseeb, U., Botting, N. and Pickles, A. (2018). Education and employment outcomes of young adults with a history of developmental language disorder. *International Journal of Language and Communication Disorders*, *53*(2), pp. 237–255.

Crane, L., Chester, J., Goddard, L., Henry, L., & Hill, E. (2016). Experiences of autism diagnosis: a survey of over 1000 parents in the United Kingdom. *The International Journal of Research and Practice, 20*(2), 153-162.

Cycyk, L. M., De Anda, S., Moore, H., & Huerta, L. (2021). Cultural and linguistic adaptations of early language interventions: Recommendations for advancing research and practice. *American Journal of Speech-Language Pathology*, 30(3), 1224-1246.

Dale, P., Price, T., Bishop, D., & Plomin, R. (2003). Outcomes of early language delay: I. Predicting persistent and transient language difficulties at 3 and 4 years. *Journal of Speech, Language and Hearing Research, 46*(3), 544-560.

Daniels, A., & Mandel, D. (2014). Explaining differences in age at autism spectrum disorder diagnosis: a critical review. *Autism, 18*(5), 583-597.

Davidson, M., Alonzo, C. & Stransky, M. (2022). Access to speech and language services and service providers for children with speech and language disorders. *American Journal of Speech-Language Pathology,* online ahead of print.

De Giacomo, A., & Fombonne, E. (1998). Parental recognition of developmental abnormalities in autism. *European Child and Adolescent Psychiatry, 7*(3), 131-136.

Deutsch, R., & Reynolds, Y. (2000). The use of dynamic assessment by educational psychologists in the UK. *Educational Psychology in Practice,* 16(3), 311-331.

Dillenburger, K., Jordan, J., McKerr, L., & Keenan, M. (2015). The millennium child with autism: early childhood trajectories for health, education and economic wellbeing. *Developmental Neurorehabilitation, 18*(1), 37-46.

Dockrell, J. E., & Marshall, C. R. (2015). Measurement issues: assessing language skills in young children. *Child and Adolescent Mental Health*, 20(2), 116-125.

Dubois, P., St-Pierre, M. C., Desmarais, C., & Guay, F. (2020). Young adults with developmental language disorder: a systematic review of education, employment, and independent living outcomes. *Journal of Speech, Language, and Hearing Research*, 63(11), 3786-3800.

Duff, F., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry, 56*(8), 848-856.

Filipek, P.A., Accardo, P.J., Baranek, G.T., Cook, E.H., Dawson, G., Gordon, B., Gravel, J.S., Johnson, C.P., Kallen, R.J., Levy, S.E., Minshew, N.J., Prizant, B.M., Rapin, I. Rogers, S.J., Stone, W.L., Teplin, S., Tuchman, R.F. & Volkmar, F. (1999). The screening and diagnosis of autism spectrum disorders. *Journal of Autism and Developmental Disorders, 29*(6), 439-484.

Fisher, E. L. (2017). A systematic review and meta-analysis of predictors of expressive-language outcomes among late talkers. Journal of Speech, Language, and Hearing Research, 60(10), 2935-2948.

Friberg, J. (2010). Considerations for test selection: how do validity and reliability impact diagnostic decisions. *Child Language Teaching and Therapy*, 26(1), 77-92.

Hanika, L., & Boyer, W. (2019). Imitation and social communication in infants. *Early Childhood Education* Journal, 47(5), 615-626.

Hanley, A., Nguyen, Q., Badawi, D., Chen, J., Ma, T. & Slopen, N. (2021). The diagnostic odyssey of autism: a cross-sectional study of 3 age cohorts of children from the 2016-2018 National Survey of Children's Health. *Child and Adolescent Psychiatry and Mental Health, 15*(58).

Harmuth, E., Silletta, E., Bailey, A., Adams, T., Beck, C., & Barbic, S. P. (2018). Barriers and facilitators to employment for adults with autism: A scoping review. *Annals of International Occupational Therapy*, *1*(1), 31-40.

Hasson, N. & Joffe, V. (2007). The case for dynamic assessment in speech and language therapy. *Child Language Teaching and Therapy, 23*(1), 9-25.

Hasson, N. & Botting, N. (2010). Dynamic assessment of children with language impairments: a pilot study. *Child Language Teaching and Therapy*, 29, 52-70.

Hasson, N., Camilleri, B., Jones, C., Smith, J., & Dodd, B. (2013). Discriminating disorder from difference: using dynamic assessment with bilingual children. *Child Language Teaching and Therapy, 29*(1), 57-75.

Hedenbro, M., & Rydelius, P. A. (2014). Early interaction between infants and their parents predicts social competence at the age of four. *Acta Paediatrica*, *103*(3), 268-274

Henry, L.A. and Botting, N. (2017). Working memory and developmental language impairments. *Child Language Teaching and Therapy*, *33*(1), 19–32.

Hollocks, M. J., Lerh, J. W., Magiati, I., Meiser-Stedman, R., & Brugha, T. S. (2019). Anxiety and depression in adults with autism spectrum disorder: a systematic review and meta-analysis. *Psychological Medicine*, 49(4), 559-572.

Huerta, M., & Lord, C. (2012). Diagnostic evaluation of autism spectrum disorders. *Pediatric Clinics,* 59(1), 103-111.

Hunt, E., Nang, C., Meldrum, S. & Armstrong, E. (2019). Can dynamic assessment identify language disorder in multilingual children? Clinical applications from a systematic review. *Language, Speech and Hearing Services in Schools, 53,* 598-625.

Hussain, S., & Woods, K. (2019). The use of dynamic assessment by educational psychologists in the early years foundation stage. *Educational Psychology in Practice,* 35(4), 424-439.

ICAN & RCSLT (2018). *Bercow: Ten years on.* London: ICAN & RCSLT.

Jullien, S. Secreening for language and speech delay in children under five years. *BMC Pediatric*s, 21 (Suppl 1), 362.

Kapantzoglou, M., Restrepo, M. A., & Thompson, M. S. (2012). Dynamic assessment of word learning skills: Identifying language impairment in bilingual children. *Language, Speech and Hearing Services in Schools, 43* (1), 81-96.

Kim, Z., Oturai G., Kiraly, I. & Knopf, M. (2015). The role of object and effects in action imitation: comparing the imitation of object-related actions vs gestures in 18-month-old infants. *Infant Behaviour and Development*, 41, 43-51.

Koo, T.K. & Li, M.Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163.

Law, J. & Roy, P. (2008). Parental report of infant language skills: a review of the development and application of the communicative development inventories. *Child and Adolescent Mental Health*, 13(4), 198-206.

Law, J., Charlton, J., Dockrell, J., Gascoigne, M., McKean, C., & Theakston, A. (2017). *Early Language Development: Needs, provision and intervention for pre-school children from socio-economically disadvantaged backgrounds*. London: Education Endowment Foundation.

Law, J., Charlton, J., McKean, C., Watson, R., Roulstone, S., Holme, C., Gilroy, V., Wilson, P. & Rush, R. (2020). *Identifying and supporting children’s early language needs*. London: Department of Education and Public Health England.

Lindsay, G., Desforges, M., Dockrell, J., Law, J., Peacey. N. & Beecham, J. (2008). *Effective and efficient use of resources in services for children and young people with speech, language and communication needs.* Warwick: DCSF.

MacLeod, A. A., & Glaspey, A. M. (2022). Dynamic assessment of multilingual children's word learning. *International Journal of Language & Communication Disorders,* Early Online.

Maenner, M.J., Shaw, K.A., Bakian, A.V., Bilder, D.A., Durkin, M.S., Esler, A., Furnier, S.M., Hallas, L., Hall-Lande, J., Hudson, A., Hughes, M.M., Patrick, M., Pierce, K., Poynter, J.N., Salinas, A., Shenouda, J., Vehorn, A., Warren, Z., Constantino, J.N., DiRienzo, M., Fitzgerald, R.T., Grzybowski, A., Spivey, M.H., Pettygrove, S., Zahorodny, W., Ali, A., Andrews, J.G., Baroud, T., Gutierrez, J., Hewitt, A., Lee, L.C., Lopez, M., Mancilla, K.C., McArthur, D., Schwenk, Y.D., Washington, A., Williams, S. & Cogswell, M.E. Prevalence and characteristics of Autism Spectrum Disorder among children aged 8 years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2018. *MMWR Surveillance Summaries,* 70(11), 1-16.

MacGregor, K. (2020). How we fail children with developmental language disorder. *Language, Speech and Hearing Services in Schools, 51*(4), 981-992.

Markus, J., Mundy, P., Morales, M., Delgado, C. E., & Yale, M. (2000). Individual differences in infant skills as predictors of child-caregiver joint attention and language. *Social Development*, 9(3), 302-315.

Meilleur, A., & Fombonne, E. (2009). Regression of language and non-language skills in pervasive developmental disorders. *Journal of Intellectual Disability Research, 53*(2), 115-124.

Messick, S. (1998). Test validity: A matter of consequence. *Social Indicators Research*, 45(1), 35-44.

Muskett, T., Body, R., & Perkins, M. (2012). Uncovering the dynamic in static assessment interaction. *Child Language Teaching and Therapy,* 28(1), 87-99.

Norbury, C., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G. & Pickles, A. (2016). The impact of non-verbal ability on prevalence and clinical presentation of language disorder: evidence from a population study. *Journal of Child Psychology and Psychiatry, 57*(11), 1247-1257.

Olswang ,L.B. & Bain, B. A. (1996). Assessment information for predicting upcoming change in language production. *Journal of Speech and Hearing Research, 39,* 414-423.

Orellana, C. I., Wada, R., & Gillam, R. B. (2019). The use of dynamic assessment for the diagnosis of language disorders in bilingual children: A meta-analysis. *American Journal of Speech-Language Pathology*, 28(3), 1298-1317.

Patterson, J. L., Rodríguez, B. L., & Dale, P. S. (2020). Dynamic assessment language tasks and the prediction of performance on year-end language skills in preschool dual language learners. *American Journal of Speech-Language Pathology*, *29*(3), 1226-1240.

Peña, E.D., Iglesias, A., & Lidz, C.S. (2001). Reducing test bias through dynamic assessment of children's word learning ability. *American Journal of Speech-Language Pathology, 10,* 138-154.

Peña, E.D., Gillam, R.B., Malek, M., Felter, R., Resendiz, M. & Fiestas, C. (2006). Dynamic assessment of children from culturally diverse backgrounds: Application to narrative assessment. *Journal of Speech, Language and Hearing Research, 49,* 1037-1057.

Peña, E., Resendiz, M. & Gillam, R. (2007). The role of clinical judgements of modifiability in the diagnosis of language impairment. *Advances in Speech-Language Pathology*, 9(4), 332-345.

Peña, E. D., Gillam, R. B., & Bedore, L. M. (2014). Dynamic assessment of narrative ability in English accurately identifies language impairment in English language learners. *Journal of Speech, Language, and Hearing Research*, *57*(6), 2208-2220.

Peña, E. D., Spaulding, T. J., & Plante, E. (2006). The composition of normative groups and diagnostic decision making: Shooting ourselves in the foot. *American Journal of Speech-Language Pathology,* 15, 247–254.

Ramos-Cabo, S., Vulchanov, V., & Vulchanova, M. (2019). Gesture and language trajectories in early development: An overview from the autism spectrum disorder perspective. *Frontiers in Psychology*, *10*, 1211.

Rannard, A., Lyons, C., & Glenn, S. (2004). Children with specific language impairment: parental accounts of the early years. *Journal of Child Health Care, 8*(2), 165-176.

Rattaz, C., Loubsersac, J., Michelon, C., Geoffray, M.M., Picot, M.C., Munir, K. & Baghdadli, A. (2022). Factors associated with age of diagnosis in children with autism spectrum disorders: report from a French cohort. *Autism,* Online ahead of print.

Reilly, S., Wake, M., Ukomunne, O., Bavin, E., Prior, M., Cini, E., Conway, L., Eadie, P. & Bretherton, L. (2010). Predicting language outcomes at 4 years of age: findings from the Early Language in Victoria Study. *Pediatrics, 126*(6), e1530-e1537.

Reilly, S., McKean, C., Morgan, A. & Wake, M. (2015). Identifying and managing common childhood language and speech impairments. *British Medical Journal*, 350, h2318.

Robins, D., Fein, D., Barton, M., & Green, J. (2001). The Modified Checklist for Autism in Toddlers: an initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 31*(2), 131-144.

Rogge, N., & Janssen, J. (2019). The economic costs of autism spectrum disorder: A literature review. *Journal of Autism and Developmental Disorders*, 49(7), 2873-2900.

Rohlfing, K. J. (2019). Learning language from the use of gestures. In J. S. Horst & J. von Koss Torkildsen (Eds.), *International Handbook of Language Acquisition* (pp. 213–233). London: Routledge.

Roman-Urrestarazu, A., van Kessel, R., Allison, C., Matthews, F.E., Brayne, C. & Baron-Cohen, S. Association of race/ethnicity and social disadvantage with autism prevalence in 7 million school children in England. *JAMA Pediatrics, 175*(6), e210054.

Roulstone, S., Marshall, J., Powell, G., Goldbart, J., Wren, Y., Coad, J., Daykin, N., Powell, J., Lascelles, L., Hollingworth, W., Emond, A., Peters, T., Pollock, J., Fernandes, C., Moultrie, J., Harding, S., Morgan, L., Hambly, H., Parker, N. & Coad, R. (2015). Evidence-based intervention for preschool children with primary speech and language impairments: Child Talk – an exploratory mixed methods study. *Programme Grants for Applied Research*, 3(5).

Rudolph, J., & Leonard, L. (2016). Early language milestones and specific language impairment. *Journal of Early Intervention, 38*(1), 41-58.

Salo, V. C., Rowe, M. L., & Reeb-Sutherland, B. C. (2018). Exploring infant gesture and joint attention as related constructs and as predictors of later language. *Infancy*, 23(3), 432-452.

Spaulding, T. J., Plante, E., & Farinella, K. A. (2006). Eligibility criteria for language impairment. *Language, Speech and Hearing Services in Schools.* 37, 61-72.

Spaulding, T., Szulga, M. & Figueroa, C. (2012). Using norm-referenced tests to determine severity of language impairment in children: disconnect between US policy makers and test developers. *Language, Speech and Hearing Services in Schools,* 43, 176-190.

Spicer-Cain, H. (2019). *Early markers of communication difficulties in children at high risk of Autism Spectrum Disorder and Language Impairment.* [Unpublished doctoral thesis]. City, University of London.

Sternberg, R. J., & Grigorenko, E. L. (2002). *Dynamic testing: The nature and measurement of learning potential*. New York: Cambridge University Press.

Szatmari, P., Chawarska, K., Dawson, G., Georgiades, S., Landa, R., Lord, C., ... & Halladay, A. (2016). Prospective longitudinal studies of infant siblings of children with autism: lessons learned and future directions. *Journal of the American Academy of Child & Adolescent Psychiatry,* 55(3), 179-187.

Thordardottir, E., Tobpas, S. & Working Group 3 of COST Action IS1406. (2021). How aware is the public of the existence, characteristics and causes of language impairment in childhood and where have they heard about it? A European survey. *Journal of Communication Disorders, 89,* 106057.

Tomblin, J., Records, N., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language and Hearing Research, 40*(6), 1345-1260.

Ukrainetz, T,A., Harpell, S., Walsh, C., & Coyle, C. (2000). A preliminary investigation of dynamic assessment with Native American kindergartners. *Language, Speech and Hearing Services in Schools, 31, 142-154.*

van ‘t Hof, M., Tisseur, C., van Berckelear-Onnes, I., van Nieuwenhuyzen, A., Daniels, A.M., Deen, M., Hoek, H.W. & Ester, W.A. (2021). Age at autism spectrum disorder diagnosis: a systematic review and meta-analysis from 2012 to 2019. *Autism, 25*(4), 862-873.

Veness, C., Prior, M., Eadie, P., Bavin, E., & Reilly, S. (2014). Predicting autism diagnosis by 7 years of age using parent report of infant social communication skills. *Journal of Paediatrics and Child Health, 50*(9), 693-700.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* Cambridge, MA.: Harvard University Press.

Wetherby, A., & Prizant, B. (2002). *Communication and Symbolic Behaviour Scales - Developmental Profile.* Baltimore: Brookes.

Zambrana, I., Pons, F., Eadie, P., & Ystrom, E. (2014). Trajectories of language delay from age 3 to 5: persistence, recovery and late onset. *International Journal of Language and Communication Disorders, 49*(3), 304-316.

Zimmerman, I., Steiner, V., & Pond, R. (2002). *Preschool Language Scales, 4th Edition.* San Antonio: The Psychological Corporation.

Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., & Szatari, P. (2005). Behavioural manifestations of autism in the first year of life. *International Journal of Developmental Neuroscience, 23*, 143-152.

Zwaigenbaum, L., Bryson, S., Lord, C., Rogers, S., Carter, A., Carver, L. C., Chawarska, K., Constantino, J., Dawson, G., Dobkins, K., Fein, D., Iverson, J., Klin, A., Landa, R., Messinger, D., Ozonoff, S., Sigman, M., Stone, W., Tager-Flusberg, H. & Yirmiya, N. (2009). Clinical assessment and management of toddlers with suspected autism spectrum disorder: insights from studies of high-risk infants. *Pediatrics, 123*(5), 1383-1391.

Table 1: Demographic information for the whole sample by group

	Whole Sample (n=92)	**Low-risk Group (n=51)**	**SHR Group (n=18)**	**LHR Group (n=23)**
Child Age (mean/SD)	13.1(3.4)	12.2 (3.0)	15.4 (3.9)	13.2(3.0)
Child Sex				
Male	42 (45.6%)	26 (51.0%)	8 (44.4%)	8 (34.8%)
Female	50 (54.4%)	25 (49.0%)	10 (55.6%)	15 (65.2%)
Birth order				
1	40 (43.5%)	30 (58.8%)	0 (0%)	10 (43.5%)
2	36 (39.1%)	16 (31.4%)	12 (66.7%)	8 (34.8%)
3 or higher	16 (17.4%)	5 (9.8%)	6 (33.3%)	5 (21.7%)
Maternal Age				
25 or younger	2 (2.6%)	2 (4.7%)	0 (0%)	0 (0%)
26-30	11 (14.3%)	7 (16.3%)	1 (8.3%)	3 (13.6%)
31-35	39 (50.6%)	21 (48.8%)	6 (50.0%)	12 (54.5%)
36 or older	25 (32.5%)	13 (30.2%)	5 (41.7%)	7 (31.8%)
Paternal Age				
25 or younger	1 (1.3%)	1 (2.3%)	0 (0%)	0 (0%)
26-30	11 (14.1%)	7 (15.9%)	2 (16.7%)	2 (9.1%)
31-35	26 (33.3%)	14 (31.8%)	5 (41.7%)	7 (31.9%)
36 or older	40 (51.3%)	22 (50.0%)	5 (41.7%)	13 (59.1%)
Maternal Education				
No formal qualifications	1 (1.3%)	0 (0%)	0 (0%)	1 (4.5%)
Level 2 (GCSE or equivalent)	2 (2.6%)	0 (0%)	2 (16.7%)	0 (0%)
Level 3 (A-Level or equivalent)	4 (5.1%)	2 (4.5%)	2 (16.7%)	0 (0%)
Degree or equivalent	30 (38.5%)	20 (45.5%)	2 (16.7%)	8 (36.4%)
Postgraduate Qualification	41 (52.6%)	22 (50.0%)	6 (50.0%)	13 (59.1%)
Paternal Education				
No formal qualifications	1 (1.3%)	0 (0%)	0 (0%)	1 (4.5%)
Level 2 (GCSE or equivalent)	8 (10.3%)	3 (6.8%)	2 (16.7%)	3 (13.6%)
Level 3 (A-Level or equivalent)	9 (11.5%)	3 (6.8%)	3 (25.0%)	3 (13.6%)
Degree or equivalent	29 (37.2%)	20 (45.5%)	4 (33.3%)	5 (22.7%)
Postgraduate Qualification	31 (39.7%)	18 (40.9%)	3 (25.0%)	10 (45.5%)
Family Income				
£14,000 or less	3 (3.9%)	1 (2.3%)	1 (8.3%)	1 (4.8%)
£14,001 to £24,000	4 (5.3%)	1 (2.3%)	2 (16.7%)	1(4.8%)
£24,001 to £42,000	15 (19.7%)	8 (18.6%)	3 (25.0%)	4 (19.0%)
£42,001 or more	54 (71.1%)	33 (76.7%)	6 (50.0%)	15 (71.4%)

Table 2: Descriptive information on DA tasks for low-risk infants

Measure	Range of Scores	Mean (SD)	Median (IQR)
Receptive Language	0-15	4.02 (4.54)	2.0 (1-7)
Motor Imitation	0-15	4.38 (4.11)	3.5 (1-6.5)
Response to Joint Attention	0-30	19.54 (9.44)	23.0 (11.75-27)
Turn Taking	0-15	2.86 (3.79)	1.0 (0-5.25)
Social Requesting	0-15	10.0 (4.59)	12.0 (76.75-14)

Table 3: Inter-rater reliability for each DA task for the whole sample and by group

Measure	Whole Sample ICC	Low-Risk ICC	SHR ICC	LHR ICC
Receptive Language	.902	.886	.966	.820
Motor Imitation	.869	.799	.918	.895
Response to Joint Attention	.958	.929	.977	.972
Turn Taking	.702	.815	.828	.386
Social Requesting	.852	.750	.971	.838

**ICC of 0.75-0.9 is considered 'good' whilst 0.9 and above is considered excellent.*

Table 4: Correlations between DA subtests for the low-risk group

	DA Receptive Language	DA Motor Imitation	DA Response to Joint attention	DA Turn Taking
DA Motor Imitation	r = .210 p = .182			
DA Response to Joint Attention	r = .240 p = .126	r = -.025 p = .875		
DA Turn Taking	r = .009 p = .954	r = -.108 p = .496	r = .140 p = .377	
DA Social Requesting	r = .115 p = .467	r = .053 p = .739	r = -.003 p = .986	r = -.090 p = .570

Table 5: Mann Whitney statistics for sex comparisons on DA tasks in the low-risk group

	Mann-Whitney *U* statistic	*P* value
Receptive Language	282.5	.449
Motor Imitation	231.5	.085
Response to Joint Attention	307.0	.776
Turn-Taking	250.5	.155
Social Requesting	292.5	.761

Table 6: Relationships between DA tasks and standardised tests for the low-risk group

	DA Receptive Language	DA Motor Imitation	DA Response to Joint Attention	DA Turn Taking	DA Social Requesting
UKCDI Receptive Vocabulary	***r* = .389** ***p* = .011**	*r* = .182 *p* = .249	*r* = .258 *p* = .100	*r* = .071 *p* = .653	*r* = .144 *p* = .362
UKCDI Expressive Vocabulary	*r* = .154 *p* = .331	*r* = .169 *p* = .285	*r* = -.080, *p* = .617	*r* = -.285, *p* = .067	***r* = .404** ***p* = .008**
UKCDI Total Gestures	*r* = .209 *p* = .183	*r* = -.164 *p* = .301	*r* = .154 *p* = .332	***r* = .315** ***p* = .042**	*r* = -.272 *p* = .082
ITC Social Subscale	*r* = .157 *p* = .321	*r* = .021 *p* = .895	*r* = .081 *p* = .610	***r* = .389** ***p* = .011**	***r* = -.307,** ***p* = .048**
ITC Speech Subscale	*r* = .223 *p* = .156	*r* = -.066 *p* = .677	*r* = -.083 *p* = .601	*r* = .083 *p* = .599	*r* = -.040 *p* = .802
ITC Symbolic Subscale	*r* = .102 *p* = .522	*r* = -.098 *p* = .536	*r* = .255, *p* = .102	***r* = .355** ***p* = .021**	*r* = -.139 *p* = .380
M-CHAT Core-6 Items Failed	*r* = -.074 *p* = .641	*r* = -.125 *p* = .431	*r* = .050 *p* = .754	*r* = -.257 *p* = .101	*r* = .272 *p* = .081
PLS-4 Auditory Comprehension Score	***r* = .477** ***p* = .001****	*r* = .184 *p* = .244	*r* = .221 *p* = .159	*r* = .252 *p* = .107	*r* = -.069 *p* = .665
PLS-4 Expressive Communication Score	***r* = .353** ***p* = .022**	*r* = .246 *p* = .117	*r* = -.056 *p* = .723	*r* = .039 *p* = .806	*r* = -.133 *p* = .400

Bold type indicates significant relations

***This relationship remains significant after adjusting for multiple comparisons using the Bonferroni method.*

Table 7: Adjusted mean scores and standard errors for each group on DA tasks

Group	Adjusted Mean Score	Standard Error	Significant differences
Receptive Language			
Low-risk	4.87	0.49	SHR<LHR=LR
SHR	1.62	0.86	
LHR	4.36	0.72	
Motor Imitation			
Low-risk	5.13	0.46	No differences across groups
SHR	3.65	0.80	
LHR	4.86	0.67	
Response to joint attention			
Low-risk	20.36	1.19	No differences across groups
SHR	17.94	2.07	
LHR	20.17	1.74	
Turn taking			
Low-risk	3.50	0.45	No differences across groups
SHR	2.72	0.78	
LHR	2.94	0.67	
Social requesting			
Low-risk	9.80	0.64	No differences across groups
SHR	10.42	1.10	
LHR	11.43	0.94	

Fig 1: DA task correlations with age for the low-risk group (separate file)

Appendix 1: Cueing and Scoring for the DA tasks

Receptive Language

The array of items was indicated and the child asked "Give me the…."	Correct response – score 3
If the child made no response, or selected the wrong item, the wrong item was returned to the array and the instruction repeated.	Correct response after repetition – score 2
If the child made no response, or selected the wrong item, three of the items were removed from the array, leaving a choice of two, and the instruction was repeated.	Correct response from choice of 2 – score 1
If the child still made no response or selected the wrong item, the correct item was indicated and the child told "Here is the….". The child was then praised.	Incorrect response after prompting – score 0

Motor Imitation

Modelling of action, preceded by "Look what I can do!" and followed by the command "X do it…"	Copying of action following model – score 3
If no response was made, the action and the command "X do it… was repeated"	Copying of action with extra verbal prompt – score 2
If the action was not copied, or the child performed another action, the action was repeated, exaggerating the action.	Copying of action after repetition – score 1
If the action was not copied following this, the child was praised and the next item presented.	No attempt to copy or another action produced – score 0

Joint Attention

The experimenter pointed at an item on the page and said “Look! A….”	Gaze shift to focus of pointing – score 3
If the child did not response to this, the instruction was repeated using the child’s name.	Gaze shift following repetition – score 2
If the child still did not shift their gaze, a light touch was used to gain their attention and direct them to the page, with the instruction “look!”	Gaze shift following physical prompt – score 1
If the child still did not attend to the focus of pointing, the next item was presented.	No gaze shift after cueing – score 0

Turn Taking

Experimenter handed ball to child following child’s turn	Child spontaneously stops and waits for experimenter to take a turn - score 3
If the child moved their ball towards the run without waiting for the experimenter to take a turn, the verbal prompt "My turn" was used	Child stops and waits for experimenter to take a turn following verbal prompt - score 2
If the child did not respond to the verbal prompt, a physical prompt of blocking the child's entry to the run was used, along with a repetition of "My turn"	Child stops and waits for experimenter to take a turn following a physical prompt - score 1
If the child persisted in taking their turn, the experimenter then allowed this, and began the procedure again for the next trial	Child does not stop and wait for experimenter's turn despite physical prompting - score 0

Social Requesting

The experimented turned off the toy and waited for a response.	Child spontaneously bids for the toy to be turned back on – score 3
If the child made no response, the verbal prompt “Would you like more?” was used.	Child bids for toy to be turned back on after verbal prompt – score 2
If the child still made no response, a light touch was used to call their attention to the toy and the verbal prompt was repeated.	Child bids for toy to be turned back on after verbal and physical prompt – score 1
If the child still made no response, the toy was turned back on, with the word “More” and the procedure was repeated.	No bid for help – score 0