



This is a repository copy of *The contribution of early language development to children's emotional and behavioural functioning at 6 years: An analysis of data from the Children in Focus sample from the ALSPAC birth cohort.*

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/95282/>

Version: Accepted Version

Article:

Clegg, J., Law, J., Rush, R. et al. (2 more authors) (2015) The contribution of early language development to children's emotional and behavioural functioning at 6 years: An analysis of data from the Children in Focus sample from the ALSPAC birth cohort. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 56 (1). pp. 67-75. ISSN 0021-9630

<https://doi.org/10.1111/jcpp.12281>

This is the peer reviewed version of the following article: Clegg, Judy, et al. "The contribution of early language development to children's emotional and behavioural functioning at 6 years: an analysis of data from the Children in Focus sample from the ALSPAC birth cohort." *Journal of Child Psychology and Psychiatry* 56.1 (2015): 67-75., which has been published in final form at <https://dx.doi.org/10.1111/jcpp.12281>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving (<http://olabout.wiley.com/WileyCDA/Section/id-820227.html>).

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

The contribution of early language development to children's emotional and behavioural functioning at 6 years: an analysis of data from the Children in Focus sample from the ALSPAC birth cohort

Short title: Children's early language development and later emotional and behavioural functioning at 6 years

J.Clegg PhD ¹ J. Law PhD² R. Rush MSc ³J. T.J. Peters PhD ⁴ and S. Roulstone PhD ⁵

¹ University of Sheffield, UK.

²Newcastle University, UK.

³ Queen Margaret University, Edinburgh, UK.

⁴ University of Bristol, UK.

⁵ University of the West of England, UK.

Short title: Children's early language development and later emotional and behavioural functioning at 6 years

Key words: children, language development, ALSPAC, emotional and behavioural functioning

Length: 6515 including all references and appendices

Address for correspondence:

Dr Judy Clegg

Human Communication Sciences

University of Sheffield

Sheffield S10 2TA

j.clegg@sheffield.ac.uk

Title: The contribution of early language development to children's emotional and behavioural functioning at 6 years: an analysis of data from the Children in Focus sample from the ALSPAC birth cohort

Abstract

Background

An association between children's early language development and their emotional and behavioural functioning is reported in the literature. The nature of the association remains unclear and it has not been established if such an association is found in a population-based cohort in addition to clinical populations.

Method

This study examines the reported association between language development and emotional and behavioural functioning in a population-based cohort. Data from 1314 children in the Children in Focus (CiF) sample from the Avon Longitudinal Study of Parents and Children (ALSPAC) were analysed. Regression models identified the extent to which early language ability at two years of age and later language ability at four years of age is associated with emotional and behavioral functioning at six years while accounting for biological and social risk and adjusting for age and performance intelligence (PIQ).

Results

A series of univariable and multivariable analyses identified a strong influence of biological risk, social risk and early and later language ability to emotional and

behavioural functioning. Interestingly, social risk dropped out of the multivariate analyses when age and PIQ were controlled for. Early expressive vocabulary at 2 years and receptive language at 4 years made a strong contribution to emotional and behavioural functioning at 6 years in addition to biological risk. The final model accounted for 11.6% of the variance in emotional and behavioural functioning at 6 years.

Conclusions

The study identified that early language ability at 2 years, specifically expressive vocabulary and later receptive language at 4 years both made a moderate, but important contribution to emotional and behavioural functioning at 6 years of age. Although children's language development is important in understanding children's emotional and behavioural functioning, the study shows that it is one of many developmental factors involved.

Key words: children, receptive and expressive language development, ALSPAC, emotional and behavioural functioning

Introduction

There has been a long standing interest in the association between language and behaviour but recently the nature of this relationship has come under renewed scrutiny (Hartas 2011; Lindsay, Dockrell and Strand 2007; Lindsay and Dockrell 2013; Van Daal, Verhoeven and van Balkom 2007). One key source has been clinical samples of children with developmental language impairment where an overlap is commonly reported with communication impairment often perceived as leading to subsequent difficulties in emotional and behavioural functioning. (Botting & Conti-Ramsden 2000; Brownlie et al., 2004; Johnson, Beitchman and Brownlie 2010; Van Daal et al., 2007). Another has been clinical samples of children with social, emotional and behavioural difficulties and again the overlap is noted with an assumption that poor communication skills contribute to the development or maintenance of the psychopathology (Cohen, et al., 1998; Clegg, Stackhouse, Finch, Murphy and Nicholls 2009; Giddan, Milling and Campbell 1996). Patterns of language impairment have also been correlated with behavioural profiles in children with social, emotional and behavioural difficulties (SEBD) (Clegg et al., 2009 Ripley and Yuill 2005; van Daal et al., 2007).

Yet, findings from these clinical studies are conflicting, with some finding a robust association (Cohen et al., 1998; Conti-Ramsden and Botting 2008; van Daal et al., 2007) and others less and more variable associations (Lindsay et al., 2007; Lindsay and Dockrell 2012). Reasons for this disparity include the small samples studied, the range of measures and differing definitions of language impairment and emotional and behavioural functioning employed, varying attention to confounding variables

and the age at which the assessments are carried out. Indeed, a recent meta-analysis confirmed that the evidence for specific language impairment as a specific risk factor for mental disorders was inconclusive (Goh Kok Yeh and O’Kearney 2013). Investigating the complex association between language and behaviour in a general population cohort may further our understanding of how this association operates over time in the general population as well as in clinical samples. Such a study will be able to account for the biological, social and developmental factors that are known to impact on emotional and behavioural functioning such as social risk (Davis, Sawyer, Lo and Wake 2010; Spratt et al., 2012), biological risk (Schoon, Sacker and Bartley 2003) and the interaction between these factors and the subsequent impact on development (Lindsay et al., 2007; Spratt et al., 2012).

While associations between language and behaviour are commonly reported, the temporal nature of the relationship has been less commonly discussed. Hartas (2011) investigated longitudinal associations between tested vocabulary development, gender and emotional and behavioural functioning in a community based sample from the Millennium Cohort Study in the UK. Moderate associations were found between 3 year and 5 year vocabulary and problem behaviour. Literacy was identified as a substantive predictor of teacher –reported behaviour difficulties at 5 years. Rescorla, Ross and McClure (2007) measured parent reported expressive vocabulary development and behaviour in children aged 18 months to 3 years who were attending child development clinics in the USA. Once neurodevelopmental delay and pervasive developmental disorders were excluded, no associations were found between language development and behaviour.

Similarly Whitehouse, Robinson and Zubrick (2012) found no association between parent report of early vocabulary development and later emotional and behaviour functioning in the Western Australian Pregnancy cohort (n = 1623 in total). At 2 years of age, caregivers completed the Language Development Survey (LDS) (Rescorla 1989), a parent report measure of expressive vocabulary and the Child Behaviour Checklist (CBCL) (Achenbach 1991), a parent report measure of child behaviour. Children with a score on the Language Development Survey (LDS) at or below the 15th percentile for their age and gender were identified as having an expressive language (namely vocabulary) delay and formed the late talkers subgroup. The CBCL was then repeated at ages 5, 8, 10, 14, and 17 years. At 2 years of age, the late talkers were more likely to have clinically significant internalising and externalising behaviour difficulties than the typical language group. However, at all the subsequent time points including the 5 year follow up, there was no difference between the late talkers and the typical language group on the CBCL suggesting that any behavioural difficulties ameliorated over time. In conclusion, the study stated that early childhood expressive vocabulary delay is not a specific risk factor for later emotional and behavioural difficulties in childhood and through adolescence.

Given the critical role played by comprehension as a risk factor for a range of negative sequelae (Beitchman et al., 2001) it is important that receptive language is included in any analyses and the relative role played by expressive and receptive language over time explored. Furthermore, it is important that relevant covariates associated with biological and social risk are included in subsequent analyses. As a number of studies have suggested there is a good case for including behaviour as a

primary outcome in the early school years as a primary indicator of “school readiness”.

Aims and research questions

The aim of this study was to explore the role played by expressive and receptive language (at two and four years) in predicting behaviour at six years of age using the Avon Longitudinal Study of Parents and Children (ALSPAC). The study asks the following research questions:

1. What is the contribution of childhood receptive and expressive language development to children’s emotional and behavioural functioning at 6 years of age relative to biological and social risk?
2. To what extent is the contribution sensitive to age within the preschool period?

Methods

ALSPAC and the Children in Focus sample

Participants

ALSPAC is a prospective population-based cohort study of all children born to mothers in an area of the west of England in the early 1990s, designed to explore the environmental and genetic factors that affect health and development. All mothers registering their pregnancy within the geographical county of Avon during

the period from 1991-1992 were invited to participate. The eligible sample consisted of 20, 248 pregnancies and the mothers of 14, 541 (71.8% pregnancies were recruited antenatally). Of these 14, 541 pregnancies, 14,062 resulted in live births of whom 13, 988 were alive at one year of age (see Boyd et al., 2013 for a detailed description). The sample was found to have some under-representation of less affluent families and fewer families from black and ethnic minority groups than is the case nationally, although the overall developmental trajectories of the children were similar to national norms for the period (Roulstone, Law, Rush, Clegg and Peters 2010). Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committee.

The Children in Focus (CiF) sample is a smaller randomly selected sample of the complete ALSPAC cohort. The purpose of the CiF sample was to collect direct observational and assessment data from the participants to both validate data collected via the parental questionnaires and reports collected in the full cohort, and also to collect direct assessment data of several important developmental abilities across cognition, speech and language, physical development and motor ability. The CiF sample was chosen at random from the last 6 months of ALSPAC births occurring in 1992. The following exclusion criteria were adhered to: 1) mothers who had moved away from Avon or were no longer contactable; 2) no consent to participate; 3) infant death and; 4) very preterm infants (born less than 33 weeks). In addition to the parent and other report measures completed by the full ALSPAC cohort, the children in the CiF sample were invited to attend for clinic examination at 4, 8 and 12 months, and then at 6 monthly intervals up to the age of 61 months. At each time point, between 994 and 1314 children attended the examination clinics.

All assessments (direct and parent report) were carried out by fully qualified and trained staff. The speech and language measures were conducted by qualified and experienced speech and language therapists (Roulstone et al., 2002).

Variables

Biological risk

The five biological risk variables were selected, as they were known to have an association with childhood development. Five biological variables were included: 1) **gender** defined as male or female; 2) **smoking** defined as mothers who smoked in the first three months of pregnancy; 3) **birthweight** defined as the weight of the infant at birth and used as a proxy measure of developmental health; 4) **parity** defined as the number of times the mother had given birth and; 5) **gestation** defined as the number of weeks at which the infant was born. These variables were all obtained from parental questionnaire data collected at birth and in the first year of life.

Social risk

This is a composite variable that was constructed from questionnaire variables identified at the end of the first year. This was based on an established procedure (Schoon, Parsons and Sacker 2004) designating binary variables as \pm to develop a broad measure of social risk comprising six variables including father's occupation (\pm skilled) and mother's education (\pm O level) – see appendix 1). A social risk score was then computed for each child with a range of 0 to 6, where 0 is a disadvantaged social background and 6 a more advantaged background (Roulstone et al., 2010).

Early Language variables (at 25 months)

The early language variables consist of a combination of parental questionnaire data and direct child assessment data collected at 25 months (2;01 years).

All Mothers in the ALSPAC cohort were asked a series of questions about their child's understanding and use of language, focusing on their vocabulary, their grammar and their ability to join words together within an utterance. The questionnaire was based on and modified from the Macarthur Communicative Development Inventory (CDI) (Fenson et al., 1993). The following variables comprised this early language questionnaire (see appendix 2 for further information).

Expressive vocabulary: number of words the child can say

Receptive vocabulary: number of words the child can understand

Expressive grammar: child's ability to join words together

RDLS Comprehension Scale: The Reynell Developmental Language Scales (RDLS) (Reynell 1977) is a standardised assessment used to assess receptive and expressive language. The RDLS comprehension scale measures a child's verbal comprehension by administering a series of activities where the child is asked to respond to and carry out a series of spoken tasks. The assessment was administered and scored according to the assessment manual.

Later Language variables (aged 49 months)

The later language variables relate to data collected from the CiF sample at 49 months (age 4 years, 1 month).

RDLS Comprehension Scale: The RDLS scale was then repeated at 49 months.

Expressive Language - The Bus Story Test (Renfrew 1997): is a standardised measure of expressive language. The assessment involves a child listening to a spoken narrative about a bus and is accompanied by pictures depicting the events that occur in the story. **The child then re-tells the story, and the child's narrative is scored for sentence length (mean sentence length of the five longest sentences), information content (number of relevant pieces of information given) and grammatical complexity (number of subordinate clauses). The first two but not the third were included in the analysis due to lack of complete available data.** Further information about the language assessments can be found in Roulstone et al., (2002).

Behaviour variable (aged 61 months)

The parent report version of the Strengths and Difficulties Questionnaire (SDQ) (Goodman 1997) was used at 61 months (5 years, 8 months). The SDQ provides a total score which is the sum of the scores for Emotional, Conduct, Hyperactivity and Peer problems subscales together with a score for the perceived impact of the difficulties experienced. There is also a score for the children's strengths – the Pro-social score – which also has a maximum score of 10 but works in reverse, with a high score indicating more pro-social behaviours. The total difficulties score of the SDQ was used in the analysis. The percentage of the CIF cohort with total behaviour scores in the abnormal range was 4.36%, 5.02% in the slightly raised/may reflect clinically significant problems range and 90.62% in the unlikely to be clinically significant range. This is comparable to other population based studies conducted at a similar time (Meltzer et al., 2000).

Age and Intelligence as Co-Variates

Two variables, age and performance intelligence (PIQ) were used in the analyses to adjust for the effects of age and intelligence. Age was taken when the children completed a school entry assessment administered by the Local Education Authority. The age of the children was 5 years, 3 months. The PIQ score was obtained for the CiF sample using the Wechsler Pre-School Scale of Intelligence Scale for Children - Revised (WPPSI-R) (Wechsler 1989). The CiF sample completed the WPPSI at age 4 years, 1 month (49 months).

Analyses

Descriptive analyses of the variables

Descriptive analyses of all the variables included in the univariable and multivariable regression modelling for the whole CiF cohort (Peters 2008) are shown in table 1.

Insert table 1 about here

Step 1

A series of univariable regression models were initially employed to identify those variables independently associated with the SDQ total score at 6 years, setting the context for the multivariable analysis. Since the outcome variable was continuous, ordinary linear regression was used. A threshold of $p < 0.1$ was used to identify those variables initially significantly associated with the SDQ score at 6 years. The threshold of $p < 0.1$ was chosen in order to further examine a range of relevant variables in the multivariable regression models. Correlations between the SDQ total

score, PIQ, social risk and the 2 year and 5 year language variables were completed using Spearman's rank order.

Step 2

Variables were then grouped into: 1) biological risk; 2) social risk; 3) early language ability (25months) and 4) later language ability (49 months). A series of multivariable analyses was then conducted with the SDQ total score as the single outcome variable. This step further identified those variables which remained independently associated with the SDQ total score ($p < 0.05$), within each grouping following identification in step 1, adjusting for other variables within the same group as well as for age and PIQ. Tests for collinearity were undertaken and all variables were within accepted limits (Tolerance < 0.2 , Variance Inflation Factor < 1 (Menard 1995)). Variables remaining significant at the $p < 0.05$ level within these multivariable analyses were carried forward to the final across-group multivariable analyses.

Step 3

This combined the earlier analysis using multivariable regression modelling across the four groups of variables and a significance threshold of 5% at each stage while adjusting for age and PIQ.

To optimise the data, any information that was missing for the ALSPAC cohort members in the CIF sample was imputed. The method of imputation employed was multiple imputation by chained equations (MICE) implemented in STATA (Royston 2005). The data set contained all variables used in the regression analyses. Five replicates of the data were created. Model estimates were averaged across these

five analyses, with their standard errors calculated according to Rubin's rule (Rubin 1987).

Results

Step 1: Univariable analysis

Table 2 shows that twelve variables were associated with the behaviour outcome variable at six years. These were: PIQ, four of the biological risk variables (gender, smoking in the first three months of pregnancy, birth weight and gestation); social risk; all four early language variables at 2 years (expressive vocabulary, receptive vocabulary, expressive grammar, and the RDLS comprehension scale) and two later language variable at 4 years (the RDLS comprehension scale and the Bus Story Information measure). Three variables were not significantly associated with the behaviour variable at six years. These were: age; one biological risk variable - parity; and the information expressive language measure at 4 years from the Bus Story Test. Significant correlations (shown in table 3) between the SDQ total score, PIQ, social risk and the 2 year and 5 year language variables (as found significant in step 1) were identified.

Insert table 2 about here

Step 2: Multivariable analysis

Model 1 Biological risk: Gender, smoking and birthweight were the three variables that were significantly associated with behaviour at 6 years (at $p < 0.05$) (models 1 and 2 in table 4). Together these accounted for 8.4% of the variance in the outcome

variable. These three variables went forward for consideration as biological risk variables in the final modelling stage.

Model 2 Social risk: The social risk variable did not go forward, once age and PIQ were controlled for, since following the univariable analysis it was not significantly associated with behaviour at 6 years ($p=0.08$ (model 3 in table 4)).

Model 3 Early language at 4 years: From the four early language variables, receptive vocabulary, expressive grammar and RDLS comprehension were not independently associated whereas expressive vocabulary accounted for 6.3% of the outcome variance (models 4 and 5 in table 4). The latter variable therefore went forward for consideration in the final modelling stage.

Model 4 Later language at 5 years: The RDLS comprehension scale was the only later language variable that was significantly associated with behaviour at 6 years, accounting for 8.5% of the outcome variance (models 6 and 7 in table 4), and it therefore went forward for the final (across group) modelling.

Insert table 4 about here

Step 3: Combining biological risk, early language and later language variables

The results of the across-group multiple regression analyses are shown in models 8 and 9 in table 4. In model 8, when expressive vocabulary at 2 years as the early language variable is added to biological risk, it no longer remains independently associated with the behaviour outcome. The addition of the RDLS comprehension scale at 4 years in model 9 increases the variance to 11.6%.

The final model (model 9 in table 4) includes biological risk, and the RDLS comprehension scale at 4 years as the later language variable. The final model accounts for 11.6% of the variance in the behaviour outcome variable. The influence of biological risk is moderate (Gender $p=0.004$; Smoking $p = 0.013$; Birthweight $p=0.021$). Later language in the form of RDLS comprehension scale at 4 years makes a stronger contribution to behaviour in addition to biological risk ($p=0.002$).

The results from the method of imputation did not differ from the multiple regression modelling.

Discussion

Returning to the two research questions it is clear that parent report of expressive language at two years and receptive (but not expressive) language at four years of age predict 6 year behaviour when taken alone. Biological risk factors (gender, smoking and birthweight) and social risk (at a univariable level) also have an important role to play. The findings suggest that this relationship is sensitive to age with expressive vocabulary at two years and receptive language at four years remaining in the final models once all the other factors are taken into consideration.

Interestingly but unexpectedly, social risk dropped out of the multivariate analyses once age and PIQ were adjusted for. Thus, suggesting that PIQ is a strong determiner of childhood emotional and behavioural functioning. Maternal characteristics such as knowledge of child development and maternal language/literacy skills are predictive of aspects of child development including IQ (Pan, Rowe, Snow & Singer 2005; Rowe 2008). The role of PIQ over and above

social risk in children's emotional and behavioural functioning is a potential area for further research.

Age differences in the contribution of receptive and expressive language development to emotional and behavioural functioning were identified. At age 2 years, only expressive vocabulary made a significant contribution but at 4 years expressive language no longer made a significant contribution and was replaced by receptive language. Here, later receptive language may be accounting for the variability in early vocabulary and also has a stronger role than early vocabulary as it is nearer the 6 year age when emotional and behavioural functioning was measured. Changes in language measures between ages 2 and 4 years may also account for the age differences. The same standardised receptive language measure (the Reynell Developmental Language Scales (Reynell 1977)) was used at both 2 years and 4 years. However, at 4 years, expressive language was measured using the Bus Story Test (Renfrew 1997) which is a narrative based standardised assessment as opposed to the parent report measure of vocabulary and expressive grammar used at 2 years.

These findings differ to those of Rescorla et al., (2007) and Whitehouse et al., (2012) with respect to the predictive role of early vocabulary. Reasons to explain this difference primarily include the different samples studied, the use of a different 2 year parent vocabulary report measure to the Language Development Survey used by Rescorla et al., and Whitehouse et al., and the inclusion of receptive language measures in addition to expressive language alone.

ALSPAC is a highly respected population-based cohort study (Sonuga-Barke 2012). In our analysis the consistency of the use of the standardised receptive language measure at 2 years and 4 years in addition to the direct language assessments in conjunction with the parent report measures at both 2 and 4 years add to the robustness of the study. Yet, the switch in measures between expressive vocabulary and expressive narrative while reflecting the increasing sophistication in the child's language may restrict the interpretation of the findings.

Focusing on the CiF sample rather than the whole ALSPAC cohort enabled the analysis of the combination of direct language assessments with parent report measures. The CiF sample consisted of between 994 and 1314 children, and in the present study the number of participants in the multivariable analyses ranged from 488 to 522. It is recognised that this size of sample may not be large enough to be representative of the ALSPAC cohort although random selection of participants and other measures to increase representativeness were addressed. As expected, the ALSPAC sample has suffered from attrition, which is common across longitudinal cohort studies. The multiple imputation analyses replicated the findings from the multivariable regression analyses thus limiting the effects of potential bias and attrition. The limitations of the language measures used in the analysis were referred to earlier in the discussion. A further point of consideration is the use of the parent report version of the SDQ to measure emotional and behaviour functioning. The SDQ is a screen completed by parents and/or carers and so does not provide detailed or in-depth profiles of children's emotional and behavioural status. An extension of the analyses could have investigated the contribution of early and later

expressive and receptive language to not only the total SDQ score but also the subscale scores. However, the rationale for such an analysis in this general population cohort was difficult to justify given that language only made a modest albeit important contribution to emotional and behavioural functioning.

This study investigated the contribution of early and later expressive and receptive language development to children's emotional and behaviour functioning at 6 years in a general population cohort. The study identified that early expressive language and later receptive language make a moderate but important contribution to emotional and behavioural functioning at 6 years of age with later receptive language making the stronger contribution. This population-based study shows that children's early language development is important in understanding children's emotional and behavioural functioning but is one of many developmental factors to consider in the association between language development and emotional and behavioural functioning.

Future research should test for other developmental mechanisms. An alternative model where early language ability mediates the relationship between early social risk and later emotional and behavioural functioning is worthy of investigation. Furthermore, defining categories of children with language impairment and emotional and behaviour difficulties in general population data will further understanding of the potential comorbidity between these diagnoses and the developmental trajectories and outcomes of these diagnostic groups. Recent population studies show that during the pre-school period, the developmental trajectories of young children with potential language impairment are fluid and these children move in and out of

categories defining typical language and language impairment (Law et al., 2012; Ukoumunne et al., 2012). Therefore, future research may need to develop more robust procedures for defining and categorising such impairments in general population data.

In conclusion, early vocabulary development is clearly an important marker in children's pre-school development and this study shows that it may be predictive of later emotional and behavioural ability. Identifying poor vocabulary development is warranted particularly when the risk for subsequent emotional and behavioural difficulties increases when later receptive language is implicated.

References

Achenbach, T.M. (1991). *Manual for the Child Behaviour Checklist/4-18 and 1991 Profile*. Burlington, VT: Department of Psychiatry, University of Vermont USA.

Botting, N., & Conti-Ramsden, G. (2000). Social and behavioural difficulties in children with language impairment. *Child Language Teaching & Therapy*, 16, 105-120.

Boyd, A., Golding, J., Macleod, J., Lawlor, D.A., Fraser, A., Henderson, J., Molloy, L., Ness, A., Ring, S., & Davey Smith, G. (2013). Cohort profile: the 'children of the 90s' – the index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology*, 42, 111-127.

Brownlie, E.B., Beitchman, J.H., Escobar, M., Young, A., Atkinson, L., Johnson, C., Wilson, B., & Douglas, L. (2004). Early language impairment and young adult delinquent and aggressive behaviour. *Journal of Abnormal Child Psychology*, 32, 4, 453-467.

Clegg, J., Stackhouse, J., Finch, K., Murphy, C., & Nicholls, S. (2009). Language abilities of secondary age pupils at risk of school exclusion: a preliminary report. *Child Language Teaching & Therapy*, 25, 123-140.

Cohen, N.J., Menna, R., Vallance, D.D., Barwick, M.A., Im, N., & Horodezky, N.B. (1998) Language, social cognitive processing and behavioural characteristics of psychiatrically disturbed children with previously identified and unsuspected language impairments. *Journal of Child Psychology and Psychiatry*, 39, 853-864.

Conti-Ramsden, G., & Botting, N. (2008). Emotional health in adolescents with and without a history of Specific Language Impairment (SLI). *Journal of Child Psychology & Psychiatry*, 49, 516-525.

Davis, E., Sawyer, M.G., Lo, S.K., & Wake, M. (2010) Socioeconomic risk factors for mental health problems in 4-5 year-old children: Australian population study *Academic Pediatrics* 10, 41-47.

Fenson, L., Dale, P.S., Reznick, J.S., Thal, D., Bates, E., Hartung, J.P., Pethick, S., & Reilly, J.S. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual*. San Diego, CA; Singular Publishing Group, USA.

Giddan, J.J., Milling, L., & Campbell, M.D. (1996). Unrecognised language and speech deficits in preadolescent psychiatric inpatients. *American Journal of Orthopsychiatry*, 66, 85-92.

Goh Kok Yew, S., & O'Kearney, R. (2013). Emotional and behavioural outcomes later in childhood and adolescence for children with specific language impairments: meta-analyses of controlled prospective studies. *Journal of Child Psychology & Psychiatry*, 54, 516-524.

Goodman R. (1997) The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, 38, 581-586.

Hartas, D. (2011). Children's language and behavioural, social and emotional difficulties and prosocial behaviour during the toddler years and at school entry. *British Journal of Special Education*, 38, 83-91.

Johnson, C.J., Beitchman, J.H., & Brownlie, E.B. (2010). Twenty-year follow up of children with and without speech-language impairments: family, educational, occupational, and quality of life outcomes. *American Journal of Speech-Language Pathology*, 19, 51-65.

Law, J., Rush, R., Anandan, C., Cox, M., & Wood, R. (2012). Predicting language change between 3 and 5 years and its implications for early identification. *Paediatrics*, 130, 1, 132-137.

Lindsay, G., & Dockrell, J. (2013). The relationship between speech, language and communication needs (SLCN) and behavioural, emotional and social difficulties (BESD). Research Report DFE-RR247-BCRP6

Lindsay, G., & Dockrell, J. (2012). Longitudinal patterns of behavioural, emotional and social difficulties and self concept in adolescents with a history of specific language impairment. *Language, Speech and Hearing Services in Schools*, early online publication DOI 10.1044/0161-1461 (2012/11-0069).

Lindsay, G., Dockrell, J., & Strand, S. (2007). Longitudinal patterns of behaviour problems in children with specific speech and language difficulties: child and contextual factors. *British Journal of Educational Psychology*, 77, 811-828.

Menard, A. (1995). Applied logistic regression analysis. Sage University paper series on quantitative applications in the social sciences, 07-106. Thousand Oaks, CA: Sage, USA.

Meltzer, H., Gatwood, R., Goodman, R., & Ford, F. (2000). Mental health of children and adolescents in Great Britain. London: The Stationery Office.

Peters, T.J. (2008) Multifarious terminology: multivariable or multivariate? Univariable or univariate? *Paediatric and Perinatal Epidemiology*, 22, 506.

Renfrew, C. (1997). *Bus Story Test*.

Rescorla, L., Ross, G.S., & McClure, S. (2007). Language delay and behavioural/emotional problems in toddlers: findings from two developmental clinics. *Journal of Speech, Language and Hearing Research*, 50, 1063-1078.

Rescorla, L. (1989). The Language Development Survey: a screening tool for delayed language in toddlers. *Journal of Speech & Hearing Disorders*, 54, 4, 587-599.

Reynell, J. (1977). *Reynell Developmental Language Scales (Revised)*. Windsor: NFER Publishing Company, UK.

Ripley, K., & Yuill, N. (2005). Patterns of language impairment and behaviour in boys excluded from school. *British Journal of Educational Psychology*, 75, 37-50.

Roulstone, S., Law, J., Rush, R., Clegg, J., & Peters, T. (2010). The role of language in children's social and educational outcomes. Department for Education (DfE), UK.

Roulstone, S., Loader, S., Northstone, K., & Beveridge, M. and the ALSPAC team, (2002). The speech and language of children aged 25 months: descriptive data from

the Avon Longitudinal Study of Parents and Children. *Early Child Development and Care*, 172, 259-268.

Royston, P. (2005). Multiple imputation of missing values: update of ICE. *Stata Journal*, 5, 4, 527-536.

Rubin, D.B. (1987). Multiple imputation for nonresponse in surveys. Wiley, New York, USA.

Schoon, I., Parsons, S., & Sacker, A. (2004). Socioeconomic adversity, educational resilience and subsequent levels of adult adaptation. *Journal of Adolescent Research*, 19, 383-404.

Schoon, I., Sacker, A., & Bartley, M (2003) Socio-economic adversity and psychosocial adjustment: a developmental-contextual perspective *Social Science & Medicine*, 57, 1001–1015.

Sonuga-Barke, E.J.S. (2012). Editorial: The developmental psychopathologist as scientist-sleuth – can large-scale longitudinal birth cohort studies provide the missing clues? *Journal of Child Psychology & Psychiatry*, 53, 619-621.

Spratt, E.G., Friendenberg, S., LaRosa, A., De Bellis, M.D., Macias, M.M., Summer, A.P., Hulsey, T.C., Runyan, D.K., & Brady, K.T. (2012). The effects of early neglect on cognitive, language and behavioural functioning in childhood. *Psychology*, 3, 175-182.

Ukoununne, O.C., Wake, H., Carlin, J., Bavin, E.I., Lum, J., Skeat, J., Williams, J., Conway, L., Cini, E., & Reilly, S. (2012). Profiles of language development in pre-school children: longitudinal latent class analysis of data from the early language in Victoria study. *Child: care, Health and Development*, 38, 3, 341-349.

Van Daal, J., Verhoeven, L., & van Balkom, H. (2007). Behaviour problems in children with language impairment. *Journal of Child Psychology & Psychiatry*, 48, 1139-1147.

Wechsler, D. (1989). Wechsler Preschool and Primary Scale of Intelligence-Revised. San Antonio, TX: The Psychological Corporation.

Whitehouse, A.J., Robinson, M., & Zubrick, S.R. (2012). Late talking and the risk for psychosocial problems during childhood and adolescence. *Paediatrics*, 128

ACKNOWLEDGEMENTS

The UK Medical Research Council, the Wellcome Trust and the University of Bristol provide core support for ALSPAC. The research presented here was initially support through funding from the Department for Education in England DFE-134

All authors declare that they have no financial interest and no potential conflict of interest.

We are grateful to all the families who have taken part in the ALSPAC study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. We would like to thank the ALSPAC speech team who collected the speech and language data and to Laura Miller who has acted as our “data-buddy” throughout this project, providing the dataset promptly and answering our various queries about the variables and coding.

Table 1 Descriptive Statistics for the Included Variables

	N	Minimum	Maximum	Mean	Standard Deviation
Age	820	49	61	55.15	4.53
SDQ total score at 6 years	757	0	31	7.31	4.68
Performance IQ	917	55	151	108.18	14.40
Biological Risk					
Gender Boys	592				
Girls	493				
Smoking Yes	181				
No	885				
Birthweight	1075	1.35	5.0	3.47	0.50
Parity	1056	0	11	.80	1.00
Gestation	1085	33.00	44.0	39.53	1.56
Social risk					
Social risk	1048	0	6	4.54	1.30
Early Language					
Expressive vocabulary	1021	0	123	61.38	34.50
Receptive vocabulary	1021	0	111	27.23	25.03
Expressive grammar	998				
-Not yet	178				
-Sometimes	299				
-Often	521				

RDLS comprehension scale	1085	-2.85	3.20	0.00	1.00
at 2 years (z scores)					
Later Language					
RDLS comprehension scale	875	-11.93	1.36	0.02	0.98
at 5 years (z scores)					
Bus Story Information	728	1	52	27.41	11.15
Bus Story Sentence length	682	3	20	9.47	2.57

Table 2 Univariable Regression Analyses with Behaviour (SDQ total score) at 6 Years as the Outcome Variable

	N	Regression coefficient	P	95% CI Lower	95% CI Upper
Age	571	-0.051	.237	-0.137	0.035
PIQ	684	-0.079	<0.000	-0.102	-0.055
Biological Risk					
Gender (girls vs boys)	757	-1.247	<0.001	-1.914	-0.581
Smoking	752	1.603	<0.001	0.696	2.51
Birthweight	751	-1.027	<0.001	-1.702	-0.352
Parity	745	0.052	.750	-0.281	0.384
Gestation	757	-0.275	.036	-0.496	-0.055
Social Risk					
Social risk	745	-0.551	<0.001	-0.813	-0.289
Early Language					
Expressive vocabulary	746	-0.026	<0.001	-0.035	-0.016
Receptive vocabulary	746	0.019	.005	0.006	0.032
Expressive grammar	731				
-Sometimes cf Not yet		-0.448	.752	-1.45	0.554
-Often cf Not yet		-1.981	<0.001	-2.891	-1.07
RDLS comprehension scale at 2 years	757	-0.957	<0.001	-1.298	-0.616
Later Language					
RDLS comprehension scale at 5 years	661	-1.172	<0.001	-1.564	-0.781

Bus Story information	558	-0.031	.042	-0.065	0.002
Bus Story sentence	516	-0.109	.165	-0.266	0.048
length					

Table 3 Correlations Between the Variables in the Multiple Regression Analysis

	SDQ	PIQ	Social risk	2 year expressive vocabulary	2 year receptive vocabulary	2 year RDLS comprehension	5 year RDLS comprehension
Performance IQ	-.205**						
Social risk	-.133**	.305**					
2 year expressive vocabulary	-.168**	.254**	.080*				
2 year receptive vocabulary	.103**	.108**	.019	.734**			
2 year RDLS comprehension	-.185**	.382**	.240**	.491**	.246**		
5 year RDLS comprehension	-.177**	.364**	.252**	.298**	.149**	.450**	
5 year Bus Story Information	-.103*	.284**	.232**	.154**	.056**	.256**	.329**

* <0.05 ; ** <0.01 ; *** <0.001

Table 4 Multivariable Regression Coefficients (95%) for Biological Risk, Social risk, Early Language and Later Language with Behaviour as the Dependent Variable

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
N	521	521	522	510	519	417	494	513	488
Child's sex	-1.106 P=0.006 (-1.894,-0.317)	-1.117 P=0.05 (-1.904,-0.331)						-1.048 P=0.01 (-1.847,-0.248)	-1.203 P=0.004 (-2.014,-0.391)
Smoking	1.224 P=0.026 (0.15,2.297)	1.206 P=0.027 (0.136,2.227)						1.137 P=0.038 (0.061,2.213)	1.421 P=0.013 (0.302,2.539)
Birth weight	-0.752 P=0.097 (-1.639,0.136)	-0.853 P=0.034 (-1.642,-0.064)						-0.83 P=0.033 (-1.615,-0.045)	-0.867 P=0.021 (-1.674,-0.061)
Gestation	-0.071 P=0.624 (-0.356,0.214)								
Social risk			-0.292 P=0.008 (-0.618,0.035)						
2 yr Expressive				-0.008 P=0.457	-0.014 P=0.024			-0.011 P=0.084	

Vocabulary									
2 yr									
Receptive									
Vocabulary									
2 yr									
Expressive									
Grammar									
Sometimes									
of Not yet									
Often of Not									
yet									
2 yr RDLS									
Comprehen									
sion									
5 yr RDLS									
Comprehen									
sion									
5 yr									
Bus Story-									
Information									
R²	0.084	0.084	0.06	0.075	0.063	0.067	0.085	0.088	0.116

All models in the table are adjusted for age and PIQ

Appendix 1

Social Risk Variable (Schoon, Parsons & Sacker 2004)

Schoon's category	Variable	Binary coding
Parental occupation	ALSPAC has both father's and mother's social class based on OPCS employment codes. In this study we have used paternal occupation.	Unskilled, partly skilled or manual occupation: 0 Skilled occupation: 1
Father education	Within ALSPAC there is considerable missing data for the partner's education; this variable was therefore excluded from the composite.	
Mother's education	Mother's education	'O' level or below (including vocational): 0 Better than 'O' level: 1
House tenure	House tenure	Rented or other housing: 0 Owner occupied: 1
Overcrowding	ALSPAC has a 'crowding index', formed by the number of people in a house, divided by the number of rooms.	More than one person per room: 0 Less than one person per room: 1
Sole use of	Not available in ALSPAC	

household amenities

Receipt of state benefits	Financial difficulties A set of questions are asked regarding the mother's ability to afford certain basic items for the baby, specifically food, clothing, heating, rent or mortgage and things she will need for the baby producing a continuous score.	Financial difficulties: 0 None or minimal financial difficulties: 1
Car ownership	Use of a car	No: 0 Yes: 1

Appendix 2

Expressive Vocabulary at 24 months item (expressive vocabulary variable) (taken from the CDI)

From a list of 123 early vocabulary words, Mothers were asked to report if their child could say the word. These words included the categories of body parts, actions, clothes, home, animal, people, places, adjectives and prepositions.

Receptive vocabulary at 24 months item (receptive vocabulary variable) (taken from the CDI)

From the same list of words in the above item, Mothers were asked to report if their child could understand the word.

Expressive grammar at 24 months (expressive grammar variable) (taken from the CDI)

Mothers were asked 'Has your child begun to combine words yet, such as 'another sweet' or 'doggie bite' Response: Not Yet, Sometimes, Often

Key Points

- An association between children's language development and their emotional and behavioural functioning is documented. However it remains unclear as to whether this association is found in the general population as well as clinical samples.
- This study identifies the contribution early and later expressive and receptive language development makes to emotional and behavioural functioning at 6 years in a prospective population based cohort.
- Early and later expressive and receptive language development makes a moderate contribution to children's emotional and behavioural functioning in the early school years while accounting for other known developmental factors.
- Children's language development is important in understanding their emotional and behavioural functioning, specifically expressive vocabulary at the age of 2 years and receptive language at school entry age.
- Identifying language difficulties in the pre-school period is important when considering children at risk of emotional and behavioural difficulties.