1

Supporting Information for Manuscript: 'Risk factors for early language delay in children within a minority ethnic, bilingual, deprived environment (Born in Bradford's Better Start): A UK community birth cohort study'

Bilingual classification
The UKBTAT (Floccia et al., 2018, doi: 10.1111/mono.12349) typically measures the amount of direct exposure to English and another language between child and parent, as well as the amount of overheard speech in English and another language. This is measured using eight questions (includes number of hours in nursery, hours exposed to relatives, hours exposed to each parent, hours of sleep per night, language spoken by each parent to the child, who speaks most to the child, followed by a six-step calculation process outlined in Floccia et al., 2018, Appendix 3). As we did not have the in-depth information required to calculate exposure using the UKBTAT due to the use of a community cohort sample, we used an *a priori* standard of direct exposure and overheard speech for English and another language for all bilingual families, assuming that bilingual children were directly spoken to equally in English and in a second language, and that they overheard their parents speaking English and a second language an equal amount.

Data linkage and sample
The paper uses data from the BiBBS pre-COVID-19 profile (Dickerson et al., 2022, doi: 10.12688/wellcomeopenres.18394.1). Note that because cohort recruitment and data collection is ongoing, the sample within the paper is a data freeze of the data available for linkage in the BiBBS cohort from all pregnancies between 1st Jan 2016 and 30th November 2019, with an estimated due date between 1st April 2016 and 8th March 2020. In order to utilise all data sources, including project data that contained the UK-CDI language measure, project data was linked using the procedures outlined in Dickerson et al., 2022. Full information on this process can be found in the Supplementary Materials for Dickerson et al., 2022 at:
https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/ZQIUNC.

Table S1. Missing data by variable used in analyses (total N = 712)

Variable	N missing cases	Percent missing (out of 712)
Ethnicity	7	0.98
Born in UK status	6	0.84
Education level	45	6.32
Employment status	27	3.79
Financial management	42	5.90
Maternal age	3	0.42
Total people in household	7	0.98
Maternal native language	28	3.93
Languages spoken at home	21	2.95
Birthweight status	1	0.14
Preterm status	6	0.84
Child sex	1	0.14
Parent-reported hearing concerns	2	0.28

Derived maternal variables

2

<u>Maternal age:</u> The original BiBBS cohort dataset lists the age at which mothers took the baseline questionnaire. This information was used and combined with the age of the child at the Talking Together screening to provide maternal age at Talking Together screening.

<u>Maternal education:</u> The original BiBBS cohort dataset has 7 sub-categories for education level as follows:1) No qualifications, 2) 5 or less GCSE (grades A-C) or equivalent, 3) 5 or more GCSE (grades A-C) or equivalent, 4) A levels or equivalent, 5) Degree or equivalent, 6) Don't know, and 7) Other.

For the purpose of analyses, the above were collapsed into five sub-categories: 1) No qualifications, 2) GCSEs (collapsing category 2 and 3 above), 3) A Levels or equivalent, 4) Degree or equivalent, 5) Missing, don't know, or other (collapsing category 6 and 7 above with missing (NA) entries).

<u>Parental employment:</u> The original BiBBS cohort dataset provides maternal and paternal employment status. For the purpose of analyses, this data was combined to provide a variable that detailed any parental employment. Where at least one parent was employed, then parental employment was coded as 1. Where no parents were employed, then parental employment was coded as 0.

Sensitivity analysis with maternal mental health variables

A sensitivity analysis was carried out, testing the value of maternal mental health variables (Patient Healthcare Questionnaire-8, PHQ-8 [Kronke et al., 2009, doi: 10.1016/j.jad.2008.06.026] and Generalised Anxiety Disorder-7, GAD-7 [Spitze et al., 2006, 10.1001/archinte.166.10.1092]) in the final model when predicting expressive vocabulary at age 2 in the population. These variables did not significantly predict expressive vocabulary. The full results can be viewed in Table S2 below.

Table S2. Sensitivity analysis results: predictors of expressive vocabulary with maternal mental health variables (*n* = 382)

	Model 1	Model 2	Model 3	Model 4	Model 5
	Estimate (*p*-value)	Estimate (*p*-value)	Estimate (*p*-value)	Estimate (*p*-value)	Estimate (*p*-value)
Maternal demographic factors					
Ethnicity (relative to White British)					
Pakistani heritage	-9.90 (.052)	-10.00 (.083)	-8.65 (.138)	-10.93 (.054)	-2.29 (.614)
Central/Eastern European	3.4960 (.792)	4.02 (.767)	5.68 (.675)	6.57 (.618)	14.60 (.164)
Other	-4.78 (.419)	-4.60 (.469)	-3.19 (.618)	-5.31 (.396)	-0.59 (.905)
Born in UK					
No	1.05 (.751)	3.77 (.387)	3.81 (.382)	3.13 (.462)	0.75 (.825)
Education level (relative to Degree)					
A Levels	-5.46 (.270)	-4.93 (.321)	-5.63 (.258)	-7.05 (.146)	-4.03 (.296)
GCSEs	-1.77 (.593)	-0.94 (.785)	-0.86 (.803)	-2.13 (.526)	-1.42 (.594)

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Paediatrics Open

3

None	-3.62 (.588)	-1.90 (.782)	-1.68 (.764)	-3.83 (.565)	-3.56 (501)
Financial security (relative to Comfortable)					
Alright	1.08 (.740)	1.37 (.676)	0.99 (.764)	-0.61 (.850)	-0.12 (.964)
Just getting by	-2.91 (.535)	-3.08 (.512)	-3.25 (.490)	-4.18 (.360)	-3.07 (.397)
Quite difficult	2.6924 (.749)	3.24 (.705)	1.67 (.846)	-0.64 (.939)	-2.50 (.707)
Very difficult	-1.24 (.941)	2.11 (.902)	1.73 (.919)	-0.50 (.976)	9.60 (.470)
Parental employment					
Neither parent employed	-1.77 (.804)	-1.68 (.815)	-0.86 (.905)	0.68 (.923)	-1.39 (.878)
Size of household (number of individuals)	1.00 (.065)	0.916 (.095)	0.85 (.123)	0.83 (.122)	0.65 (.127)
Maternal age at screening	-0.17 (.572)	-0.21 (.490)	-0.19 (.517)	-0.08 (.772)	0.04 (.538)
Maternal mental health					
GAD-7 Score	-0.88 (.062)	-0.85 (.073)	-0.75 (.115)	-0.71 (.125)	-0.23 (.526)
PHQ-8 Score	0.42 (.367)	0.35 (.447)	0.26 (.583)	0.11 (.806)	-0.08 (.814)
Maternal language factors					
First language (relative to English)					
Bilingual-English		-2.29 (.583)	-2.67 (.524)	-1.24 (.761)	-2.00 (.538)
Not English		-3.75 (.499)	-4.44 (.424)	-3.09 (.568)	-2.39 (.578)
Languages at home (relative to English)					
Bilingual-English		3.84 (.348)	4.62 (.262)	2.81 (.485)	0.63 (842)
Not English		-1.60 (.756)	-0.89 (864)	-1.42 (.777)	-1.45 (.717)
Perinatal factors					
Birthweight (kg)			3.48 (.356)	5.73 (.121)	4.29 (.144)
Gestation at birth (weeks)			0.64 (.578)	0.18 (.874)	-0.19 (.830)
Distal child factors					
Child Sex					
Male*				-12.13 (<.001)	-7.89 (.001)
Child age (months)*				3.46 (.002)	2.38 (.008)
Proximal child factors					

4

Receptive vocabulary					0.73 (<.001)
Parent-reported hearing concerns					
Yes					-18.10 (.014)
R²	3.66%	4.38%	5.13%	11.34%	44.37%
Δ R²		+0.72%	+0.75%	+6.21%	+33.03%
F for change in R² **(*p*-value)**		0.68 (.605)	1.42 (.244)	12.51 (<.001)	105.36 (<.001)

Receptive vocabulary

An exploratory analysis of receptive vocabulary on the Oxford-CDI: Short as a continuous outcome variable was conducted using all predictors on a multiple imputation dataset as per the main analysis. Results can be viewed in Table S3 below.

Table S3. Predictors of receptive vocabulary as a continuous variable (*N* = 712).

	Model 1	*Model 2*	*Model 3*	*Model 4*	*Model 5*
	Est (p-value)	*Est (p-value)*	*Est (p-value)*	*Est (p-value)*	*Est (p-value)*
Ethnicity (relative to White British)					
*Pakistani heritage**	-10.72 (.001)	-10.92 (.002)	-10.24 (.003)	-10.88 (.002)	-10.97 (.001)
Central/Eastern European	-9.57 (.125)	-9.53 (.137)	-8.43 (.190)	-8.77 (.168)	-8.38 (.188)
Other	-6.33 (.069)	-6.35 (.089)	-5.66 (.130)	-6.24 (.092)	-6.26 (.090)
Born in UK					
No	4.65 (.014)	4.56 (.065)	4.17 (.092)	3.87 (.114)	3.57 (.146)
Education level (relative to Degree)					
A Levels	1.62 (.561)	1.73 (.536)	1.19 (.671)	0.62 (.824)	0.60 (.828)
GCSEs	0.17 (.932)	0.06 (.977)	0.02 (.993)	-0.17 (.931)	-0.23 (.907)
None	-0.01 (.998)	0.15 (.965)	0.21 (.952)	-0.31 (.926)	-0.33 (.922)
Financial security (relative to Comfortable)					
Alright	-0.85 (.655)	-0.81 (.672)	-0.91 (.633)	-1.52 (.422)	-1.60 (.396)
Just getting by	-0.96 (.710)	-1.11 (.669)	-1.46 (.576)	-1.67 (.517)	-1.73 (.502)
Quite difficult	-1.47 (.729)	-1.39 (.747)	-1.60 (.709)	-1.96 (.645)	-1.87 (.661)
Very difficult	-8.43 (.323)	-8.48 (.323)	-9.21 (.283)	-8.80 (.299)	-7.71 (.365)

Supplemental material   BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)   *BMJ Paediatrics Open*

5

Parental employment					
Neither parent employed	3.59 (.249)	3.44 (.271)	3.75 (.230)	4.02 (.193)	4.12 (.181)
Size of household (number of individuals)	0.47 (.164)	0.45 (.192)	0.43 (.210)	0.43 (.206)	0.43 (.203)
Maternal age at screening	-0.17 (.286)	-0.21 (.203)	-0.16 (.332)	-0.14 (.395)	-0.14 (.391)
First language (relative to English)					
Bilingual-English		0.72 (.773)	0.45 (.857)	1.14 (.645)	1.24 (.617)
Not English		2.79 (.371)	2.33 (.454)	2.80 (.364)	3.02 (.327)
Languages at home (relative to English)					
Bilingual-English		-0.12 (.960)	0.20 (.933)	-0.55 (.809)	-0.59 (.796)
Not English		-3.39 (.221)	-2.93 (.293)	-3.21 (.245)	-3.27 (.235)
Birthweight (kg)			-0.43 (.829)	0.54 (.784)	0.45 (.818)
Gestation at birth (weeks)			1.00 (.084)	0.84 (.145)	0.82 (.153)
Child Sex					
Male*				-6.21 (<.001)	-6.24 (<.001)
Child age (months)*				1.69 (.009)	1.69 (.008)
Parent-reported hearing concerns Yes					-9.16 (.120)
R^2	3.42%	3.82%	4.43%	7.05%	7.38%
Δ R^2	NA	+0.40%	+0.61%	+2.62%	+0.33%
F for change in R^2 (*p*-value)	1.68 (.052)	0.68 (.606)	2.23 (.107)	9.94 (<.001)	2.48 (.115)

*p <.05