

Early language competence, but not general cognitive ability, predicts children's recognition of emotion from facial and vocal cues (#45198)

1

First submission

Guidance from your Editor

Please submit by **1 Mar 2020** for the benefit of the authors (and your \$200 publishing discount) .



Structure and Criteria

Please read the 'Structure and Criteria' page for general guidance.



Custom checks

Make sure you include the custom checks shown below, in your review.



Raw data check

Review the raw data.



Image check

Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files

Download and review all files from the [materials page](#).

3 Figure file(s)

1 Table file(s)

3 Raw data file(s)

! Custom checks

Human participant/human tissue checks

- ! Have you checked the authors [ethical approval statement](#)?
- ! Does the study meet our [article requirements](#)?
- ! Has identifiable info been removed from all files?
- ! Were the experiments necessary and ethical?



Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. BASIC REPORTING
2. EXPERIMENTAL DESIGN
3. VALIDITY OF THE FINDINGS
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready [submit online](#).

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your [guidance page](#).

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to [Peerj standards](#), discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see [Peerj policy](#)).

EXPERIMENTAL DESIGN

- Original primary research within [Scope of the journal](#).
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Speculation is welcome, but should be identified as such.
- Conclusions are well stated, linked to original research question & limited to supporting results.



The best reviewers use these techniques

Tip

Example

Support criticisms with evidence from the text or from other sources

Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.

Give specific suggestions on how to improve the manuscript

Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).

Comment on language and grammar issues

The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 - the current phrasing makes comprehension difficult.

Organize by importance of the issues, and number your points

1. Your most important issue
2. The next most important item
3. ...
4. The least important points

Please provide constructive criticism, and avoid personal opinions

I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC

Comment on strengths (as well as weaknesses) of the manuscript

I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.

Early language competence, but not general cognitive ability, predicts children's recognition of emotion from facial and vocal cues

Sarah Griffiths ^{Corresp., 1}, Shaun Kok Yew Goh ^{1,2}, Courtenay Fraiser Norbury ^{1,3}

¹ Psychology and Language Sciences, University College London, University of London, London, London, United Kingdom

² Centre for Research in Child Development, Office of Educational Research, National Institute of Education, Nanyang Technological University, Singapore, Singapore

³ Department of Special Needs Education, University of Oslo, Oslo, Norway

Corresponding Author: Sarah Griffiths
Email address: sarah.griffiths@ucl.ac.uk

The ability to accurately identify and label emotions in the self and others is crucial for successful social interactions and good mental health. In the current study we tested the longitudinal relationship between early language skills and recognition of facial and vocal emotion cues in a representative UK population cohort with diverse language and cognitive skills (N = 369), including a large sample of children that met criteria for Developmental Language Disorder (DLD, N = 97). Language skills, but not non-verbal cognitive ability, at age 5-6 predicted emotion recognition at age 10-12. Children that met the criteria for DLD showed a large deficit in recognition of facial and vocal emotion cues. The results highlight the importance of language in supporting identification of emotions from non-verbal cues. Impairments in emotion identification may be one mechanism by which language disorder in early childhood predisposes children to later adverse social and mental health outcomes.

1

2 **Early language competence, but not general cognitive ability, predicts**3 **children's recognition of emotion from facial and vocal cues**

4 Running head: Language competence predicts emotion recognition

5

6 Sarah Griffiths¹, Shaun Kok Yew Goh^{1,2}, Courtenay Frazier Norbury^{1,3} & the SCALES team

7

8 1 Psychology and Language Sciences, University College London, London, UK

9 2 Centre for Research in Child Development, Office of Educational Research, National Institute
10 of Education, Nanyang Technological University, Singapore

11 3 Department of Special Needs Education, University of Oslo, Oslo, Norway

12

13 Corresponding Author:

14 Sarah Griffiths

15 Address: Chandler House, 2 Wakefield Street, London, WC1N 1PF

16 Email address: sarah.griffiths@ucl.ac.uk

17 **Abstract**

18 The ability to accurately identify and label emotions in the self and others is crucial for
19 successful social interactions and good mental health. In the current study we tested the
20 longitudinal relationship between early language skills and recognition of facial and vocal
21 emotion cues in a representative UK population cohort with diverse language and cognitive skills
22 (N = 369), including a large sample of children that met criteria for Developmental Language
23 Disorder (DLD, N = 97). Language skills, but not non-verbal cognitive ability, at age 5-6
24 predicted emotion recognition at age 10-12. Children that met the criteria for DLD showed a
25 large deficit in recognition of facial and vocal emotion cues. The results highlight the importance
26 of language in supporting identification of emotions from non-verbal cues. Impairments in
27 emotion identification may be one mechanism by which language disorder in early childhood
28 predisposes children to later adverse social and mental health outcomes.

29 **Keywords:** developmental language disorder, emotion recognition, facial expression, vocal
30 expression, longitudinal cohort study, language development.

31

32 **Introduction**

33 Recognition of emotional cues, such as facial and verbal expressions, is an important social skill.
34 It provides us with information about other people's internal emotional states and helps us to
35 interpret and predict their behaviour. Children have typically acquired the vocabulary for basic
36 emotions by 4-6 years of age (Baron-Cohen, Golan, Wheelwright, Granader, & Hill, 2010;
37 Ridgeway, Waters, & Kuczaj, 1985), but accuracy in identifying non-verbal emotional cues
38 continues to improve into late adolescence (Grosbras, Ross, & Belin, 2018; Herba & Phillips,
39 2004; Rodger, Vizioli, Ouyang, & Caldara, 2015). Accurate emotion identification has been
40 linked to positive outcomes later in development, including academic success (Denham et al.,
41 2012; Izard et al., 2001), social integration (Sette, Spinrad, & Baumgartner, 2017) and good
42 mental health (Ciarrochi, Scott, Deane, & Heaven, 2003).

43 A critical part of learning to identify emotions is developing emotional concepts that align
44 precisely with the emotional concepts held by other people. The Theory of Constructed Emotion
45 (TCE; Gendron & Barrett, 2018) proposes that language is crucial for acquiring nuanced
46 emotional concepts. Verbal labels provide a framework to organise highly variable input from
47 the environment into coherent emotion concepts (Gendron & Barrett, 2018; Lindquist, 2017).
48 Critically, the TCE suggests that the role of language in supporting emotion recognition goes
49 beyond acquisition of emotion vocabulary. Precise conceptual alignment is achieved through
50 communication with others. If an individual has less opportunity to learn about emotion concepts
51 through language, their conceptual alignment would be compromised, which would lead to less
52 accurate emotion identification. Previous research has shown that parent-child discourse about
53 emotions predicts children's emotion identification accuracy months later (Dunn, Brown,

54 Slomkowski, Tesla, & Youngblade, 1991), consistent with the theory that language aids learning
55 about emotions.

56 In the current study, we test the hypothesis that language supports development of accurate
57 emotion identification by studying a population that have reduced opportunity to learn about
58 emotion concepts through language. Children with Developmental Language Disorder (DLD;
59 previously known as Specific Language Impairment; Bishop, Snowling, Thompson, Greenhalgh,
60 & The Catalise Consortium, 2017) have difficulties with receptive and/or expressive language
61 that cannot be explained by a sensory deficit or neurological impairments (American Psychiatric
62 Association, 2013). Unlike children with autism spectrum disorders (ASD), children with DLD
63 do not have primary social or emotional difficulties, so any problems with emotion recognition
64 are likely to be a consequence of difficulties acquiring language. If language is necessary for
65 emotion conceptual alignment, children with DLD should have persistent difficulties with
66 emotion identification, due to reduced opportunity to learn about emotion concepts through
67 language.

68 A few small studies have found that school-aged children with DLD have difficulty labelling and
69 categorising facial (Bakopoulou & Dockrell, 2016; Taylor, Maybery, Grayndler, & Whitehouse,
70 2015) and verbal (Boucher, Lewis, & Collis, 2000; Taylor et al., 2015) emotional expressions.
71 However, other studies have reported contradictory results (Creusere, Alt, & Plante, 2004;
72 Loukusa, Mäkinen, Kuusikko-Gauffin, Ebeling, & Moilanen, 2014; Trauner, Ballantyne, Chase,
73 & Tallal, 1993). This equivocal evidence is likely due to variable diagnostic criteria, the
74 heterogeneity of the tasks used, and reduced statistical power due to the small sample sizes. The
75 estimated effect size for emotion recognition deficits in ASD, which we would assume to be
76 larger than the size of any deficit in DLD (due to the primary of their social-emotional

77 difficulties), is estimated to be 0.41 (Uljarevic & Hamilton, 2013). Power calculation suggests a
78 sample size of 135 participants in each group is needed to reliably detect an effect of this size
79 (Uljarevic & Hamilton, 2013). Therefore, much larger studies are required to determine whether
80 children with DLD do have difficulties with emotion identification.

81 A number of cross-sectional studies in the typically developing population have found
82 associations between language competence and the ability to label and match emotion emotional
83 facial expressions in early childhood (Beck, Kumschick, Eid, & Klann-Delius, 2012; Pons,
84 Lawson, Harris, & De Rosnay, 2003; Rosenqvist, Lahti-Nuuttila, Laasonen, & Korkman, 2014),
85 although other studies have failed to find this relationship (Herba, Landau, Russell, Ecker, &
86 Phillips, 2006; Herba & Phillips, 2004). Concurrent relationships between emotion recognition
87 performance and language competence in early childhood may be the result of children not
88 having the vocabulary to meet the language demands in the task. Stronger support for a role of
89 language in refining emotional concepts would come from demonstrating a longitudinal
90 relationship between language competence in early childhood and later accuracy in applying
91 labels to emotion cues, at an age when children have acquired basic emotion vocabulary.

92 In the current study, we use data from a well-characterised longitudinal population cohort that
93 includes children with the full spectrum of language abilities, including a large number of
94 children with DLD. We tested the relationship between early language competence (age 5-6) and
95 the ability to match facial and vocal emotion cues to basic emotion labels in middle childhood
96 (age 10-12) controlling for children's non-verbal cognitive ability. Additionally, we tested
97 whether emotion recognition is poorer in children that meet the criteria for DLD at age 5-6,
98 compared to children with typical language. The analysis plan for this study was preregistered on
99 the Open Science Framework (osf.io/pwcms).

100

101 **Materials and Methods**102 **Sample description**

103 Data are taken from the Surrey Communication and Language in Education Study (SCALES).

104 This study has followed a cohort of children who entered state-maintained schools in the county

105 of Surrey in the United Kingdom in September 2011. Language and communication skills were

106 assessed at school entry via a teacher report questionnaire (Children's Communication Checklist-

107 Short; CCC-S; Norbury et al., 2016). Based on screening, children were classified as having (1)

108 no phrase speech (NPS) (2) high risk for DLD (3) low risk for DLD. Children were classified

109 NPS if their teacher responded 'no' to the question 'is the child combining words into phrases or

110 sentences?' The cut-off between high and low risk status was based on age and sex specific cut-

111 on the CCC-S derived from the entire screened population (N=7267) Norbury et al.

112 2016 for details).

113 Stratified random sampling identified a subset of 636 children from the screened population who

114 were invited to take part in in-depth assessments. Children attending special schools for children

115 with severe intellectual or physical disability, and/or English as a second language were excluded

116 from sampling. All remaining children identified as being NPS (N=48) were invited, as were 233117 low risk and 355 high risk children. 529 monolingual children were assessed in Year 1, and

118 of these were assessed in Year 6. This final assessment included the emotion recognition tasks.

119 Language assessments in Year 1 were used to calculate composite scores for expressive

120 language, receptive language, vocabulary, grammar and narrative skills (Norbury et al., 2016).

121 Children were identified as meeting Developmental Language Disorder (DLD) if they scored -

122 1.5 SD on at least 2 out of 5 of these composite scores in Year 1. Scores from block design and
123 matrix reasoning (WPPSI-III; Wechsler, 2003) in Year 1 were used to calculate a non-verbal IQ
124 composite. This was used to identify children with intellectual disability, defined as a non-verbal
125 IQ composite score of less than -2 SD. Children that met DLD criteria in Year 1 were
126 additionally classified as having DLD with no known associated biomedical condition or DLD
127 with a known associated biomedical condition (hereafter termed LD+). Inclusion criteria for
128 ‘known associated biomedical condition’ was intellectual disability based on non-verbal IQ
129 assessments and/or parent/teacher reported diagnosis of an associated condition such as autism
130 (Norbury et al., 2016).

131 **Consent**

132 Consent procedures and study protocol were developed in consultation with Surrey County
133 Council and approved by the Royal Holloway Ethics Committee (in where the study started) in
134 Year 1 and the UCL Research Ethics Committee in Year 6 (9733/002). Informed consent was
135 collected from parents before in-depth assessments in Year 1 and Year 6. Informed assent was
136 collected from children prior to the Year 6 assessment. Children were given certificates and
137 small prizes at the end of each assessment session.

138 **Sample size and power calculations**

139 We conducted a priori sensitivity analyses in G-Power based on a sample size estimate of 399
140 participants (assuming a retention rate of 80% from previous assessment time-point in Year 3).
141 Sensitivity analysis suggested we would have 90% power to detect small ($r = .15$) associations
142 between language and emotion recognition accuracy in the whole sample (Cohen, 2013). We
143 also conducted a sensitivity analysis for assessing the group difference between DLD group and

144 the rest of the sample. Assuming equal attrition we estimated that we would have 103 children in
145 Year 6 that had met the DLD criteria in Year 1, including 70 with DLD with no additional
146 diagnosis. Sensitivity analysis suggest that this would provide 90% power to detect a small-
147 medium group differences $d = 0.34$ for comparison with full DLD group and $d = 0.38$ for
148 comparison excluding children with DLD with known origin) (Cohen, 2013).

149 Assessment procedure

150 Year 1 Language

151 In Year 1 children completed 6 tasks to assess receptive and expressive language skills. These
152 were; (1-2) Receptive/Expressive One word Picture Vocabulary Test (R/EOWPVT-4; Martin &
153 Brownell, 2000), (3) Test of Reception of Grammar – Short Form (TROG-S), (4) School-Age
154 Sentence Imitation Test (SASIT E32), (5-6) Assessment of Comprehension and Expression 6-11
155 (ACE 6-11; Adams et al., 2001). We combined scores on all of these tasks into a single language
156 composite score by averaging the Z-scores.

157 Year 1 Non-verbal IQ assessment

158 In Year 1 children completed two test of non-verbal IQ (NVIQ); (1) Wechsler Preschool and
159 Primary Scale of Intelligence^{3rd} Edition Block Design and (2) Matrix Reasoning subtests
160 (WPPSI-III; Wechsler, 2003). Performance on these two tasks has been combined to create a
161 NVIQ composite score (Norbury et al., 2016).

162 Year 6 Emotion recognition

163 In Year 6 children completed two emotion recognition tasks; one to measure recognition of
164 emotion from faces and one to measure emotion recognition from voices. Each task consisted of

165 60 trials in which children were presented with photos of faces or recordings of vocal sounds
166 corresponding to one of 6 emotions (happy, sad, angry, surprised, scared and disgusted). For the
167 facial expression task, stimuli were photos of 10 actors (5 female and 5 male) selected from the
168 Radboud Faces Database (Langner et al., 2010). For the vocal expression task, non-verbal sound
169 stimuli were selected from a validated set (Sauter, Eisner, Calder, & Scott, 2010) that have
170 previously been used in research with adults with developmental disorders (Jones et al., 2011).
171 The sounds are made by 4 actors (2 male and 2 female) that contribute one or more vocal
172 expression to each emotion category. In both tasks, participants were shown a fixation cross for
173 500ms, followed by face stimuli for 2 seconds or the audio clip accompanied by a cartoon image
174 of a listening man, followed by the 6 emotion labels presented in a circular formation on the
175 screen. The labels remained until the participant made a response. The order of the emotion
176 labels on the screen was randomised between participants and tasks, but kept the same between
177 trials for each participant. Total accuracy scores were calculated out of 60 for each task
178 separately.

179 Before completing the task, we checked children's understanding of the 6 emotion words by
180 asking them to read the labels aloud and describe or imitate that emotion. If the child was unable
181 to describe or imitate one or more of the emotions, the assessment was terminated as it was
182 assumed they did not have the basic emotion vocabulary. A very small number of children were
183 not able to read the labels but could describe or imitate the emotion when the word was said
184 aloud. For these children the researcher asked them give their response verbally during the task
185 and inputted their responses for them.

186 Analysis plan

187 Standardisation of scores

188 Test scores from each of the six language assessments and the two NVIQ assessments in Year 1
189 were standardised using the LMS method (Vamvakas, Norbury, Vitoratou, Gooch, & Pickles,
190 2019). LMS is a method of standardisation based on the Box-Cox transformation that converts
191 scale raw scores to normality. The resulting scores reflect standardised scores adjusted for age,
192 with a mean of 0 and a standard deviation of 1. We planned to standardise emotion recognition
193 scores using the same method but this was not necessary as performance was not correlated with
194 age in our sample (faces = .05, $p = .37$; voices = .002, $p = .97$).

195 Sampling weights and missing data

196 Sampling weights were included in all analyses to account for study design and any bias in
197 attrition. This adjustment means that estimates are representative of the screened sample of 6,459
198 mono-lingual children in state-maintained schools. Sampling weights were produced by
199 multiplying the inverse of the predicted probability of two logistic regression models that predict
200 inclusion in the sample. The first regression model estimates a child's likelihood of being
201 initially invited into the study. This was fitted to the entire population of 6,459 mono-lingual
202 children in mainstream schools that were screened at school entry. The covariates in this model
203 are those that determined selection into the study due to the stratified sampling method. These
204 are total number of children assessed per school, and whether a child was identified as at risk for
205 DLD based on CCC-S teacher ratingsth (86th percentile or above for sex and age group). The second
206 regression model was fitted to the 636 children invited into the study. This model includes any
207 variable that might be predictive of inclusion in the sample included in the current analysis. This
208 included individual characteristics such as sex, income deprivation score, special education
209 needs, free school meals, English as additional language, CCC-2 score, language in Year 1,
210 season of birth, SDQ total difficulties, and school characteristics such as number of pupils on

211 role, percentage of girls, percentage with SEN, and percentage with free school meals. These
212 variables were tested in a stepwise elimination process and included in the model if they predict
213 inclusion above a cut-off point of .2.

214 Statistical analysis

215 Statistical analyses were conducted in R version 3.5.3 and M-Plus. Structural Equation Models
216 (SEM) were built under robust maximum likelihood estimator which is robust to deviations from
217 normality. To test the hypothesis that language competence in Year 1 predicts emotion
218 recognition from faces and voices in Year 6, path analysis was used to model the association
219 between children's composite language scores in Year 1 and their scores on the facial expression
220 and vocal expressions tasks in Year 6. Additionally, because one previous study had suggested
221 that children with DLD may be more impaired in recognition of emotion cues from voices rather
222 than faces (Trauner, Ballantyne, Chase, & Tallal, 1993) we compared the strength of the
223 pathways between language and performance on the facial expression task and vocal expression
224 task using Wald test of parameter constraints. Finally, we then entered Year 1 NVIQ composite
225 into the model to assess whether language scores continue to predict emotion recognition after
226 accounting for variation in non-verbal cognitive ability.

227 We then compared children in the DLD group to the rest of the sample on total accuracy from
228 the emotion recognition tasks separately. We did not control for NVIQ in this analysis because
229 low NVIQ is not an exclusion criteria for DLD (Bishop, Snowling, Thompson, Greenhalgh, &
230 The Catalise Consortium, 2017) and language severity is associated with NVIQ (Norbury et al.,
231 2017). This means 'controlling' for group differences in NVIQ would 'control' for relevant and
232 non-random differences between the two groups (Dennis et al. 2009). We conducted this analysis
233 both with and without removing children with additional diagnoses, to determine if there was

234 still a group difference after removing children with co-occurring conditions that have also been
235 associated with problems with emotion recognition (e.g. autism and/or severe intellectual
236 disability).

237 Results

238 Participants

239 Of the 384 participants were seen for assessment in Year 6, 362 (including 67 with DLD and 29
240 with LD+ additional diagnoses) completed the facial emotion recognition task and 359 (63 with
241 DLD and 27 with LD+ additional diagnoses) completed the vocal emotion recognition task. 369
242 completed at least one task (67 with DLD and 30 with LD+ additional diagnoses) so were
243 included in the analysis. Of the 15 children that did not complete either task, six met criteria for
244 language disorder in Year 1. These children did not complete the task because they did not have
245 the basic emotion vocabulary or were otherwise unable to engage in the task. The other nine
246 children did not have DLD did not complete the tasks due to technical issues. Table 1 provides
247 descriptive statistics for all variables for the total sample, and DLD, LD+, and typical language
248 groups separately.

249 Attrition was slightly higher than we had anticipated when we conducted our a priori sensitivity
250 analysis. However, our achieved sample size still gave us 90% power to detect small associations
251 between language and emotion recognition ($d = 0.15$). We also still had 90% power to detect
252 small-medium size group differences in emotion recognition accuracy between the DLD group
253 and typical language group ($d = 0.35$), even after excluding those with LD+ additional diagnoses
254 ($d = 0.41$).

255 Does early language competence predict later emotion recognition accuracy?

256 Path analysis supported our prediction that early language competence is positively associated
257 with emotion recognition ability in late childhood. There were moderate prospective
258 relationships between language competence in Year 1 and emotion recognition from vocal
259 expressions ($\beta = .40$, S.E = .06, 95% CI [.28,.51]) and facial expressions ($\beta = .42$, S.E = .06, 95%
260 CI [.30,.55]) in Year 6.

261 The relationship between language and emotion recognition was similar for faces and voices.
262 Wald's test of parameter constraints did not provide evidence for a difference in path strengths
263 ($X^2(1) = 2.51$, $p = .11$). We had planned on combining the two emotion recognition scores into a
264 single composite score if there was no evidence for a difference in path strengths. However, the
265 correlation between the two outcomes estimated in the model was not sufficient to justify this (r
266 = .37, S.E = .06, 95% CI [.25, .50]).

267 When NVIQ in Year 1 was entered into the path model as a predictor, the relationships between
268 language and performance on the two emotion recognition tasks remained, and there was no
269 statistical evidence for a prospective relationship between NVIQ and performance on either
270 emotion recognition task (see Fig. 1 for the standardised regression coefficients and confidence
271 intervals for these paths). These results support our hypothesis that early language is positively
272 associated with emotion recognition accuracy, even after adjustment for non-verbal IQ.

273 **Do children with DLD have poorer emotion recognition skills than their peers with typical** 274 **language?**

275 Figure 2 illustrates the distributions of raw accuracy scores on the facial and vocal emotion
276 recognition task for children diagnosed with DLD and children with language in the typical
277 range. Weighted t-tests provided clear statistical evidence for a large group difference in

278 recognition of emotions from faces; $t(360) = 4.06, p < .001, d = .90$, and voices $t(353) = 4.24, p$
279 $< .001, d = 0.89$, between these groups. When children with LD+ additional diagnoses were
280 removed, the effect sizes reduced slightly but there was still evidence for a medium-large group
281 difference for recognition of emotion from faces; $t(331) = 2.72, p = .007, d = .72$, and voices;
282 $t(326) = 2.87, p < .001, d = .78$. These findings support our hypothesis that children who meet
283 the criteria for DLD in early childhood are poorer at recognising emotions from both faces and
284 voices in late childhood than their peers with typical language.

285 **Do children with DLD make similar errors in emotion recognition tasks to their peers with** 286 **typical language?**

287 In order to explore possible differences in the kinds of errors made by children with DLD and
288 those without DLD, we created confusion matrices for each task for each group of children (Fig.
289 3.). From these it can be seen that in general the pattern of errors is very similar across the two
290 groups. The most commonly misidentified emotion in the facial emotion recognition task was
291 disgust in the DLD group and fear in the typical language group and the least commonly
292 misidentified emotion was happiness in both groups. For the vocal emotion recognition task, the
293 most commonly misidentified emotion in both groups was surprised and the least commonly
294 misidentified emotion was happiness in the DLD group and disgust in the typical language
295 group.

296 **Discussion**

297 In the present study we examined the prospective relationship between language competence in
298 early childhood and identification of non-verbal emotion cues in middle childhood in a large
299 population-derived cohort of children with diverse language and cognitive skills. We found

300 evidence for a moderate relationship between language competence at age 5-6 and recognition of
301 facial and vocal emotional cues at age 10-12. This relationship held when adjusting for non-
302 verbal cognitive ability, suggesting language selectively predicts emotion recognition ability in
303 middle childhood. Furthermore, we found that children with DLD, who are less able to learn
304 about emotions through language, have a large deficit in emotion recognition ability. This is the
305 first longitudinal evidence to support the hypothesis from the TCE that language competence
306 plays a role in supporting accurate identification non-verbal emotion cues (Gendron & Barrett,
307 2018).

308 These results help clarify contradictory literature on whether children with DLD have deficits in
309 emotion recognition (Bakopoulou & Dockrell, 2016; Boucher, Lewis, & Collis, 2000; Creusere,
310 Alt, & Plante, 2004; Loukusa, Mäkinen, Kuusikko-Gauffin, Ebeling, & Moilanen, 2014; Taylor,
311 Maybery, Grayndler, & Whitehouse, 2015; Trauner et al., 1993). The majority of previous
312 studies have been small ($N < 20$ children with DLD) and have therefore lacked the statistical
313 power to detect the expected medium-small effect size (Uljarevic & Hamilton, 2013). Ninety-
314 seven children in our cohort that met the DSM criteria for DLD when they were 5-6 years old,
315 based on rigorous linguistic and cognitive testing, completed at least one emotion recognition
316 task in Year 6, giving us sufficient statistical power. We found strong evidence for a large
317 difference in emotion recognition ability at age 10-12 between those that met the criteria for
318 DLD at age 5-6 compared to those with language in the typical range at age 5-6. When children
319 with other diagnoses that associate with both language and emotion recognition difficulties were
320 excluded from the DLD group (e.g. autism or intellectual disability), the group difference
321 attenuated somewhat, but the statistical evidence for differences between children with DLD and
322 their peers remained strong. These findings provide the best evidence to date that children with

323 DLD have emotion recognition deficits, and that these are directly linked to language
324 proficiency.

325 Our emotion recognition task was verbal in the sense that children had to match non-verbal
326 emotion cues to verbal labels. It could therefore be argued that the verbal demands of the task
327 explain the relationship between language competence and emotion recognition performance.
328 However, the labels were basic emotion words that are highly frequent and well within the
329 vocabulary range of children aged 10-12 (Baron-Cohen et al., 2010), even the vast majority of
330 those with DLD. We checked that children understood the emotion words before completing the
331 assessment. A very small number of children with DLD (N=6) lacked the vocabulary to engage
332 in the task, so were not included in the study. Non-verbal tasks, such as a facial expression
333 matching task (Taylor et al., 2015), can be completed using visual features alone, without any
334 comprehension of the underlying emotion and so do not truly test emotion identification. We
335 argue that performance on an emotion labelling is associated with early language competence not
336 just because it involves a verbal label, but because language is involved in developing nuanced
337 emotion concepts through communication with others throughout childhood (Gendron & Barrett,
338 2018).

339 The ability to recognise and label emotions in the self and others is an important component of
340 social problem solving. The finding that this ability is compromised in DLD may elucidate why
341 children with DLD are at increased risk of internalising, externalising and ADHD symptoms
342 (Yew & O’Kearney, 2013). The causal pathway between DLD and poor mental health outcomes
343 is unclear, but one possibility is that language problems interfere with aspects of social-
344 emotional processing (such as emotion concept development), which in turn leads to negative
345 social, emotional and mental health outcomes. Im-Bolter, Cohen, and Farnia (2013) found that

346 adolescents referred to mental health services had poorer structural and figural language than
347 peers recruited from the community and were poorer at social problem solving. The findings in
348 the current study raise the possibility that emotion identification may be one pathway in which
349 poor language in early childhood compromises social functioning and mental health in children
350 with and without DLD.

351 Although our findings are consistent with a causal relationship between language competence
352 and emotion recognition, they cannot provide proof of causality. One way to investigate whether
353 this relationship is truly causal would be to test whether interventions aimed at improving
354 language have positive, cascading effects on emotion recognition skills later in development. To
355 date there has been one preliminary study (N= 208) investigating whether a nine-week
356 intervention focusing specifically on improving language related to emotion through storybooks
357 improves other emotional skills in typically developing 7-9 year old children. The intervention
358 group showed improvements in emotional vocabulary, emotion knowledge and recognition of
359 masked emotions from vignettes compared to a ‘treatment as usual’ control group straight after
360 the intervention ($\beta = 1.05-1.32$; Kumschick et al., 2014). Future research is needed to determine
361 whether a similar intervention could improve emotion recognition and other social-emotional
362 competencies in children with and without DLD.

363 **Conclusions**

364 In conclusion, this study provides the first longitudinal evidence that early language skills
365 specifically predict later emotion recognition from both facial and vocal cues. These findings
366 support the hypothesis that language plays a role in supporting emotion identification (Barrett,
367 Lindquist, & Gendron, 2007; Gendron & Barrett, 2018; Lindquist, 2017). Children with DLD are

368 therefore especially vulnerable to difficulties recognising their own and others' emotional states.

369 We propose that this deficit may be one causal mechanism that underpins the reported

370 relationship between early language skills and later adverse mental health.

371

372 **Acknowledgements**

373 We thank Surrey County Council for facilitating the data collection process and the children,
374 parents, schools and teachers for taking part in the study. We also thank the other members of the
375 SCALES team: Debbie Gooch, Gillian Baird, Tony Charman, Andrew Pickles and Emily
376 Simonoff for their advice. Finally, we thank Dorothy Bishop for permission to develop the
377 Children's Communication Checklist-Short and allowing us access to the standardization data.
378 The views expressed in this article are those of the authors and not necessarily those of the
379 Wellcome Trust, the ESRC, the British Academy or Surrey County Council.

380 **Open Practices Statement**

381 The hypotheses and analysis plan for this study was preregistered on the Open Science
382 Framework (osf.io/pwcms). Twelve participant's parents did not consent to Open Data sharing
383 so we are unable to share the full dataset. However, our code and a dataset with these 12
384 participants removed is available on the OSF page for this study.

385 **References**

386

387 Bakopoulou, I., & Dockrell, J. E. (2016). The role of social cognition and prosocial behaviour in
388 relation to the socio-emotional functioning of primary aged children with specific
389 language impairment. *Research in Developmental Disabilities, 49-50*, 354-370.

390 doi:<https://doi.org/10.1016/j.ridd.2015.12.013>

391 Baron-Cohen, S., Golan, O., Wheelwright, S., Granader, Y., & Hill, J. (2010). Emotion word
392 comprehension from 4 to 16 years old: a developmental survey. *Frontiers in evolutionary
393 neuroscience, 2*, 109-109. doi:10.3389/fnevo.2010.00109

394 Barrett, L. F., Lindquist, K. A., & Gendron, M. (2007). Language as context for the perception of
395 emotion. *Trends in cognitive sciences, 11(8)*, 327-332.

396 doi:<https://doi.org/10.1016/j.tics.2007.06.003>

397 Beck, L., Kumschick, I. R., Eid, M., & Klann-Delius, G. (2012). Relationship between language
398 competence and emotional competence in middle childhood. *Emotion, 12(3)*, 503-514.

399 doi:10.1037/a0026320

400 Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & The Catalise
401 Consortium. (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi
402 consensus study of problems with language development: Terminology. *Journal of Child
403 Psychology and Psychiatry, 58(10)*, 1068-1080. doi:10.1111/jcpp.12721

404 Boucher, J., Lewis, V., & Collis, G. M. (2000). Voice Processing Abilities in Children with
405 Autism, Children with Specific Language Impairments, and Young Typically Developing
406 Children. *Journal of Child Psychology and Psychiatry, 41(7)*, 847-857.

407 doi:10.1111/1469-7610.00672

408 Ciarrochi, J., Scott, G., Deane, F. P., & Heaven, P. C. L. (2003). Relations between social and
409 emotional competence and mental health: a construct validation study. *Personality and
410 Individual Differences, 35(8)*, 1947-1963. doi:[https://doi.org/10.1016/S0191-](https://doi.org/10.1016/S0191-8869(03)00043-6)

411 [8869\(03\)00043-6](https://doi.org/10.1016/S0191-8869(03)00043-6)

412 Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*: Routledge.

413 Creusere, M., Alt, M., & Plante, E. (2004). Recognition of vocal and facial cues to affect in
414 language-impaired and normally-developing preschoolers. *Journal of Communication
415 Disorders, 37(1)*, 5-20. doi:[https://doi.org/10.1016/S0021-9924\(03\)00036-4](https://doi.org/10.1016/S0021-9924(03)00036-4)

416 Denham, S. A., Bassett, H. H., Way, E., Mincic, M., Zinsser, K., & Graling, K. (2012).
417 Preschoolers' emotion knowledge: self-regulatory foundations, and predictions of early
418 school success. *Cogn Emot, 26(4)*, 667-679. doi:10.1080/02699931.2011.602049

419 Dunn, J., Brown, J., Slomkowski, C., Tesla, C., & Youngblade, L. (1991). Young children's
420 understanding of other people's feelings and beliefs: individual differences and their
421 antecedents. *Child Dev, 62(6)*, 1352-1366.

422 Gendron, M., & Barrett, L. F. (2018). Emotion Perception as Conceptual Synchrony. *Emotion
423 Review, 10(2)*, 101-110. doi:10.1177/1754073917705717

424 Grosbras, M.-H., Ross, P. D., & Belin, P. (2018). Categorical emotion recognition from voice
425 improves during childhood and adolescence. *Scientific Reports, 8(1)*, 14791.

426 doi:10.1038/s41598-018-32868-3

427 Herba, C., Landau, S., Russell, T., Ecker, C., & Phillips, M. (2006). The development of
428 emotion-processing in children: effects of age, emotion, and intensity. *Journal of Child
429 Psychology and Psychiatry, 47(11)*, 1098-1106. doi:10.1111/j.1469-7610.2006.01652.x

- 430 Herba, C., & Phillips, M. (2004). Annotation: Development of facial expression recognition from
431 childhood to adolescence: behavioural and neurological perspectives. *Journal of Child*
432 *Psychology and Psychiatry*, 45(7), 1185-1198. doi:10.1111/j.1469-7610.2004.00316.x
- 433 Im-Bolter, N., Cohen, N. J., & Farnia, F. (2013). I thought we were good: social cognition,
434 figurative language, and adolescent psychopathology. *J Child Psychol Psychiatry*, 54(7),
435 724-732. doi:10.1111/jcpp.12067
- 436 Izard, C., Fine, S., Schultz, D., Mostow, A., Ackerman, B., & Youngstrom, E. (2001). Emotion
437 knowledge as a predictor of social behavior and academic competence in children at risk.
438 *Psychol Sci*, 12(1), 18-23. doi:10.1111/1467-9280.00304
- 439 Jones, C. R., Pickles, A., Falcaro, M., Marsden, A. J., Happe, F., Scott, S. K., . . . Charman, T.
440 (2011). A multimodal approach to emotion recognition ability in autism spectrum
441 disorders. *J Child Psychol Psychiatry*, 52(3), 275-285. doi:10.1111/j.1469-
442 7610.2010.02328.x
- 443 Kumschick, I. R., Beck, L., Eid, M., Witte, G., Klann-Delius, G., Heuser, I., . . . Menninghaus,
444 W. (2014). READING and FEELING: the effects of a literature-based intervention
445 designed to increase emotional competence in second and third graders. *Frontiers in*
446 *Psychology*, 5(1448). doi:10.3389/fpsyg.2014.01448
- 447 Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A.
448 (2010). Presentation and validation of the Radboud Faces Database. *Cognition and*
449 *Emotion*, 24(8), 1377-1388. doi:10.1080/02699930903485076
- 450 Lindquist, K. A. (2017). The role of language in emotion: existing evidence and future
451 directions. *Current Opinion in Psychology*, 17, 135-139.
452 doi:<https://doi.org/10.1016/j.copsyc.2017.07.006>
- 453 Loukusa, S., Mäkinen, L., Kuusikko-Gauffin, S., Ebeling, H., & Moilanen, I. (2014). Theory of
454 mind and emotion recognition skills in children with specific language impairment,
455 autism spectrum disorder and typical development: group differences and connection to
456 knowledge of grammatical morphology, word-finding abilities and verbal working
457 memory. *International Journal of Language & Communication Disorders*, 49(4), 498-
458 507. doi:10.1111/1460-6984.12091
- 459 Norbury, C. F., Vamvakas, G., Gooch, D., Baird, G., Charman, T., Simonoff, E., & Pickles, A.
460 (2017). Language growth in children with heterogeneous language disorders: a
461 population study. *Journal of Child Psychology and Psychiatry*, 58(10), 1092-1105.
462 doi:10.1111/jcpp.12793
- 463 Pons, F., Lawson, J., Harris, P. L., & De Rosnay, M. (2003). Individual differences in children's
464 emotion understanding: Effects of age and language. *Scandinavian Journal of*
465 *Psychology*, 44(4), 347-353. doi:10.1111/1467-9450.00354
- 466 Ridgeway, D., Waters, E., & Kuczaj, S. A. (1985). Acquisition of emotion-descriptive language:
467 Receptive and productive vocabulary norms for ages 18 months to 6 years. *Dev Psychol*,
468 21(5), 901.
- 469 Rodger, H., Vizioli, L., Ouyang, X., & Caldara, R. (2015). Mapping the development of facial
470 expression recognition. *Developmental Science*, 18(6), 926-939. doi:10.1111/desc.12281
- 471 Rosenqvist, J., Lahti-Nuutila, P., Laasonen, M., & Korkman, M. (2014). Preschoolers'
472 recognition of emotional expressions: Relationships with other neurocognitive capacities.
473 *Child Neuropsychology*, 20(3), 281-302. doi:10.1080/09297049.2013.778235

- 474 Sauter, D. A., Eisner, F., Calder, A. J., & Scott, S. K. (2010). Perceptual Cues in Nonverbal
475 Vocal Expressions of Emotion. *Quarterly Journal of Experimental Psychology*, *63*(11),
476 2251-2272. doi:10.1080/17470211003721642
- 477 Sette, S., Spinrad, T. L., & Baumgartner, E. (2017). The Relations of Preschool Children's
478 Emotion Knowledge and Socially Appropriate Behaviors to Peer Likability. *Int J Behav*
479 *Dev*, *41*(4), 532-541. doi:10.1177/0165025416645667
- 480 Taylor, L. J., Maybery, M. T., Grayndler, L., & Whitehouse, A. J. O. (2015). Evidence for shared
481 deficits in identifying emotions from faces and from voices in autism spectrum disorders
482 and specific language impairment. *International Journal of Language & Communication*
483 *Disorders*, *50*(4), 452-466. doi:10.1111/1460-6984.12146
- 484 Trauner, D. A., Ballantyne, A., Chase, C., & Tallal, P. (1993). Comprehension and expression of
485 affect in language-impaired children. *Journal of Psycholinguistic Research*, *22*(4), 445-
486 452. doi:10.1007/bf01074346
- 487 Uljarevic, M., & Hamilton, A. (2013). Recognition of Emotions in Autism: A Formal Meta-
488 Analysis. *Journal of Autism and Developmental Disorders*, *43*(7), 1517-1526.
489 doi:10.1007/s10803-012-1695-5
- 490 Vamvakas, G., Norbury, C. F., Vitoratou, S., Gooch, D., & Pickles, A. (2019). Standardizing test
491 scores for a target population: The LMS method illustrated using language measures from
492 the SCALES project. *PLoS ONE*, *14*(3), e0213492. doi:10.1371/journal.pone.0213492
- 493 Wechsler, D. (2003). Wechsler intelligence scale for children—Fourth Edition (WISC-IV). *San*
494 *Antonio, TX: The Psychological Corporation*.
- 495 Yew, S. G. K., & O’Kearney, R. (2013). Emotional and behavioural outcomes later in childhood
496 and adolescence for children with specific language impairments: meta-analyses of
497 controlled prospective studies. *Journal of Child Psychology and Psychiatry*, *54*(5), 516-
498 524. doi:doi:10.1111/jcpp.12009
- 499

Table 1 (on next page)

Descriptive statistics for the full sample and each language group separately.

Language and NVIQ are standard scores based on population norms estimated using sample weights. Emotion recognition scores are raw total accuracy scores on each task.

1 Table 1:

2 **Descriptive statistics for the full sample and each language group separately.**

3 Language and NVIQ are standard scores based on population norms estimated using sample
4 weights. Emotion recognition scores are raw total accuracy scores on each task.

	Full sample N = 369	Typical language N = 272	DLD N = 67	LD+ N = 30
	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>
Age Year 1 (Months)	72 (4.69)	72 (4.84)	71 (4.05)	74 (4.06)
Age Year 6 (Months)	134 (4.06)	134 (4.04)	133 (4.29)	135 (3.69)
Male (%)	50.14	47.43	55.22	63.33
Language composite Year 1	-0.59 (1.06)	-0.09 (0.85)	-1.78 (0.49)	-2.50 (0.67)
NVIQ composite Year 1	-0.39 (1.07)	-0.13 (0.98)	-0.83 (0.65)	-1.81 (1.19)
ER faces Year 6 ^a	0.76 (0.12)	0.79 (0.10)	0.71 (0.14)	0.61 (0.13)
ER voices Year 6 ^b	0.75 (0.11)	0.78 (0.08)	0.71 (0.12)	0.59 (0.16)

5

6 NVIQ = non-verbal IQ; ER = Emotion recognition

7 a Based on 362 total, 272 TL, 67 DLD, 29 LD+

8 b Based on 359 total, 278 TL, 63 DLD, 30 LD+

9

Figure 1

Path model showing prospective relationships from language (Lang) and non-verbal IQ (NVIQ) in Year 1 to emotion recognition from faces (ER faces) and voices (ER voices) in Year 6.

Significant paths are solid lines while insignificant paths are dashed line Significant paths are solid lines while insignificant paths are dashed lines.

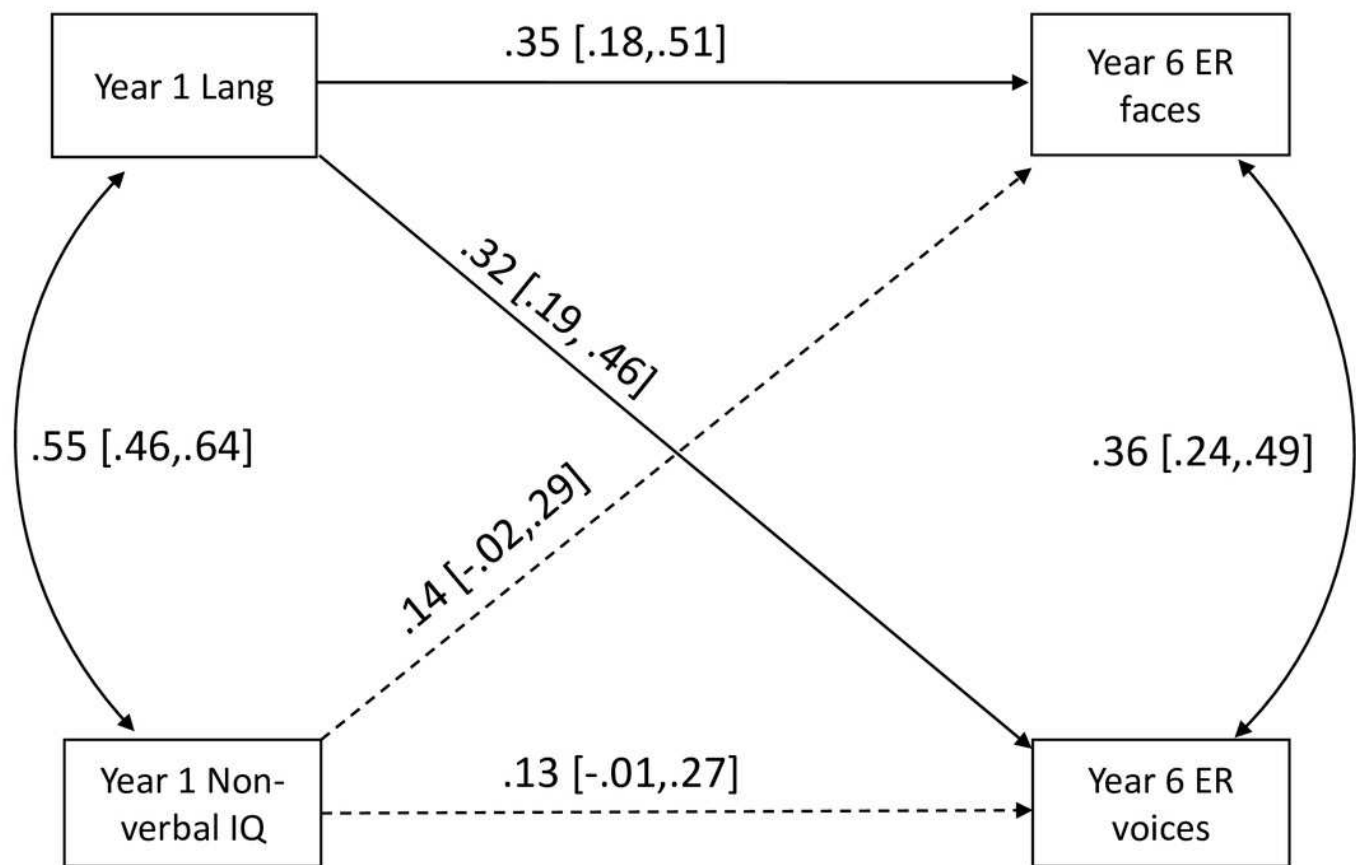


Figure 2

Pirate plot showing distribution of total scores on (A) the vocal emotion recognition task and (B) the facial emotion recognition task for group with DLD and the typically language group.

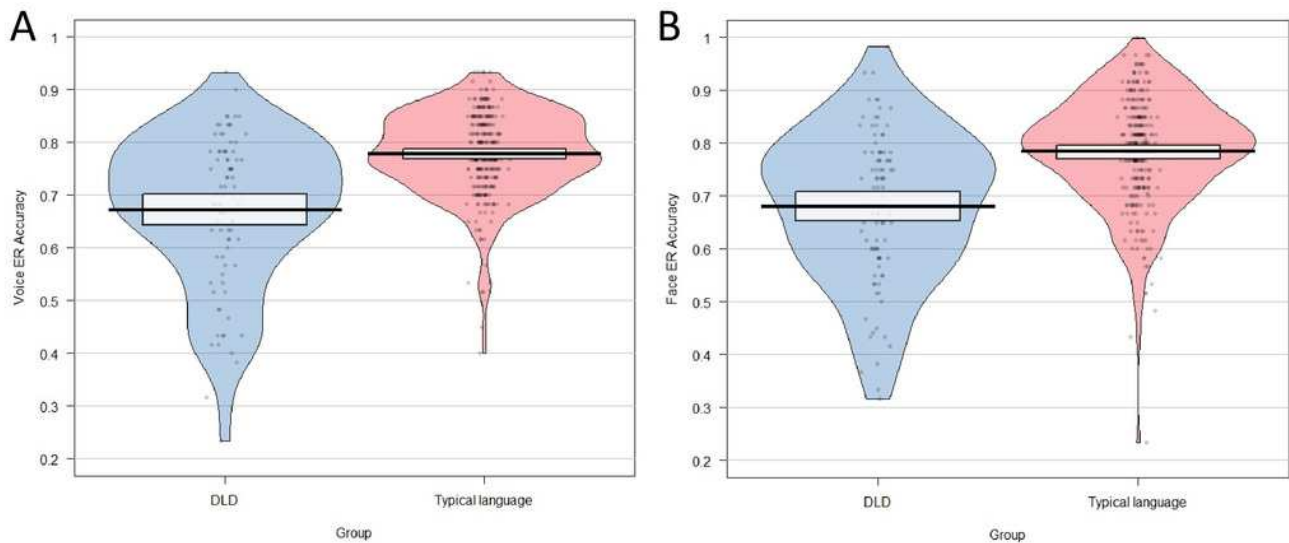
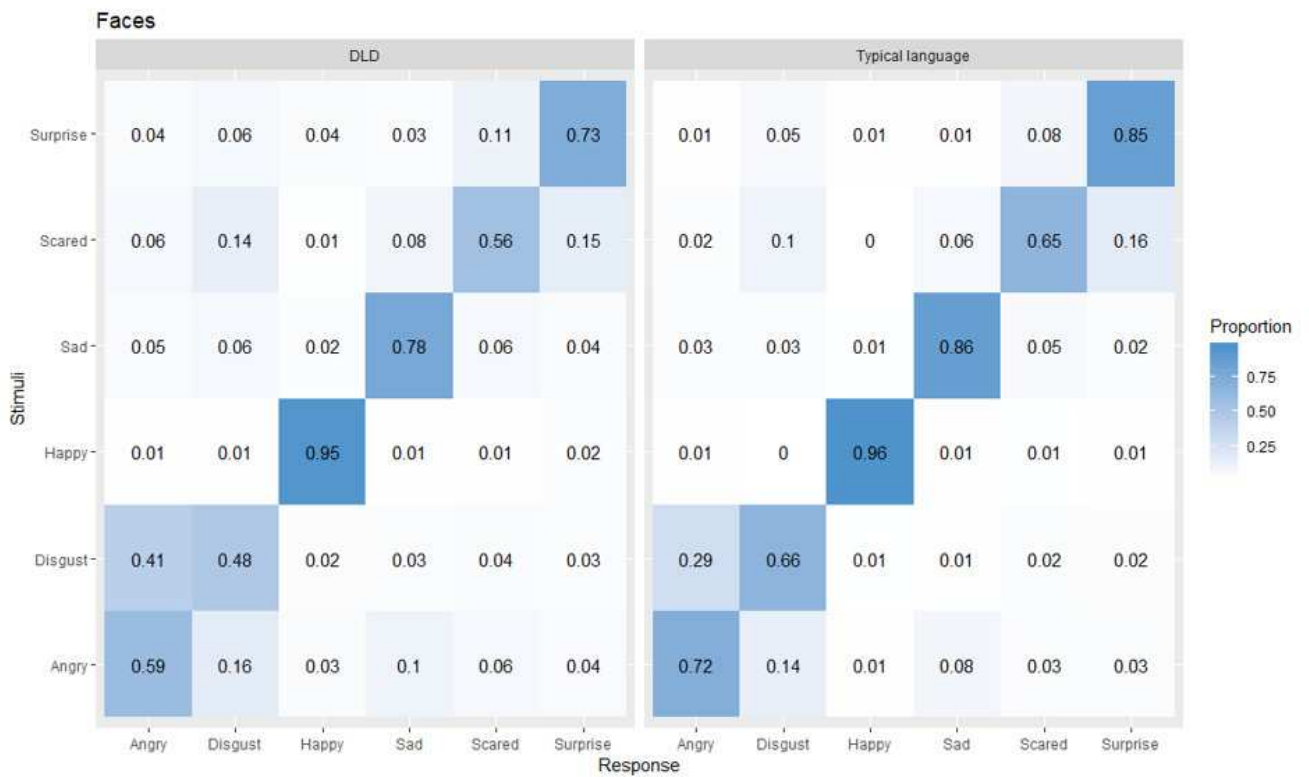


Figure 3

Confusion matrices showing proportion of responses in each category for each presented emotion separately for group with DLD (left) and with typical language (right) for (a) faces and (b) voices.

A



B

