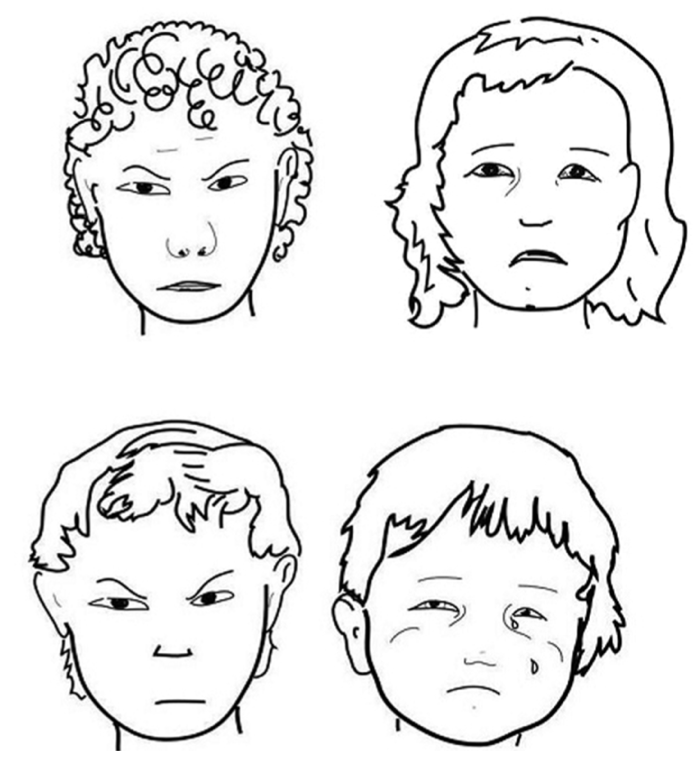
**Supplementary files**



Supplementary Figure 1. Examples of facial expressions used in this study. From left to right: angry facial expressions and sad facial expressions.

Supplementary Table 1. Means, standard deviations (SD), and *t*-test comparisons of group performances in the emotion-discrimination task at each time point.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***Autistic*** | | | ***Non-autistic*** | | |  | |
|  | **Mean** | ***SD*** | ***N*** | **Mean** | ***SD*** | ***N*** | ***p*** | |
| **Testing tasks Positive vs. Negative (0-3)** | | | | | | | |  |
| Time 1 | 1.94 | 1.06 | 60 | 2.19 | .84 | 121 .089 | |  |
| Time 2 | 2.30 | .96 | 45 | 2.64 | .74 | 51 .060 | |  |
| Time 3 | 2.49 | .83 | 43 | 2.86 | .41 | 49 .011 | |  |
| Time 4 | 2.59 | .81 | 41 | 2.89 | .31 | 47 .025 | |  |
| **Negative vs. negative (0-3)** | | | | | | | |  |
| Time 1 | 1.60 | 1.06 | 60 | 1.87 | .91 | 121 .077 | |  |
| Time 2 | 1.97 | .95 | 45 | 2.17 | .94 | 52 .287 | |  |
| Time 3 | 2.17 | .94 | 43 | 2.56 | .75 | 49 .034 | |  |
| Time 4 | 2.33 | .90 | 40 | 2.73 | .52 | 47 .014 | |  |

Supplementary Table 2. Means, standard deviations (SD), and *t*-test comparisons of group performances in the emotion-identification task at each time point.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | ***Autistic*** | | | ***Non-autistic*** | | |
|  | **Mean** | ***SD*** | ***N*** | **Mean** | ***SD*** | ***N p*** |
| **Positive (0-2)** |  |  |  |  |  |  |
| Time 1 | 1.41 | .86 | 61 | 1.75 | .60 | 121 .007 |
| Time 2 | 1.67 | .74 | 45 | 1.98 | .14 | 52 .007 |
| Time 3 | 1.88 | .45 | 43 | 1.94 | .24 | 49 .458 |
| Time 4 | 1.93 | .35 | 41 | 2.00 | .00 | 47 .183 |
| **Negative (0-2)** |  |  |  |  |  |  |
| Time 1 | 1.16 | .81 | 61 | 1.50 | .60 | 121 .004 |
| Time 2 | 1.47 | .71 | 45 | 1.85 | .26 | 52 .001 |
| Time 3 | 1.76 | .55 | 43 | 1.95 | .14 | 49 .030 |
| Time 4 | 1.85 | .43 | 41 | 1.96 | .14 | 47 .096 |

Supplementary Table 3. Means, standard deviations (SD), and *t*-test comparisons of group performances in the emotion-attribution task at each time point.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Autistic** | | | **Non-autistic** | | |  |
|  | **Mean** | ***SD*** | ***N*** | **Mean** | ***SD*** | ***N*** | ***p*** |
| **Verbal condition** | | | | | | | |
| **Positive emotions (0-2)** |  |  |  |  |  |  |  |
| Time 1 | 1.46 | .77 | 61 | 1.59 | .68 | 121 | .742 |
| Time 2 | 1.53 | .76 | 45 | 1.81 | .39 | 52 | .033 |
| Time 3 | 1.56 | .70 | 43 | 1.78 | .47 | 49 | .089 |
| Time 4 | 1.61 | .63 | 41 | 1.83 | .38 | 47 | .055 |
| **Negative emotions (0-2)** |  |  |  |  |  |  |  |
| Time 1 | 1.20 | .65 | 61 | 1.11 | .53 | 116 | .388 |
| Time 2 | 1.30 | .61 | 45 | 1.29 | .31 | 52 | .891 |
| Time 3 | 1.45 | .55 | 43 | 1.28 | .38 | 49 | .076 |
| Time 4 | 1.33 | .55 | 41 | 1.30 | .45 | 47 | .756 |
| **Visual condition** | | | | | | | |
| **Positive emotions (0-2)** |  |  |  |  |  |  |  |
| Time 1 | 1.13 | .89 | 61 | 1.59 | .68 | 121 | <.001 |
| Time 2 | 1.58 | .78 | 45 | 1.88 | .32 | 52 | .017 |
| Time 3 | 1.60 | .69 | 43 | 1.78 | .47 | 49 | .177 |
| Time 4 | 1.63 | .58 | 41 | 1.86 | .35 | 43 | .035 |
| **Negative emotions (0-2)** |  |  |  |  |  |  |  |
| Time 1 | .96 | .73 | 60 | 1.13 | .49 | 115 | .096 |
| Time 2 | 1.34 | .62 | 45 | 1.33 | .30 | 52 | .946 |
| Time 3 | 1.47 | .57 | 43 | 1.27 | .39 | 49 | .053 |
| Time 4 | 1.34 | .53 | 41 | 1.34 | .45 | 45 | .956 |

Supplementary Table 4. Descriptive statistics of the background and study variables measured at Time 1, and the *t*-test comparisons of the subsamples with and without missing data of the outcome variables at Time 2.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Subsample with missing data | | Subsample without missing data | | *t*-test statistics | |
|  | ***N*** | ***Mean (SD)*** | ***N*** | ***Mean (SD)*** | ***t*** | ***p*** |
| IQ | 27 | 104.30 (13.39) | 86 | 106.02 (17.38) | .47 | .637 |
| Education mother | 17 | 3.12 (.78) | 55 | 3.29 (.69) | .88 | .381 |
| Education father | 18 | 3.39 (.61) | 55 | 3.27 (.91) | .50 | .616 |
| Family income | 35 | 3.29 (.1.41) | 71 | 3.62 (1.07) | .1.24 | .221 |
| Emotion vocabulary | 60 | 1.53 (.38) | 91 | 1.47 (.45) | .87 | .386 |
| Language expression | 55 | .89 (.15) | 88 | .88 (.21) | .41 | .685 |
| Language comprehension | 54 | .89 (.16) | 85 | .83 (.22) | 1.65 | .101 |
| Emotion discrimination | | | | | | |
| *Positive vs. Negative* | 86 | 2.03 (.91) | 95 | 2.17 (.94) | 1.01 | .315 |
| *Negative vs. Negative* | 86 | 1.78 (1.01) | 95 | 1.77 (.92) | .08 | .938 |
| Emotion identification | | | | | | |
| *Positive emotions* | 86 | 1.65 (.68) | 96 | 1.63 (.74) | .25 | .806 |
| *Negative emotions* | 86 | 1.39 (.67) | 96 | 1.38 (.71) | .09 | .926 |
| Emotion attribution | | | | | | |
| *Verbal positive* | 86 | 1.21 (.80) | 96 | 1.23 (.82) | .21 | .835 |
| *Verbal negative* | 86 | .92 (.56) | 96 | .99 (.61) | .85 | .395 |
| *Visual positive* | 86 | 1.42 (.79) | 95 | 1.44 (.78) | .20 | .841 |
| *Visual negative* | 86 | 1.05 (.56) | 96 | 1.10 (.61) | .62 | .533 |

Supplementary Table 5. Sample size justification.

|  |  |
| --- | --- |
| Aspect | Explanation |
| Power analysis for the larger project | An a priori power analysis was conducted for the larger research project that embedded this study. It showed that to observe a medium-sized effect (effect size = .35, power = .80, alpha = .05), a total sample size of 216 children would be needed for analyses with four repeated measures and two groups. Note that this analysis was done for the larger project and based on a repeated measure ANOVA design. We used Linear Mixed Models in the current study because it better accounts for the dependency within the data and is robust in handling missing and unbalanced data. |
|  |  |
| Power analysis for the present study | We could not conduct an a priori power analysis for this study because the study used the existing data. Yet, to understand the sample size needed for detecting the group effect in mixed models, we referred to Green (1991), who provided a rule of thumb of N > 104 + number of independent variable (IV) for estimating sample size required for multiple linear regression. In addition, Harris (1985) suggested to have a minimum of 10 participants per IV for regression using six or more IVs. A sample size of 115 children was considered necessary (with 3 predicting variables and 8 interactions). This study used data from a sample of 182 children. |

Supplementary Table 6. Pearson’s correlations between the outcomes of the emotion-recognition tasks measured at Time 1 and the outcomes of the parent questionnaires measuring emotion recognition, emotion talk with parents, prosocial behaviors, and peer relationship at Time 1. The results showed that the behavioral tasks had good concurrent validity. Namely, better performances in the emotion-recognition tasks were related to better emotion recognition, better communication about emotions with parents, more prosocial behaviors, and more positive peer relationship as evaluated by parents.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **Emotion discrimination** | | **Emotion identification** | | **Emotion attribution** | | | |
|  |  |  |  |  |  | *Verbal condition* | | *Visual condition* | |
|  |  | *Positive vs. negative* | *Negative vs. negative* | *Positive* | *Negative* | *Positive* | *Negative* | *Positive* | *Negative* |
| **EEQ1** | *Emotion Recognition* | .26\*\* | .18\* | .32\*\*\* | .23\*\* | .26\*\* | .26\*\* | .26\*\* | .24\*\* |
|  | *Emotion Talk* | .29\*\*\* | .30\*\*\* | .37\*\*\* | .35\*\*\* | .22\*\* | .24\*\* | .25\*\* | .28\*\* |
| **EmQue2** | *Prosocial Action* | .16† | .24\*\* | .20\* | .22\*\* | .18\* | .16† | .21\* | .16† |
| **SDQ3** | *Prosocial Behavior* | .23\*\* | .27\*\* | .29\*\*\* | .26\*\* | .22\*\* | .19\* | .30\*\*\* | .23\*\* |
|  | *Peer Relationship* | .18\* | .14 | .24\*\* | .25\*\* | .12 | .08 | .18\* | .12 |

1 The Emotion Expression Questionnaire (EEQ; Rieffe et al., 2010) is a 35-item parent questionnaire that measures children’s emotion expressions. To check the concurrent validity of the behavioral tasks, we used two scales: (1) the Emotion Recognition Scale asks parents to rate the extent to which children can recognize emotions in their parents (6 items; e.g., “Does your child know when you are happy?”); (2) the Emotion Talk Scale asks parents to rate how well they could talk with their children during or after an emotion episode (5 items; e.g., “Can you talk with your child about his/her emotional reaction during or after the reaction?”). Parents rated each item on a 5-point scale ranging from “1 = (almost) never” to “5 = (almost) always”.

2 The Empathy Questionnaire (EmQue; Rieffe et al., 2010) asks parents to evaluate the extent to which their children showed empathic reactions when seeing another person in distress over the past two months. Parents rated each item on a three-point scale: 0 = not at all applicable; 1 = a little or sometimes applicable; 2 = clearly or often applicable. To check the concurrent validity of the behavioral tasks, we used the Prosocial Action Scale (6 items; e.g., “When another child starts to cry, my child tries to comfort him/her”).

3 The Strengths and Difficulties Questionnaire (SDQ, Goodman, 1997) is a brief behavior screening questionnaire for completion by parents or teachers. We used the parent-report version. For the validity check, we used the Peer Relationship Scale (5 items; e.g., “Generally liked by other children”), and the Prosocial Behavior Scale (5 items; e.g., “Considerate of other people’s feelings”).

† *p* ⩽.06, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001

Supplementary Table 7. Cronbach’s alphas of the emotion-recognition tasks in the whole sample, non-autistic, and autistic sample.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Total sample** | | **Non-autistic** | | **Autistic** | |
|  | **N** | **Cronbach’s ⍺** | **N** | **Cronbach’s ⍺** | **N** | **Cronbach’s ⍺** |
| **Emotion discrimination** (4 items) | 180 | .81 | 121 | .75 | 59 | .89 |
| **Emotion identification** (4 items) | 182 | .82 | 121 | .72 | 61 | .89 |
| **Emotion attribution verbal** (8 items) | 176 | .89 | 115 | .84 | 61 | .94 |
| **Emotion attribution visual** (8 items) | 172 | .87 | 114 | .79 | 58 | .93 |

Supplementary Table 8. Eight vignettes depicting emotion-provoking situations in the emotion attribution task.

|  |
| --- |
| Vignette content |
| 1. The boy is building a tower; someone knocks it down. 2. The boy receives an ice cream. 3. Someone is pulling at the boy’s shirt. 4. The boy falls off from the bicycle. 5. The boy receives a present. 6. The boy sees a dog. 7. The spade of the boy is broken. 8. The boy sees a crocodile. |

Supplementary Table 9. Model fit indices of the best fitting models for predicting performances in the emotion-recognition tasks.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Emotion discrimination** | | | | |
|  | **Positive vs. negative** | | | |
|  | AIC | BIC | -2LL | *χ2* statistics |
| ***Null model*** | 1134.07 | 1146.44 | 1128.07 | - |
| ***Best age model*:** age (linear; fixed & random), group | 1011.56 | 1040.27 | 997.56 | *χ2* (4) = 130.57, *p<*.001 |  |
|  | **Negative vs. negative** | |  | |
|  | AIC | BIC | -2LL | *χ2* statistics |
| ***Null model*** | 1235.20 | 1247.57 | 1229.20 | - |
| ***Best age model:*** age (linear; fixed), group | 1111.54 | 1132.05 | 1101.54 | *χ2* (2) = 127.67, *p<*.001 |
| **Emotion identification** | | | | |
|  | **Positive emotion** | | |  |
|  | AIC | BIC | -2LL | *χ2*statistics |
| ***Null model*** | 702.87 | 715.26 | 696.87 |  |
| ***Best age model:*** age (linear; fixed & random), group, age x group | 538.30 | 517.14 | 522.30 | *χ2* (5) = 174.57, *p<*.001 |
|  | **Negative emotion** | |  |  |
|  | AIC | BIC | -2LL | *χ2*statistics |
| ***Null model*** | 762.21 | 774.60 | 756.21 |  |
| ***Best age model:*** age (linear; fixed & random), group, age x group | 558.33 | 591.17 | 542.33 | *χ2* (5) = 213.88, *p<*.001 |
| **Emotion attribution** | | | | |
|  | ***Positive emotions (Verbal)*** | | | |
|  | AIC | BIC | -2LL | *χ2*statistics |
| ***Null model*** | 770.95 | 783.34 | 764.95 |  |
| ***Best age model:*** age (linear; fixed) | 752.69 | 769.11 | 744.69 | *χ2* (1) = 20.26,  *p <* .001 |
|  | ***Positive emotions (Visual)*** | | | |
|  | AIC | BIC | -2LL | *χ2*statistics |
| ***Null model*** | 885.84 | 898.20 | 879.84 |  |
| ***Best age model:*** age (linear; fixed & random), group | 776.72 | 805.39 | 762.72 | *χ2* (4) = 117.12, *p <* .001 |
|  | ***Negative emotions (Verbal)*** | | | |
|  | AIC | BIC | -2LL | *χ2*statistics |
| ***Null model*** | 532.68 | 545.04 | 526.68 |  |
| ***Best age model:*** age (linear; fixed) | 515.92 | 532.30 | 507.92 | *χ2* (1) = 18.76,  *p <* .001 |
|  | ***Negative emotions (Visual)*** | | | |
|  | AIC | BIC | -2LL | *χ2*statistics |
| ***Null model*** | 679.11 | 691.44 | 673.11 |  |
| ***Best age model:*** age (linear; fixed & random) | 607.56 | 632.07 | 595.56 | *χ2* (3) = 77.55, *p <* .001 |

NOTE. Models removed during the formal model-fitting procedures were not presented here. The χ2 statistics present the comparisons of the -2LL values between the best fitting models and the null models. *P*-values smaller than .05 indicate a better model fit of the best fitting models than the null models.

Supplementary Table 10. Model fit indices of the predicting models with the means score of SRS as the predictor on emotion recognition in autistic children.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Emotion discrimination** | | | | |
| **Positive vs. negative** | | | | |
|  | AIC | BIC | -2LL | *χ2* statistics |
| *Age-only model* | 481.50 | 500.76 | 469.50 | - |
| *Model with SRS mean* | 455.94 | 478.02 | 441.94 | *χ2*(1) = 27.56,  *p <* .001 |  |
|  |  |  | **Negative vs. negative** | |  |
|  | AIC | BIC | -2LL | *χ2* statistics |  |
| *Age-only model* | 478.63 | 491.47 | 470.63 | - |  |
| *Model with SRS mean* | 446.58 | 462.35 | 436.58 | *χ2*(1) = 34.05,  *p <* .001 |  |
| **Emotion identification** | | | | |
| **Positive** | | | | |
|  | AIC | BIC | -2LL | *χ2*statistics |
| *Age-only model* | 265.27 | 284.56 | 253.27 | - |
| *Model with SRS mean* | 241.66 | 263.77 | 227.66 | *χ2*(1) = 25.61,  *p <* .001 |
|  |  |  |  | **Negative** |
|  | AIC | BIC | -2LL | *χ2*statistics |
| *Age-only model* | 297.11 | 316.40 | 285.11 | - |
| *Model with SRS mean* | 276.09 | 298.21 | 262.09 | *χ2*(1) = 23.02,  *p <* .001 |
| **Emotion attribution – positive emotions** | | | | |
| **Verbal condition** | | | | |
|  | AIC | BIC | -2LL | *χ2*statistics |
| *Age-only model* | 340.48 | 353.34 | 332.48 | - |
| *Model with SRS mean* | 302.68 | 318.47 | 292.68 | *χ2*(1) = 39.80,  *p <* .001 |
|  |  |  |  | **Visual condition** |
|  | AIC | BIC | -2LL | *χ2*statistics |
| *Age-only model* | 366.63 | 385.92 | 354.63 | - |
| *Model with SRS mean* | 343.05 | 365.17 | 329.05 | *χ2*(1) = 25.58,  *p <* .001 |
| ***Emotion attribution - negative emotions*** | | | | |
|  |  | **Verbal condition** | | |
|  | AIC | BIC | -2LL | *χ2*statistics |
| *Age-only model* | 237.87 | 250.73 | 229.87 | - |
| *Model with SRS mean* | 208.35 | 224.14 | 198.35 | *χ2*(1) = 31.52,  *p <* .001 |
|  |  |  |  | **Visual condition** |
|  | AIC | BIC | -2LL | *χ2*statistics |
| *Age-only model* | 314.13 | 333.38 | 302.13 | - |
| *Model with SRS mean* | 294.36 | 316.43 | 280.36 | *χ2*(1) = 21.77,  *p <* .001 |

NOTE. The χ2 statistics present the comparisons of the -2LL values of the age-only models and the models with SRS mean and change scores as predictors. *P*-values smaller than .05 indicate a better model fit of the models with SRS mean than the age-only model.

Supplementary Table 11. Fixed and random effects of the best fitting models for predicting performances in the practice tasks for emotion discrimination.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***Cars vs. Flowers*** | | | | ***Faces with a hat vs. Faces with glasses*** | | | |
| **Fixed effects** | **Estimates** | ***SE*** | ***CI* [low, high]** |  | **Estimates** | ***SE*** | ***CI* [low, high]** |  |
| *Intercept* | 2.78 | .07 | [2.65, 2.92] |  | 2.36 | .09 | [2.19, 2.54] |  |
| *Age* | .004 | .001 | [.002, .007] |  | .001 | .002 | [.007, .01] |  |
| *Group* | -.12 | .05 | [-.23, -.02] |  | -.20 | .07 | [-.35, -.06] |  |
| **Random effects** | **Estimates** | ***SE*** | ***CI* [low, high]** | **Wald’s *Z*** | **Estimates** | ***SE*** | ***CI* [low, high]** | **Wald’s *Z*** |
| *Residual* | .09 | .01 | [.07, .11] | 10.01 | .21 | .02 | [.18, .25] | 10.59 |
| *Intercept* | .49 | .08 | [.36, .67] | 6.44 | .63 | .13 | [.42, .93] | 4.95 |
| *Age* | .0001 | .00003 | [.0001, 0002] | 4.60 | .0001 | .00006 | [.00006, .0003] | 2.36 |

Supplementary Table 12. Coefficients SRS baseline for predicting emotion-recognition abilities of autistic children.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **SRS baseline** |  |  |  |
|  | **Coefficient** | **Standard error** | ***CI* [low, high]** | ***p-*value** |
| **Emotion discrimination** | | | | |
| *Positive vs. negative* | -.007 | .003 | [-.01, -.0005] | .035 |
| *Negative vs. negative* | -.005 | .004 | [-.01, .003] | .190 |
| **Emotion identification** | | | | |
| *Positive emotion* | -.0006 | .003 | [-.006, .005] | .821 |
| *Negative emotion* | -.005 | .003 | [-.01, .003] | .064 |
| **Emotion attribution** | | | | |
| *Positive emotion verbal* | -.002 | .005 | [-.01, .01] | .704 |
| *Positive emotion visual* | -.001 | .003 | [-.007, .005] | .697 |
| *Negative emotion verbal* | -.003 | .003 | [-.009, .002] | .248 |
| *Negative emotion visual* | -.005 | .003 | [-.01, .005] | .077 |

Supplementary Table 13. The means (standard deviations) of the three language indices and the *t*-test statistics of group comparisons.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Autistic** |  | **Non-autistic** | | *t*-test statistics | |
|  | N | Mean (SD) | N | Mean (SD) | *t* | *p* |
| **Child Development Inventorya (Time 1)** |  |  |  |  |  |  |
| Language expression | 53 | .78 (.25) | 90 | .95 (.10) | 4.64 | <.001 |
| Language comprehension | 50 | .73 (.25) | 89 | .93 (.12) | 5.30 | <.001 |
| **Emotion Vocabularyb** |  |  |  |  |  |  |
| Time 1 | 56 | 1.22 (.51) | 95 | 1.66 (.26) | 6.09 | <.001 |
| Time 2 | 47 | 1.31 (.51) | 46 | 1.80 (.14) | 6.44 | <.001 |
| Time 3 | 41 | 1.53 (.42) | 40 | 1.88 (.12) | 5.20 | <.001 |
| Time 4 | 28 | 1.67 (.51) | 31 | 1.93 (.10) | 2.59 | .015 |

a The Child Development Inventory (CDI; Ireton & Glascoe, 1995) assesses the developmental levels of children aged 1 to 6 years. We used parents’ ratings on two scales as indices of children’s language skills: the Language Expression Scale (e.g., “(My child) poses questions with the words ‘why’, ‘when’, or ‘how’”) and the Language Comprehension Scale (e.g., “(My child) understands what ‘upwards’ and ‘downwards’ mean”). Each scale consists of 50 items and parents were asked to indicate whether this does or does not apply to their child (0 = no, 1 = yes). The data of the two scales were collected only at Time 1.

b The Emotion Vocabulary Questionnaire (EV; Ketelaar et al., 2015) asks parents to rate the extent to which their children know and use 20 words that refer to emotional and mental states (e.g., happiness, disappointed, thinking). Parents give the score “0” when their children do not know or use the word, “1” when the parents are not sure if their children are simply repeating the word or understand the word, and “2” when their children know and use the word. The data of EV were collected at four time points.

Supplementary Table 14. Fixed and random effects of the models with the language variables added for predicting *emotion discrimination*.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***Positive vs. Negative*** | | | | ***Negative vs. Negative*** | | | |
| **Fixed effects** | **Estimates** | ***SE*** | ***CI* [low, high]** |  | **Estimates** | ***SE*** | ***CI* [low, high]** |  |
| *Intercept* | .60\* | .28 | [1.04, 1.15] |  | .35 | .30 | [-.24, .93] |  |
| *Age* | .01\*\* | .003 | [.01, .02] |  | .02\*\* | .003 | [.01, .03] |  |
| *Group* | .04 | .11 | -.12, .25] |  | .09 | .12 | [-.15, .34] |  |
| *Emotion vocabulary* | .29 | .16 | [-.03, .62] |  | .41\* | .18 | [.06, .76] |  |
| *Language expressa* | .49 | .61 | [-.73, 1.72] |  | -.48 | .66 | [-1.79, .83] |  |
| *Language compreb* | .44 | .56 | [-.68, 1.57] |  | .92 | .62 | [-.31, 2.14] |  |
| **Random effects** | **Estimates** | ***SE*** | ***CI* [low, high]** | **Wald’s *Z*** | **Estimates** | ***SE*** | ***CI* [low, high]** | **Wald’s *Z*** |
| *Residual* | .43\*\* | .04 | [.34, .54] | 8.43 | .49\*\* | .05 | [.40, .60] | 9.60 |
| *Intercept* | .45 | .19 | [.16, 1.24] | 1.93 | .12\* | .05 | [.05, .27] | 2.46 |
| *Age* | .0001 | .0001 | [.00001, 0001] | .85 | - | - | - | - |

a Language expression; b Language comprehension

\**p* <.05; \*\* *p* < .001

Supplementary Table 15. Fixed and random effects of the models with language variables added for predicting *emotion identification*.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***Positive emotions*** | | | | ***Negative emotions*** | | | |
| **Fixed effects** | **Estimates** | ***SE*** | ***CI* [low, high]** |  | **Estimates** | ***SE*** | ***CI* [low, high]** |  |
| *Intercept* | .56\* | .19 | [.19, .93] |  | .31 | .19 | [-.08, .69] |  |
| *Age* | .003 | .002 | [-.001, .007] |  | .01\*\* | .002 | [.005, .01] |  |
| *Group* | -.16 | .16 | [-.48, .17] |  | -.30 | .16 | [-.61, .01] |  |
| *Age x group* | .005 | .003 | [-.001, .01] |  | .007\* | .003 | [.001, .01] |  |
| *Emotion vocabulary* | .23\* | .09 | [.04, .41] |  | .23\* | .10 | [.03, .44] |  |
| *Language expressa* | 1.04\*\* | .36 | [.31, 1.77] |  | .81\* | .39 | [.007, 1.61] |  |
| *Language compreb* | -.22 | .33 | [-.88, .44] |  | -.09 | .36 | [-.83, .65] |  |
| **Random effects** | **Estimates** | ***SE*** | ***CI* [low, high]** | **Wald’s *Z*** | **Estimates** | ***SE*** | ***CI* [low, high]** | **Wald’s *Z*** |
| *Residual* | .11\*\* | .01 | [.09, .14] | 9.52 | .14\*\* | .01 | [.11, .17] | 9.32 |
| *Intercept* | .40\*\* | .02 | [.32, .67] | 6.25 | .27\*\* | .08 | [.15, .48] | 3.37 |

a Language expression; b Language comprehension

\**p* <.05; \*\* *p* < .001

Supplementary Table 16. Fixed and random effects of the models with language variables added for predicting *emotion attribution*.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Positive emotions** | ***Verbal condition*** | | | | ***Visual condition*** | | | |
| **Fixed effects** | **Estimates** | ***SE*** | ***CI* [low, high]** |  | **Estimates** | ***SE*** | ***CI* [low, high]** |  |
| *Intercept* | -.07 | .18 | [-.42, .29] |  | -.04 | .23 | [-.49, .41] |  |
| *Age* | .006\* | .002 | [.002, .01] |  | .01\* | .002 | [.003, .01] |  |
| *Group* | - | - | - |  | .11 | .08 | [-.05, .28] |  |
| *Emotion vocabulary* | -.12 | .11 | [-.35, .10] |  | .11 | .13 | [-.14, .35] |  |
| *Language expressa* | 2.70\*\* | .50 | [1.72, 3.69] |  | 1.53\* | .49 | [.56, 2.50] |  |
| *Language compreb* | -.88 | .46 | [-1.78, .03] |  | -.24 | .45 | [-1.12, .65] |  |
| **Random effects** | **Estimates** | ***SE*** | ***CI* [low, high]** | **Wald’s *Z*** | **Estimates** | ***SE*** | ***CI* [low, high]** | **Wald’s *Z*** |
| *Residual* | .16\*\* | .02 | [.13, .20] | 9.05 | .19\*\* | .03 | [.14, .25] | 6.77 |
| *Intercept* | .12\*\* | .03 | [.08, .20] | 4.10 | .59\* | .17 | [.34, 1.04] | 3.48 |
| *Slope* | - | - | - | - | .0002 | .0001 | [.0001, .0004] | 2.34 |
| **Negative emotions** | ***Verbal condition*** | | | | ***Visual condition*** | | | |
| **Fixed effects** | **Estimates** | ***SE*** | ***CI* [low, high]** |  | **Estimates** | ***SE*** | ***CI* [low, high]** |  |
| *Intercept* | .19 | .17 | [-.15, .54] |  | .01 | .14 | [-.27, .30] |  |
| *Age* | .003\* | .001 | [.001, .007] |  | .01\*\* | .002 | [.005, .01] |  |
| *Emotion vocabulary* | -.04 | .10 | [-.22, .15] |  | .05 | .10 | [-.15, .26] |  |
| *Language expressa* | 1.66\* | .47 | [.73, 2.59] |  | 1.12\* | .39 | [.35, 1.89] |  |
| *Language compreb* | -.62 | .44 | [-1.48, .24] |  | -.15 | .36 | [-.86, .55] |  |
| **Random effects** | **Estimates** | ***SE*** | ***CI* [low, high]** | **Wald’s *Z*** | **Estimates** | ***SE*** | ***CI* [low, high]** | **Wald’s *Z*** |
| *Residual* | .09\*\* | .01 | [.07, .11] | 9.47 | .14\*\* | .02 | [.11, .19] | 7.08 |
| *Intercept* | .14\*\* | .02 | [.10, .20] | 5.68 | .16 | .10 | [.05, .51] | 1.72 |
| *Slope* | - | - | - | - | .0001\* | .0001 | [.00005, .0003] | 1.96 |

a Language expression; b Language comprehension

\**p* <.05; \*\* *p* < .001

References:

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586. <https://doi.org/10.1111/j.1469-7610.1997.tb01545.x>

Green, S. B. (1991). How many subjects does it take to do a regression analysis. *Multivariate Behavioral Research*, *26*(3), 499-510. <https://doi.org/10.1207/s15327906mbr2603_7>

Harris, P. (1985). Testing for variance homogeneity of correlated variables. *Biometrika*, *72*(1), 103-107. <https://doi.org/10.1093/biomet/72.1.103>

Ireton, H., & Glascoe, F. P. (1995). Assessing children’s development using parents’ reports. The Child Development Inventory Clin- ical Pediatrics, 34(5), 248–255. https://doi.org/10.1177/000992289 503400504

Ketelaar, L., Rieffe, C., Wiefferink, C. H., & Frijns, J. H. M. (2013). Social competence and empathy in young children with cochlear implants and with normal hearing. *The Laryngoscope*, *123*, 518–523. <https://doi.org/10.1002/lary.23544>

Rieffe, C., Ketelaar, L., & Wiefferink, C. H. (2010). Assessing empathy in young children: Construction and validation of an Empathy Questionnaire (EmQue). *Personality and Individual Differences*, *49*(5), 362–367. <https://doi.org/10.1016/j.paid.2010.03.046>