*behavioral sciences*

MDPI

*Article*

# Social Cognition in Bilingual and Monolingual Adolescents: Task-Specific Patterns in Emotion Recognition and Theory of Mind

**Tugce Turan Kaya [1], Mutlu Muhammed Ozbek [2], Esra Sizer [3] and Remzi Ogulcan Ciray [4,*]**

[1] Child and Adolescent Psychiatry, Merzifon Kara Mustafa Pasa State Hospital, Merzifon 05300, Amasya, Türkiye; drtugceturan@gmail.com
[2] Department of Child and Adolescent Psychiatry, Faculty of Medicine, Yalova University, Cinarcik 77200, Yalova, Türkiye; mutlu.ozbek@yalova.edu.tr
[3] Department of Child and Adolescent Psychiatry, Faculty of Medicine, Mardin Artuklu University, Artuklu 47200, Mardin, Türkiye; esrasizer@artuklu.edu.tr
[4] Department of Child and Adolescent Psychiatry, Faculty of Medicine, Dokuz Eylul University, Balcova 35340, Izmir, Türkiye
* Correspondence: ogulcan.ciray@deu.edu.tr

**Abstract**

The extent to which bilingual–monolingual differences vary across social-cognitive domains during adolescence remains unclear. This cross-sectional study aimed to compare mental-state inference, facial emotion recognition, and faux pas understanding in 60 bilingual and 60 monolingual adolescents aged 12–18 years from Türkiye. Bilingualism was defined as mean use of 30–70% for each language across five Bilingual Language Profile contexts. Monolingual participants used only Turkish. Assessments included the Reading the Mind in the Eyes Test (RMET), a facial emotion recognition task, and the Faux Pas Test. Bilingual adolescents scored lower on the RMET (Cohen's $d = -1.11$) and facial emotion recognition ($d = -1.18$), with no significant difference in Faux Pas scores. These differences remained significant after adjustment for age, sex, intelligence quotient, household income, and maternal education. Emotion-specific analyses showed lower scores in the bilingual group for sadness, anger, disgust, and surprise, higher scores for fear, and no difference for happiness. The group-by-task interaction was significant (Wald $\chi^2$ (2) = 22.43, $p < 0.001$). Correlations were small and nonsignificant after controlling for language group. The findings suggest limited convergence and task-dependent differences in performance within this sample. Their linguistic, cultural, and social-cognitive explanations remain unresolved.

**Keywords:** bilingualism; social cognition; theory of mind; facial emotion recognition; reading the mind in the eyes test; faux pas test; neuropsychological assessment; adolescence

Check for updates

Academic Editor: Michele Roccella

Received: 27 September 2026
Revised: 26 October 2026
Accepted: 28 October 2026
Published: 24 November 2026

**Copyright:** © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

## 1. Introduction

Social cognition encompasses multiple cognitive and affective processes involved in understanding others' mental states, emotions, and beliefs (Schurz et al., 2021). A central component of social cognition is theory of mind (ToM), which refers to the capacity to attribute mental states, such as beliefs and intentions, to others and to use these inferences to understand and predict their behavior (Premack & Woodruff, 1978).

Bilingualism refers broadly to the use of two languages in daily life, without implying equal proficiency in both (Birdsong, 2014; Feng et al., 2024). The bilingual advantage hypothesis proposes that bilingual children may outperform monolingual children on ToM

tasks (Schroeder, 2018). Sociolinguistic accounts suggest that adapting language to others' knowledge and perspectives may contribute to this association. Metalinguistic accounts emphasize recognizing that language, like beliefs, may misrepresent reality, while executive-function accounts emphasize controlling competing languages and setting aside one's own knowledge. Empirical support for these explanations remains partial (Yu et al., 2021). Thus, the relationship between bilingual experience and social-cognitive performance may vary across tasks with distinct vocabulary, emotion-label, and narrative demands (Feng et al., 2024; Griffiths et al., 2020; Kittel et al., 2022; Yu et al., 2021).

Bilingual children may differ in proficiency, language dominance, age of acquisition, and language use across home and school contexts (Birdsong, 2014; Feng et al., 2024). Age of acquisition and patterns of exposure may shape which language becomes dominant and the proficiency attained in each language (Birdsong, 2014; Unsworth et al., 2018). These differences may affect understanding of the mental-state vocabulary and verbal information required by social-cognitive assessments (Buac & Kaushanskaya, 2020; Grazzani & Ornaghi, 2012; Kittel et al., 2022).

Findings may vary according to how bilingualism is defined, the characteristics of the samples, and the social-cognitive tasks used (Cox et al., 2016; Feng et al., 2024). In particular, the literature has relied heavily on false-belief tasks, with relatively limited attention to social and emotional aspects of ToM (Feng et al., 2024). The extent of this measurement diversity is illustrated by a scoping review that identified 96 different ToM tasks across 53 publications (Feng et al., 2024). Different ToM tasks have shown weak associations with one another, suggesting that they may assess partly distinct abilities (Warnell & Redcay, 2019). The Reading the Mind in the Eyes Test (RMET) is commonly used to assess affective ToM and the social-perceptual component of ToM (Meinhardt-Injac et al., 2020; Rosso & Riolfo, 2020). Its performance is associated with both emotion perception and vocabulary, suggesting that it involves perceptual and language-related processes (Kittel et al., 2022). Facial emotion recognition represents another important component of social cognition and involves identifying emotions from facial cues (Griffiths et al., 2020; Kessels et al., 2014). It is treated here as an ability related to ToM, rather than as a direct measure of reasoning about beliefs or knowledge. Unlike these visually based measures, the Faux Pas Test assesses social-cognitive ToM through story-based scenarios and requires narrative comprehension, mental-state attribution, and pragmatic judgment (Baron-Cohen et al., 1999; Meinhardt-Injac et al., 2020). Together, these measures allow group performance to be compared across tasks emphasizing mental-state decoding, basic emotion recognition, and narrative social reasoning.

Evidence regarding bilingualism across these domains remains limited and mixed. Monolingual and bilingual preschool children did not differ significantly in facial emotion identification (Champoux-Larsson et al., 2019), whereas Turkish monolingual children outperformed sequential bilingual peers on both the RMET and Faux Pas measures (Asan & Yıldız, 2026). These studies differ in age, language background, and task design, limiting direct comparisons. Examining multiple measures within the same adolescent sample is therefore relevant, particularly as mental-state decoding and social reasoning continue to develop during this period (Meinhardt-Injac et al., 2020).

Bilingual experience occurs within broader family, educational, socioeconomic, and cultural contexts that may also be relevant to social-cognitive performance (Asan & Yıldız, 2026; Calvo & Bialystok, 2014; Selcuk et al., 2023). Family discussions of beliefs and emotions and access to educational resources are associated with ToM development (Selcuk et al., 2023). Educational experiences may influence familiarity with the assessment language (Asan & Yıldız, 2026), while cultural familiarity may affect the interpretation of facial expressions and social situations (Elfenbein & Ambady, 2002; Selcuk et al., 2023). Studies

of Turkish–Arabic and Turkish–Kurdish bilingual communities suggest that language-use patterns may differ between home and school settings and vary across local sociocultural contexts (Çetintaş Yıldırım, 2020; Guldenoglu et al., 2026). Socioeconomic status and bilingualism have been independently associated with children's language and cognitive performance, suggesting that observed differences between language groups should not be attributed to bilingualism alone (Calvo & Bialystok, 2014).

Taken together, it remains unclear whether bilingual–monolingual differences are consistent across social-cognitive domains or vary by task, particularly during adolescence. Evidence is also limited from studies examining multiple domains within the same sample while accounting for relevant individual and socioeconomic factors. The present study compares clinically screened bilingual and monolingual adolescents in Türkiye across three complementary social-cognitive outcomes: mental-state inference from eye-region cues, facial emotion recognition, and faux pas understanding. The primary aim is to determine whether the groups differ in their overall social-cognitive profiles after accounting for age, sex, general intellectual ability, household income, and maternal education. We expect the magnitude of group differences to vary across tasks rather than show a uniform bilingual advantage or disadvantage. Secondary analyses compare group performance for individual facial emotions. Exploratory correlations describe associations among tasks before and after accounting for language group.

## 2. Materials and Methods

### *2.1. Participants and Procedure*

This cross-sectional study included 120 adolescents aged 12–18 years who were recruited from Child and Adolescent Psychiatry outpatient clinics at two state hospitals located in eastern and southeastern Türkiye between January and June 2021. The sample comprised 60 bilingual and 60 monolingual participants. Each site contributed 30 bilingual and 30 monolingual participants. The groups were comparable in terms of age, sex, IQ, and socioeconomic status.

For the purposes of this study, bilingualism was defined as a mean language-use distribution ranging from 30% to 70% for each language across the five contexts of the Bilingual Language Profile (BLP) Language Use module (see Section 2.2). The criterion was established during study design, before participant recruitment. Eligibility and language-group classification were determined using the BLP Language Use module before the social-cognitive tasks were administered. Participants meeting this criterion for Turkish–Arabic or Turkish–Kurdish language use were classified as bilingual. Participants who reported speaking only Turkish were classified as monolingual. Among the 71 participants who self-identified as bilingual and consented to participate, 11 did not meet the predefined BLP language-use criterion for bilingual classification and were therefore excluded from the study. These participants were not reassigned to the monolingual group. Accordingly, the final bilingual sample consisted of 60 participants. The 11 excluded participants were not included in the analyses.

All participants were assessed by a child and adolescent psychiatrist using the Turkish version of the Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Lifetime Version, DSM-5 (K-SADS-PL DSM-5). Participants who met the diagnostic criteria for any psychiatric disorder were excluded. Presentations related to situational concerns, subthreshold symptoms, or developmental and adolescent-related difficulties were not considered psychiatric disorders unless diagnostic criteria were met. Adolescents with a serious neurological or systemic disorder, an illness severe enough to interfere with test administration, or difficulty complying with the assessment procedures were excluded. General intellectual ability was assessed using the Turkish version of the

Wechsler Intelligence Scale for Children–Revised (WISC-R) for participants aged 16 years and 11 months or younger (Savaşır & Şahin, 1995). Participants who exceeded the WISC-R age limit completed the Wechsler Adult Intelligence Scale–Revised (WAIS-R) (Sezgin et al., 2014). The WISC-R was administered to 43 bilingual and 40 monolingual participants, and the WAIS-R to 17 bilingual and 20 monolingual participants. Total IQ scores were used as indicators of general intellectual ability in the analyses.

The three social-cognitive tasks were administered sequentially in Turkish during a single session by psychologists. At both sites, all participants completed the tasks using the same instructions and procedures, in the same fixed order: RMET, facial emotion recognition, and the Faux Pas Test. Participants were allowed to ask questions about the instructions and unfamiliar words, which were explained by the examiner. For the RMET, unfamiliar response-option terms were clarified using the test glossary. Before the facial emotion recognition task, participants' understanding of the six emotion labels was checked. For the Faux Pas Test, clarifications were limited to unfamiliar words in the instructions or stories. Explanations were limited to word meanings, without interpreting the social situation. Requests for clarification were not recorded. The psychologists were not informed of participants' language-group classification, although language background may have become apparent during testing. There were no omitted or invalid responses to the administered questions.

### *2.2. Measures*

#### 2.2.1. Sociodemographic Form

Sociodemographic information, including age, sex, current educational level, household income, and maternal education, was collected using a form prepared for this study. Monthly household income was categorized as low, middle, or high, with cutoffs of USD 1000 and USD 2000. Maternal education was recorded according to the highest educational level completed.

#### 2.2.2. Schedule for Affective Disorders and Schizophrenia for School-Age Children (6–18 Years)—Present and Lifetime Version, DSM-5 (K-SADS-PL-DSM-5)

This semi-structured interview was used to identify current and lifetime psychiatric disorders in children and adolescents (Kaufman et al., 1997). Diagnostic evaluation integrates information obtained through interviews with the child or adolescent and their caregiver, along with other available sources. The Turkish validity and reliability of the scale have been established (Ünal et al., 2019).

#### 2.2.3. Bilingual Language Profile (BLP)

The instrument comprises 19 self-report items covering four domains of language experience: language history, language use, language proficiency, and language attitudes (Birdsong et al., 2012). The BLP provides a continuous measure of language dominance, and its modules can also be used separately. For the present study, the Turkish translation of the five-item Language Use module available through the official BLP website was used. The module assesses the percentage of each language used with friends, with family, at school or work, during self-talk, and while counting. Reported percentages total 100% within each context.

The 30–70% range was used to identify adolescents reporting substantial use of both languages for comparison with adolescents reporting exclusive use of Turkish. Similar minimum exposure thresholds have been reported in previous bilingual research (Feng et al., 2024; Nguyen & Astington, 2014). The criterion was applied to mean language use across the five contexts. It reflected language use rather than proficiency in either language.

Proficiency and vocabulary knowledge in both languages were not directly assessed. Age of acquisition and simultaneous or sequential bilingual acquisition were not recorded.

#### 2.2.4. Reading the Mind in the Eyes Test (RMET)

This task assesses the ability to infer complex mental states from cues restricted to the eye region (Baron-Cohen et al., 2001). The present study used the 36-item revised version, which has also been used in adolescent samples (Palmer et al., 2024). Each item presents a photograph of the eye region accompanied by four mental-state terms, from which participants select the term that best describes what the depicted person is thinking or feeling. Instructions were delivered orally. The response options were presented in writing and read aloud by the examiner. Correct responses received one point, and incorrect responses received zero points. Total scores range from 0 to 36, with higher scores indicating better mental-state inference (Baron-Cohen et al., 2001). The items can also be grouped into positive, negative, and neutral valence (Harkness et al., 2005). Psychometric studies have questioned these groupings because of poor internal consistency and evidence that item valence varies along a continuum rather than discrete categories (Hudson et al., 2020; Olderbak et al., 2015). Therefore, only the total RMET score was included in the analyses. The Turkish version has demonstrated adequate validity and reliability (Cronbach's $\alpha = 0.71$; Girli, 2014). In the present study, Cronbach's $\alpha$ was 0.75.

#### 2.2.5. Facial Emotion Recognition Task

This task assesses recognition of six basic emotions using photographs from the Pictures of Facial Affect series (Ekman & Friesen, 1976). It comprises 60 photographs, with 10 depicting each of six emotions: happiness, sadness, anger, fear, disgust, and surprise. Instructions were delivered orally, and the six emotion labels remained visible on the screen throughout the task. Correct responses received one point, and incorrect responses received zero points. Each emotion score represents the number of correctly identified expressions and ranges from 0 to 10. Total scores range from 0 to 60, with higher scores indicating better facial emotion recognition (Young et al., 2002). Previous studies have reported Cronbach's alpha values ranging from 0.52 to 0.81 for the Ekman 60-Faces Test (Assendelft et al., 2026). In the present study, Cronbach's $\alpha$ was 0.77.

#### 2.2.6. Faux Pas Test

The Faux Pas Test assesses advanced ToM through socially inappropriate situations (Baron-Cohen et al., 1999). A Turkish version adapted from the adult Faux Pas Recognition Test (Stone et al., 1998) was administered (Yazgeç et al., 2023). The same form was administered to all participants, providing a common set of stories and scoring procedures across the age range. The version used in this study comprises 10 stories, each followed by five scored questions. Stories without a faux pas were not administered. The first two assess faux pas recognition by asking whether anyone said something inappropriate or awkward and who said it. The remaining three assess understanding of the inappropriate remark, false-belief reasoning, and empathy, respectively. Each story also includes two comprehension control questions, which do not contribute points to the total score.

The examiner read each story aloud while participants followed a printed copy. Each correct response received one point, and each incorrect response received zero points. If the participant answered "no" to the first recognition question, the remaining scored questions were skipped and scored zero. Following a "yes" response, all subsequent scored questions were administered. If the speaker was incorrectly identified in the second question, that question and the remaining scored questions were scored zero. The examiner then administered the comprehension control questions for every study. Stories with an incor-

rect response to either comprehension control question were excluded from subsequent social-cognitive scoring. All participants answered comprehension control questions correctly. Each story was scored from 0 to 5, yielding a total score ranging from 0 to 50, with higher scores indicating better ToM performance. In the present study, the Faux Pas Test demonstrated acceptable internal consistency (Cronbach's $\alpha = 0.79$).

### 2.3. Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics, version 26.0 (IBM Corp., Armonk, NY, USA). Categorical variables were summarized as frequencies and percentages and continuous variables as means and standard deviations. A sensitivity power analysis for a two-sided independent-samples $t$ test indicated that, with a total sample size of 120 participants (60 per group), a two-sided $\alpha$ level of 0.05, and 80% statistical power, the study was adequately powered to detect between-group differences corresponding to approximately Cohen's $d = 0.51$ or larger.

Participant characteristics were compared between bilingual and monolingual groups using independent-samples $t$ tests or Pearson's chi-square tests, as appropriate. Unadjusted group differences in RMET, facial emotion recognition, and Faux Pas total scores were examined using independent-samples $t$ tests.

The primary analysis used multivariate analysis of covariance (MANCOVA), with RMET, facial emotion recognition, and Faux Pas total scores as dependent variables. Language group, sex, household income, and maternal education were entered as factors, and age and IQ as covariates. Age was included to account for developmental variation in social-cognitive performance during adolescence (Lawrence et al., 2015; Meinhardt-Injac et al., 2020). The primary models assumed an approximately linear association between age and each outcome, with a common age slope across language groups. Pillai's trace was used to evaluate multivariate effects. Adjusted group differences for each outcome were subsequently examined using linear regression with heteroskedasticity-consistent (HC3) robust standard errors. Monolingual participants were the reference group. Holm correction was applied across the three total-score comparisons. Adjusted unstandardized coefficients (B), HC3 standard errors, and Holm-adjusted $p$ values were reported. Unadjusted Cohen's $d$ values were also calculated, with 95% confidence intervals reported for $B$ and $d$.

Differences in group contrasts across tasks were examined using a generalized estimating equation (GEE) model with a Gaussian distribution, identity link, unstructured working correlation, and participant-clustered robust standard errors. Each total score was standardized using its mean and sample standard deviation across all 120 participants (Supplementary Table S1). Task was a three-level within-participant factor. The model included language group, task, and the group × task interaction, with adjustment for age, sex, IQ, household income, and maternal education. Interactions between task and each of these five adjustment variables were also included. The group × task interaction was evaluated using a Wald test with two degrees of freedom. Following a significant interaction, group contrasts within tasks and pairwise differences between these contrasts were estimated. Holm correction was applied separately to each set of three comparisons.

Secondary analyses compared the six facial emotion scores using the same HC3 models, with Holm correction across the six comparisons. Exploratory Pearson correlations and partial correlations controlling for language group were examined among the three total scores. The three pooled and three partial correlations were corrected separately using the Holm procedure, and 95% confidence intervals were reported.

Model residuals were examined using Q–Q and residual-versus-fitted plots (Supplementary Figures S1 and S2) and Shapiro–Wilk tests. Levene tests compared residual

variances between language groups. Group × age and group × IQ interactions were tested to assess whether covariate slopes differed between language groups. Two observations were flagged using an absolute externally studentized residual threshold > 3, one in the RMET model and one in the facial emotion recognition model (Supplementary Table S2). Both were retained in the primary analyses and excluded together from the MANCOVA, all three total-score HC3 models, and GEE in the exclusion sensitivity analysis ($n = 118$). Separate sensitivity analyses added a quadratic age term, allowed group-specific age and IQ slopes, or restricted the sample to participants assessed with the WISC-R (ages 12–16 years; $n = 83$). GEE models were also fitted with exchangeable and independent working correlations. Additional sensitivity analyses adjusted for recruitment site.

There were no missing data in the primary analyses. All tests were two-tailed, with statistical significance set at $p < 0.05$.

## 3. Results

### 3.1. Participant Characteristics

The final sample comprised 120 adolescents, including 60 bilingual and 60 monolingual participants. The mean age was 15.10 years ($SD = 2.00$), and 62 participants (51.7%) were male. The mean IQ score was 102.07 ($SD = 13.56$). Participant characteristics by language group are presented in Table 1. No statistically significant group differences were detected in age, $t\ (118) = -0.64$, $p = 0.525$, or IQ score, $t\ (118) = -1.27$, $p = 0.207$. Pearson's chi-square tests also showed no significant group differences in sex, $\chi^2\ (1) = 0.13$, $p = 0.715$; school stage, $\chi^2\ (1) = 0.03$, $p = 0.852$; household income, $\chi^2\ (2) = 2.88$, $p = 0.237$; or maternal education, $\chi^2\ (3) = 3.07$, $p = 0.381$. No significant differences in participant characteristics were detected between recruitment sites (all $p$s $\geq 0.05$; Supplementary Table S3).

**Table 1.** Participant characteristics by language group.

| Participant Characteristics | Bilingual ($n = 60$) | Monolingual ($n = 60$) | $p$ |
|---|---|---|---|
| Age, mean ($SD$) [a] | 14.98 (2.07) | 15.22 (1.94) | 0.525 |
| Sex, % ($n$) [b] | | | 0.715 |
| Male | 53.3 (32) | 50.0 (30) | |
| Female | 46.7 (28) | 50.0 (30) | |
| Education level, % ($n$) [b] | | | 0.852 |
| Middle school | 40.0 (24) | 38.3 (23) | |
| High school | 60.0 (36) | 61.7 (37) | |
| Household income, % ($n$) [b] | | | 0.237 |
| Low | 25.0 (15) | 28.3 (17) | |
| Middle | 43.3 (26) | 53.3 (32) | |
| High | 31.7 (19) | 18.3 (11) | |
| Maternal education, % ($n$) [b] | | | 0.381 |
| Primary school | 13.3 (8) | 21.7 (13) | |
| Middle school | 31.7 (19) | 21.7 (13) | |
| High school | 23.3 (14) | 30.0 (18) | |
| University | 31.7 (19) | 26.7 (16) | |
| IQ score, mean ($SD$) [a] | 100.50 (13.99) | 103.63 (13.05) | 0.207 |

Note. $SD$: standard deviation; IQ: intelligence quotient. [a] Independent-samples $t$ test. [b] Pearson's chi-square test.

### 3.2. Group Differences in Social-Cognitive Performance

In unadjusted comparisons, bilingual adolescents scored lower than monolingual adolescents on the RMET, $t\ (118) = -6.09$, $p < 0.001$, and facial emotion recognition, $t\ (118) = -6.44$, $p < 0.001$. No statistically significant group difference was detected for

Faux Pas scores, $t$ (118) = −0.17, $p$ = 0.864. The primary MANCOVA showed a significant multivariate effect of language group, Pillai's trace = 0.395, $F$ (3, 108) = 23.46, $p < 0.001$. No other multivariate effects were significant (all $p$s ≥ 0.05). In the planned outcome-specific adjusted analyses, the bilingual group had significantly lower RMET and facial emotion recognition scores than the monolingual group. Both differences remained significant after Holm correction, whereas no adjusted significant group difference was found for Faux Pas scores (Table 2).

**Table 2.** Unadjusted scores and adjusted language-group differences in social-cognitive outcomes.

| Outcome | Bilingual M (*SD*) | Monolingual M (*SD*) | Cohen's *d* [95% CI] | Adjusted *B* (HC3 SE) | 95% CI for B | Holm-Adjusted *p* |
|---|---|---|---|---|---|---|
| RMET total | 22.18 (3.07) | 25.35 (2.61) | −1.11 [−1.50, −0.73] | −3.04 (0.57) | [−4.18, −1.91] | **<0.001** |
| Facial emotion recognition total | 46.63 (2.95) | 50.10 (2.94) | −1.18 [−1.56, −0.79] | −3.54 (0.57) | [−4.68, −2.40] | **<0.001** |
| Happiness | 9.18 (0.62) | 9.30 (0.72) | −0.17 [−0.53, 0.19] | −0.09 (0.14) | [−0.36, 0.18] | 0.498 |
| Sadness | 7.33 (1.26) | 8.45 (1.02) | −0.98 [−1.35, −0.60] | −1.11 (0.23) | [−1.57, −0.65] | **<0.001** |
| Anger | 7.10 (1.17) | 8.53 (1.26) | −1.18 [−1.57, −0.79] | −1.40 (0.23) | [−1.86, −0.94] | **<0.001** |
| Fear | 8.02 (1.31) | 6.97 (1.59) | 0.72 [0.35, 1.09] | 1.04 (0.29) | [0.46, 1.63] | **0.001** |
| Disgust | 7.07 (1.39) | 8.22 (1.21) | −0.88 [−1.26, −0.51] | −1.18 (0.26) | [−1.70, −0.65] | **<0.001** |
| Surprise | 7.93 (1.21) | 8.63 (0.94) | −0.65 [−1.01, −0.28] | −0.80 (0.21) | [−1.22, −0.38] | **<0.001** |
| Faux Pas total | 39.78 (3.34) | 39.88 (3.05) | −0.03 [−0.39, 0.33] | 0.01 (0.63) | [−1.23, 1.25] | 0.986 |

Note. M: mean; *SD*: standard deviation; RMET: Reading the Mind in the Eyes Test; *B*: unstandardized regression coefficient; SE: standard error; CI: confidence interval; IQ: intelligence quotient; HC3 SE: heteroscedasticity-consistent standard error estimator. Descriptive statistics and Cohen's *d* values are unadjusted. Regression coefficients were adjusted for age, sex, IQ score, household income, and maternal education. *B* and Cohen's *d* were calculated as bilingual minus monolingual. Bold values indicate statistical significance ($p < 0.05$).

For facial emotion recognition, the bilingual group had lower sadness, anger, disgust, and surprise scores but a higher fear score. No significant group difference was observed for happiness. The maximum happiness score (10/10) was obtained by 18 bilingual participants (30.0%) and 26 monolingual participants (43.3%). All five group differences remained significant after Holm correction (Table 2).

All participants in both language groups passed the Faux Pas comprehension checks, and no story received a zero score due to comprehension-control errors. In an exploratory analysis of the faux pas recognition score, no significant difference was observed between the bilingual (16.13 ± 2.73) and monolingual (16.48 ± 2.61) groups ($t$ (118) = −0.72, $p$ = 0.474; mean difference = −0.35, 95% CI [−1.32, 0.62]).

Unadjusted group means for the three standardized total scores and six facial emotion recognition scores are shown in Figure 1.

### 3.3. Task-Specific Group Differences

The group × task interaction was significant, Wald $\chi^2$ (2) = 22.43, $p < 0.001$, indicating that the magnitude of the bilingual–monolingual difference varied across the three tasks. Task-specific contrasts showed lower standardized RMET and facial emotion recognition scores in the bilingual group, whereas no statistically significant group difference was observed for Faux Pas scores. Pairwise comparisons indicated that the bilingual–monolingual contrasts for both RMET and facial emotion recognition differed significantly from that for the Faux Pas Test. The RMET and facial emotion recognition contrasts did not differ significantly from one another (Table 3).

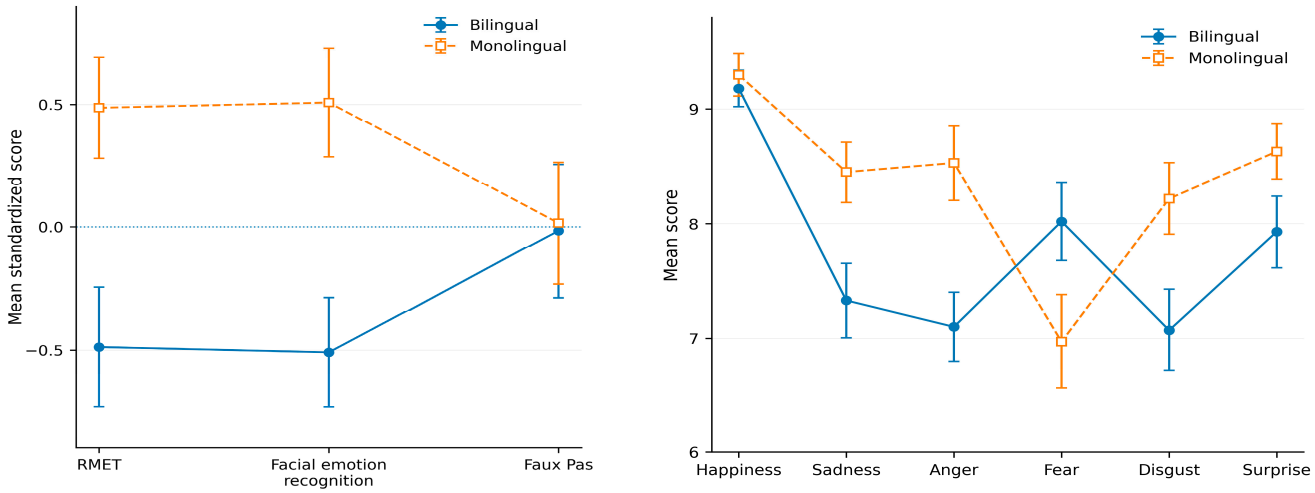


**Figure 1.** Social-cognitive performance in bilingual and monolingual adolescents. Points represent unadjusted group means, and error bars indicate 95% confidence intervals. (**a**) Total scores were standardized separately across the full sample. (**b**) Facial emotion recognition scores are shown for the six individual emotions, each ranging from 0 to 10. RMET, Reading the Mind in the Eyes Test.

**Table 3.** Adjusted task-specific language-group contrasts and pairwise differences from the GEE model.

| Contrast | Estimate (Robust SE) | 95% CI | Holm-Adjusted *p* |
|---|---|---|---|
| Panel A. Task-specific group contrasts | | | |
| RMET | −0.94 (0.16) | [−1.25, −0.62] | **<0.001** |
| Facial emotion recognition | −1.04 (0.15) | [−1.34, −0.73] | **<0.001** |
| Faux Pas | 0.00 (0.18) | [−0.35, 0.36] | 0.984 |
| Panel B. Differences between task-specific group contrasts | | | |
| RMET vs. Faux Pas | −0.94 (0.24) | [−1.40, −0.48] | **<0.001** |
| Facial emotion recognition vs. Faux Pas | −1.04 (0.24) | [−1.51, −0.57] | **<0.001** |
| RMET vs. facial emotion recognition | 0.10 (0.22) | [−0.33, 0.53] | 0.646 |

Note. RMET: Reading the Mind in the Eyes Test; GEE: generalized estimating equations; SE: standard error; CI: confidence interval. Panel A shows bilingual-minus-monolingual contrasts. Panel B compares these contrasts across tasks, calculated as the first-listed task minus the second-listed task. Bold values indicate statistical significance ($p < 0.05$).

### *3.4. Correlation Analyses*

In the full sample, Pearson correlation analyses showed a small positive association between RMET and facial emotion recognition total scores, $r = 0.243$, 95% CI [0.067, 0.405], Holm-adjusted $p = 0.022$. No significant correlations were observed between RMET and Faux Pas scores, $r = 0.070$, 95% CI [−0.111, 0.246], Holm-adjusted $p = 0.864$, or between facial emotion recognition and Faux Pas scores, $r = -0.072$, 95% CI [−0.248, 0.108], Holm-adjusted $p = 0.864$.

After controlling for language group, the correlation between RMET and facial emotion recognition was close to zero, partial $r = -0.008$, 95% CI [−0.188, 0.172]. Partial correlations with Faux Pas scores were 0.072 for RMET, 95% CI [−0.110, 0.248], and −0.094 for facial emotion recognition, 95% CI [−0.269, 0.088]. None of the three partial correlations was significant (all Holm-adjusted $p$ values = 0.934) (Figure 2).

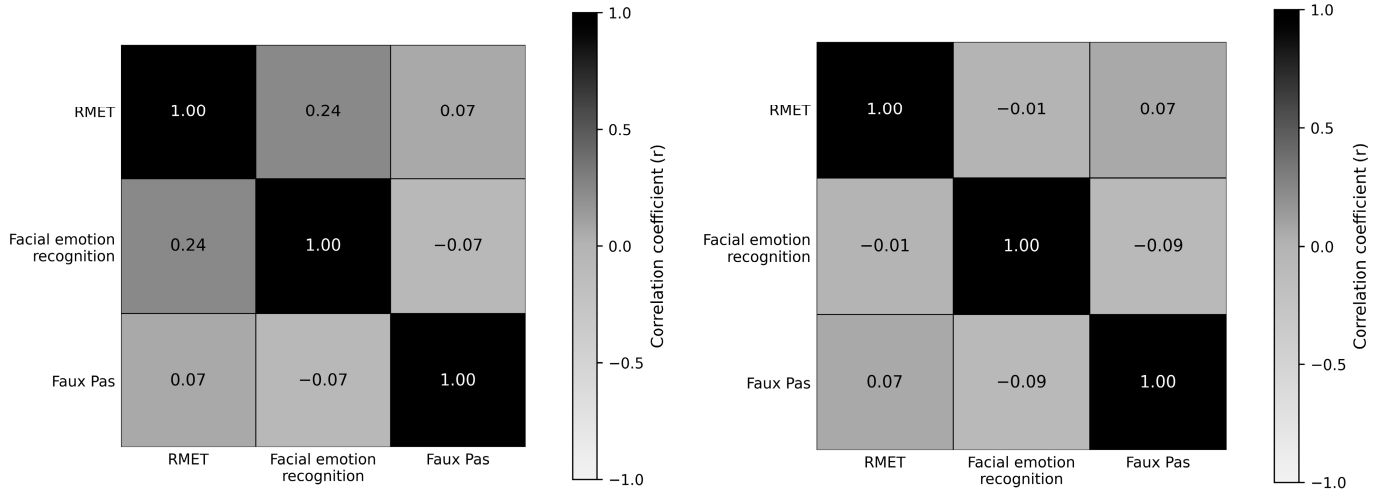


**Figure 2.** Correlations among social-cognitive measures. (**a**) Pearson correlations among RMET, facial emotion recognition, and Faux Pas total scores. (**b**) Partial correlations among the same measures controlling for language group. RMET, Reading the Mind in the Eyes Test.

### *3.5. Model Diagnostics and Sensitivity Analyses*

RMET residuals departed from normality (Shapiro–Wilk $p = 0.022$). Levene tests did not detect between-group differences in residual variance (all $p$s $> 0.05$) (Supplementary Table S4). Group × age and group × IQ interactions were nonsignificant in the multivariate and outcome-specific models (all $p$s $> 0.05$) (Supplementary Table S5). The multivariate language-group effect and the GEE group × task interaction remained significant across sensitivity analyses. Pairwise comparisons of task-specific group contrasts showed the same pattern as the primary analysis (Supplementary Tables S6–S9). Further adjustment for recruitment site did not materially change the language-group differences or the group × task interaction (Supplementary Tables S6–S9).

## 4. Discussion

Bilingual adolescents obtained lower scores than monolingual adolescents on the RMET and facial emotion recognition task, whereas no significant group difference was detected in Faux Pas performance. The differences in RMET and facial emotion recognition remained after adjustment for age, sex, IQ, household income, and maternal education. The significant group × task interaction indicated that the magnitude of the bilingual–monolingual difference varied across the three social-cognitive tasks. Importantly, this task-specific pattern does not support a uniform bilingual advantage or disadvantage in social cognition. The marked differences in RMET and facial emotion recognition alongside the absence of a statistically significant group difference in Faux Pas performance, suggest that bilingual–monolingual differences may be task-dependent.

The lower RMET scores among bilingual adolescents parallel the findings of Asan and Yıldız (2026), who reported lower RMET performance among Syrian Arabic–Turkish sequential bilingual children than among Turkish monolingual children. However, their monolingual advantage in Faux Pas performance was not observed in the present study. In their sample, assessment language was associated with RMET performance, whereas Turkish school-language attainment predicted Faux Pas performance. This pattern may partly reflect the specific perceptual, linguistic, and sociocultural demands of different social-cognitive measures, although these explanations were not directly tested in this study. These findings are consistent with a broader literature in which studies examining a

bilingual ToM advantage have produced inconsistent findings, including divergent results across different tasks within the same study (Schroeder, 2018). Although these studies differ from the present study in their samples and task designs, the variability across their findings provides a broader context for the present task-specific pattern.

RMET requires matching eye-region cues to nuanced mental-state terms, whereas facial emotion recognition requires matching facial expressions to basic emotion labels. Both provide limited situational context. While Faux Pas requires narrative comprehension, its stories provide contextual cues about events and characters' knowledge that may support social inference. In this study, all participants correctly answered the comprehension control questions, and no significant group difference was observed in the exploratory faux pas recognition analysis. However, these findings do not establish equivalent Turkish language proficiency or rule out subtle differences in narrative comprehension between groups. Interpreting neuropsychological test performance in bilingual children also requires consideration of their linguistic and educational background (Canas et al., 2020). Language proficiency and age of acquisition were not measured in this study. The available data cannot separate the contributions of bilingual experience, language proficiency, and sociocultural context in this pattern. The findings therefore describe differences between the sampled language groups and should not be attributed to bilingual experience alone. Lower RMET and facial emotion recognition scores also do not, by themselves, establish lower underlying social-cognitive ability in bilingual adolescents.

The Faux Pas confidence interval remains compatible with differences favoring either group. The confidence interval comparing the RMET and facial emotion recognition group differences also leaves uncertainty about which task shows the larger difference. No formal equivalence analysis was conducted. The nonsignificant comparisons therefore do not establish either equivalent Faux Pas performance between groups or equally large group differences across RMET and facial emotion recognition.

The group difference in facial emotion recognition was not uniform across emotions. Monolingual adolescents scored higher for sadness, anger, disgust, and surprise, no statistically significant group difference was observed for happiness, and bilingual adolescents scored higher for fear. Happiness scores were close to the maximum in both groups, consistent with the high recognition accuracy for happy expressions reported across childhood and adolescence (Lawrence et al., 2015). The high scores and restricted variability may have limited the ability to detect a group difference in happiness. The higher fear score is notable because it contrasts with the lower overall facial emotion recognition score in the bilingual group. This suggests that the total score does not capture the direction of every emotion-specific group difference. In an adult bilingual sample, fear and surprise were the most frequently confused emotions (Szreder, 2026), highlighting the need to examine errors alongside correct responses. In our study, the higher fear score cannot distinguish better discrimination from a greater tendency to select the fear label. The higher fear score in the bilingual group should therefore be considered an emotion-specific finding requiring replication rather than evidence of a general bilingual advantage in fear recognition.

When both language groups were analyzed together, RMET and facial emotion recognition scores showed a weak positive correlation. After controlling for language group, the partial correlation between these scores was close to zero. The pooled correlation may therefore partly reflect between-group differences in mean scores on both tasks. Faux Pas scores were not significantly correlated with either measure, and all three partial correlations were small and nonsignificant. The limited convergence among our measures is consistent with previous reports of weak associations across different ToM tasks (Warnell & Redcay, 2019). However, the confidence intervals for our partial correlations remain compatible with small

associations. These exploratory analyses therefore do not establish that the underlying abilities are independent.

Educational and sociocultural characteristics may also be relevant to the interpretation of these findings. In a sample of 40,248 participants aged 10–70 years, education, race, and ethnicity were associated with more pronounced differences in RMET performance than in the other cognitive tasks examined (Dodell-Feder et al., 2020). In the present study, the group differences in RMET and facial emotion recognition persisted after adjustment for household income, maternal education, and the other covariates. These indicators represent only part of the broader socioeconomic context. In Kyuchuk's (2025) study, Turkish–Bulgarian bilingual preschool children scored lower than monolingual children on classical ToM tasks, but this pattern was not observed on a nonverbal ToM task. The author discussed community-level social-class differences and limited educational support for the minority language as possible contextual contributors to this pattern. These factors were not assessed in our study. In clinical assessment, these scores should be considered together with other social-cognitive measures and a detailed linguistic, educational, and sociocultural history (O'Donald et al., 2025; Olson & Jacobson, 2015).

The results of this study should be interpreted in light of several limitations. The cross-sectional design does not permit causal conclusions about bilingual experience and social-cognitive performance. Participants were recruited from two child and adolescent psychiatry outpatient clinics. Although psychiatric disorders were excluded, clinic-based recruitment may have introduced selection effects and limited generalizability to adolescents in the general population. Also, subthreshold symptoms may have influenced test performance.

Language-group classification was based on self-reported language use. The classification of bilingualism does not infer proficiency, dominance, or acquisition patterns. The 30–70% range does not necessarily indicate balanced proficiency. Averaging language-use percentages across five contexts may also have obscured differences between settings, such as home and school. Proficiency and vocabulary knowledge in both languages were not directly assessed. Language dominance, age of acquisition, bilingual acquisition pattern (simultaneous or sequential), and familiarity with the Turkish mental-state and emotion terms required by the tasks were also not assessed. All tasks were administered in Turkish. The bilingual and monolingual groups may therefore differ in test performance because of differences in Turkish proficiency or familiarity with mental-state and emotion terminology rather than differences in social cognition. Adjustment for total IQ does not rule out group differences in Turkish proficiency or familiarity with task-specific vocabulary. Formal measurement invariance testing across language groups was not conducted. Arabic–Turkish and Kurdish–Turkish bilingual adolescents were also analyzed together, and their subgroup composition was not recorded. Analyzing these linguistic groups together may have obscured differences between them.

Household income and maternal education captured only selected aspects of socioeconomic status. Other potentially relevant factors, including household conditions, school resources, neighborhood characteristics, migration history, and broader cultural and educational experiences, were not assessed. Their possible contributions to the findings could therefore not be evaluated. Adjustment for household income and maternal education does not rule out confounding by these unmeasured contextual factors.

Although the findings were consistent across the alternative age specifications examined, these analyses do not establish linear developmental trajectories. Finally, inter-rater reliability was not assessed for tasks.

## 5. Conclusions

Among adolescents recruited from outpatient clinics in Türkiye, bilingual participants obtained lower RMET and facial emotion recognition scores than monolingual participants, whereas no significant group difference was detected in Faux Pas performance. The findings suggest task-dependent differences in performance and limited convergence among the measures within this sample. Turkish proficiency was not assessed, and the two bilingual language groups were analyzed together. The linguistic, cultural, and social-cognitive explanations for these differences remain unresolved. Future studies should characterize bilingual experience along a continuum and assess language proficiency and social-cognitive performance in both languages. Arabic–Turkish and Kurdish–Turkish bilingual groups should also be examined separately, with more detailed assessment of their cultural, educational, and socioeconomic contexts.

**Supplementary Materials:** The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/bs16122231/s1, Figure S1: Q–Q plots of primary model residuals; Figure S2: Primary model residuals against fitted values; Table S1: Full sample standardization constants; Table S2: Observations flagged by residual screening; Table S3: Participant characteristics by recruitment site; Table S4: Residual distribution and variance diagnostics; Table S5: Tests of common age and IQ slopes; Table S6: Multivariate language-group effects; Table S7: Adjusted unstandardized group differences; Table S8: Group contrasts and interaction tests; Table S9: Differences between task-specific group contrasts.

**Author Contributions:** Conceptualization, M.M.O., E.S. and R.O.C.; Methodology, T.T.K. and R.O.C.; Validation, T.T.K.; Formal Analysis, T.T.K.; Investigation, E.S., M.M.O. and R.O.C.; Resources, M.M.O.; Data Curation, E.S. and R.O.C.; Writing—Original Draft Preparation, T.T.K.; Writing—Review and Editing, R.O.C.; Visualization, T.T.K.; Supervision, R.O.C.; Project Administration, R.O.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding. This research received no external funding. The article processing charge (APC) was funded by Dokuz Eylul University.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki and was approved by the Kafkas University Faculty of Medicine Ethics Committee on 24 December 2020 (session no. 15; document no. 80576354-050-99/331).

**Informed Consent Statement:** Informed consent was obtained from the parents or legal guardians of all participants, and assent was obtained from all participating adolescents.

**Data Availability Statement:** The raw data supporting the conclusions of this article will be made available by the authors on request.

**Acknowledgments:** We would like to thank the participants for their voluntary participation and contribution to the study During the preparation of this manuscript, the authors used OpenAI ChatGPT (GPT-5.6 Sol) for language editing. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

Asan, M. N., & Yıldız, T. (2026). Advanced theory of mind in Syrian Arabic–Turkish sequential bilinguals and Turkish monolinguals: Assessment language, school-language attainment, and the ecology of bilingualism. *First Language*. Advance online publication. [CrossRef]

Assendelft, L., Poos, J. M., & van den Berg, E. (2026). Psychometric properties of the Ekman 60-Faces Test in clinical populations: A systematic review and meta-analysis. *Journal of Neuropsychology*. Advance online publication. [CrossRef] [PubMed]

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders*, *29*(5), 407–418. [CrossRef] [PubMed]

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, *42*(2), 241–251. [CrossRef]

Birdsong, D. (2014). Dominance and age in bilingualism. *Applied Linguistics*, *35*(4), 374–392. [CrossRef]

Birdsong, D., Gertken, L. M., & Amengual, M. (2012). *Bilingual language profile: An easy-to-use instrument to assess bilingualism*. COERLL, University of Texas at Austin. Available online: https://blp.coerll.utexas.edu/ (accessed on 25 September 2026).

Buac, M., & Kaushanskaya, M. (2020). Predictors of theory of mind performance in bilingual and monolingual children. *International Journal of Bilingualism*, *24*(2), 339–359. [CrossRef]

Calvo, A., & Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition*, *130*(3), 278–288. [CrossRef] [PubMed]

Canas, A., Bordes Edgar, V., & Neumann, J. (2020). Practical considerations in the neuropsychological assessment of bilingual (Spanish-English) children in the United States: Literature review and case series. *Developmental Neuropsychology*, *45*(4), 211–231. [CrossRef] [PubMed]

Champoux-Larsson, M.-F., Dylman, A. S., Örnkloo, H., & Esteves, F. (2019). Identification of facial expressions of emotion by 4-year-old children from different linguistic environments. *International Journal of Bilingualism*, *23*(5), 1208–1219. [CrossRef]

Cox, S. R., Bak, T. H., Allerhand, M., Redmond, P., Starr, J. M., Deary, I. J., & MacPherson, S. E. (2016). Bilingualism, social cognition and executive functions: A tale of chickens and eggs. *Neuropsychologia*, *91*, 299–306. [CrossRef] [PubMed]

Çetintaş Yıldırım, F. (2020). Language choice and identity: An investigation based on the comparison of language attitudes from two different localities. *Journal of Language and Linguistic Studies*, *16*(2), 1032–1042. [CrossRef]

Dodell-Feder, D., Ressler, K. J., & Germine, L. T. (2020). Social cognition or social class and culture? On the interpretation of differences in social cognitive performance. *Psychological Medicine*, *50*(1), 133–145. [CrossRef] [PubMed]

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Consulting Psychologists Press.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, *128*(2), 203–235. [CrossRef] [PubMed]

Feng, J., Cho, S., & Luk, G. (2024). Assessing theory of mind in bilinguals: A scoping review on tasks and study designs. *Bilingualism: Language and Cognition*, *27*(4), 531–545. [CrossRef]

Girli, A. (2014). Psychometric properties of the Turkish child and adult form of "Reading the Mind in the Eyes Test". *Psychology*, *5*(11), 1321–1337. [CrossRef]

Grazzani, I., & Ornaghi, V. (2012). How do use and comprehension of mental-state language relate to theory of mind in middle childhood? *Cognitive Development*, *27*(2), 99–111. [CrossRef]

Griffiths, S., Goh, S. K. Y., Norbury, C. F., & The SCALES Team. (2020). Early language competence, but not general cognitive ability, predicts children's recognition of emotion from facial and vocal cues. *PeerJ*, *8*, e9118. [CrossRef] [PubMed]

Guldenoglu, B., Erfidan, G., & Alatlı, R. (2026). The reading process in bilingual and monolingual children: Fluency as a mediating mechanism in a morphologically complex and transparent language. *Reading and Writing*. Advance online publication. [CrossRef]

Harkness, K. L., Sabbagh, M. A., Jacobson, J. A., Chowdrey, N. K., & Chen, T. (2005). Enhanced accuracy of mental state decoding in dysphoric college students. *Cognition & Emotion*, *19*(7), 999–1025. [CrossRef]

Hudson, C. C., Shamblaw, A. L., Harkness, K. L., & Sabbagh, M. A. (2020). Valence in the Reading the Mind in the Eyes task. *Psychological Assessment*, *32*(7), 623–634. [CrossRef] [PubMed]

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., Williamson, D., & Ryan, N. (1997). Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(7), 980–988. [CrossRef] [PubMed]

Kessels, R. P. C., Montagne, B., Hendriks, A. W., Perrett, D. I., & de Haan, E. H. F. (2014). Assessment of perception of morphed facial expressions using the Emotion Recognition Task: Normative data from healthy participants aged 8–75. *Journal of Neuropsychology*, *8*(1), 75–93. [CrossRef] [PubMed]

Kittel, A. F. D., Olderbak, S., & Wilhelm, O. (2022). Sty in the mind's eye: A meta-analytic investigation of the nomological network and internal consistency of the "Reading the Mind in the Eyes" Test. *Assessment*, *29*(5), 872–895. [CrossRef] [PubMed]

Kyuchuk, H. S. (2025). Language and "theory of mind" development of bilingual and monolingual children in Bulgaria. *Frontiers in Human Neuroscience*, *19*, 1522507. [CrossRef] [PubMed]

Lawrence, K., Campbell, R., & Skuse, D. (2015). Age, gender, and puberty influence the development of facial emotion recognition. *Frontiers in Psychology*, *6*, 761. [CrossRef] [PubMed]

Meinhardt-Injac, B., Daum, M. M., & Meinhardt, G. (2020). Theory of mind development from adolescence to adulthood: Testing the two-component model. *British Journal of Developmental Psychology*, *38*(2), 289–303. [CrossRef] [PubMed]

Nguyen, T.-K., & Astington, J. W. (2014). Reassessing the bilingual advantage in theory of mind and its cognitive underpinnings. *Bilingualism: Language and Cognition*, *17*(2), 396–409. [CrossRef]

O'Donald, F., Calia, C., & Franzen, S. (2025). Cross-cultural paediatric neuropsychological assessment: Key considerations. *Health Science Reports*, *8*(6), e70959. [CrossRef] [PubMed]

Olderbak, S., Wilhelm, O., Olaru, G., Geiger, M., Brenneman, M. W., & Roberts, R. D. (2015). A psychometric analysis of the Reading the Mind in the Eyes Test: Toward a brief form for research and applied settings. *Frontiers in Psychology*, *6*, 1503. [CrossRef] [PubMed]

Olson, K., & Jacobson, K. (2015). Cross-cultural considerations in pediatric neuropsychology: A review and call to attention. *Applied Neuropsychology: Child*, *4*(3), 166–177. [CrossRef] [PubMed]

Palmer, S. J., Fanucci-Kiss, A., Kipervassar, E., Jalnapurkar, I., Hodge, S. M., Frazier, J. A., & Cochran, D. (2024). Effect of emotional valence on emotion recognition in adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *54*(4), 1494–1506. [CrossRef] [PubMed]

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526. [CrossRef]

Rosso, A. M., & Riolfo, A. (2020). A further look at Reading the Mind in the Eyes-Child Version: Association with fluid intelligence, receptive language, and intergenerational transmission in typically developing school-aged children. *Frontiers in Psychology*, *11*, 586065. [CrossRef] [PubMed]

Savaşır, I., & Şahin, N. (1995). *Wechsler Çocuklar İçin Zeka Ölçeği uygulama kitapçığı* [Wechsler intelligence scale for children administration manual]. Türk Psikologlar Derneği.

Schroeder, S. R. (2018). Do bilinguals have an advantage in theory of mind? A meta-analysis. *Frontiers in Communication*, *3*, 36. [CrossRef]

Schurz, M., Radua, J., Tholen, M. G., Maliske, L., Margulies, D. S., Mars, R. B., Sallet, J., & Kanske, P. (2021). Toward a hierarchical model of social cognition: A neuroimaging meta-analysis and integrative review of empathy and theory of mind. *Psychological Bulletin*, *147*(3), 293–327. [CrossRef] [PubMed]

Selcuk, B., Gonultas, S., & Ekerim-Akbulut, M. (2023). Development and use of theory of mind in social and cultural context. *Child Development Perspectives*, *17*(1), 39–45. [CrossRef]

Sezgin, N., Baştuğ, G., Yargıcı Karaağaç, S., & Yılmaz, B. (2014). Wechsler Yetişkinler İçin Zekâ Ölçeği Gözden Geçirilmiş Formu (WAIS-R) Türkiye standardizasyonu: Ön çalışma [Turkish standardization of Wechsler adult intelligence scale-revised (WAIS-R): Pilot study]. *Ankara Üniversitesi Dil ve Tarih-Coğrafya Fakültesi Dergisi*, *54*(1), 451–480. Available online: https://dergipark.org.tr/en/pub/dtcfdergisi/article/1044484 (accessed on 27 October 2026).

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience*, *10*(5), 640–656. [CrossRef] [PubMed]

Szreder, M. (2026). The relationship between language and emotion recognition in bilinguals is not robust to cultural and linguistic differences. *International Journal of Bilingualism*, *30*(5), 1427–1451. [CrossRef]

Unsworth, S., Chondrogianni, V., & Skarabela, B. (2018). Experiential measures can be used as a proxy for language dominance in bilingual language acquisition research. *Frontiers in Psychology*, *9*, 1809. [CrossRef] [PubMed]

Ünal, F., Öktem, F., Çetin Çuhadaroğlu, F., Çengel Kültür, S. E., Akdemir, D., Foto Özdemir, D., Çak, H. T., Ünal, D., Tıraş, K., Aslan, C., Kalaycı, B. M., Aydos, B. S., Kütük, F., Taşyürek, E., Karaokur, R., Karabucak, B., Karakök, B., Karaer, Y., & Artık, A. (2019). Okul çağı çocukları için duygulanım bozuklukları ve şizofreni görüşme çizelgesi–şimdi ve yaşam boyu şekli–DSM-5 Kasım 2016–Türkçe uyarlamasının (ÇDŞG-ŞY-DSM-5-T) geçerlik ve güvenirliği [Reliability and validity of the schedule for affective disorders and schizophrenia for school-age children-present and lifetime version, DSM-5 November 2016-Turkish adaptation (K-SADS-PL-DSM-5-T)]. *Türk Psikiyatri Dergisi*, *30*(1), 42–50. [CrossRef]

Warnell, K. R., & Redcay, E. (2019). Minimal coherence among varied theory of mind measures in childhood and adulthood. *Cognition*, *191*, 103997. [CrossRef] [PubMed]

Yazgeç, E., Bora, E., İngeç, S., Çıray, R. O., Bağcı, B., & Alkın, T. (2023). Social cognition in opioid use disorder. *Substance Use & Misuse*, *58*(8), 996–1003. [CrossRef] [PubMed]

Young, A., Perrett, D., Calder, A., Sprengelmeyer, R., & Ekman, P. (2002). *Facial expressions of emotion: Stimuli and tests (FEEST)*. Thames Valley Test Company.

Yu, C.-L., Kovelman, I., & Wellman, H. M. (2021). How bilingualism informs theory of mind development. *Child Development Perspectives*, *15*(3), 154–159. [CrossRef] [PubMed]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.