International Journal of Assessment Tools in Education
2024, Vol. 11, No. 3, 481–506
https://doi.org/10.21449/ijate.1366478
journal homepage: https://dergipark.org.tr/en/pub/ijate
Research Article

The validity and reliability study of the theory of mind inventory-2 (TOMI-2) Turkish version

Canan Keleş Ertürk[1*], **Kezban Tepeli**[2]

[1]KTO Karatay University, Vocational School of Health Services, Child Development Program, Konya, Türkiye
[2]Selçuk University, Faculty of Health Sciences, Child Development, Konya, Türkiye

ARTICLE HISTORY

Received: Sep. 26, 2023

Accepted: June 18, 2024

Keywords:
Theory of mind,
Preschool period,
False belief,
Validity,
Reliability.

Abstract: This study aims to conduct the Turkish adaptation, validity, and reliability study of the Theory of Mind Inventory-2 (TOMI-2) developed by Hutchins and Prelock (2016) for 3-5-year-old children. The study group consists of 310 mothers with children in the 3-5 age group in Konya city center. Personal Information Form and Theory of Mind Inventory-2 (TOMI-2) were used as data collection tools in the study. After the TOMI-2 was translated into Turkish, the normality assumption was checked with the "Shapiro-Wilk" test. The relationship between two continuous variables was evaluated with the Pearson Correlation Coefficient. Exploratory Factor Analysis, Confirmatory Factor Analysis, Content Validity, Criterion Validity, and Reliability analyses were also used in the study. The findings of the analyses show that the Turkish version of the TOMI-2 is a valid and reliable measurement tool for children aged 3-5, with 60 items in the original form.

1. INTRODUCTION

Theory of mind (ToM), defined as the ability to predict and explain people's behavior, is considered an important milestone in social cognitive development (Slaughter & Repacholi, 2003). Theory of mind, which also means the capacity to interpret, predict, and explain the behaviors of others according to their underlying mental states, begins to develop from early childhood (Scholl & Leslie, 1999). As theory of mind involves both explaining one's actions and interpreting and predicting the actions of other individuals, it forms the basis for understanding human behavior (Astington & Dack, 2008). Theory of mind refers not only to a cognitive tool used to predict and explain action but also to a system of ideas about mental states and activities (Sodian, 2005, p.112).

Different views and theories on the development of the theory of mind have been developed. According to the theory, the believed situation creates a biased effect on perception and the experience shapes the theory of mind (Flavell, 1999). According to the modular theory, theory of mind is acquired through neurological processes, but performance and experience are not ignored (Sodian & Kristen, 2010). Simulation theory, on the other hand, focuses on knowledge

*CONTACT: Canan KELEŞ ERTÜRK ✉ canankeles90@gmail.com KTO Karatay University, Vocational School Of Health Services, Child Development Program, Konya, Türkiye

e-ISSN: 2148-7456

about how to perform practical skills (Ratcliffe, 2007). There are two main views of the Theory of Mind. According to the traditional view, ToM is a unifying construct about the cognitive aspects of knowing what another person knows. Assessments here focus on understanding cognition, thinking about what someone thinks, knows, or believes. In the modern view, the theory of mind is a construct that is closely related to language. However, it is not a unifying construct. With the increase in brain imaging studies, there is evidence that ToM has different dimensions such as cognitive ToM, emotional-cognitive ToM, and emotional empathy. In addition, the interpersonal theory of mind, which is explained as thinking about others' thoughts and emotions, and the personal theory of mind, which is explained as thinking about one's thoughts and feelings, involve different neurophysiology and different skill groups (Westby & Robinson, 2014).

A developed theory of mind enables an individual to understand that behavior can be guided by mental states (such as desires, knowledge, and beliefs). Theory of mind is also accepted as a fundamental skill for social cognition (interacting with other individuals) and reading comprehension. In this respect, the development of the theory of mind, which is a comprehensive concept, proceeds in certain stages (Table 1) (Tucci, 2016).

Table 1. *Developmental sequence.*

Stage	Age of Mastery (Months)	Task Description
Diverse Desires Stage	36-48 months	The child is given a choice of two snacks. The child picks a favorite snack. Another character chooses the opposing snack as his/her favorite. The child is asked what the character will choose to eat. The child must inhibit his/her desire and choose the opposing snack.
Diverse Beliefs Stage	36-48 months	The child is given a choice of two locations for a missing cat. The child picks the location where he thinks the cat is hiding. Another character chooses the opposing location. The child is asked where the character will look for the cat. The child must inhibit his/her desire and choose the opposing location.
Social Pretend Stage	48-54 months	The child and assessor pretend to paint a blue cup green. The child is asked what color another character thinks the cup is. The child should say the initial color of the cup.
Knowledge Access	55 months	The child is shown a nondescript box. A toy is hidden inside the box. The child is asked what the character thinks is inside the box. The child must say the character doesn't know.
False Belief-Unexpected Contents	60 months	The child is shown a candy box and when asked what is inside, is expected to answer candy. When the inside of the box is shown, it is understood that it is something different. The child is asked what another character thinks is inside the box and is expected to answer candy.

Babies can distinguish between the movements of animate and inanimate objects around 6 months. Perception-goal psychology, the most basic form of theory of mind, emerges around 9 months. It allows individuals to understand that they may have different perceptual perspectives and different goals and act accordingly. At around 12 months, the development of joint attention begins, and by 18 months, the theory of mind manifests itself in the ability of joint attention. Improvement of basic theory of mind skills occurs between 1-3 years of age. From the age of 4, it emerges in the form of belief-desire psychology. This is associated with the subjective representation of others' ways of seeing the world, which may be inaccurate and incompatible with one's subjective view. Higher-level theory of mind is improved until adulthood (Rakoczy, 2022; Şahin et al. 2019).

In this gradual development process of the theory of mind, evaluation is an important issue. In the assessment of the theory of mind, the false belief task, which involves obtaining accurate

predictions about another person's behavior by referring to that person's mental (false) representation (i.e. their false belief), is used (Slaughter & Repacholi, 2003). The false belief task is implemented in two general forms: unexpected content and unexpected location. The unexpected location task is related to the changing location of the object. The unexpected content is about the contents of a box. Here, the child is shown that there is a very different content (such as money, ribbon) in a box that typically belongs to one type of content (e.g., candy, paint box) and is asked what another person who has not seen the inside of the box thinks is in the box (Miller, 2016, p. 9). In the process of evaluating the Theory of Mind (ToM), it should be taken into consideration that ToM is influenced by social conversation experiences, interactions with siblings, participation in imaginary games, and secure attachment, in short, environmental factors and genetic foundations are also involved in this process (Zufferey, 2010, p.39; Wellman, 2014, p.107; Slaughter et al., 2015; Wellman, 2017; McElwain, et al., 2019).

In the literature, scale development and adaptation studies have been carried out to evaluate the Theory of Mind skills of 3-5-year-old children. Gözün Kahraman (2012) conducted a Turkish adaptation study of the Theory of Mind Scale developed by Wellman and Liu (2004). The scale consists of 6 tasks. Each task is presented with scenarios written for children, small toy figures and pictures are used, and then the child is asked the relevant question. Kılıç Tülü and Ergül (2022) developed the "Theory of Mind Test for 3-5 Year Old Children". The test includes 27 items for the 3 and 4-year-olds and 26 items for the 5-year-olds. The test is applied by telling short stories about the skills to be measured, showing the photographs, asking the relevant questions, and getting answers from the child. Altıntaş (2014) and Keleş Ertürk & Tepeli (2023) carried out the Turkish adaptation study of the Theory of Mind Task Battery (TOMTB), which was developed and revised by Hutchins and Prelock in 2010. The TOMTB is in booklet format with a test of 15 basic questions, colorful pictures, and accompanying text. The text is read and the child is asked to find the picture showing the correct answer. When the scales developed and adapted in the national literature are examined, it is seen that the assessment of ToM is a developing subject, and scale applications are limited to 3–5-year-old children. This study was planned based on the idea that evaluating the theory of mind across a wider age range and developmental stages, with input from parents, would be more useful. For this reason, the adaptation study of the TOMI-2 will provide a more detailed evaluation of ToM and provide guiding information for national inventory development.

1.1. Present Study

When the developed and adapted measurement tools are examined, it is seen that theory of mind is an emerging topic in the literature and the measurement tools are limited to children. This study is an inventory adaptation study that emerged due to the need for the evaluation of the theory of mind in the field. The TOMI-2, whose Turkish adaptation, validity, and reliability study was conducted within the scope of this study, examines ToM skills in a wider age range and progressively based on parental opinions. This detailed examination is provided by the 6 subscales of the TOMI-2. The Early subscale assesses ToM abilities that typically emerge in late infancy and childhood. The Basic subscale assesses ToM abilities that typically emerge during the preschool years. The Advanced subscale assesses ToM abilities that typically emerge in late childhood but persist into adolescence. The Emotion Recognition subscale focuses on the ability to recognize various emotions. The Comprehension of Mental State Terms subscale provides an understanding of mental state terms. The Pragmatics subscale provides an understanding of the pragmatic and metalinguistic aspects of language (Hutchins & Prelock, 2016).

With the TOMI-2, both the ToM skills of children can be determined, and individually delayed or advanced ToM skills can be revealed and suggestions and points that need to be developed can be determined individually. From this point of view, the TOMI-2 can assess the ToM of 3-5-year-old children more comprehensively and can also be used in clinical assessment.

In light of this information, this study aimed to conduct a scientifically accurate Turkish adaptation, validity, and reliability study of the Theory of Mind Inventory-2 (TOMI-2) developed by Hutchins and Prelock (2016) for 3-5-year-old children. In line with this general purpose, the following sub-goals were tested.

I. Does the Theory of Mind Inventory-2 (TOMI-2) provide content validity for 3–5-year-old children?
II. Does the Theory of Mind Inventory-2 (TOMI-2) provide construct validity for 3–5-year-old children?
III. Does the Theory of Mind Inventory-2 (TOMI-2) provide reliability for 3–5-year-old children?
IV. Does the Theory of Mind Inventory-2 (TOMI-2) provide criterion validity for 3to 5-year-old children?

2. METHOD

This study aimed to adapt, validate, and test the reliability of the Theory of Mind Inventory-2 (TOMI-2) for Turkish children aged 3-5, originally developed by Hutchins and Prelock (2016), using the general survey model. There are steps to be followed for a measurement tool developed in one language to be used in another language. According to Hambleton and Patsula (1998), when the purpose of the adapted test is cross-cultural or international assessment, an adapted test is the most effective way to produce an equivalent test in a second language. Considering this principle, the adaptation study of the TOMI-2 was planned.

As stated by Hambleton and Patsula (1999), there are basic principles that should be followed in the process of adapting a measurement tool. To ensure linguistic equivalence, people who are fluent in both languages and have knowledge of the subject should be selected and forward and backward translation techniques should be used. A different group of translators should then review the adapted test. After the linguistic process, a pilot study should be conducted with a small group. After all the arrangements are completed, the application should be made in the sample group and the necessary analyses should be made. In this study, an adaptation study was conducted according to the principles determined by Hambleton and Patsula (1999).

2.1. Participants

The sample of the study was determined by the Appropriate Case Study Group, which is one of the Purposeful Study Groups. A Convenient Case Study Group is the easy selection of individuals and groups to be researched (Sönmez & Alacapınar, 2018: 175). Based on this, the study group of the research consists of a total of 310 children in the 3-5 age group and their mothers attending 5 kindergartens affiliated with the Ministry of National Education in Konya city center. The sample size was estimated based on relative criteria such as the number of items or factors. The sample size for factor analysis was reported as 100=*poor*, 200=*adequate*, 300=*good*, 500=*very good*, and 1000 and above=*excellent*. Bryman and Cramer's sample size recommendation is to apply the number obtained by multiplying the number of items by 5 or 10 (Çokluk et al. 2018). Therefore, the sample in this study was determined as 310 people. Descriptive statistics regarding the personal characteristics of the children are presented in Table 2, and descriptive statistics regarding the characteristics of the parents are presented in Table 3.

Table 2 shows that the average age of the children of the parents who participated in the study was 55.95 months. Of the children, 151 (%48.7) were girls and 159 (%51.3) were boys. 160 (%51.6) were the first child and 72 (%23.2) were the only child. The duration of preschool attendance was less than 6 months for 161 (%51.9) children.

Table 2. *Descriptive statistics on the characteristics of children.*

	Statistics
Age (Month)	
Mean±SD	55.95±8.28
Min-Max	37-70
Age (Month) category	
37-48 Month	66 (%21.3)
49-60 Month	100 (%32.25)
61-70 Month	144 (%46.45)
Gender	
Female	151 (%48.7)
Male	159 (%51.3)
Birth order	
First child	160 (%51.6)
Middle child or one of the middle children	38 (%12.3)
Last Child	112 (%36.1)
Number of Siblings	
0	72 (%23.2)
1	164 (%52.9)
2	56 (%18.1)
3 and more	18 (%5.8)
Duration of Preschool Education	
0-6 months	161 (%51.9)
7-12 months	47 (%15.2)
13-18 months	35 (%11.3)
19-24 months	36 (%11.6)
More than two years	31 (%10)

Summary statistics are given as *mean ± standard and Median (minimum. maximum)* for numerical data and *Number (Percentage)* for categorical data.

Table 3 shows that while the mothers of 70 (%22.6) children are 29 years old or younger, there are 30 (%9.7) children whose fathers are 29 years old and younger. There are 25 (%8.1) children whose mothers have postgraduate degrees and 50 (%16.1) children whose fathers have postgraduate degrees. In addition, there are 162 (%52.3) children whose mothers are working and 304 (%98.1) whose fathers are working. Of the 88 mothers who selected others (%28.4), 52 were health personnel, 11 were lawyers and 25 were engineers. Of the 132 fathers who selected Other (%42.6), 47 were health personnel, 7 were lawyers, 22 were security personnel, 43 were merchants and 13 were engineers.

Table 3. *Descriptive statistics of the characteristics of the parents.*

	Statistics
Mother's age	
29 years and below	70 (%22.6)
30-39 years	198 (%63.9)
40-49 years	42 (%13.5)
Father's age	
29 years and below	30 (%9.7)
30-39 years	191 (%61.6)
40-49 years	81 (%26.1)
50 years and older	8 (%2.6)
Mother's Education	
Primary and secondary school	36 (%11.6)
High School	53 (%17.1)
University	196 (%63.2)
Postgraduate	25 (%8.1)
Father's Education	
Primary and secondary school	21 (%6.8)
High School	50 (%16.1)
University	189 (%61)
Postgraduate	50 (%16.1)
Mother's employment status	
Working	162 (%52.3)
Not working	148 (%47.7)
Father's employment status	
Working	304 (%98.1)
Not working	6 (%1.9)
Mother's occupation	
Housewife	132 (%42.6)
Officer	79 (%25.5)
Worker	6 (%1.9)
Self-employed	5 (%1.6)
Other	88 (%28.4)
Father's occupation	
Officer	86 (%27.7)
Worker	22 (%7.1)
Self-employed	70 (%22.6)
Other	132 (%42.6)

Summary statistics are given as *Number (Percentage)* values.

2.2. Data Collection Tools

2.2.1. *Personal ınformation form*

In the study, the "Personal Information Form" prepared by the researcher was used to determine the demographic characteristics of the parents of children in the 3-5 age group. This form consists of multiple-choice questions about the child's gender, birth order, date of birth, number of siblings, duration of preschool attendance, socio-economic level of the family, parent's age, education level, occupation, and employment status.

2.2.2. *Theory of mind inventory-2 (TOMI-2)*

The Theory of Mind Inventory is designed to assess social cognitive states. The inventory is completed by parents or individuals primarily responsible for the care of typically developing children between the ages of 2 and 12 and individuals diagnosed with autism spectrum disorder. The first version of the TOMI consists of 42 items. Each item is answered with a 20 cm continuum supporting the statements "definitely no, probably no, undecided, probably, definitely". The participant is asked to read the item and mark the appropriate point on the 20 cm line. The validity and reliability study of the first version of the TOMI was conducted with the participation of 124 participants. The test-retest r=0.89; standard error of measurement 1.50; internal consistency Cronbach's Alpha value a=0.98; criterion validity r=0.73 were calculated for the first version of TOMI. As a result of the analyses, the first version of the TOMI was found to be a valid and reliable measurement tool. Then, the number of items was increased and the 60-item TOMI-2 was created. The norm study of TOMI-2 consists of 802 participants. In the analysis conducted for the structural validity of TOMI-2, the Pearson correlation was found to be $r = 0.67$ ($p < 0.001$). TOMI-2 explains 80% of the total variance. The correlation between TOMI-2 and TOMI is r=0.89. TOMI-2 Cronbach Alpha reliability coefficient is $a = 0.98$. For TOMI-2, the standard error of measurement (SEM) was 2.12 for the composite score ($M = 100$, $SD = 15$) and 1.4 for the subscale scores ($M = 50$, $SD = 10$). As a result of the analyses, it was determined that TOMI-2 is a reliable and valid measurement tool. The ToMI-2 consists of 6 subscales and a total of 60 items. Each of the 60 items that make up the ToMI-2 belongs to one of 6 empirically derived subscales (Early, Basic, Advanced, Emotion Recognition, Mental State Term Understanding, and Pragmatic) that reflect a progression in ToM development. The Early subscale focuses on ToM abilities that typically emerge during late infancy and toddlerhood. The Basic subscale includes ToM abilities that typically emerge during the preschool years. The Advanced subscale includes ToM abilities that typically emerge in late childhood but persist into adolescence. The Emotion Recognition subscale includes the ability to recognize various emotions. The Understanding Mental State Terms subscale includes an understanding of mental state terms such as 'want', 'think', and 'know'. The Pragmatics subscale includes understanding the pragmatic and metalinguistic aspects of language. Both manual (paper and pencil) scoring and computer-based scoring can be done with the ToMI-2. In manual (paper-and-pencil) scoring, each of the 60 items that make up the ToMI-2 is scored using the ruler on the last page of the scale. The 20-centimeter ruler gives possible scores ranging from 0 to 20 for each item. Computer-based scoring is accomplished by entering the scores obtained on the TOMI-2 online. By scoring the TOMI-2, raw scores, percentiles, and standard scores can be generated. The examination of raw scores can be useful when the user is interested in individual item-level and/or subscale-level analyses. Percentiles are also obtained in computer-based scoring. Percentiles are a type of ordinal norm-referenced scores. For example, for very young children in early developing ToM capacity, a 2-point difference may result in a relatively large change in percentile rank, whereas for an older child, a 2-point difference may result in a very small change in percentile rank. A standard score is a raw score converted into a scale with known characteristics (e.g., a specific mean and standard deviation). The ToMI-2 uses two different standard scores: the standard score for the composite (overall) score has a mean of 100 and a standard deviation of 15, and the standard scores for the six subscale scores (Early, Basic, Advanced, Emotion Recognition, Mental State Term Understanding, Pragmatics) have a mean of 50 and a standard deviation of 10 (i.e., these are T scores) (Hutchins & Prelock, 2016; Prelock, Hutchins & Bonazinga Bouyea, 2016). Adaptation studies of the original TOMI and TOMI-2 have been conducted with different samples. The adaptation study of the French version of the original TOMI was conducted by Houssa, Mazzone, & Nader-Grosbois (2014) with 107 typically developing children aged 3-5 years. The factor validity study of the TOMI-2 was conducted by Lee et al. (2023) with 420 typically developing children aged 3-7 years in a Taiwanese sample.

2.3. Procedure and Data Analysis

In the translation of TOMI-2 from English to Turkish, forward and backward translation procedure was applied and language equivalence was ensured. Then, field experts were consulted to evaluate the content of the TOMI-2.

In the study, the SPSS software was used to conduct an explanatory factor analysis on the collected data set. In exploratory factor analysis, the dimensions obtained as a linear combination of observed variables are called factors. The factors are hypothetical variables formed by observed variables (Rencher, 2002). To evaluate the suitability of the data for factor analysis, the correlation matrix should be examined. If a significant portion of the coefficients in the correlation matrix is not greater than 0.30, the application of factor analysis may not be appropriate (Hair et al., 1998). The rejection of the basic hypothesis indicates that the variables are suitable for factor analysis.

In addition, the Kaiser-Meyer-Olkin (KMO) criterion, which is obtained by using correlation and partial correlation coefficients, is important in evaluating the suitability of the data for factor analysis. KMO, which is the sample adequacy criterion, takes a value between 0-1. If the KMO value is less than 0.5, the data set is not suitable for factor analysis (Cerney & Kaiser, 1997). In the study, the principal components method was used to obtain the factors. In determining the appropriate number of factors, factor selection criteria as much as the number of eigenvalues greater than one were taken into account. In addition, factor rotation was performed to clarify the variables contributing to the formation of each common factor. The Varimax method was applied to this process. Confirmatory factor analysis was also applied to test the suitability of the factors obtained by exploratory factor analysis to hypothetical or theoretical factor structures. Exploratory factor analysis is generally applied before measurement tool development and construct validity testing.

Confirmatory factor analysis, on the other hand, is used to confirm the structure obtained as a result of explanatory factor analysis or the theoretical factor structure (Brown, 2015). In explanatory factor analysis, the appropriate number of factors to define the basic structure is revealed based on the data matrix, while in confirmatory factor analysis, the number of factors is known a priori. SPSS and Amos package programs were used for confirmatory factor analysis in the study.

Descriptive statistics for the variables in the study were given as number of units (*n*), percentage (*%*), mean ± standard deviation, median (*M*), minimum (*min*), and maximum (*max*) values. In addition, the normality assumption, one of the prerequisites of parametric tests, was examined with the "Shapiro-Wilk" test. The relationship between two continuous variables was evaluated with Pearson Correlation Coefficient. $p<0.05$ level was considered statistically significant.

2.4. Ethical Principles

This study was conducted by scientific ethical principles. First of all, the developers of the TOMI-2 were contacted and the necessary permissions were obtained. The informed consent form was given to the participants of the study and their participation was ensured voluntarily. It was approved with decision number 2023/043 of KTO Karatay University Non-Pharmaceutical and Medical Device Research Ethics Committee that the study could be carried out.

3. FINDINGS

The mean scores of the items in the TOMI-2 are presented in Figure 1 and the descriptive statistics are in Table 4. When Figure 1 is analyzed, it is seen that 60 items in TOMI-2 have a value between 0 and 20 points. The mean values of the items are shown in the figure. While the 8th item has the highest mean, the 19th item has the lowest mean.

According to Table 4, the mean of the Early subscale in the first part of the TOMI-2 was 14.36±2.44, the mean of the Basic subscale was 13.38±2.46 and the mean of the Advanced

subscale was 11.00±2.43 points. The mean of the Emotion Recognition Subscale in the second part was 13.35±2.52, the mean of the Mental State Term Comprehension Subscale was 14.3±3.23 and the mean of the Pragmatic Subscale was 11.34±2.72 points. There are high-level statistically significant relationships between the Early, Basic, Advanced, Emotion Recognition, Mental State Term Comprehension, and Pragmatics subscales in the first and second parts.

Figure 1. *Mean score table of the items in TOMI-2.*

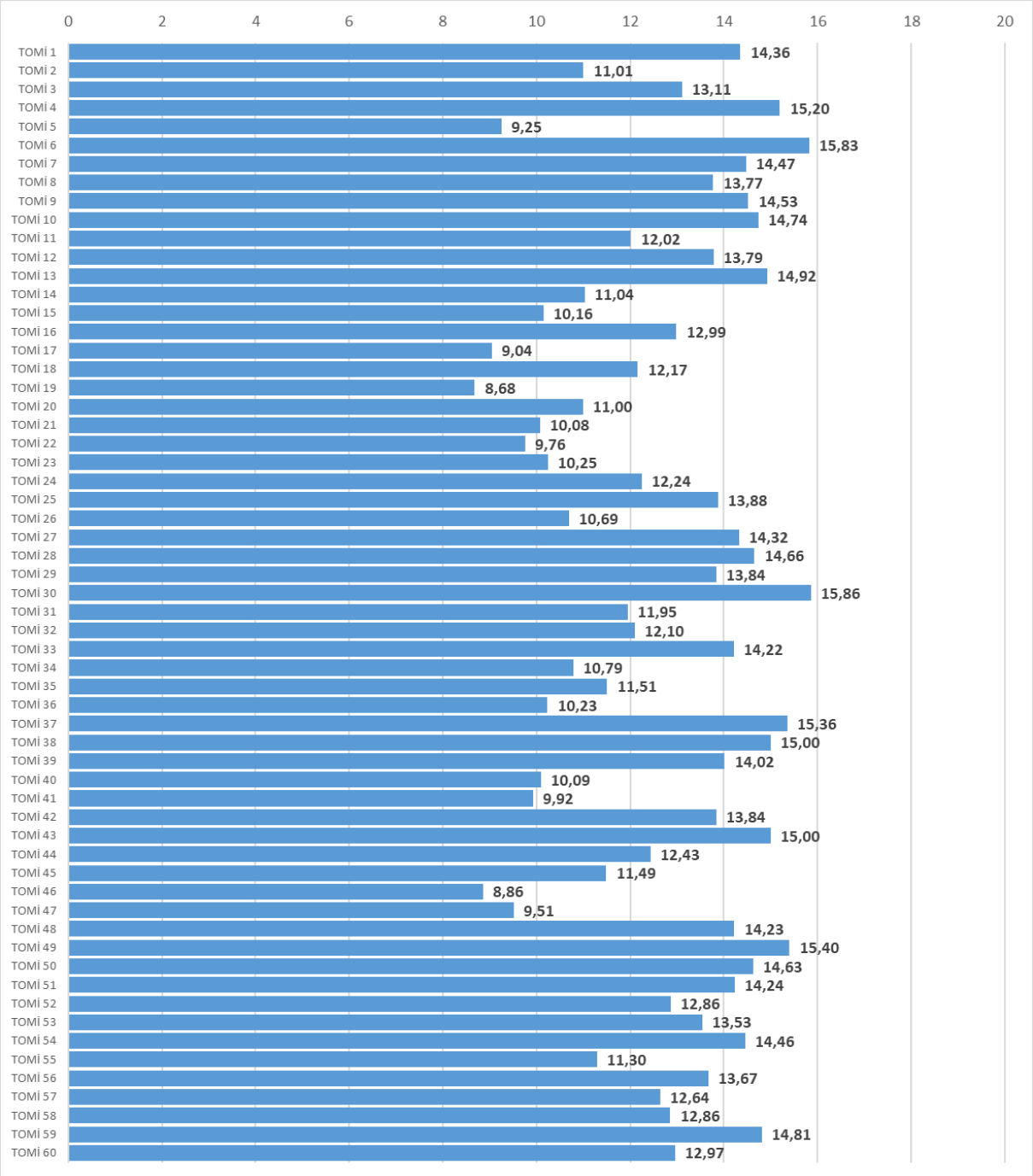


Table 4. *Descriptive statistics of TOMI-2.*

	Statistics	Early Subscale	Basic Subscale	Advanced Subscale	Emotion recognition Subscale	Mental State Term Comprehension Subscale
Early						
Mean±SD	14.36±2.44					
M (*min-max*)	14.9 (4-19)					
Basic						
Mean±SD	13.38±2.46	*rho*=0.871 *p*<0.001				
M (*min-max*)	13.6 (4-20)					
Advanced						
Mean±SD	11.00±2.43	*rho*=0.643 *p*<0.001	*rho*=0.744 *p*<0.001			
M (*min-max*)	10.9 (3-18)					
Emotion recognition						
Mean±SD	13.35±2.52	*rho*=0.882 *p*<0.001	*rho*=0.826 *p*<0.001	*rho*=0.763 *p*<0.001		
M (*min-max*)	13.5 (4-18)					
Mental State Term Comprehension		*rho*=0.789 *p*<0.001	*rho*=0.877 *p*<0.001	*rho*=0.566 *p*<0.001	*rho*=0.728 *p*<0.001	
Mean±SD	14.3±3.23					
M (*min-max*)	15.1 (4-36)					
Pragmatics						
Mean±SD	11.34±2.72	*rho*=0.612 *p*<0.001	*rho*=0.73 *p*<0.001	*rho*=0.899 *p*<0.001	*rho*=0.659 *p*<0.001	*rho*=0.558 *p*<0.001
M (*min-max*)	11.5 (3-18)					

rho: Pearson Correlation Coefficient; Summary statistics are given as *mean* ± *standard* value. Bolded sections are statistically significant ($p<0.05$).

3.1. Small Group Practice

The Theory of Mind Inventory-2 (TOMI-2) was first administered face-to-face to 20 participants aged 29-49 with children aged 3-5 years. The participants were asked whether the items in the inventory were clearly understood. All participants who participated in the small group application stated that all items in the inventory were clearly expressed and that there was no need for correction.

3.2. Content Validity

For the content validity of the TOMI-2, for which validity and reliability analyses were conducted for parents with 3-5-year-old children, expert opinions were obtained from 5 academicians (2 of them have a bachelor's degree in preschool teaching), 1 with a doctorate in guidance and counseling and 4 with a doctorate in child development and education. All experts reported that the items in the TOMI-2 were necessary and appropriate. Therefore, all items in the original form were used in the data collection process.

3.3. Exploratory Factor Analysis

In this study, exploratory factor analysis was first conducted to assess the construct validity of the TOMI-2. Table 5 shows that the TOMI-2 consists of 2 sections. The first part includes early, basic, and advanced subscales, while the second part includes emotion recognition, mental state term comprehension, and pragmatic subscales. As seen in Table 5, the three-factor structure in the first part explains 64.71% of the total variance, while the three-factor structure in the second part explains 54.91%. The Cronbach Alpha reliability coefficients of the whole inventory and its subscales are also high. The KMO value between 0.90 and 1.00 evaluates the sample adequacy as "very good" (Alpar, 2022, p.625). The Kaiser Meyer Olkin coefficient (KMO) of the TOMI-2 was calculated as 0.94 and the sample was found to be adequate. As a result of the explanatory factor analysis, it is seen that the TOMI-2 is a valid and reliable measurement tool.

Table 5. *Exploratory factor analysis results of TOMI-2.*

Part 1							Part 2						
Factor	Item No	Factor Loads			Explained Variance %	Cronbach Alpha	Factor	Item No	Factor Loads			Explained Variance %	Cronbach Alpha
		1	2	3					1	2	3		
Early	3	0.498			14.65	0.930	Emotion Recognition	6	0.643			21.15	0.904
	6	0.683						17	0.481				
	24	0.478						25	0.607				
	25	0.650						48	0.734				
	28	0.624						49	0.775				
	37	0.638						50	0.872				
	38	0.688						32	0.551				
	43	0.746						51	0.741				
	44	0.513						52	0.645				
	48	0.735						55	0.499				
	49	0.803					Mental State Term Comprehension	7		0.736		17.03	0.895
	50	0.778						10		0.736			
	54	0.672						39		0.691			
	59	0.680						33		0.681			
Basic	1		0.641		17.22	0.949		54		0.689			
	4		0.693					53		0.636			
	7		0.707				Pragmatics	2			0.548	16.74	0.806
	8		0.590					13			0,520		
	9		0.617					14			0,501		
	10		0.685					18			0.608		
	11		0.710					19			0.763		
	12		0.572					20			0.662		
	15		0.563					35			0.562		
	16		0.550					36			0.568		
	26		0.528					45			0.521		
	29		0.604									Scale 54.91	0.929
	30		0.678				*KMO*=0.941 *Df*=300 χ^2=4173.421 *p*<0.001						
	31		0.548										
	32		0.578										
	33		0.612										

	35	0.552		
	39	0.648		
	42	0.645		
	51	0.632		
	53	0.574		
	57	0.539		
	60	0.545		
Advan-ced	2	0.525	14.84	0.918
	5	0.626		
	13	0.512		
	14	0.502		
	17	0.545		
	18	0.479		
	19	0.753		
	20	0.496		
	21	0.614		
	22	0.647		
	23	0.710		
	27	0.472		
	34	0.597		
	36	0.513		
	40	0.533		
	41	0.587		
	45	0.466		
	46	0.620		
	47	0.556		
	52	0.525		
	55	0.524		
	56	0.517		
	58	0.508		
		Scale	64.71	0.964
KMO=0.941 *Df*=1770 χ^2=11183.489 *p*<0.001				

KMO: Kaiser–Meyer–Olkin test; *Df*: Degrees of Freedom

3.4. Confirmatory Factor Analysis

Confirmatory Factor Analysis was conducted with the data obtained from each subscale for the construct validity of the TOMI-2. The boundary values in CFA analysis (Schumacker & Lomax, 2004; Hu & Bentler, 1999; Thompson, 2004; Kline 2015) were evaluated according to Table 6.

Table 6. *Boundary values in CFA analysis.*

Indices	Boundary Values
χ^2/SD	Perfect ≤ 3≤ Good ≤ 5
RMSEA	Perfect ≤ 0.05 ≤ Good ≤ 0.08
SRMR	Perfect ≤ 0.05 ≤ Good ≤ 0.08
CFI	Perfect ≥ 0.95 ≥ Good ≥ 0.90
NNFI	Perfect ≥ 0.95 ≥ Good ≥ 0.90
GFI	Perfect ≥ 0.95 ≥ Good ≥ 0.90
AGFI	Perfect ≥ 0.95 ≥ Good ≥ 0.90

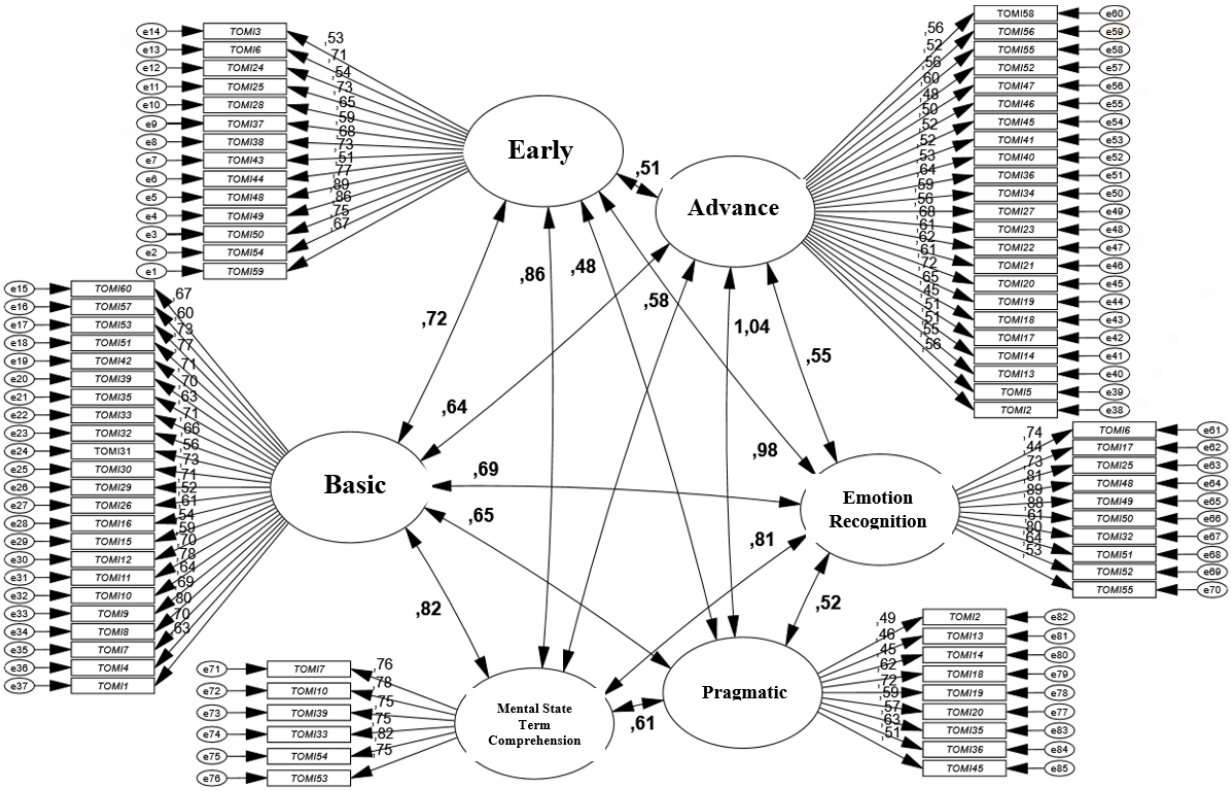
The model (χ^2=20043.471 *df*=3450) obtained as a result of the factor analysis explained in Table 7 includes a total of 6 subscales of the TOMI-2. The fit indices show that the model is an acceptable fit. The first part of the TOMI-2 consists of 60 items and 3 subscales, while the second part consists of 25 items and 3 subscales. The interactions between the two parts and the model created for the TOMI-2 are presented visually in Figure 2.

Table 7. *Statistical values for the model fit of TOMI-2.*

Measurement	(χ^2/SD)	RMSEA	IFI	CFI	GFI	SRMR
Early	2.184	0.062	0.965	0.964	0.928	0.043
Basic	1.817	0.051	0.947	0.953	0.901	0.047
Advanced	2.209	0.063	0.902	0.901	0.877	0.060
Emotion recognition	3.050	0.080	0.963	0.963	0.938	0.057
Mental State Term Comprehension	2.136	0.061	0.991	0.991	0.981	0.030
Pragmatics	0.964	0.001	0.992	0.999	0.982	0.042

Figure 2 shows the factor loadings of the 6 subscales of the TOMI-2. Accordingly, factor loadings ranged between 0.51 and 0.89 in the Early Subscale, 0.52 and 0.80 in the Basic Subscale, 0.45 and 0.72 in the Advanced Subscale, 0.44 and 0.89 in the Emotion Recognition Subscale, 0.75 and 0.82 in the Mental State Term Comprehension Subscale, and 0.45 and 0.72 in the Pragmatic Subscale.

Figure 2. *Confirmatory factor analysis model for the TOMI-2.*



3.5. Findings Regarding Criterion Validity

Theory of Mind Task Battery (TOMTB), which is used to evaluate children's theory of mind skills, was used for criterion validity. TOMTB is a 15-item battery that evaluates the theory of mind in 3 subscales: early, basic, and advanced (Keleş Ertürk & Tepeli, 2023). According to Table 8, there is a highly statistically significant relationship between the Early, Basic, Advanced, Emotion Recognition, Mental State Term Comprehension, and Pragmatics subscales of the TOMI-2, the Theory of Mind Task Battery (TOMTB) Early, Basic, Advanced subscales and the TOMTB total score.

Table 8. *Findings of the criterion validity of the TOMI-2.*

	Early	Basic	Advanced	TOMTB
Early	*rho*=0.327 *p*<0.001	*rho*=0.227 *p*<0.001	*rho*=0.193 *p*<0.001	*rho*=0.295 *p*<0.001
Basic	*rho*=0.299 *p*<0.001	*rho*=0.190 *p*<0.001	*rho*=0.204 *p*<0.001	*rho*=0.276 *p*<0.001
Advanced	*rho*=0.188 *p*<0.001	*rho*=0.158 *p*=0.005	*rho*=0.161 *p*=0.005	*rho*=0.215 *p*<0.001
Emotion Recognition	*rho*=0.291 *p*<0.001	*rho*=0.218 *p*<0.001	*rho*=0.142 *p*=0.012	*rho*=0.254 *p*<0.001
Mental State Term Comprehension	*rho*=0.252 *p*<0.001	*rho*=0.183 *p*=0.001	*rho*=0.252 *p*<0.001	*rho*=0.293 *p*<0.001
Pragmatics	*rho*=0.218 *p*<0.001	*rho*=0.180 *p*=0.001	*rho*=0.196 *p*<0.001	*rho*=0.253 *p*<0.001

3.6. Reliability and Item Analyses of the TOMI-2

The Cronbach Alpha coefficients of the TOMI-2 were calculated as 0.930 for the Early Subscale, 0.904 for the Emotion Recognition Subscale, 0.895 for the Mental State Term Comprehension Subscale, 0.949 for the Basic Subscale, 0.806 for the Pragmatics Subscale and 0.918 for the Advanced Subscale. For the test-retest reliability of the TOMI-2, 30 participants were interviewed again 3 weeks later. According to the results of the analysis, the test-retest reliability (Table 9) ranged between 0.76 and 0.98 (p<0.05). In this case, it can be said that the reliability of the measurements obtained in terms of consistency is good and very good.

Table 9. *Test-Retest results of TOMI-2 on a subscale basis.*

	Test-retest reliability
Early	0.963
Basic	0.938
Advanced	0.979
Emotion recognition	0.960
Mental State Term Comprehension	0.758
Pragmatics	0.961

CR-Composite Reliability values should be examined for the reliability of the CFA model, and convergent and discriminant validity should be examined for its validity (Çalık et al., 2013). Since the Composite Reliability (CR) value for each factor should exceed 0.7, it can be concluded that the reliability of the CFA model is ensured (Hair et al., 2018). If the CR value is higher than 0.7, it is accepted that the AVE value is greater than 0.4 and it is stated that convergent validity is not impaired (Huang et al, 2013; Fornel & Larcker, 1981; Karadeniz & Kocamaz, 2020; Biçer & Kılıç, 2022). Accordingly, in Table 10, the CR values of the CFA model are between 0.927 and 0.960, while the AVE values are between 0.410 and 0.589. The CR and AVE values prove that the measurement model shows good fit validity.

Table 10. *Findings on CR-Composite reliability values.*

	N	AVE	CR
Early	14	0.485	0.928
Basic	29	0.458	0.960
Advanced	33	0.410	0.958
Emotion recognition	20	0.519	0.954
Mental State Term Comprehension	12	0.589	0.945
Pragmatics	18	0.419	0.927

The effects between the TOMI-2 items and its subscales are given in in Appendix (see Table A1). Table A1 shows that each of the path coefficients of the subscales in the first part of 60 items is statistically significant (p<0.05). Accordingly, the Early subscale consists of item 3, 6, 24, 25, 28, 37, 38, 43, 44, 48, 49, 50, 54 and 59. The Basic subscale consists of item 1, 4, 7, 8, 9, 10, 10, 11, 12, 15, 16, 26, 29, 30, 31, 32, 33, 35, 39, 42, 51, 53, 57 and 60. The Advanced subscale consists of item 2, 5, 13, 14, 17, 18, 18, 19, 20, 21, 22, 23, 27, 34, 36, 40, 41, 45, 46, 47, 52, 55, 56 and 58. Each of the path coefficients of the subscales in the second part on 25 items is statistically significant (p<0.05). Accordingly, the Early subscale consists of item 3, 6, 24, 25, 28, 37, 38, 43, 44, 48, 49, 50, 54 and 59. The Emotion Recognition subscale consists of items 6, 17, 25, 48, 49, 50, 32, 51, 52 and 55. The Mental State Term Comprehension subscale consists of items 7, 10, 39, 33, 54, and 53. The Pragmatic subscale consists of item 2, 13, 14, 18, 19, 20, 35, 36 and 45. All subscales have a highly statistically significant effect on the item.

The findings regarding the evaluation of the effects between the subscales of the TOMI-2 are given in Table 11. When Table 11 was analyzed, it was found that the relationships between the Early, Basic, Advanced, Emotion Recognition, Mental State Term Comprehension, and Pragmatics subscales of the TOMI-2 were statistically significant.

Item-total correlations for TOMI-2 were also calculated and are given in the appendix (see Table A2). Table A2 shows the item total correlations for the TOMI-2 which range between 0.325 and 0.603. According to Tavşancıl (2002), item-test correlations for the items in the scale are recommended to be 0.30 and above. The values obtained for the TOMI-2 also meet this criterion.

Table 11. *Evaluation of the effects between the subscales of the TOMI-2.*

			zβ	*β*	*se*	*t*	*p*
Early	<->	Basic	0.719	5.286	0.712	7.422	<0.001
Basic	<->	Emotion Recognition	0.688	5.439	0.722	7.536	<0.001
Advanced	<->	Emotion Recognition	0.553	3.97	0.634	6.257	<0.001
Basic	<->	Advanced	0.639	5.283	0.814	6.492	<0.001
Advanced	<->	Pragmatics	1.044	7.689	1.079	7.126	<0.001
Early	<->	Advanced	0.508	3.388	0.577	5.867	<0.001
Basic	<->	Mental State Term Comprehension	0.821	7.197	0.883	8.151	<0.001
Advanced	<->	Mental State Term Comprehension	0.583	4.635	0.723	6.413	<0.001
Early	<->	Mental State Term Comprehension	0.858	6.069	0.729	8.325	<0.001
Mental State Term Comprehension	<->	Pragmatics	0.608	4.75	0.712	6.673	<0.001
Basic	<->	Pragmatics	0.646	5.239	0.785	6.671	<0.001
Early	<->	Pragmatics	0.481	3.147	0.55	5.727	<0.001
Emotion Recognition	<->	Pragmatics	0.516	3.637	0.596	6.1	<0.001
Early	<->	Emotion Recognition	0.979	6.252	0.727	8.594	<0.001
Emotion Recognition	<->	Mental State Term Comprehension	0.808	6.149	0.725	8.485	<0.001

β: Regression coefficient, se: Standard error, zβ: Standardized regression coefficient. Bolded sections are statistically significant (p<0.05).

4. DISCUSSION and CONCLUSION

Theory of mind forms the basis of the ability to interpret people's communication and actions and also includes the understanding that there are different perspectives (Astington, 2020). The individual also uses the theory of mind when considering the feelings and thoughts of others (Astington & Edward, 2010). Especially in the preschool period, interactions with parents and siblings, and cultural-social-speech experiences shape ToM and can cause significant differences in the developmental stages of ToM (Wellman, 2014; Slaughter et al., 2015). This study was planned to evaluate and support the development process of ToM in preschool and daily life. In this study, the Turkish adaptation, validity, and reliability study of the Theory of Mind Inventory-2 (TOMI-2) developed by Hutchins and Prelock (2016) for 3-5-year-old children was conducted.

It is seen that there is no comprehensive, progressive, and up-to-date assessment of the theory of mind in the national literature, but the national literature also focuses on the development of

the theory of mind. Therefore, it was concluded that adapting the Theory of Mind Inventory-2 to Turkish culture was appropriate. Within the scope of the study, first of all, the necessary permissions for the use of the TOMI-2 were obtained and the process started with its translation into Turkish. Forward and backward translation techniques were used in the translation process, and a different group of translators examined the adapted test. The comprehensibility of the statements was also ensured by conducting a pilot study with a small group. Field experts were also consulted and feedback was received that no item should be removed from the inventory. In the analysis of the data, the normality assumption was examined with the "Shapiro-Wilk" test. The relationship between two continuous variables was evaluated with Pearson Correlation Coefficient. Exploratory Factor Analysis, Confirmatory Factor Analysis, Content Validity, Criterion Validity, and Reliability analyses were also evaluated.

Content validity is determined by applying expert opinion to determine whether the items in the measurement tool are appropriate for measurement (Karasar, 2017, p.195). For this purpose, the opinions of field experts were obtained for the content validity of TOMI-2 and no changes were deemed necessary in the original form.

Kaiser Meyer Olkin coefficient (KMO) was calculated to test the sample adequacy and it was found to be 0.94. A KMO value between 0.90 and 1.00 evaluates the sampling adequacy as "very good" (Alpar, 2022, p.625). When Table 5 is examined, it is seen that TOMI-2 consists of 2 sections. The first part includes early, basic, and advanced subscales, while the second part includes emotion recognition, mental state term comprehension, and pragmatic subscales. The three-factor structure in the first part explains 64.71% of the total variance, while the three-factor structure in the second part explains 54.91%. Generally, an explained variance between 0.50 and 0.70 is considered sufficient. In social sciences, an explained variance between 0.40 and 0.60 is considered acceptable (Alpar, 2022, p.633). In this case, it is possible to evaluate the explained variance of TOMI-2 as sufficient.

The fit indices show that the model has an acceptable level of fit. The factor loadings of the TOMI-2 ranged between 0.51 and 0.89 in the Early Subscale; 0.52 and 0.80 in the Basic Subscale; 0.45 and 0.72 in the Advanced Subscale; 0.44 and 0.89 in the Emotion Recognition Subscale; 0.75 and 0.82 in the Mental State Term Comprehension Subscale; and 0.45 and 0.72 in the Pragmatic Subscale. Factor loadings of 0.60 and above are considered to be high, while loadings between 0.30-0.59 are considered to be moderate (Büyüköztürk, 2002). Accordingly, it can be said that the factor loadings of the TOMI-2 are at high and medium levels.

Criterion validity involves comparing a test that is believed to measure performance, skill, etc., against a standard or another test that measures the same characteristic. (Alpar, 2022, p.536). For this purpose, the Theory of Mind Task Battery (TOMTB) was used for criterion validity, and it was found that there was a highly statistically significant relationship between the TOMI-2 and the Theory of Mind Task Battery (TOMTB).

Reliability, which is also expressed as the stability between independent measurements, can be tested in different ways (Thanasegaran, 2009; Alpar, 2022, p.532). Internal consistency, test-retest reliability, and composite reliability were calculated to test the reliability of the TOMI-2. The Cronbach Alpha reliability coefficients of the TOMI-2 ranged between 0.806 and 0.949, and the test-retest reliability ranged between 0.76 and 0.98. When the Cronbach Alpha reliability coefficient is between 0.60 and 0.79, it is interpreted as highly reliable; when it is between 0.80 and 1.00, it is interpreted as highly reliable (Karagöz, 2019, p.1003). In this case, it can be said that the reliability of the measurements obtained in terms of stability is also highly reliable and highly reliable.

The combined reliability (CR-Composite Reliability) values for the reliability of the CFA model are between 0.927 and 0.960, while the AVE values are between 0,.410, and 0.589. The CR

and AVE values prove that the measurement model shows fit validity. Since the CR (Combined Reliability) value should be greater than 0.7 for each factor, it can be said that the reliability of the CFA model is ensured (Hair et al., 2018). If the CR value is greater than 0.7, the AVE value is accepted to be greater than 0.4, and convergent validity is not impaired (Huang cd., 2013; Fornel & Larcker, 1981; Karadeniz & Kocamaz, 2020; Biçer & Kılıç, 2022).

The item-total correlations of the TOMI-2 ranged between 0.325 and 0.603. According to Tavşancıl (2002), item-test correlations for the items in the scale are recommended to be 0.30 and above. The values obtained for TOMI-2 also meet this criterion.

As a result of the analyses, it was determined that the validity and reliability values of the Turkish version of the ToM Inventory-2 with 60 items and 6 subscales by the original model showed an acceptable fit and can be used in Turkish culture. Accordingly, the TOMI-2 can be used to assess children's ToM skills according to parents' views. ToM is a concept that has the power to affect both the social and cognitive development of the individual. The use of the TOMI-2 in children's ToM skills provides a detailed evaluation of ToM. The scores obtained from the inventory reveal at which stage the child is in ToM, and according to the results obtained, guidance can be provided to prepare a supportive environment. However, a limitation of this study is that it was restricted to children aged 3-5 and involved a relatively small sample size. Therefore, it is recommended to conduct validity and reliability analyses of the inventory with different age groups and a larger sample.

Declaration of Conflicting Interests and Ethics

The authors declare no conflict of interest. This research study complies with research publishing ethics. The scientific and legal responsibility for manuscripts published in IJATE belongs to the authors **Ethics Committee Number**: KTO Karatay University Non-Pharmaceutical and Medical Device Research Ethics Committee, 2023/043.

Contribution of Authors

Canan Keleş Ertürk: Investigation, Resources, Visualization, Software, Formal Analysis, and Writing-original draft. **Kezban Tepeli**: Methodology, Supervision, and Validation.

Orcid

Canan Keleş Ertürk https://orcid.org/0000-0001-6247-0073
Kezban Tepeli https://orcid.org/0000-0003-3403-3890

REFERENCES

Alpar, R. (2022). *Spor sağlık ve eğitim bilimlerinden örneklerle uygulamalı istatistik ve geçerlik güvenirlik SPSS de çözümleme adımları ile birlikte* [*Applied statistics and validity and reliability with examples from sports, health and education sciences with analysis steps in SPSS*]. Detay Publishing.

Altıntaş, M. (2014). Çocuklar için Zihin Kuramı Test Bataryası'nın 4-5 yaş türk çocuklarına uyarlanması, geçerlik güvenirlik çalışması [Adaptation of Theory of Mind Task Battery for Children to 4-5 year old Turkish children, validity and reliability study]. *Master Thesis*, Haliç University.

Astington, J.W. (2020). The developmental interdependence of theory of mind and language. In N.J. Enfield, & S.C. Levinson (Eds.), *Roots of Human Sociality Culture, Cognition and Interaction* (p.179- 206). Routledge.

Astington, J.W., & Dack, L.A. (2008). Theory of mind. In M.M. Haith, & J.B. Benson (Eds.), *Encyclopedia of Infant and Early Childhood Development Volume 1* (p. 343-356). Elsevier Inc.

Astington, J.W., & Edward, M.J. (2010). The development of theory of mind in early childhood. *Social Cognition*.

Biçer, M., & Kılıç, K.C. (2022). Yönetici Davranışları Ölçeğinin Türkçe'ye uyarlanması: geçerlilik ve güvenilirlik uyarlaması [Adaptation of the Managerial Behaviors Scale into Turkish: validity and reliability adaptation]. *Karamanoglu Mehmetbey University Journal of Social and Economic Research, 24*(42), 277-291.

Brown, T.A. (2015). *Confirmatory factor analysis for applied research.* Guilford publications.

Büyüköztürk, Ş. (2002). Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı *[*Factor analysis: Basic concepts and use in scale improvement*]. Educational Administration in Theory and Practice, 22*, 470-483.

Cerny, B.A., & Kaiser, H.F. (1977). A study of a measure of sampling adequacy for factor-analytic correlation matrices. *Multivariate Behavioral Research, 12*(1), 43-47.

Çalık, M., Altunışık, R., & Sütütemiz, N. (2013). Bütünleşik pazarlama iletişimi, marka performansı ve pazarlama performansı ilişkisinin incelenmesi [Analyzing the relationship between integrated marketing communication, brand performance and marketing performance]. *International Journal of Management Economics and Business*, *9*(19), 137162.

Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2018). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve Lisrel uygulamaları* (5. Baskı) [*Multivariate statistics for social sciences: SPSS and Lisrel applications* (5th Edition)]. Pegem Academy Publishing.

Flavell, J.H. (1999). Cognitive development: Children's knowledge about the mind. *Annual Reviews Psychology, 50*, 21-45.

Fornel, C., & Larcker, D.F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50.

Gözün Kahraman, Ö. (2012). *Zihin kuramına dayalı eğitim programının 48-60 aylık çocukların bilişsel bakış açısı becerileri ve prososyal davranışları üzerindeki etkisinin incelenmesi* [*Investigating the effect of a theory of mind-based education program on cognitive perspective skills and prosocial behaviors of 48-60 month old children*] [Doctoral Thesis], Gazi University.

Hair, J.F., Black, W.C., Babin, B.J., & Anderson, R.E. (2018). *Multivariate data analysis* (8. Edition). Cangage Learning.

Hair, J.F., Anderson, R.E., Tatham, R.L., & Black, W.C., (1998). *Multivariate data analysis*, Prentice Hall.

Hambleton, R.K., & Patsula, L. (1998). Adapting tests for use in multiple languages and cultures. *Social Indicators Research*, 45, 153-171.

Hambleton, R.K., & Patsula, L. (1999). Increasing the validity of adapted tests: Myths to be avoided and guidelines for improving test adaptation practices. *Journal of Applied Testing Technology, 1*, 1-30.

Houssa, M., Mazzone, S., & Nader-Grosbois, N. (2014). Validation d'une version francophone de l'inventaire de la Théorie de l'Esprit (ToMI-vf). *European Review of Applied Psychology, 64*, 169-179. http://dx.doi.org/10.1016/j.erap.2014.02.002

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55.

Huang, C.C., Wang, Y.M., Wu, T.W., & Wang, P.A. (2013). An empirical analysis of the antecedents and performance consequences of using the moodle platform. *International Journal of Information and Education Technology, 3*(2), 217-221.

Hutchins, T.L., & Prelock, P. (2010). *Technical manual for the Theory of Mind Task Battery*. Unpublished Copyrighted Manuscript. Available at: theoryofmindinventory.com

Hutchins, T.L., & Prelock, P.A. (2016). *Technical manual for the Theory of Mind Inventory-2*. Unpublished Copyrighted Manuscript. Available at: theoryofmindinventory.com.

Karadeniz, M., & Kocamaz, İ. (2020). An investigation of post-purchase cognitive dissonance and its determinants in online shopping. *Journal of Yasar University, 15*, 307-315.

Karagöz, Y. (2019). *SPSS-AMOS-META uygulamalı istatistiksel analizler* [*SPSS-AMOS-META applied statistical analysis*]. Nobel Publishing.

Karasar, N. (2017). *Bilimsel araştırma yöntemi: Kavramlar ilkeler teknikler* (32. Basım) [*Scientific research method: Concepts principles techniques* (32nd Edition)]. Nobel Publishing.

Keleş Ertürk, C., & Tepeli, K (2023). Validity and reliability of Theory of Mind Task Battery (TOMTB) for 3-5 year old children. *International Anatolian Journal of Social Sciences*, *7*(3), 627-639. https://doi.org/10.47525/ulasbid.1322818

Kılıç Tülü, B., & Ergül, C. (2022). 3-5 yaş grubu çocuklara yönelik Zihin Kuramı Testi: Geçerlik güvenirlik çalışması [Theory of Mind Test for 3-5 year old children: Validity and reliability study]. *Marmara University Atatürk Education Faculty Journal of Educational Sciences, 55*, 31-61. https://doi.org/10.15285/maruaebd.966350

Kline, R.B. (2015). *Principles and practice of structural equation modeling* (4. Edition). Guilford Publications.

Lee, S.C., Fu, I.N., Liu, M.R., Yu, T.Y., & Chen, K.L. (2023). Factorial validity of the Theory of Mind Inventory-2 in typically developing children. *Journal of Autism and Developmental Disorders*, 53, 310-318. https://doi.org/10.1007/s10803-022-05426-0

McElwain N.L, Ravindran N., Emery H.T., & Swartz R. (2019). Theory of mind as a mechanism linking mother–toddler relationship quality and child–friend interaction during the preschool years. *Social Development, 28*, 998-1015. https://doi.org/10.1111/sode.12377

Miller, S.A. (2016). *Parenting and theory of mind.* Oxford University Press.

Prelock, P.A., Hutchins, T.L., & Bonazinga Bouyea, L. (2016). *Theory of Mind Inventory-2 guide to clinical decision-making.* Unpublished Copyrighted Manuscript.

Rakoczy, H. (2022). Foundations of theory of mind and its development in early childhood. *Nature Reviews Psychology, 1*(4), 223-235. https://doi.org/10.1038/s44159-022-00037-z

Ratcliffe, M. (2007). *Rethinking commonsense psychology: A critique of folk psychology, theory of mind and simulation*. Palgrave Macmillan.

Rencher, A.C. (2002). *Methods of multivariate analysis* (2. Edition), John Wiley & Sons, Inc.

Scholl, B.J., & Leslie A.M. (1999). Modularity, development, and 'theory of mind. *Mind & Language, 14*(1), 131-153.

Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling (2. Edition)*. Lawrence Erlbaum Associates.

Slaughter, V., & Repacholi, B. (2003). Introduction: Individual differences in theory of mind what are we investigating?. In B. Repacholi & Virginia Slaughter (Eds.), *Individual Differences in Theory of Mind: Implications for Typical and Atypical Development* (p. 1-13). Psychology Press.

Slaughter, V., Imuta, K., Peterson, C.C., & Henry, J.D. (2015). Meta-Analysis of theory of mind and peer popularity in the preschool and early school years. *Child Development, 86*(4), 1159-1174. https://doi.org/10.1111/Cdev.12372

Sodian, B. (2005). Theory of mind—the case for conceptual development. In W. Schneider, R. Schumann-Hengsteler & B. Sodian (Eds.), *Young Children's Cognitive Development Interrelationships Among Executive Functioning, Working Memory, Verbal Ability, And Theory of Mind* (p. 95-130). Lawrence Erlbaum Associates.

Sodian, B., & Kristen, S. (2010). Theory of mind. *In* B.M. Glatzeder, V. Goel & A.V. Müller (Eds.), *Towards A Theory of Thinking Building Blocks for A Conceptual Framework* (p. 189-202). Springer.

Sönmez, V., & Alacapınar, F.G. (2018). *Örneklendirilmiş bilimsel araştırma yöntemleri* (Genişletilmiş 6. Baskı) [*Exemplified scientific research methods* (Expanded 6th Edition)]. Anı Publishing.

Şahin, B., Bozkurt, A., Usta, M.B., Aydın, M., Çobanoğlu, C., & Karabekiroğlu, K. (2019). Zihin kuramı: Gelişim, nörobiyoloji, ilişkili alanlar ve nörogelişimsel bozukluklar [Theory of mind: Development, neurobiology, related areas and neurodevelopmental disorders]. *Current Approaches in Psychiatry*, *11*(1), 24-41. https://doi.org/10.18863/pgy.390629

Tavşancıl, E. (2002). Tutumların ölçülmesi ve spss ile veri analizi [Measurement of attitudes and data analysis with spss]. Nobel Publishing.

Thanasegaran, G. (2009). Reliability and Validity issues in research. *Integration & Dissemination, 4*, 35-40.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications (1. Edition).* Psychology Press.

Tucci, S.L. (2023, November 11). *Theory of mind development and children who are deaf and hard of hearing.* Victorian Deaf Education Institute. https://www.deafeducation.vic.edu.au/Resource/fact-sheets/Theory of Mind Summary.pdf

Wellman, H.M. (2017). The development of theory of mind: Historical reflections. *Child Development Perspectives, 11*(3), 207-214. https://doi.org/10.1111/Cdep.12236

Wellman, H.M., & Liu, D. (2004). Scaling of theory of mind tasks. *Child Development, 75*(2), 523 – 541.

Wellman, H.M. (2014). *Making minds how theory of mind develops.* Oxford University Press.

Westby, C., & Robinson, L. (2014). A developmental perspective for promoting theory of mind. *Topics In Language Disorders, 34*(4), 362-382. https://doi.org/10.1097/TLD.0000000000000035

Zufferey, S. (2010). *Lexical pragmatics and theory of mind the acquisition of connectives. Volume 201*. Benjamins Publishing Company.

APPENDIX

Table A1. *Evaluation of the effects between the items and subscales in the TOMI-2.*

			zβ	*β*	*se*	*t*	*p*
TOMI2_59	←	Early	0.673	1.000			
TOMI2_54	←	Early	0.752	1.178	0.096	12.260	<0.001
TOMI2_50	←	Early	0.857	1.214	0.088	13.748	<0.001
TOMI2_49	←	Early	0.885	1.128	0.080	14.139	<0.001
TOMI2_48	←	Early	0.771	1.138	0.091	12.543	<0.001
TOMI2_44	←	Early	0.514	0.855	0.099	8.633	<0.001
TOMI2_43	←	Early	0.730	0.981	0.082	11.944	<0.001
TOMI2_38	←	Early	0.681	0.911	0.081	11.219	<0.001
TOMI2_37	←	Early	0.590	0.850	0.086	9.829	<0.001
TOMI2_28	←	Early	0.654	0.920	0.085	10.819	<0.001
TOMI2_25	←	Early	0.730	1.052	0.088	11.944	<0.001
TOMI2_24	←	Early	0.541	0.736	0.081	9.070	<0.001
TOMI2_6	←	Early	0.713	0.839	0.072	11.694	<0.001
TOMI2_3	←	Early	0.533	0.695	0.078	8.933	<0.001
TOMI2_51	←	Basic	0.772	1.057	0.086	12.348	<0.001
TOMI2_42	←	Basic	0.707	1.020	0.089	11.431	<0.001
TOMI2_39	←	Basic	0.704	0.997	0.087	11.398	<0.001
TOMI2_35	←	Basic	0.627	0.980	0.095	10.264	<0.001
TOMI2_33	←	Basic	0.709	0.989	0.086	11.459	<0.001
TOMI2_32	←	Basic	0.656	0.984	0.092	10.693	<0.001
TOMI2_31	←	Basic	0.558	0.920	0.100	9.222	<0.001
TOMI2_30	←	Basic	0.730	1.030	0.088	11.764	<0.001
TOMI2_29	←	Basic	0.713	1.029	0.089	11.516	<0.001
TOMI2_26	←	Basic	0.521	0.806	0.093	8.650	<0.001
TOMI2_16	←	Basic	0.613	0.950	0.094	10.059	<0.001
TOMI2_15	←	Basic	0.536	0.785	0.088	8.879	<0.001
TOMI2_12	←	Basic	0.590	0.984	0.101	9.700	<0.001
TOMI2_11	←	Basic	0.698	0.981	0.087	11.303	<0.001
TOMI2_10	←	Basic	0.779	1.140	0.092	12.450	<0.001
TOMI2_9	←	Basic	0.644	0.899	0.086	10.516	<0.001
TOMI2_8	←	Basic	0.694	0.973	0.086	11.255	<0.001
TOMI2_60	←	Basic	0.668	1.000			
TOMI2_57	←	Basic	0.605	0.901	0.079	11.388	<0.001
TOMI2_53	←	Basic	0.732	1.078	0.091	11.795	<0.001
TOMI2_1	←	Basic	0.629	0.848	0.082	10.297	<0.001
TOMI2_4	←	Basic	0.704	0.909	0.080	11.386	<0.001
TOMI2_7	←	Basic	0.795	1.113	0.088	12.674	<0.001
TOMI2_60	←	Basic	0.668	1.000			
TOMI2_57	←	Basic	0.605	0.901	0.079	11.388	<0.001
TOMI2_53	←	Basic	0.732	1.078	0.091	11.795	<0.001
TOMI2_1	←	Basic	0.629	0.848	0.082	10.297	<0.001
TOMI2_4	←	Basic	0.704	0.909	0.080	11.386	<0.001
TOMI2_7	←	Basic	0.795	1.113	0.088	12.674	<0.001

TOMI2_13	←	Advanced	0.509	1.003	0.132	7.595	<0.001
TOMI2_14	←	Advanced	0.508	1.000	0.132	7.585	<0.001
TOMI2_17	←	Advanced	0.449	0.775	0.112	6.892	<0.001
TOMI2_18	←	Advanced	0.648	1.084	0.120	9.019	<0.001
TOMI2_19	←	Advanced	0.716	1.090	0.114	9.601	<0.001
TOMI2_20	←	Advanced	0.608	0.992	0.115	8.636	<0.001
TOMI2_21	←	Advanced	0.617	0.978	0.112	8.727	<0.001
TOMI2_22	←	Advanced	0.609	0.970	0.112	8.643	<0.001
TOMI2_23	←	Advanced	0.682	1.030	0.111	9.317	<0.001
TOMI2_27	←	Advanced	0.555	0.957	0.118	8.106	<0.001
TOMI2_34	←	Advanced	0.590	0.905	0.107	8.456	<0.001
TOMI2_36	←	Advanced	0.636	1.095	0.123	8.907	<0.001
TOMI2_40	←	Advanced	0.533	0.776	0.099	7.869	<0.001
TOMI2_41	←	Advanced	0.523	0.839	0.108	7.758	<0.001
TOMI2_45	←	Advanced	0.521	1.052	0.136	7.731	<0.001
TOMI2_46	←	Advanced	0.495	0.767	0.103	7.440	<0.001
TOMI2_47	←	Advanced	0.481	0.756	0.104	7.281	<0.001
TOMI2_52	←	Advanced	0.597	0.943	0.110	8.535	<0.001
TOMI2_55	←	Advanced	0.565	0.880	0.107	8.202	<0.001
TOMI2_58	←	Advanced	0.562	0.930	0.114	8.171	<0.001
TOMI2_56	←	Advanced	0.525	0.888	0.114	7.775	<0.001
TOMI2_2	←	Advanced	0.556	1.000			
TOMI2_5	←	Advanced	0.546	0.845	0.087	9.754	<0.001
TOMI2_41	←	Advanced	0.523	0.839	0.108	7.758	<0.001
TOMI2_45	←	Advanced	0.521	1.052	0.136	7.731	<0.001
TOMI2_46	←	Advanced	0.495	0.767	0.103	7.440	<0.001
TOMI2_47	←	Advanced	0.481	0.756	0.104	7.281	<0.001
TOMI2_52	←	Advanced	0.597	0.943	0.110	8.535	<0.001
TOMI2_55	←	Advanced	0.565	0.880	0.107	8.202	<0.001
TOMI2_58	←	Advanced	0.562	0.930	0.114	8.171	<0.001
TOMI2_56	←	Advanced	0.525	0.888	0.114	7.775	<0.001
TOMI2_2	←	Advanced	0.556	1.000			
TOMI2_5	←	Advanced	0.546	0.845	0.087	9.754	<0.001
TOMI2_6	←	Emotion recognition	0.737	1.000			
TOMI2_17	←	Emotion recognition	0.440	0.861	0.112	7.692	<0.001
TOMI2_25	←	Emotion recognition	0.726	1.038	0.080	13.049	<0.001
TOMI2_48	←	Emotion recognition	0.813	1.120	0.076	14.765	<0.001
TOMI2_49	←	Emotion recognition	0.892	1.061	0.065	16.375	<0.001
TOMI2_50	←	Emotion recognition	0.879	1.165	0.072	16.114	<0.001
TOMI2_32	←	Emotion recognition	0.612	1.005	0.093	10.858	<0.001
TOMI2_51	←	Emotion recognition	0.796	1.150	0.080	14.421	<0.001
TOMI2_52	←	Emotion recognition	0.638	1.018	0.090	11.346	<0.001
TOMI2_55	←	Emotion recognition	0.534	0.879	0.093	9.398	<0.001
TOMI2_6	←	Emotion recognition	0.737	1.000			
TOMI2_17	←	Emotion recognition	0.440	0.861	0.112	7.692	<0.001
TOMI2_25	←	Emotion recognition	0.726	1.038	0.080	13.049	<0.001
TOMI2_48	←	Emotion recognition	0.813	1.120	0.076	14.765	<0.001

TOMI2_49	←	Emotion recognition	0.892	1.061	0.065	16.375	<0.001
TOMI2_50	←	Emotion recognition	0.879	1.165	0.072	16.114	<0.001
TOMI2_32	←	Emotion recognition	0.612	1.005	0.093	10.858	<0.001
TOMI2_51	←	Emotion recognition	0.796	1.150	0.080	14.421	<0.001
TOMI2_52	←	Emotion recognition	0.638	1.018	0.090	11.346	<0.001
TOMI2_55	←	Emotion recognition	0.534	0.879	0.093	9.398	<0.001
TOMI2_7	←	Mental State Term Comprehension	0.764	1.000			
TOMI2_10	←	Mental State Term Comprehension	0.777	0.995	0.069	14.328	<0.001
TOMI2_39	←	Mental State Term Comprehension	0.748	1.011	0.074	13.689	<0.001
TOMI2_33	←	Mental State Term Comprehension	0.746	0.986	0.072	13.657	<0.001
TOMI2_54	←	Mental State Term Comprehension	0.818	1.108	0.073	15.213	<0.001
TOMI2_53	←	Mental State Term Comprehension	0.748	1.081	0.079	13.695	<0.001
TOMI2_7	←	Mental State Term Comprehension	0.764	1.000			
TOMI2_10	←	Mental State Term Comprehension	0.777	0.995	0.069	14.328	<0.001
TOMI2_39	←	Mental State Term Comprehension	0.748	1.011	0.074	13.689	<0.001
TOMI2_33	←	Mental State Term Comprehension	0.746	0.986	0.072	13.657	<0.001
TOMI2_54	←	Mental State Term Comprehension	0.818	1.108	0.073	15.213	<0.001
TOMI2_53	←	Mental State Term Comprehension	0.748	1.081	0.079	13.695	<0.001
TOMI2_20	←	Pragmatics	0.594	1.000			
TOMI2_19	←	Pragmatics	0.719	1.111	0.105	10.567	<0.001
TOMI2_18	←	Pragmatics	0.624	1.046	0.110	9.533	<0.001
TOMI2_14	←	Pragmatics	0.450	0.882	0.121	7.312	<0.001
TOMI2_13	←	Pragmatics	0.456	0.809	0.109	7.406	<0.001
TOMI2_2	←	Pragmatics	0.494	0.941	0.119	7.921	<0.001
TOMI2_35	←	Pragmatics	0.567	0.981	0.111	8.850	<0.001
TOMI2_36	←	Pragmatics	0.635	1.118	0.116	9.659	<0.001
TOMI2_45	←	Pragmatics	0.513	1.010	0.124	8.172	<0.001
TOMI2_20	←	Pragmatics	0.594	1.000			
TOMI2_19	←	Pragmatics	0.719	1.111	0.105	10.567	<0.001
TOMI2_18	←	Pragmatics	0.624	1.046	0.110	9.533	<0.001
TOMI2_14	←	Pragmatics	0.450	0.882	0.121	7.312	<0.001
TOMI2_13	←	Pragmatics	0.456	0.809	0.109	7.406	<0.001
TOMI2_2	←	Pragmatics	0.494	0.941	0.119	7.921	<0.001
TOMI2_35	←	Pragmatics	0.567	0.981	0.111	8.850	<0.001
TOMI2_36	←	Pragmatics	0.635	1.118	0.116	9.659	<0.001
TOMI2_45	←	Pragmatics	0.513	1.010	0.124	8.172	<0.001

β: Regression coefficient, se: Standard error, zβ: Standardized regression coefficient. Bolded sections are statistically significant (p<0.05).

Table A2. *Item total correlations for the TOMI-2.*

Factor	Item No	Item Total Correlation
Early	3	0.468
	6	0.540
	24	0.559
	25	0.682
	28	0.633
	37	0.534
	38	0.609
	43	0.650
	44	0.482
	48	0.567
	49	0.634
	50	0.583
	54	0.630
	59	0.619
Basic	1	0.432
	4	0.549
	7	0.630
	8	0.587
	9	0.502
	10	0.652
	11	0.538
	12	0.416
	15	0.413
	16	0.520
	26	0.254
	29	0.589
	30	0.603
	31	0.450
	32	0.591
	33	0.584
	35	0.541
	39	0.350
	42	0.582
	51	0.639
	53	0.605
	57	0.528
	60	0.593
Advanced	2	0.430
	5	0.435
	13	0.589
	14	0.450
	17	0.337
	18	0.567
	19	0.462
	20	0.467

Advanced	21	0.533
	22	0.509
	23	0.587
	27	0.551
	34	0.542
	36	0.570
	40	0.496
	41	0.488
	45	0.448
	46	0.383
	47	0.442
	52	0.630
	55	0.606
	56	0.583
	58	0.608
Emotion Recognition	6	0.575
	17	0.388
	25	0.690
	48	0.615
	49	0.686
	50	0.657
	32	0.622
	51	0.692
	52	0.627
	55	0.567
Mental State Term Comprehension	7	0.639
	10	0.639
	39	0.371
	33	0.557
	54	0.662
	53	0.643
Pragmatics	2	0.347
	13	0.555
	14	0.325
	18	0.530
	19	0.375
	20	0.436
	35	0.515
	36	0.478
	45	0.378