Examining the relationship between theory of mind and counting skills in children aged 48–72 months

Furkan Albayrak[1] | **Türker Sezer**[2*]

1 Department of Childcare and Youth Services, European Vocational School, Kocaeli University of Health and Technology, Kocaeli, Türkiye | 2 Department of Primary Education, Faculty of Education, Bolu Abant İzzet Baysal University, Bolu, Türkiye

Received: 21 May 2026 | **Revised:** 11 September 2026 | **Accepted:** 28 September 2026

ABSTRACT

Early number and counting skills constitute an important foundation for later mathematical development, while theory of mind reflects children's ability to understand beliefs, intentions, and perspectives. Although both domains involve representational thinking, empirical evidence on their association in Turkish preschool samples remains limited. This study aimed to examine the relationships between early numeracy skills and theory of mind in children aged 48–72 months. The study was designed as a cross-sectional correlational survey. The participants were 102 preschool children, including 48 girls and 54 boys, attending independent preschools in Kocaeli, Türkiye. Data were collected individually using the Early Numeracy Test and four theory of mind tasks: the Sally-Anne, Bonibon, Chocolate Bar, and Ice Cream Truck tasks. Because the data did not meet normality assumptions, Mann-Whitney U tests were used for age- and gender-based comparisons, and Spearman correlation analysis was conducted to examine associations among variables. Children aged 61–72 months scored significantly higher than children aged 48–60 months on comparison, classification, resultative counting, general number knowledge, and total early numeracy scores. Gender differences were limited, with girls outperforming boys only in classification. In theory of mind, age-related differences emerged only for the Ice Cream Truck task, whereas no gender-related differences were found. Positive and significant associations were observed between several early numeracy dimensions and the Sally-Anne, Bonibon, and Chocolate Bar tasks, while the Ice Cream Truck task was not significantly associated with early numeracy scores. The findings suggest that early quantitative thinking and social-cognitive representational processes may develop in partly related ways. Early childhood practices may therefore benefit from integrated activities that support counting, perspective-taking, and representational reasoning together.

KEYWORDS: Early numeracy; False belief; Social cognition; Quantitative survey; Early childhood education

DOI: https://doi.org/10.29329/pedper.2026.189 | Vol. 5, No. 2 (2026) | pp. 419-436

1. Introduction

Early number and counting skills constitute a foundational component of children's mathematical development. Longitudinal evidence shows that preschool mathematics competencies, particularly symbolic number knowledge and early counting-related skills, are associated with later mathematical achievement (Hawes et al., 2019; Nguyen et al., 2016). Number sense has also been

Open access under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/)
Corresponding author: sezer_t@ibu.edu.tr

identified as a robust indicator of children's subsequent mathematics performance (Jordan et al., 2010). In this respect, the preschool period represents a critical developmental window for understanding how children acquire, coordinate, and use numerical representations.

Counting is not limited to reciting number words. Meaningful counting requires children to coordinate one-to-one correspondence, stable order, cardinality, abstraction, and order irrelevance (Gelman & Gallistel, 1978; Hannula-Sormunen et al., 2015). These principles allow children to connect number words with quantities and to understand that the last number word used in a count indicates the cardinal value of a set. Accordingly, early numeracy is better understood as a multi-component construct that includes comparison, classification, matching, seriation, counting, resultative counting, general number knowledge, and estimation (Aunio & Räsänen, 2016; Purpura & Lonigan, 2013; Van de Walle et al., 2018).

From a developmental perspective, early numeracy requires children to move between concrete objects, spoken number words, and abstract quantitative relations. A child who counts a group of objects must understand that each object is counted once, that number words follow a conventional order, and that the final word represents the size of the whole set. This coordination of symbols and quantities is cognitively demanding because it requires children to hold multiple representations in mind at the same time. For this reason, early number and counting skills provide a suitable context for examining representational development in the preschool years (Geary & vanMarle, 2016; Howell & Kemp, 2010; Salminen et al., 2018).

Children's early mathematics performance is shaped by both contextual and individual factors. Home learning support and the home mathematics environment are associated with mathematical development (Zhao & Gibson, 2023), while classroom experiences can provide systematic opportunities for numerical reasoning. At the child level, age, language, working memory, and executive functions have been linked to early mathematics outcomes (De Smedt et al., 2009; ten Braak et al., 2022; Vukovic & Lesaux, 2013). These findings suggest that early numerical development should be examined in relation to broader cognitive and social-cognitive processes rather than as an isolated academic skill.

The present focus on counting skills is therefore not a narrow focus on rote performance. Counting is an entry point into children's broader mathematical reasoning because it allows children to compare sets, determine quantity, monitor change, and construct relations among objects. When children understand that number words can stand for quantities, they begin to use symbols to organize experience. This symbolic function is developmentally relevant to the present study because theory of mind also depends on children's ability to understand that internal representations may stand for events, objects, or situations in the world.

Theory of mind refers to children's ability to understand that people have mental states such as beliefs, desires, intentions, and emotions, and that these mental states may differ from reality or from one's own perspective (Gopnik & Wellman, 1992; Wellman & Liu, 2004). During the preschool years, children gradually become more successful in false-belief tasks and in understanding that others may act on the basis of mistaken beliefs (Hughes, 2011). Language development, executive functioning, effortful control, and social experiences are among the factors associated with this developmental progression (Aubuchon et al., 2023; Devine & Hughes, 2014; Grazzani et al., 2018).

The possible association between early numeracy and theory of mind is theoretically meaningful because both domains require children to understand that one reality can be represented in different ways. In numerical tasks, a set of five objects can be represented through counting, matching,

comparing, or a written numeral. In false-belief tasks, the same event can be represented differently by the child, by a character who has correct knowledge, and by a character who has an inaccurate belief. This shared demand for coordinating alternative representations suggests that the relationship between counting and theory of mind should be examined empirically rather than assumed to be separate developmental pathways (Perner et al., 2021; Wellman & Liu, 2004).

A growing body of research indicates that theory of mind and early academic skills may be developmentally related. Preschool false-belief understanding has been found to predict later mathematics achievement (Devine & Hughes, 2014; Lockl et al., 2017), and longitudinal findings suggest that theory of mind and executive function can contribute to mathematics and reading abilities over time (Kloo et al., 2022). One explanation for this association is that both early numeracy and theory of mind require representational flexibility: in counting, children coordinate quantities, number words, and objects; in theory of mind, they coordinate reality with another person's belief representation (Perner et al., 2021; Sarnecka & Wright, 2013). Despite this conceptual overlap, empirical evidence from Turkish preschool samples remains limited. Therefore, the present study focuses on the relationship between early numeracy and theory of mind in children aged 48–72 months.

1.1. The Current Study

The purpose of this study was to examine age- and gender-related differences in early numeracy and theory of mind scores and to determine the associations between these two developmental domains in preschool children. The study was conducted with children aged 48–72 months because this period is developmentally important for both counting principles and false-belief understanding. By examining these variables together, the study aims to contribute to a more integrated understanding of children's quantitative and social-cognitive development.

The study is also important because the preschool years are characterized by rapid changes in both numerical and social-cognitive reasoning. Children in the 48–60-month age range may be consolidating basic counting principles and first-order false-belief understanding, whereas children in the 61–72-month age range may be more prepared for complex counting and second-order belief tasks. Examining these two age groups together may therefore reveal whether the two domains show similar developmental patterns or whether some tasks differentiate children more strongly than others.

The study was guided by the following research questions:

(1) Do early numeracy skills differ according to children's age group and gender?

(2) Do theory of mind task scores differ according to children's age group and gender?

(3) What is the relationship between children's early numeracy skills and theory of mind task scores?

2. Method

2.1. Research Design

This study employed a cross-sectional correlational design. This design was considered appropriate because the study aimed to examine age- and gender-related differences in early numeracy and theory of mind scores and to determine the associations between these developmental domains at a single point in time. No intervention was implemented, and no causal inference was intended (Cohen et al., 2007; Wang & Cheng, 2020).

2.2. Participants

The participants were 102 children aged 48–72 months who were attending preschool education during the 2023–2024 academic year in six independent preschools located in the İzmit, Başiskele, and Gölcük districts of Kocaeli, Türkiye. The sample included 48 girls and 54 boys. Participants were selected through convenience sampling based on accessibility, school permission, parental consent, and children's willingness to participate. Inclusion criteria were being between 48 and 72 months of age, attending an independent preschool, and having written parental consent. Children who were unwilling, distracted, or uncomfortable during the individual session were not forced to continue.

Table 1 ***Demographic characteristics of the participants***

Characteristic	*f*	%
Age group		
48–60 months	49	48.0
61–72 months	53	52.0
Total	**102**	**100.0**
Gender		
Girls	48	47.1
Boys	54	52.9
Total	**102**	**100.0**

As Table 1 shows, 49 children (48%) were in the 48–60-month group and 53 (52%) were in the 61–72-month group, with 48 girls (47.1%) and 54 boys (52.9%).

2.3. Data Collection Instruments

Data were collected using a participant information form, the Early Numeracy Test, and four theory of mind tasks: Sally-Anne, Bonibon, Chocolate Bar, and Ice Cream Truck. The participant information form included items on children's gender, age group, parental education, and family income.

The Early Numeracy Test was developed by van Luit et al. (1994) and adapted into Turkish by Önkol (2012). The test is suitable for preschool and first-grade children and assesses early mathematical proficiency and early number knowledge. It includes nine subdimensions: comparison, classification, matching, seriation, counting, structural counting, resultative counting, general number knowledge, and estimation. Each correct response is scored as 1, and each incorrect response is scored as 0. In the Turkish adaptation study, the test-retest correlation was .84 for Form A, Cronbach's alpha was .94, and the Kuder-Richardson 20 reliability coefficient was .91 (Önkol, 2012).

The use of a multidimensional early numeracy measure was important for this study because the research question did not focus on counting accuracy alone. The Early Numeracy Test makes it possible to examine whether theory of mind is associated only with counting-related subdimensions or whether the relationship extends to broader numerical skills such as classification, seriation,

general number knowledge, and estimation. This structure allowed the study to identify differentiated patterns rather than relying solely on a single total mathematics score.

Theory of mind was assessed using four tasks adapted by Girli and Tekin (2010). The Sally-Anne task assesses first-order false-belief understanding through a location-change scenario originally developed by Baron-Cohen et al. (1985). The Bonibon task assesses unexpected-content false belief. The Chocolate Bar task involves a belief scenario in which one character observes a transfer of an object while another character's belief is inferred. The Ice Cream Truck task, originally developed as a second-order false-belief task by Perner and Wimmer (1985), requires children to understand what one character thinks about another character's belief. In the Turkish adaptation study, KR-21 reliability values were reported as .78 for the Sally-Anne task, .82 for the Bonibon task, .84 for the Chocolate Bar task, and .84 for the Ice Cream Truck task; the KR-21 value for the overall task set was .83 (Girli & Tekin, 2010).

2.4. Data Collection Process

After the required permissions were obtained, information and consent forms were sent to parents. Only children whose parents provided written informed consent were included. Data were collected individually in a quiet room within the preschool setting during the spring term of the 2023–2024 academic year. Before the tasks, the researcher engaged in brief rapport-building activities. The measurement instruments were administered one-to-one in approximately 30 minutes per child. Children were seated opposite the researcher at a table, and the tasks were administered in a child-appropriate manner. No audio recording, video recording, or photographs were used.

2.5. Data Analysis

The data were analyzed using a statistical package program. Normality was examined using Kolmogorov-Smirnov and Shapiro-Wilk tests because the sample size and score distributions required careful assessment of distributional assumptions (Razali & Wah, 2011). Because the variables did not meet the normality assumption, nonparametric analyses were conducted. Mann-Whitney U tests were used for two-group comparisons by age group and gender. Spearman's rho correlation coefficients were calculated to examine the associations between early numeracy scores and theory of mind task scores. For Mann-Whitney U tests, effect sizes were calculated using r = |z| / sqrt(N) and interpreted as small around .10, medium around .30, and large around .50 (Field, 2013). The significance level was set at $p < .05$.

Table 2 ***Results of the normality test of the data***

Variable	Age group	Kolmogorov–Smirnov			Shapiro–Wilk		
		Statistic	*df*	*p*	Statistic	*df*	*p*
Comparison	48–60 months	.390	49	< .001	.674	49	< .001
	61–72 months	.485	53	< .001	.460	53	< .001
Classification	48–60 months	.207	49	< .001	.880	49	< .001
	61–72 months	.340	53	< .001	.780	53	< .001
Matching	48–60 months	.219	49	< .001	.854	49	< .001

Variable	Age group	Kolmogorov–Smirnov			Shapiro–Wilk		
		Statistic	*df*	*p*	Statistic	*df*	*p*
	61–72 months	.200	53	< .001	.856	53	< .001
Seriation	48–60 months	.252	49	< .001	.885	49	< .001
	61–72 months	.159	53	.002**	.885	53	< .001
Counting	48–60 months	.214	49	< .001	.926	49	< .001
	61–72 months	.175	53	< .001	.882	53	< .001
Structural counting	48–60 months	.213	49	< .001	.888	49	< .001
	61–72 months	.210	53	< .001	.856	53	< .001
Resultative counting	48–60 months	.147	49	.010*	.924	49	.004**
	61–72 months	.217	53	< .001	.847	53	< .001
General number knowledge	48–60 months	.191	49	< .001	.910	49	.001**
	61–72 months	.165	53	< .001	.895	53	< .001
Estimation	48–60 months	.298	49	< .001	.796	49	< .001
	61–72 months	.379	53	< .001	.711	53	< .001
Sally-Anne task	48–60 months	.323	49	< .001	.763	49	< .001
	61–72 months	.341	53	< .001	.729	53	< .001
Bonibon task	48–60 months	.445	49	< .001	.557	49	< .001
	61–72 months	.463	53	< .001	.505	53	< .001
Chocolate Bar task	48–60 months	.235	49	< .001	.810	49	< .001
	61–72 months	.253	53	< .001	.802	53	< .001
Ice Cream Truck task	48–60 months	.289	49	< .001	.757	49	< .001
	61–72 months	.246	53	< .001	.809	53	< .001

Note. *p < .05. **p < .01.

2.6. Ethical Considerations

This study was conducted in accordance with international ethical principles for research involving human participants (e.g., the Declaration of Helsinki) and received ethics committee approval on 09 April 2023 (approval number: 2023/164). Permission was also obtained from the Provincial Directorate of National Education. Written informed consent was obtained from the parents or legal

guardians of all participating children before data collection. Children's willingness was monitored during each session, and children who did not wish to continue were not forced to participate.

3. Findings

The findings are presented in three parts. First, age- and gender-related differences in early numeracy scores are reported. Second, age- and gender-related differences in theory of mind task scores are presented. Third, the relationships between early numeracy dimensions and theory of mind task scores are reported.

3.1. Age and Gender Differences in Early Numeracy

Table 3 ***Early numeracy test scores by age group***

Early Numeracy Test	Age group	*N*	MR	SR	*U*	*z*	*p*
Comparison	48–60 months	49	46.61	2284.00	1059.00	2.101	.036*
	61–72 months	53	56.02	2969.00			
	Total	**102**					
Classification	48–60 months	49	44.28	2169.50	944.50	-2.510	.012*
	61–72 months	53	58.18	3083.50			
	Total	**102**					
Matching	48–60 months	49	50.54	2476.50	1251.50	-0.333	.739
	61–72 months	53	52.39	2776.50			
	Total	**102**					
Seriation	48–60 months	49	50.51	2475.00	1250.00	-0.333	.739
	61–72 months	53	52.42	2778.00			
	Total	**102**					
Counting	48–60 months	49	46.81	2293.50	1068.50	-1.576	.115
	61–72 months	53	55.84	2959.50			
	Total	**102**					
Structural counting	48–60 months	49	47.24	2315.00	1090.00	-1.423	.155
	61–72 months	53	55.43	2938.00			
	Total	**102**					
Resultative counting	48–60 months	49	39.83	1951.50	726.50	-3.904	< .001

Early Numeracy Test	Age group	*N*	MR	SR	*U*	*z*	*p*
	61–72 months	53	62.29	3301.50			
	Total	**102**					
General number knowledge	48–60 months	49	44.62	2186.50	961.50	-2.306	.021*
	61–72 months	53	57.86	3066.50			
	Total	**102**					
Estimation	48–60 months	49	50.02	2451.00	1226.00	-0.551	.582
	61–72 months	53	52.87	2802.00			
	Total	**102**					
ENT total	48–60 months	49	43.98	2155.00	930.00	-2.471	.013*
	61–72 months	53	58.45	3098.00			
	Total	**102**					

Note. MR = mean rank; SR = sum of ranks. *p < .05. **p < .01.

As shown in Table 3, the Mann-Whitney U test indicated significant age-group differences in comparison, classification, resultative counting, general number knowledge, and total early numeracy scores. In all significant comparisons, the mean ranks favored children aged 61–72 months. The effect sizes were small for comparison (r = .21), classification (r = .25), general number knowledge (r = .23), and total early numeracy (r = .24), and moderate for resultative counting (r = .39). No significant age-group differences were found for matching, seriation, counting, structural counting, or estimation.

Table 4 ***Early numeracy test scores by gender***

Early Numeracy Test	Gender	*N*	MR	SR	*U*	*z*	*p*
Comparison	Girls	48	50.86	2441.50	1265.50	-0.268	.789
	Boys	54	52.06	2811.50			
	Total	**102**					
Classification	Girls	48	59.53	2857.50	910.50	-2.736	.006**
	Boys	54	44.36	2395.50			
	Total	**102**					
Matching	Girls	48	53.16	2551.50	1216.50	-0.551	.581
	Boys	54	50.03	2701.50			

Early Numeracy Test	Gender	*N*	MR	SR	*U*	*z*	*p*
	Total	**102**					
Seriation	Girls	48	52.77	2533.00	1235.00	-0.419	.675
	Boys	54	50.37	2720.00			
	Total	**102**					
Counting	Girls	48	54.58	2620.00	1148.00	-1.015	.310
	Boys	54	48.76	2633.00			
	Total	**102**					
Structural counting	Girls	48	54.68	2624.50	1143.50	-1.042	.297
	Boys	54	48.68	2628.50			
	Total	**102**					
Resultative counting	Girls	48	53.53	2569.50	1198.50	-0.666	.505
	Boys	54	49.69	2683.50			
	Total	**102**					
General number knowledge	Girls	48	55.68	2672.50	1095.50	-1.373	.170
	Boys	54	47.79	2580.50			
	Total	**102**					
Estimation	Girls	48	52.00	2496.00	1272.00	-0.182	.855
	Boys	54	51.06	2757.00			
	Total	**102**					
ENT total	Girls	48	55.64	2670.50	1097.50	-1.332	.183
	Boys	54	47.82	2582.50			
	Total	**102**					

Note. MR = mean rank; SR = sum of ranks. *p < .05. **p < .01.

As shown in Table 4, no statistically significant gender differences were found in the early numeracy subdimensions or total score, except for classification. Girls had significantly higher classification scores than boys, U = 910.50, z = -2.74, p = .006, r = .27. The effect size indicated a small-to-moderate difference.

3.2. Age and Gender Differences in Theory of Mind

Table 5 *Theory of mind scores by age group*

ToM task	Age group	*N*	MR	SR	*U*	*z*	*p*
Sally-Anne	48–60 months	49	51.33	2515.00	1290.00	-0.063	.950
	61–72 months	53	51.66	2738.00			
	Total	**102**					
Bonibon	48–60 months	49	50.33	2466.00	1241.00	-0.527	.598
	61–72 months	53	52.58	2787.00			
	Total	**102**					
Chocolate Bar	48–60 months	49	48.93	2397.50	1172.50	-0.914	.361
	61–72 months	53	53.88	2855.50			
	Total	**102**					
Ice Cream Truck	48–60 months	49	43.88	2150.00	925.00	-2.724	.006**
	61–72 months	53	58.55	3103.00			
	Total	**102**					

Note. MR = mean rank; SR = sum of ranks. *p < .05. **p < .01.

As shown in Table 5, no significant age-group differences were found in the Sally-Anne, Bonibon, or Chocolate Bar tasks. A significant age-group difference was found only for the Ice Cream Truck task, U = 925.00, z = -2.72, p = .006, r = .27, with mean ranks favoring the 61–72-month group.

Table 6 *Theory of mind scores by gender*

ToM task	Gender	*N*	MR	SR	*U*	*z*	*p*
Sally-Anne	Girls	48	55.58	2668.00	1100.00	-1.453	.146
	Boys	54	47.87	2585.00			
	Total	**102**					
Bonibon	Girls	48	52.17	2504.00	1264.00	-0.294	.769
	Boys	54	50.91	2749.00			
	Total	**102**					
Chocolate Bar	Girls	48	56.14	2694.50	1073.50	-1.616	.106
	Boys	54	47.38	2558.50			

ToM task	Gender	*N*	MR	SR	*U*	*z*	*p*
	Total	**102**					
Ice Cream Truck	Girls	48	55.43	2660.50	1107.50	-1.376	.169
	Boys	54	48.01	2592.50			
	Total	**102**					

Note. MR = mean rank; SR = sum of ranks.

As shown in Table 6, no statistically significant gender differences were found in any of the four theory of mind task scores.

Table 7 ***Relationships between ToM and ENT scores***

ENT dimension		Sally-Anne	Bonibon	Chocolate Bar	Ice Cream Truck
Comparison	*r*	.027	.086	.033	-.042
	p	.788	.393	.741	.673
	N	102	102	102	102
Classification	*r*	.203*	.224*	.235*	.012
	p	.040	.024	.018	.906
	N	102	102	102	102
Matching	*r*	.140	.160	.225*	.007
	p	.161	.108	.023	.945
	N	102	102	102	102
Seriation	*r*	.303**	.168	.231*	-.044
	p	.002	.091	.020	.661
	N	102	102	102	102
Counting	*r*	.256**	.313**	.501**	.138
	p	.010	.001	< .001	.165
	N	102	102	102	102
Structural counting	*r*	.154	.226*	.349**	.042
	p	.122	.023	< .001	.678
	N	102	102	102	102

ENT dimension		Sally-Anne	Bonibon	Chocolate Bar	Ice Cream Truck
Resultative counting	*r*	.164	.312**	.379**	.165
	p	.099	.001	< .001	.098
	N	102	102	102	102
General number knowledge	*r*	.322**	.488**	.465**	-.024
	p	.001	< .001	< .001	.812
	N	102	102	102	102
Estimation	*r*	.021	-.012	.013	.066
	p	.830	.903	.895	.511
	N	102	102	102	102
ENT total	*r*	.289**	.340**	.448**	.059
	p	.003	< .001	< .001	.553
	N	102	102	102	102

Note. ToM = theory of mind; ENT = Early Numeracy Test. *p < .05. **p < .01.

As shown in Table 7, positive and significant correlations were observed between several early numeracy dimensions and the Sally-Anne, Bonibon, and Chocolate Bar tasks. Sally-Anne was positively associated with classification, seriation, counting, general number knowledge, and total early numeracy. Bonibon was positively associated with classification, counting, structural counting, resultative counting, general number knowledge, and total early numeracy. Chocolate Bar showed the strongest and most consistent pattern, correlating positively with all early numeracy subdimensions except comparison and estimation, as well as with total early numeracy. The Ice Cream Truck task was not significantly correlated with any early numeracy subdimension or with total early numeracy.

4. Discussion

This study examined age- and gender-related differences in early numeracy and theory of mind and investigated the relationships between these two developmental domains in children aged 48–72 months. The findings showed that older children performed better than younger children in several early numeracy dimensions and in the Ice Cream Truck task. Gender differences were limited. Most importantly, several early numeracy dimensions were positively associated with first-order theory of mind tasks, whereas the second-order Ice Cream Truck task stood outside this pattern.

The finding that children aged 61–72 months scored higher on comparison, classification, resultative counting, general number knowledge, and total early numeracy is consistent with developmental accounts of early mathematics. As children grow older, they have more opportunities to engage in numerical activities, use number words, compare quantities, and coordinate symbolic and nonsymbolic representations (Aunio & Räsänen, 2016; Nguyen et al., 2016). The moderate effect

observed for resultative counting is especially meaningful because this skill requires children not only to count but also to understand the cardinal meaning of the final number word, a principle that is central to meaningful counting (Gelman & Gallistel, 1978; Sarnecka & Wright, 2013).

The age-related findings should also be interpreted in relation to children's accumulated learning opportunities. Older preschool children may have had more sustained exposure to counting routines, classroom materials, number games, peer interaction, and adult-guided mathematical language. These experiences can support children's ability to compare quantities, classify objects according to relevant attributes, and use number knowledge flexibly. Therefore, the observed age differences may reflect both maturation and cumulative educational experience. Because the study did not measure preschool attendance duration or classroom mathematics quality, this explanation should be treated as a theoretically plausible interpretation rather than a directly tested conclusion.

Not all early numeracy dimensions differed by age group. Matching, seriation, counting, structural counting, and estimation did not show significant age-related differences. This pattern may indicate that some tasks assessed skills that were already accessible to both age groups or that had been frequently practiced in preschool settings. It may also reflect variability in the difficulty level of the subtests. Estimation, for example, may require more advanced number-line reasoning and representational precision than simple counting or matching tasks (Booth & Siegler, 2006; Siegler & Booth, 2004). Future studies with item-level analyses may clarify whether particular tasks are too easy, too difficult, or insufficiently sensitive for specific age ranges.

Gender differences in early numeracy were minimal. The only significant difference was observed in classification, where girls had higher mean ranks than boys. This finding should be interpreted cautiously because no gender difference emerged in total early numeracy or in the other subdimensions. The broader literature generally suggests that girls and boys do not differ intrinsically in early numerical abilities (Aunio et al., 2004; Kersey et al., 2018). Therefore, the classification difference in this sample may reflect group-specific experiences, task characteristics, or unmeasured variables rather than a stable gender-based difference.

The absence of broad gender differences is also practically important. In early childhood classrooms, assumptions about gender-based mathematical capacity may influence teachers' expectations, the kinds of tasks offered to children, and the feedback children receive. The present findings do not support such assumptions. Instead, they suggest that both girls and boys should be offered rich, challenging, and equally encouraging opportunities to engage in counting, classification, comparison, and explanation. The isolated classification difference should therefore be considered a finding that requires further examination rather than a basis for differentiated expectations.

In theory of mind, age-related differences emerged only for the Ice Cream Truck task. This result is theoretically plausible because the Ice Cream Truck task is a second-order false-belief task and requires children to understand what one character thinks about another character's belief (Perner & Wimmer, 1985). It involves a longer narrative sequence, greater working-memory demands, and more complex linguistic processing than first-order false-belief tasks. The absence of age differences in the Sally-Anne, Bonibon, and Chocolate Bar tasks may indicate that many children in both age groups had already reached a basic level of first-order false-belief understanding. Similar developmental patterns have been discussed in research showing that theory of mind performance varies according to task type and linguistic complexity (Rai & Mitchell, 2004; Wellman & Liu, 2004).

The findings also highlight the importance of considering task difficulty. Tasks that appear to assess the same broad domain may differ substantially in their cognitive and linguistic demands. For

example, a child may succeed in a simple location-change false-belief task but struggle when the task requires tracking multiple characters, a changing location, and a second-order belief. Similarly, a child may recite number words accurately but experience difficulty when asked to use counting to determine a result or to make a numerical judgment. This variability suggests that future studies should examine not only total scores but also the structure and difficulty of individual subdimensions and tasks.

Gender did not significantly affect theory of mind task scores. This result differs from some studies reporting girls' advantages in theory of mind performance (Calero et al., 2013; Sundqvist et al., 2018). However, gender effects in theory of mind are not always consistent across cultural contexts, task formats, and age ranges. Cultural and linguistic experiences may shape children's opportunities to talk about mental states and to practice perspective-taking (Shahaeian et al., 2011). Therefore, the absence of gender differences in the present sample should be interpreted as evidence of comparable performance within this particular group rather than as a universal developmental pattern.

The central finding of the study was the positive relationship between several early numeracy dimensions and three of the four theory of mind tasks. This result is consistent with studies suggesting developmental links between false-belief understanding, executive function, and academic skills (Cantin et al., 2016; Devine & Hughes, 2014; Kloo et al., 2022; Lecce & Devine, 2022; Lockl et al., 2017). One possible explanation is that both early numeracy and theory of mind require representational flexibility. In early numeracy, children must understand that number words can represent quantities and that the same set can be described through different numerical relations. In theory of mind, children must understand that beliefs can represent reality accurately or inaccurately. Thus, children's ability to coordinate different representations may support performance in both domains (Perner et al., 2021; Sarnecka & Wright, 2013).

Another issue concerns the potential role of unmeasured shared variables. Language development may support both domains because children need to understand task instructions, relational expressions, mental-state verbs, and quantitative language. Working memory may also contribute because children must remember the sequence of events in theory of mind stories and maintain counting information during numerical tasks. Executive functions may help children inhibit their own knowledge of reality in false-belief tasks and suppress irrelevant strategies in numerical tasks. Since these variables were not directly assessed, the correlations reported in this study should be interpreted as associations rather than evidence of a direct developmental pathway.

The lack of significant correlations between the Ice Cream Truck task and early numeracy dimensions deserves particular attention. Rather than contradicting the general relationship between early numeracy and theory of mind, this finding may indicate that the Ice Cream Truck task measured a more advanced and linguistically demanding aspect of theory of mind. Second-order false-belief tasks may require narrative comprehension, memory for multiple characters' knowledge states, and the integration of nested mental representations. These demands may have reduced the task's sensitivity to the early numeracy skills measured in this study. Future research should include language, working memory, and executive function measures to clarify whether these variables explain the association between theory of mind and early numeracy.

4.1. Limitations

This study has several limitations. First, the cross-sectional correlational design does not allow causal conclusions about the developmental direction of the relationship between early numeracy

and theory of mind. Second, the sample was limited to 102 children attending six independent preschools in one province, which restricts the generalizability of the findings. Third, variables known to influence both early numeracy and theory of mind, such as language development, working memory, executive functions, effortful control, preschool attendance duration, and socioeconomic status, were not included in the analysis. Fourth, the Ice Cream Truck task may have imposed higher linguistic and narrative demands than the other tasks, which may have affected children's performance independently of their theory of mind competence.

A further limitation is that the study did not include descriptive variables such as duration of preschool attendance, classroom mathematics exposure, or children's language scores. These variables may help explain why some early numeracy dimensions and some theory of mind tasks were associated while others were not. Moreover, because the analyses were based on nonparametric group comparisons and bivariate correlations, the findings do not show whether the observed associations remain significant after controlling for age or other developmental variables.

4.2. Implications and Recommendations

The findings suggest that early childhood education practices may benefit from integrated activities that support numerical reasoning and social-cognitive representation together. Activities that ask children to count, compare, classify, predict another person's action, explain a character's mistaken belief, or discuss alternative perspectives may provide opportunities to strengthen both quantitative thinking and perspective-taking. Teachers can design play-based and story-based tasks in which children solve numerical problems while also reasoning about what different characters know, think, or expect.

For classroom practice, the findings point to the value of designing activities that naturally combine mathematical and social-cognitive talk. For instance, teachers may use story-based problems in which a character miscounts objects, hides a set of materials, or makes an incorrect prediction about quantity. Children can then be asked not only how many objects there are but also what the character thinks, why the character is mistaken, and how the character could solve the problem. Such activities may help children connect counting, explanation, justification, and perspective-taking within meaningful play contexts.

Future studies should use larger and more diverse samples and should include measures of language development, executive functions, working memory, and socioeconomic characteristics. Longitudinal studies are needed to determine whether theory of mind predicts later early numeracy, whether early numeracy contributes to social-cognitive development, or whether both are explained by shared underlying cognitive mechanisms. Intervention studies may also examine whether integrated activities improve both counting skills and theory of mind performance.

Policy and program-level implications should also be considered cautiously. The study does not provide evidence for a specific intervention model, but it supports the idea that preschool programs should avoid treating cognitive, mathematical, and social-emotional domains as disconnected areas. Developmentally appropriate early childhood curricula can create integrated learning experiences in which children count, compare, explain, negotiate, predict, and reason about others' perspectives. Such integration may be especially valuable in classrooms where children's language, symbolic thinking, and peer interaction are developing simultaneously.

5. Conclusion

This study showed that early numeracy and theory of mind are related developmental domains in the preschool period. Older children performed better in selected early numeracy dimensions and in the second-order theory of mind task, while gender differences were minimal. Significant positive associations between early numeracy and first-order theory of mind tasks suggest that quantitative thinking and social-cognitive representation may develop in partially connected ways. The findings support the value of holistic early childhood practices that address children's numerical reasoning, symbolic thinking, and perspective-taking together.

6. Declarations

6.1. Conflict of interest

The authors declare that they have no conflict of interest to disclose.

6.2. Funding

This research received no specific funding.

6.3. Author contributions (CRediT)

Furkan Albayrak: Conceptualization; Methodology; Formal analysis; Investigation; Writing—original draft; Writing—review & editing

Türker Sezer: Conceptualization; Methodology; Writing—review & editing; Supervision

6.4. Data availability statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

6.5. Ethics approval

This study was conducted in accordance with international ethical principles for research involving human participants (e.g., the Declaration of Helsinki) and received ethics committee approval on 09 April 2023 (approval number: 2023/164). Permission was also obtained from the Provincial Directorate of National Education. Written informed consent was obtained from the parents or legal guardians of all participating children.

6.6. Use of artificial intelligence (AI) tools

Generative AI tools (ChatGPT, GPT-5.5, OpenAI) were used to support manuscript drafting, literature review organization, and Turkish–English language editing. The authors reviewed, verified, and approved all AI-assisted content. Generative AI tools were not used for data collection, data analysis, interpretation of findings, or authorship.

6.7. Acknowledgements

This study is based on a master's thesis prepared by the first author under the supervision of the second author. The authors would like to thank the participating children, their families, and the preschool institutions that supported the data collection process.

References

Aubuchon, O., Libenstein, J., Moenner, M., Seguin, M., Bellerose, J., Bernier, A., & Beauchamp, M. H. (2023). What contributes to false belief understanding in childhood? A multidimensional approach. *Cognitive Development, 68*, Article 101382. https://doi.org/10.1016/j.cogdev.2023.101382

Aunio, P., & Räsänen, P. (2016). Core numerical skills for learning mathematics in children aged five to eight years: A working model for educators. *European Early Childhood Education Research Journal, 24*(5), 684–704.

Aunio, P., Ee, J., Lim, S. E. A., Hautamäki, J., & Van Luit, J. E. H. (2004). Young children's number sense in Finland, Hong Kong and Singapore. *International Journal of Early Years Education, 12*(3), 195–216. https://doi.org/10.1080/0966976042000268681

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a theory of mind? *Cognition, 21*(1), 37–46.

Booth, J., & Siegler, R. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 42*(1), 189–201. https://doi.org/10.1037/0012-1649.41.6.189

Calero, C. I., Salles, A., Semelman, M., & Sigman, M. (2013). Age and gender dependent development of theory of mind in 6- to 8-year-old children. *Frontiers in Human Neuroscience, 7*, Article 281. https://doi.org/10.3389/fnhum.2013.00281

Cantin, R. H., Gnaedinger, E. K., Gallaway, K. C., Hesson-McInnis, M. S., & Hund, A. M. (2016). Executive functioning predicts reading, mathematics, and theory of mind during the elementary years. *Journal of Experimental Child Psychology, 146*, 66–78. https://doi.org/10.1016/j.jecp.2016.01.014

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). Routledge.

De Smedt, B., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquière, P. (2009). Working memory and individual differences in mathematics achievement: A longitudinal study from first grade to second grade. *Journal of Experimental Child Psychology, 103*(2), 186–201. https://doi.org/10.1016/j.jecp.2009.01.004

Devine, R. T., & Hughes, C. (2014). Relations between false belief understanding and executive function in early childhood: A meta-analysis. *Child Development, 85*(5), 1777–1794. https://doi.org/10.1111/cdev.12237

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage.

Geary, D., & vanMarle, K. (2016). Young children's core symbolic and nonsymbolic quantitative knowledge in the prediction of later mathematics achievement. *Developmental Psychology, 52*(12), 2130–2144. https://doi.org/10.1037/DEV0000214

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Harvard University Press.

Girli, A., & Tekin, D. (2010). Investigating false belief levels of typically developed children and children with autism. *Procedia-Social and Behavioral Sciences, 2*(2), 1944–1950. https://doi.org/10.1016/j.sbspro.2010.03.261

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind & Language, 7*, 145–171. https://doi.org/10.1111/j.1468-0017.1992.tb00202.x

Grazzani, I., Ornaghi, V., Conte, E., Pepe, A., & Caprin, C. (2018). The relation between emotion understanding and theory of mind in children aged 3 to 8: The key role of language. *Frontiers in Psychology, 9*, Article 724. https://doi.org/10.3389/fpsyg.2018.00724

Hannula-Sormunen, M., Lehtinen, E., & Räsänen, P. (2015). Preschool children's spontaneous focusing on numerosity, subitizing, and counting skills as predictors of their mathematical performance seven years later at school. *Mathematical Thinking and Learning, 17*, 155–177. https://doi.org/10.1080/10986065.2015.1016814

Hawes, Z., Nosworthy, N., Archibald, L., & Ansari, D. (2019). Kindergarten children's symbolic number comparison skills relate to first-grade mathematics achievement: Evidence from a two-minute paper-and-pencil test. *Learning and Instruction, 59*, 21–33. https://doi.org/10.1016/J.LEARNINSTRUC.2018.09.004

Howell, S. C., & Kemp, C. R. (2010). Assessing preschool number sense: Skills demonstrated by children prior to school entry. *Educational Psychology, 30*(4), 411–429. https://doi.org/10.1080/01443411003695410

Hughes, C. (2011). *Social understanding and social lives: From toddlerhood through to the transition to school*. Psychology Press.

Jordan, C. N., Glutting, J., & Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. *Learning and Individual Differences, 20*(2), 82–88. https://doi.org/10.1016/j.lindif.2009.07.004

Kersey, A., Braham, E., Csumitta, K., Libertus, M., & Cantlon, J. (2018). No intrinsic gender differences in children's earliest numerical abilities. *npj Science of Learning, 3*, Article 12. https://doi.org/10.1038/s41539-018-0028-7

Kloo, D., Osterhaus, C., Kristen-Antonow, S., & Sodian, B. (2022). The impact of theory of mind and executive function on math and reading abilities: A longitudinal study. *Infant and Child Development, 31*(6), Article e2356. https://doi.org/10.1002/icd.2356

Lecce, S., & Devine, R. T. (2022). Understanding the role of theory of mind in children's academic achievement: New directions for research. *Child Development Perspectives, 16*(2), 81–87. https://doi.org/10.1111/cdep.12450

Lockl, K., Ebert, S., & Weinert, S. (2017). Predicting school achievement from early theory of mind: Differential effects on achievement tests and teacher ratings. *Learning and Individual Differences, 53*, 93–102. https://doi.org/10.1016/j.lindif.2016.11.007

Nguyen, T., Watts, T. W., Duncan, G. J., Clements, D. H., Sarama, J. S., Wolfe, C., & Spitler, M. E. (2016). Which preschool mathematics competencies are most predictive of fifth grade achievement? *Early Childhood Research Quarterly, 36*, 550–560. https://doi.org/10.1016/J.ECRESQ.2016.02.003

Önkol, F. L. (2012). *Erken Sayı Testi'nin uyarlanması ve erken sayı gelişim programı'nın altı yaş çocukların sayı gelişimlerine etkisinin incelenmesi* [Adaptation of the Early Numeracy Test and the effect of the early numeracy development program on the number development of six-year-old children] [Unpublished doctoral dissertation]. Marmara University.

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that..." Attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology, 39*, 437–471. https://doi.org/10.1016/0022-0965(85)90051-7

Perner, J., Aichhorn, M., Tholen, M., & Schurz, M. (2021). Mental files and teleology. In M. Gilead & K. Ochsner (Eds.), *The neural bases of mentalizing* (pp. 257–281). Springer.

Purpura, D. J., & Lonigan, C. J. (2013). Informal numeracy skills: The structure and relations among numbering, relations, and arithmetic operations in preschool. *American Educational Research Journal, 50*(1), 178–209. https://doi.org/10.3102/0002831212465332

Rai, R., & Mitchell, P. (2004). Five-year-old children's difficulty with false belief when the sought entity is a person. *Journal of Experimental Child Psychology, 89*, 112–126. https://doi.org/10.1016/j.jecp.2004.05.003

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics, 2*(1), 21–33.

Salminen, J., Koponen, T., & Tolvanen, A. (2018). Individuality in the early number skill components underlying basic arithmetic skills. *Frontiers in Psychology, 9*, Article 1056. https://doi.org/10.3389/fpsyg.2018.01056

Sarnecka, B. W., & Wright, C. E. (2013). The idea of an exact number: Children's understanding of cardinality and equinumerosity. *Cognitive Science, 37*, 1493–1506. https://doi.org/10.1111/cogs.12043

Shahaeian, A., Peterson, C. C., Slaughter, V., & Wellman, H. M. (2011). Culture and the sequence of steps in theory of mind development. *Developmental Psychology, 47*(5), 1239–1247. https://doi.org/10.1037/a0023899

Siegler, R., & Booth, J. (2004). Development of numerical estimation in young children. *Child Development, 75*(2), 428–444. https://doi.org/10.1111/J.1467-8624.2004.00684.X

Sundqvist, A., Holmer, E., Koch, F. S., & Heimann, M. (2018). Developing theory of mind abilities in Swedish pre-schoolers. *Infant and Child Development, 27*(4), Article e2090. https://doi.org/10.1002/icd.2090

ten Braak, D., Lenes, R., Purpura, D. J., Schmitt, S. A., & Størksen, I. (2022). Why do early mathematics skills predict later mathematics and reading achievement? The role of executive function. *Journal of Experimental Child Psychology, 214*, Article 105306. https://doi.org/10.1016/j.jecp.2021.105306

van Luit, J. E. H., van de Rijt, B. A. M., & Pennings, A. H. (1994). *Utrechtse Getalbegrip Toets: Handleiding en toets* [Utrecht Early Numeracy Test: Manual and test]. Graviant.

Van de Walle, J. A., Karp, K. S., & Bay-Williams, J. M. (2018). *Elementary and middle school mathematics: Teaching developmentally*. Pearson.

Vukovic, R. K., & Lesaux, N. K. (2013). The language of mathematics: Investigating the ways language counts for children's mathematical development. *Journal of Experimental Child Psychology, 115*(2), 227–244. https://doi.org/10.1016/j.jecp.2013.02.002

Wang, X., & Cheng, Z. (2020). Cross-sectional studies: Strengths, weaknesses, and recommendations. *Chest, 158*(1), S65–S71. https://doi.org/10.1016/j.chest.2020.03.012

Wellman, H. M., & Liu, D. (2004). Scaling of theory of mind tasks. *Child Development, 75*(2), 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

Zhao, Y. V., & Gibson, J. L. (2023). Early home learning support and home mathematics environment as predictors of children's mathematical skills between age 4 and 6: A longitudinal analysis using video observations and survey data. *Child Development, 94*(6), e377-e392. https://doi.org/10.1111/cdev.13971