



Construct Validation of the Preschool Semantic Comprehension Test (TPSP): Development and Psychometric Evaluation

Rexsy Taruna^{1✉}, Anggawati Imanniyah², Dina Putri Dewi³, Hanna Salsabila⁴, Dwi Cinta Putri Maizaki⁵

Universitas MERCUBAKTIJAYA, Indonesia^(1,2,3,4,5)

DOI: [10.31004/obsesi.v9i5.7431](https://doi.org/10.31004/obsesi.v9i5.7431)

Abstract

This study aimed to examine the construct validity of the Preschool Semantic Comprehension Test (TPSP) as a measure of semantic comprehension in preschool-aged children. A total of 72 children aged 3–6 years participated and were classified into two groups: typically developing and atypically developing children. Each child completed an individual testing session using TPSP, lasting approximately 10 to 15 minutes. TPSP is a standardized Indonesian assessment tool designed to measure semantic comprehension in preschool-aged children. The test is administered by speech therapists. Data were analyzed using independent samples t-tests, ANCOVA, and effect size estimation (Cohen's *d*). Results revealed a highly significant difference in semantic comprehension scores between the two groups ($p < .001$, $d = -2.631$). Covariance analysis indicated that this difference remained significant even after controlling for age. These findings support the construct validity of TPSP, demonstrating its sensitivity in distinguishing developmental language profiles. The practical implications highlight TPSP's potential as an effective early assessment tool in Indonesian speech-language therapy settings. Future research is recommended to evaluate TPSP's longitudinal sensitivity and application in multilingual contexts.

Keywords: *semantic comprehension; preschool; test; construct validity; Indonesia*

Copyright (c) 2025 Rexsy Taruna, et al.

✉ Corresponding author: Rexsy Taruna

Email Address: rexsytaruna.mercubaktijaya.ac.id (Padang, Indonesia)

Received 3 August 2025, Accepted 21 September 2025, Published 21 September 2025

Introduction

Language assessment plays a pivotal role in evaluating and monitoring children's language development (Barry et al., 2024; Feltner et al., 2024). It enables professionals, particularly speech-language pathologists, to systematically observe and measure various dimensions of linguistic proficiency (McAllister & Miller, 2013). These include vocabulary acquisition, syntactic structure, narrative ability, and overall communicative competence—each of which contributes to a holistic understanding of a child's language skills.

The insights gained through such assessments are instrumental in designing individualized and evidence-based intervention strategies. These strategies are especially critical for children exhibiting signs of language delay or impairment, as early and targeted support can significantly enhance their developmental trajectory and communication outcomes. In 2023, it was reported that between 0.1% and 0.4% of Indonesian children over the age of one exhibited various form of language disorders, often linked to developmental delays (0.4%), autism spectrum disorder (0.2%),

Asperger syndrome (0.1%), attention deficit hyperactivity disorder (ADHD) (0.1%), cerebral palsy (0.1%), dyslexia (0.1%), intellectual disabilities (0.3%), and Down syndrome (0.5%) (Kemenkes, 2023). Additionally, between 0.5% and 1% of children aged 5 to 17 were identified as having disabilities that significantly affect language and communication abilities. These include intellectual disabilities (1%), sensory impairments (0.2%), and communication disorders (0.5%) (Kemenkes, 2023).

Given Indonesia's large and diverse population, the country faces distinct challenges in the domains of education and healthcare. One critical yet under-researched area is language assessment in children. The absence of accessible and reliable language testing tools poses a substantial risk for the early identification of language development issues. Without timely screening, many children with language impairments may go unnoticed during critical stages of growth, potentially impacting their academic and social functioning (Jerome et al., 2002; Ziegenfusz et al., 2022).

Language forms the foundation for learning and social interaction, and undetected language disorders can hinder a child's ability to read, write, and communicate effectively with peers. Without appropriate and timely intervention, these difficulties can lead to reduced self-esteem, diminished academic motivation, and long-term challenges in educational and social contexts (Wadman et al., 2008). Early detection and intervention, therefore, are essential components in supporting the developmental needs of children at risk of language and communication impairments (Chen et al., 2022).

The development of child language assessment tools in Indonesia faces several critical limitations. One of the primary challenges is the scarcity of standardized and validated language tests that are specifically tailored to the Indonesian context. Many of the assessments currently in use are adaptations of tools originally developed in other countries. As a result, these imported instruments may not fully capture the cultural and linguistic nuances relevant to Indonesian children, thereby limiting their effectiveness and validity.

A further complication arises from Indonesia's rich linguistic diversity. With hundreds of regional languages spoken across the archipelago, designing a universally applicable language test is particularly complex. Children who grow up speaking local or indigenous languages may struggle to perform well on assessments based solely on standardized Indonesian. This discrepancy can lead to inaccurate results, potentially underestimating a child's true linguistic capabilities and misinforming subsequent intervention strategies.

These limitations highlight the urgent need for culturally sensitive and linguistically appropriate language assessment tools (Shipley & McAfee, 2023). The absence of such instruments not only hampers accurate diagnosis but also affects the quality of support provided to children with language development needs. Without valid and reliable tools, early detection and intervention efforts may be compromised, affecting children's educational and social outcomes (Paul & Roth, 2011).

Consequently, the development of context-specific, standardized language assessments is essential to promote equitable access to language development resources. Such efforts would ensure that all Indonesian children, regardless of linguistic background, are assessed fairly and receive the appropriate support necessary for their optimal growth and learning. The Preschool Semantic Comprehension Test (TPSP) is a language assessment instrument specifically developed by Rexsy Taruna in 2024 (Taruna, 2024). Designed for children aged four to six years, the TPSP aims to evaluate their semantic comprehension abilities. The test has demonstrated high reliability across different age groups, with coefficients ranging from $r = .90$ to $r = .95$. Moreover, the TPSP shows evidence of concurrent validity through a statistically significant correlation with the vocabulary subtest of the WPPSI ($r = 0.746$; $p < .001$).

While concurrent validity is an important indicator of a test's quality, it alone is insufficient to fully establish the validity of an assessment tool. Psychometric standards emphasize the necessity of evaluating construct validity to determine whether a test truly measures the theoretical construct it purports to assess. In this context, construct validity refers to the degree to which TPSP accurately reflects the domain of semantic comprehension in preschool-aged children.

Given these considerations, the present study seeks to investigate the construct validity of the Preschool Semantic Comprehension Test. This involves examining whether the TPSP appropriately captures the underlying theoretical construct of semantic understanding among young children, rather than merely correlating with other measures. Such validation is essential to ensure the test's effectiveness in clinical and educational settings.

Methods

Participants

A total of 72 preschool-aged children participated in this study. The participants were classified into two groups based on their developmental status: typically developing children and atypically developing children.

Instrument and Procedure

The primary assessment tool was the Preschool Semantic Comprehension Test (Tes Pemahaman Semantik Prasekolah), which evaluates children's understanding of word meanings and semantic relationships. The TPSP specifically evaluates preschoolers' semantic comprehension in various language domains, such as quality, temporal, spatial (positional), negation, comparative, and verbal analogy concepts. The test includes 22 items and utilizes a picture book as the primary stimulus material. Scoring is binary, with 1 point awarded for accurate responses given without cues, and 0 for incorrect answers. The test was administered individually in a single session lasting approximately 10 to 15 minutes. All sessions were conducted in quiet, controlled environments to minimize distractions and optimize the validity of responses. The TPSP has demonstrated evidence of validity, as it is significantly correlated with verbal intelligence and verbal memory. Its reliability falls within the high range, with coefficients between $r = .90$ and $r = .95$.

Data Analysis

Data were analyzed using the latest version of JASP software. The primary statistical analysis was conducted using an independent sample t-test to assess group differences in mean scores. Additional analyses included calculation of effect size (Cohen's d) to estimate the magnitude of the group difference, and analysis of covariance (ANCOVA) to control for age as a covariate. Statistical significance was determined at the $p < .05$ level.

Result and Discussion

Gender Distribution

The distribution of respondents based on gender is presented in Table 1. In the atypical group, the majority of participants were male (71,875%), while females accounted for only 28,125%. In contrast, the typical group showed a more balanced distribution, with females comprising 52,273% and males 47,727%. This indicates that the atypical group was predominantly male, whereas the gender distribution in the typical group was relatively even.

Table 1: Frequencies for Gender

Category	Gender	Percent
Atypical	Female	28,125%
	Male	71,875%
Typical	Female	52,273%
	Male	47,727%

Age Distribution

Table 2 displays the age distribution of respondents within each group. In the atypical group, most children were 5 years old (50%), followed by 6 years old (25%), 4 years old (18,750%), and 3 years old (6,250%). Similarly, in the typical group, the majority of participants were also 5 years old (75%), followed by 6 years old (15,909%) and 4 years old (9,091%). No 3-year-old children were present in the typical group. These findings suggest that the majority of children involved in this study were between the ages of 4 and 6, with 5 years old being the most common age in both groups.

Table 2. Frequencies for Age

Category	Age (Years)	Percent
Atypical	3-year-old child	6,250%
	4-year-old child	18,750%
	5-year-old child	50%
	6-year-old child	25%
Typical	3-year-old child	0%
	4-year-old child	9,091%
	5-year-old child	75%
	6-year-old child	15,909%

Descriptive Statistics of Semantic Comprehension

Descriptive statistics regarding semantic comprehension abilities are presented in Table 3. The mean semantic comprehension score in the typical group was substantially higher ($M = 14,250$; $SD = 3.635$) compared to the atypical group ($M = 3,656$; $SD = 4.512$). The median score in the typical group was 14, while in the atypical group it was 2, indicating a striking difference in central tendency. The score ranges were relatively similar – 15 in the typical group and 16 in the atypical group. However, the higher standard deviation in the atypical group indicates a greater variability in scores, reflecting more pronounced disparities in semantic comprehension abilities within this group.

Table 3. Semantic Comprehension

Category	Mean	Media	Std. Deviation	Range
Atypical	3,656	2,000	4,512	16
Typical	14,250	14,000	3,635	15

Between-Group Comparison

To determine whether the difference in semantic comprehension scores between typically and atypically developing preschool children was statistically significant, an independent sample t-test was conducted. The analysis revealed a highly significant difference between the two groups, with a test statistic of $t(74) = -11.326$, $p < .001$. The very small p-value ($< .001$) indicates that the likelihood of this difference occurring by chance is extremely low.

Effect Size

The effect size was calculated using Cohen's d , yielding a value of -2.631 . According to Cohen's conventions, this represents a very large effect, indicating that the difference in semantic comprehension between the typical and atypical groups is not only statistically significant but also substantially meaningful in practical and clinical terms. This suggests a strong influence of developmental status (typical vs. atypical) on children's semantic comprehension performance.

Table 4. Between-Group Comparison

	t	df	p	Cohen's d
Raw Score	-11.326	74	< .001	-2.631

Covariance Analysis (ANCOVA)

To determine whether the difference in semantic comprehension scores between typically and atypically developing children remained significant after controlling for age, an analysis of covariance (ANCOVA) was conducted. The results indicated that the group category (typical vs. atypical) had a statistically significant effect on semantic comprehension scores, $F(1, 73) = 126.193$, $p < .001$. This finding demonstrates that even after accounting for age, a substantial and significant difference exists between the two groups.

In contrast, the covariate age did not have a significant effect, $F(1, 73) = 2.578$, $p = 0.113$, suggesting that age did not contribute meaningfully to the differences in scores. Therefore, developmental status (typical vs. atypical) emerged as the primary factor influencing semantic comprehension performance.

Table 5. Covariance Analysis

Cases	Sum of Squares	df	Mean Square	F	p
Category	2002.757	1	2002.757	126.193	< .001
Age	40.918	1	40.918	2.578	0.113
Residuals	1158.551	73	15.871		

The present study revealed a highly significant difference in semantic comprehension between typically and atypically developing preschool children. Children with typical development scored substantially higher on the Preschool Semantic Comprehension Test (TPSP) compared to their atypical peers. Inferential analysis through an independent samples t-test yielded a p-value of $< .001$ and a very large effect size (Cohen's $d = -2.631$), indicating that developmental status is a primary factor influencing children's semantic comprehension—beyond random variation or chronological age alone.

Further support was provided by the results of the covariance analysis (ANCOVA), which demonstrated that the difference in scores remained significant even after controlling for age as a covariate. Age had no significant influence on semantic comprehension scores ($p = .113$), while developmental status remained a highly significant factor ($p < .001$). These findings confirm that the developmental condition of a child plays a critical role in their semantic competence and support the construct validity of TPSP as a reliable and sensitive instrument for distinguishing developmental profiles.

These results align with prior literature showing that children with language disorders or atypical developmental conditions often exhibit deficits in semantic comprehension and usage (Foster-Cohen et al., 2023). This includes children diagnosed with Autism Spectrum Disorder (ASD), Attention Deficit Hyperactivity Disorder (ADHD), language disorder (LD), intellectual disability (ID), and global developmental delay (GDD), each of which presents unique semantic challenges. These commonly manifest as difficulty understanding word relationships, naming objects, and organizing words into semantic categories.

Children with ASD often show marked difficulties in both pragmatic and semantic aspects of language, including limitations in interpreting ambiguous meanings, metaphors, and context-dependent word use (Eigsti et al., 2011; Geurts & Embrechts, 2008; Kwok et al., 2015; Norbury, 2005; Tager-Flusberg, 2006). A study by Kwok et al. (2015) found that even when syntactic skills were relatively preserved, children with ASD showed significantly lower performance in semantic relationship comprehension. This suggests that semantic deficits are distinct and not merely the result of general language impairment.

In children with ADHD, semantic difficulties are frequently secondary to deficits in attention and executive functioning, which affect the integration of meaning in complex language tasks (El Sady et al., 2013; Väisänen et al., 2014). Kim and Kaiser (2000) reported that while children with ADHD may understand individual word meanings, they often struggle to apply them appropriately in contextualized speech, such as narrative construction and inferential reasoning.

For children with intellectual disabilities, semantic delays tend to be more global. Longitudinal studies have shown that the development of conceptual vocabulary and semantic classification progresses much more slowly than in typically developing peers, even when matched for mental age (Abbeduto et al., 2007). These limitations result in restricted lexical networks and impaired semantic generalization.

Language disorder (LD), by contrast, is characterized by core deficits in semantic processing. Children with LD often struggle to acquire new vocabulary, understand nuanced meanings, and categorize words meaningfully (Alt & Plante, 2006; Geurts & Embrechts, 2008; Nudel et al., 2023; Rinaldi et al., 2021; Verbeek et al., 2023). These challenges necessitate intervention strategies specifically targeting semantic expansion and lexical flexibility.

Children with GDD face multi-domain delays, and semantic comprehension is often hampered by co-occurring cognitive, linguistic, and social limitations. Their vocabulary learning tends to be passive and tied to concrete experiences, and they often require repetitive, scaffolded learning strategies (Saban-Bezalel, 2025). These developmental conditions highlight the necessity of precise and sensitive assessment tools.

The TPSP was shown to significantly distinguish these semantic differences, thus strengthening its empirical validity and potential utility in both clinical and educational contexts. As an instrument developed and validated in the Indonesian sociolinguistic context, TPSP addresses the urgent need for culturally responsive and linguistically appropriate language assessment tools. This is especially critical in Indonesia, a nation with high linguistic diversity and limited access to standardized language tests tailored to local needs.

This study offers several notable strengths. First, it employed an assessment instrument that was developed locally and validated psychometrically in the Indonesian context. This enhances both ecological validity and practical applicability. Second, the use of robust statistical analyses – independent t-test, effect size calculations, and ANCOVA – provided comprehensive insights into group differences and helped isolate the developmental status as a key predictor of semantic ability.

However, the study is not without limitations. The sample size, while sufficient for initial validation, is limited and does not represent the full diversity of the Indonesian preschool population. Additionally, sociodemographic variables such as home language, communication environment, and educational background were not included, which limits the generalizability of findings.

From a clinical speech-language therapy perspective, the TPSP demonstrates promising potential as a screening tool for early detection of language disorders. Speech-language pathologists may incorporate TPSP as part of an initial assessment battery to inform semantic-focused intervention planning. Its application could be particularly impactful in underserved regions lacking standardized language assessment tools. This aligns with national and international calls for more inclusive and context-specific approaches in child health and educational services.

Conclusion

Overall, this study provides strong evidence supporting the validity and effectiveness of the Preschool Semantic Comprehension Test (TPSP) as a developmental screening instrument. The significant differences identified in semantic comprehension between typical and atypical children were not only statistically robust but also clinically meaningful. The findings highlight the importance of developing culturally grounded assessment tools to facilitate early detection and intervention in children with language difficulties. Future research is recommended to examine the longitudinal sensitivity of TPSP, expand its use across multilingual populations, and explore its integration into broader child development screening programs in Indonesia.

Funding Statement

This research was supported by Kementerian Pendidikan Tinggi, Sains, dan Teknologi Republik Indonesia [017/LL10/DT.05.00/PL/2025].

References

- Abbeduto, L., Warren, S. F., & Conners, F. A. (2007). Language development in Down syndrome: From the prelinguistic period to the acquisition of literacy. *Mental Retardation and Developmental Disabilities Research Reviews*, 13(3), 247–261.
- Alt, M., & Plante, E. (2006). Factors that influence lexical and semantic fast mapping of young children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 49(5). [https://doi.org/10.1044/1092-4388\(2006/068\)](https://doi.org/10.1044/1092-4388(2006/068))
- Barry, M. J., Nicholson, W. K., Silverstein, M., Chelmow, D., Coker, T. R., Davis, E. M., Donahue, K. E., Jaén, C. R., Li, L., Mangione, C. M., Ogedegbe, G., Rao, G., Ruiz, J. M., Stevermer, J., Tsevat, J., Underwood, S. M., & Wong, J. B. (2024). Screening for Speech and Language Delay and Disorders in Children: US Preventive Services Task Force Recommendation Statement. In *JAMA* (Vol. 331, Issue 4). <https://doi.org/10.1001/jama.2023.26952>
- Chen, S. Q., Ge, L. J., Chen, G. Q., He, W. J., Guo, J. J., Li, Y., Shi, L., & Cheng, Y. L. (2022). Identification, diagnosis, and early intervention of children with developmental language disorder in Ningxia. *Translational Pediatrics*, 11(2). <https://doi.org/10.21037/tp-21-601>
- Eigsti, I. M., De Marchena, A. B., Schuh, J. M., & Kelley, E. (2011). Language acquisition in autism spectrum disorders: A developmental review. In *Research in Autism Spectrum Disorders* (Vol. 5, Issue 2). <https://doi.org/10.1016/j.rasd.2010.09.001>
- El Sady, S. R., Nabeih, A. A. S., Mostafa, E. M. A., & Sadek, A. A. (2013). Language impairment in attention deficit hyperactivity disorder in preschool children. *Egyptian Journal of Medical Human Genetics*, 14(4). <https://doi.org/10.1016/j.ejmhg.2013.05.001>
- Feltner, C., Wallace, I. F., Nowell, S. W., Orr, C. J., Raffa, B., Middleton, J. C., Vaughan, J., Baker, C., Chou, R., & Kahwati, L. (2024). Screening for Speech and Language Delay and Disorders in Children 5 Years or Younger: Evidence Report and Systematic Review for the US Preventive Services Task Force. In *JAMA* (Vol. 331, Issue 4). <https://doi.org/10.1001/jama.2023.24647>
- Foster-Cohen, S., Macrae, T., & Newbury, J. (2023). Variation in morpho-lexical development within and between diagnoses in children with neurodevelopmental disorders. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.968408>
- Geurts, H. M., & Embrechts, M. (2008). Language profiles in ASD, SLI, and ADHD. *Journal of Autism and Developmental Disorders*, 38(10). <https://doi.org/10.1007/s10803-008-0587-1>
- Jerome, A. C., Fujiki, M., Brinton, B., & James, S. L. (2002). Self-esteem in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 45(4). [https://doi.org/10.1044/1092-4388\(2002/056\)](https://doi.org/10.1044/1092-4388(2002/056))
- Kemenkes. (2023). *Survei Kesehatan Indonesia*. Kemenkes.
- Kim, O. H., & Kaiser, A. P. (2000). Language Characteristics of Children with ADHD. *Communication Disorders Quarterly*, 21(3). <https://doi.org/10.1177/152574010002100304>
- Kwok, E. Y. L., Brown, H. M., Smyth, R. E., & Oram Cardy, J. (2015). Meta-analysis of receptive and expressive language skills in autism spectrum disorder. *Research in Autism Spectrum Disorders*, 9. <https://doi.org/10.1016/j.rasd.2014.10.008>
- McAllister, J., & Miller, J. E. (2013). *Introductory Linguistics for Speech and Language Therapy practice*. John Wiley & Sons.
- Norbury, C. F. (2005). The relationship between theory of mind and metaphor: Evidence from children with language impairment and autistic spectrum disorder. *British Journal of Developmental Psychology*, 23(3). <https://doi.org/10.1348/026151005X26732>
- Nudel, R., Christensen, R. V., Kalnak, N., Schwinn, M., Banasik, K., Dinh, K. M., Erikstrup, C., Pedersen, O. B., Burgdorf, K. S., Ullum, H., Ostrowski, S. R., Hansen, T. F., & Werge, T. (2023). Developmental language disorder – a comprehensive study of more than 46,000 individuals. *Psychiatry Research*, 323. <https://doi.org/10.1016/j.psychres.2023.115171>

- Paul, D., & Roth, F. P. (2011). Guiding principles and clinical applications for speech-language pathology practice in early intervention. *Language, Speech, and Hearing Services in Schools, 42*(3). [https://doi.org/10.1044/0161-1461\(2010/09-0079\)](https://doi.org/10.1044/0161-1461(2010/09-0079))
- Rinaldi, S., Caselli, M. C., Cofelice, V., D'Amico, S., De Cagno, A. G., Corte, G. Della, Di Martino, M. V., Di Costanzo, B., Levorato, M. C., Penge, R., Rossetto, T., Sansavini, A., Vecchi, S., & Zoccolotti, P. (2021). Efficacy of the treatment of developmental language disorder: A systematic review. In *Brain Sciences* (Vol. 11, Issue 3). <https://doi.org/10.3390/brainsci11030407>
- Saban-Bezalel, R. (2025). Communicative Profiles of Children with Developmental Delay Compared to Age- or Language-Matched Typically Developing Peers. *Journal of Autism and Developmental Disorders*. <https://doi.org/https://doi.org/10.1007/s10803-025-06882-0>
- Shipley, K. G., & McAfee, J. G. (2023). *Assessment in Speech-Language Pathology: A Resource Manual*. Plural Publishing.
- Tager-Flusberg, H. (2006). Defining language phenotypes in autism. *Clinical Neuroscience Research, 6*(3-4). <https://doi.org/10.1016/j.cnr.2006.06.007>
- Taruna, R. (2024). Indonesian Language Normative Test: Preschool Semantic Comprehension Test. *Journal of Speech Language and Communication Research, 1*(1), 33-38.
- Väisänen, R., Loukusa, S., Moilanen, I., & Yliherva, A. (2014). Language and pragmatic profile in children with ADHD measured by children's communication checklist 2nd edition. *Logopedics Phoniatrics Vocology, 39*(4). <https://doi.org/10.3109/14015439.2013.784802>
- Verbeek, L., Vissers, C., Kleemans, T., Scheper, A., & Verhoeven, L. (2023). Early intervention of language and behavior in monolingual and bilingual preschoolers with Developmental Language Disorders. *Early Childhood Research Quarterly, 64*. <https://doi.org/10.1016/j.ecresq.2023.02.007>
- Wadman, R., Durkin, K., & Conti-Ramsden, G. (2008). Self-Esteem, shyness, and sociability in adolescents with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research, 51*(4). [https://doi.org/10.1044/1092-4388\(2008/069\)](https://doi.org/10.1044/1092-4388(2008/069))
- Ziegenfusz, S., Paynter, J., Flückiger, B., & Westerveld, M. F. (2022). A systematic review of the academic achievement of primary and secondary school-aged students with developmental language disorder. In *Autism and Developmental Language Impairments* (Vol. 7). <https://doi.org/10.1177/23969415221099397>