Validity and Reliability Test Of Gross Motor Development-3 For Children Aged 4-5 In Indonesia

Frida Deswita Putri Kemal Pasha[1], M.E. Winarno[2*], Mu'arifin[3]

Universitas Negeri Malang (Indonesia)
*Corresponding Author: m.e.winarno.fik@um.ac.id

Frida Deswita Putri Kemal Pasha, M.E. Winarno, Mu'arifin

Abstract. This study aims to test the empirical validity, reliability and objectivity of the TGMD-3. The population in this study amounted to 45 children with an age range of 4-5 years conducted in three schools. The sampling technique in this study used total sampling so that the entire study population would be the research sample. Data analysis uses empirical validity test using Pearson's simple formula, reliability test using Cronbach's Alpha coefficient and objectivity test using Pearson's simple formula. The results of the empirical validity test obtained a result of 0.504 with an r table of 0.294. So that it is obtained 0.504> 0.294, then the instrument can be said to be valid. The results of the process assessment reliability test obtained a Cronbach's Alpha reliability coefficient of 0.954 with "Excellent" criteria and the product assessment reliability test obtained a Cronbach's Alpha reliability coefficient of 0.895 with "High" criteria. The objectivity test results show that the process assessment conducted by the first examiner and the second examiner has a result of 0.812 with the criteria "High" and the objectivity test results of the product assessment conducted by the first examiner and the second examiner have a result of 0.999 with the criteria "Excellent". This shows that the TGMD-3 really measures children's performance in accordance with the results achieved, has constancy between assessors and avoids bias. So it can be concluded if the TGMD-3 is proven to be empirically valid, reliable and objective to measure gross motor skills of early childhood with an age range of 4-5 years.
Keywords: TGMD-3, validity, reliability and objectivity, gross motor, early childhood, 4-5 years old, Indonesia

Introduction

Basically, gross and fine motor skills are the foundation for children's ability to carry out physical activities and daily routines such as walking, running, writing, playing (Bangun, 2016; Lestariningrum, Iftitah, & Pangastuti, 2022) developing cognitive abilities that have a strong correlation with children's gross motor levels with cognitive progress (Syafi'i, Marfiyanto, & Rodiyah, 2018). Children who have adequate motor skills tend to do better in terms of concentration, problem solving, and learning academic skills such as reading and writing (Jaenudin & Sahroni, 2021). Children's independence is also encouraged through the development of good motor skills, so that it can make it easier for children to carry out daily activities (Hanief, Nanda, Yulingga, & Sugito, 2017).

Gross motor skills in early childhood include large muscle activities that are important for children's physical development and motor skills. This involves controlling body movements that involve coordination of large muscles. This is supported by the theory put forward by Gallahue, Ozmun, & Goodway (2012); Papalia, Olds, & Feldman (2009); Santrock (2010), which states that motor development involves controlling body movements through large muscle activity, including three basic movements, namely locomotor basic movements, non-locomotor basic movements, and manipulative basic movements.

Based on needs analysis and observations, it was found that there are obstacles that occur, namely the lack of teacher focus in learning gross motor skills in early childhood so that teachers tend to emphasize the development of fine motor skills such as drawing, coloring, and sticking. The lack of physical activities carried out by the teacher, where children's time in learning should be balanced with gross motor and fine motor. Because of the tendency to develop children's fine motor skills, physical activities that support children's development in gross motor skills are often limited. Another statement was expressed by the teacher who mentioned that if the teacher carried out gross motor activities for children, it was only limited to getting children to move, no assessment was carried out due to the lack of more specific instruments, where the instrument for assessing gross motor skills in early childhood was limited and the lack of understanding of teachers regarding assessment techniques, making it difficult to choose and use the right assessment techniques to measure gross motor development in early childhood.

One of the instruments regarding gross motor, namely the Test of Gross Motor Development-3 (TGMD-3) developed by Ulrich (2019) which examines basic locomotor movements and ball skills. The TGMD-3 is the latest version of the TGMD published in 1985 by Ulrich, so researchers used the latest version, the TGMD-3 published in 2019. In the instrument, there are 13 kinds of tests that will be given to children. This includes 6 tests on basic locomotor movements and 7 tests on ball skills. TGMD-3 is declared as a valid instrument in measuring children's motor skills.

Previous research conducted by Matthias O. Wagner found that the TGMD-3 was valid and reliable enough to measure children's abilities in Germany (Wagner, Kipling Webster, & Ulrich, 2017), Finland (Rintala,

Sääkslahti, & Livonen, 2017), Peru (Mamani-Ramos et al., 2023), Iran (Mohammadi, Bahram, Khalaji, Ulrich, & Ghadiri, 2019), Slovenia and Canada (Maïano et al., 2021; Marinšek, Bedenik, & Kovac, 2023). There are previous studies that state that the TGMD-3 can be used to examine Fundamental Motor Skill (FMS) as a whole so as to get a better picture to evaluate the level of Fundamental Motor Skill (FMS) in children (Garn & Webster, 2021; Rey, Carballo-Fazanes, Varela-Casal, & Abelairas-Gómez, 2020; Temple & Foley, 2016).

The various studies above explain that the TGMD-3 is an instrument that can measure motor skills validly and reliably in various countries and provide a satisfactory picture or results in terms of measuring children's motor skills in these countries. This makes the measurement of children's motor skills very important so that the assessment must be carried out properly.

The impact that will occur if the assessment is not carried out properly, such as late identification, where children who experience difficulties in gross motor development will most likely not be detected early so that the child's opportunity to get appropriate intervention is limited. The learning program prepared by the teacher becomes ineffective because the program is not in accordance with the needs of the child. The lack of attention to gross motor development can hinder children's holistic development. These consequences can certainly be overcome if instrument development is carried out validly and reliably.

This study uses the TGMD-3 instrument by taking several tests of locomotor basic movements and ball skills. This is done to adapt the test items to the characteristics of children in Indonesia. The tests that will be carried out include 4 tests on locomotor movements, including: run, gallop, skip, and horizontal jump, while in ball skills there are 4 tests, including: two-hand catch, kick a stationary ball, overhand throw, and underhand throw. The researcher's consideration in adopting the TGMD-3 is because there has been no study of its validity and reliability for early childhood in Indonesia. The research examining validity and reliability was conducted in (Duncan et al., 2022), Brazil (Valentini, Zanella, & Webster, 2017), and Spain (Estevan et al., 2017) so that validity and reliability have not been detected in Indonesia.

Based on the explanation above, there is a gap between the expectations that must be met and the actual conditions in the field. The results of the needs analysis indicate that the implementation of early childhood gross motor skills learning has been carried out but no assessment has been carried out. The novelty in this research lies in the research instrument using the Test of Gross Motor Development: Third Edition developed by Dale A. Ulrich. Of course this is an innovation for researchers to develop assessment instruments regarding early childhood gross motor skills according to teacher needs.

Materials and methods

Research Design

This study uses a research design in the form of instrument development which aims to develop measuring instruments that can be used effectively in measuring gross motor skills of early childhood. The development process in this study involves several stages, namely the needs analysis stage, identification of important components, selecting instrument items, completing facilities and infrastructure, making test guides and test rubrics, determining research subjects, conducting trials, instrument validation tests and instrument reliability tests (Miller, Linn, & Gronlund, 2009).

Data Type

This research uses quantitative data collected in the form of numbers from the measurement of gross motor skills of early childhood. Furthermore, the data will be interpreted qualitatively so as to provide a clearer and more in-depth description and explanation (Sugiyono, 2019).

Research Population

The population in this study involved early childhood aged 4-5 years at Dharma Wanita 1 Ketawang Kindergarten totaling 17 children, Dharma Wanita Wonokerto Kindergarten totaling 18 children, and Dharma Wanita Jantok Kindergarten totaling 10 children. So, the total population obtained from the three schools amounted to 45 children.

Research Sample and Sampling Technique

This study uses a total sampling technique, in which the entire research population will be the research sample (Budiwanto, 2017). So, in this study the total sample was 45 children.

Research Instruments

The research instrument used is a test instrument that includes an assessment of early childhood gross motor skills that adopts the TGMD-3. Data collection techniques are observation which is useful for knowing the conditions that occur in the field, interviews conducted with teachers to strengthen the findings from observations and confirm the phenomena that occur, questionnaires include test instruments that will be carried out by early childhood related to gross motor and documentation.

Validity Test

Validity in an instrument is useful for providing an accuracy if the instrument used is confirmed to measure

what will be measured. In this study, the validity used is empirical validity, where data will be obtained from the results of testing the research sample. The data analysis technique for the empirical validity test in this study uses the Pearson simple correlation formula. Calculation of the validity test using SPSS software with the following criteria: (1) If r count> r table then the statement can be said to be valid, (2) If r count < r table then the question can be said to be invalid.

Reliability Test

A good instrument is an instrument that has reliability. Reliability can also be referred to as constancy, where the instrument can measure steadily. Reliability can be obtained from the fixed results of the test carried out on the first trial test with the second trial test, and so on. The reliability test in this study, namely by using the test-retest method, the reliability test analysis was carried out by correlating the first trial test with the second trial test conducted by the first examiner and the second examiner (Winarno, 2013). The reliability test analysis uses Cronbach's alpha so that the reliability value obtained can be categorized using the reliability criteria limit table revealed by Kirkendall, Gruber, & Johnson (1980).

The calculation of the reliability test uses the help of SPSS software with the limitations of the criteria in the reliability test of a test revealed by Kirkendall, Gruber & Johnson (1980) as follows:

Table 1. Limitation of Reliability Criteria

Limitation of Reliability Criteria	
0,00 – 0,67	Unacceptable
0,68 – 0,87	Average
0,88 – 0,92	High
0,93 – 1,00	Excellent

Source: Kirkendall, Gruber, & Johnson (1980)

Objectivity Test

Objectivity is an important component, where the assessment is carried out by at least two or more assessors. The objectivity test is carried out to determine the objectivity coefficient by correlating the results of the assessment conducted by the first examiner with the second examiner. The formula used is using Pearson's simple correlation which will be categorized according to the following criteria:

Table 2. Limitation of Objectivity Criteria

Limitation of Objectivity Criteria	
0,00 – 0,57	Unacceptable
0,68 – 0,77	Average
0,78 – 0,87	High
0,88 – 1,00	Excellent

Source: Kirkendall, Gruber, & Johnson (1980)

Results

Validity Result

The results of the validity test use Pearson's simple formula which calculates the correlation between process assessment and product assessment, so as to produce an instrument validity coefficient. The validity test includes empirical validity obtained from analyzing the data from the assessment results by the sample. The following is a table of the results of the validity test calculation between process assessment and product assessment using SPSS software:

Table 3. Empirical Validity Test Results Using Pearson Simple Correlation Formula Correlations

Correlations			
		Total Process	Total Product
Total Process	Pearson Correlation	1	.504**
	Sig. (2-tailed)		.000
	N	45	45
Total Product	Pearson Correlation	.504**	1
	Sig. (2-tailed)	.000	
	N	45	45

From the test results above, it can be seen that the Pearson correlation value is 0.504. Referring to the validity criteria which states that if r count> r table then the instrument can be said to be valid, this is evident from the validity test results where the results of r count are 0.504> r table of 0.294 which is obtained from a total sample of 45 with degree of freedom (N-2). So, it can be concluded that the instrument used includes empirical validity, so that the instrument is suitable for measuring gross motor skills of 4-5 year old children.

Reliability Result

The results of the reliability test that calculates the constancy between the process assessment on the first trial and the second trial, so as to produce the Cronbach's Alpha reliability coefficient. The following table shows the results of the reliability test calculation between the process assessment in the first experiment and the second experiment using SPSS software:

Table 4. Reliability Test Results of Process Assessment Using Cronbach's Alpha Coefficient

Reliability Statistics	
Cronbach's Alpha	N of Items
.954	2

From the test results above, it can be seen that the Cronbach's Alpha reliability coefficient is 0.954. Referring to the criteria limits expressed by Kirkendall, Gruber, & Johnson (1980), the reliability of the process assessment in the first experiment with the second experiment can be categorized as "Excellent" criteria.

This shows that the reliability of the instrument is very good, so that the process assessment on the instrument is said to be reliable and feasible to use to measure the truth of motion from the gross motor skills of 4-5 year old children.

Next, namely calculating the reliability between the product assessment on the first trial and the second trial, so as to produce the Cronbach's Alpha reliability coefficient. The following is a table of the results of the calculation of the reliability test between the product assessment in the first experiment and the second experiment using SPSS software:

Table 5. Results of Product Assessment Reliability Test Using Cronbach's Alpha Coefficient

Reliability Statistics	
Cronbach's Alpha	N of Items
.895	2

From the test results above, it can be seen that the Cronbach's Alpha reliability coefficient is 0.895. Referring to the criteria limits expressed by Kirkendall, Gruber, & Johnson (1980), the reliability of the process assessment in the first experiment with the second experiment can be categorized as "High" criteria. This shows that the reliability of the instrument is very high, so that the product assessment on the instrument is said to be reliable and feasible to use to measure the results of gross motor skills of 4-5 year old children.

Objectivity Result

The objectivity test results using Pearson's simple formula are used to determine the objectivity coefficient by correlating the results of the assessment conducted by the first examiner with the second examiner. The following is a table of the results of the objectivity test calculation between the results of the process assessment by the first examiner and the second examiner using SPSS software:

Table 6. Results of the Process Assessment Objectivity Test Using the Pearson Simple Correlation Formula

Correlations			
		First Tester Process	Second Tester Process
First Tester Process	Pearson Correlation	1	.812**
	Sig. (2-tailed)		.000
	N	45	45
Second Tester Process	Pearson Correlation	.812**	1
	Sig. (2-tailed)	.000	
	N	45	45

From the table above, it can be seen that the objectivity coefficient between the results of the process assessment carried out by the first examiner and the second examiner is 0.812. The results of the objectivity coefficient will be categorized according to the criteria expressed by Kirkendall, Gruber, & Johnson (1980) and get the criteria "High". This shows that the process assessment instrument used is very clear and unbiased, which is indicated by the results of the process assessment of the two examiners whose results are relatively the same.

Next, namely conducting an objectivity test on the results of the product assessment conducted by the first examiner with the second examiner. The formula used is Pearson's simple formula. The following is a table of the results of the objectivity test calculation between the results of the product assessment by the first examiner and the second examiner using SPSS software:

Table 7. Product Assessment Objectivity Test Results Using Pearson Simple Correlation Formula

Correlations			
		First Tester Product	Second Tester Product
First Tester Product	Pearson Correlation	1	.999**
	Sig. (2-tailed)		.000
	N	45	45
Second Tester Product	Pearson Correlation	.999**	1
	Sig. (2-tailed)	.000	
	N	45	45

From the table above, it can be seen that the objectivity coefficient between the results of the product assessment conducted by the first tester and the second tester is 0.999. The results of the objectivity coefficient will be categorized according to the criteria expressed by Kirkendall, Gruber, & Johnson (1980) and get the criteria "Excellent". This shows that the process assessment instrument used is very clear and unbiased, which is indicated by the results of the product assessment of the two examiners whose results are relatively the same.

Discussion

Gross motor in early childhood is a development carried out by children in physical coordination involving nerve centers and muscles (Gallahue et al., 2012). The movement involves large muscles in carrying out various activities. According to Samsudin (2008), elements in early childhood gross motor include the ability of locomotor movement, non-locomotor movement and manipulative movement. The stages in early childhood gross motor development are certainly not the same between children, due to influencing factors, such as the maturity of the child's muscles and nerves (Gallahue,

2017). The series of gross motor development begins with the head, then continues on the trunk, hands and ends at the feet (Hurlock, 1972). So, it can be concluded that if a child's motor development, muscle and nerve maturity is a major factor in the success of children performing gross motor skills.

The importance of gross motor for early childhood is that it can support their health and physical fitness, as a place to channel the energy held in the child's body so that the child's body will be free from tension, anxiety and despair (Gallahue et al., 2012). Children who can perform gross motor skills well will have high self-confidence and can provide opportunities for children to learn social skills (Suyanto, 2005). This will have an effect on the child's life where the more mature the child will be the better in doing body movements and social skills.

Based on the explanation above, we can know if gross motor skills in early childhood will affect the quality of life of children, so there is a need for early assessment. The assessment aims to evaluate the obstacles that occur in children in their motor development, so that teachers or parents can provide interventions to improve the child's gross motor (Winarno, 2004). The assessment carried out must include the principle of objectivity, which means that every child gets the same opportunity to do the test, accuracy in giving values or scores, has the principle of reliability, which repeats at least twice in doing the test and the principle of validity that really measures the gross motor development of early childhood (Winarno, 2014). So, it is important to do an assessment using appropriate and good instruments including, valid, reliable and objective.

Movement skills in children emerge and develop during preschool and elementary school. The physical growth and movement experiences that children have greatly influence their movement patterns. If parents and educators are unable to diagnose and correct disorders in basic movement development, it is feared that these movement problems will continue into adulthood. Webster & Ulrich (2017) from the University of Michigan, have compiled a test to evaluate basic movement skills called the (Test of Gross Motor Development). TGMD has three editions, TGMD in 1985, TGMD-2 in 2000 and TGMD-3 in 2019. TGMD is a test used to measure an individual's gross motor skills that develop early in life.

Therefore, the researcher will adopt the gross motor skills test instrument developed by Dale A. Ulrich entitled Test of Gross Motor Development: Thrid Edition (TGMD-3). The use of TGMD-3 in this study is certainly carried out with consideration, namely (1) TGMD-3 has gone through a careful and rigorous development process, involving a fairly extensive field trial so as to produce an instrument with high validity and reliability. So, the TGMD-3 is a reliable instrument, (2) The TGMD-3 instrument has been empirically tested and proven to be able to measure children's gross motor development accurately, (3) The TGMD-3 has a very standardized implementation procedure, from sample selection to scoring, (4) The TGMD-3 covers various aspects of gross motor development, including locomotor movements and control objects, this provides a comprehensive picture of children's motor abilities, (5) The TGMD-3 manual is arranged in detail, clearly and easily so that it makes it easier for researchers to take measurements.

This research aims to test the validity, reliability, and objectivity of the TGMD-3 as a measurement tool for gross motor skills in early childhood, specifically focusing on the characteristics of children in Indonesia aged 4-5 years. Validity testing consists of several aspects, which is content validity, empirical validity, and criterion validity (Budiwanto, 2017). Content validity in an instrument is useful to ensure that the instrument can represent all aspects of the ability to be measured, so that the testing objectives are achieved. The authors emphasize the need for data to support the content validity of the measuring instrument. A valid measuring instrument will provide accurate measurement results, while an invalid measuring instrument will produce data that does not match the initial objectives (Alim, Tomoliyus, Refiater, & Gani, 2024; Colio & Donato, 2022). However, in this study, the authors only used empirical validity testing conducted on the research sample. This is supported by Budiwanto (2017: 190) who states that testing content validity is not through statistical testing procedures, but through rational analysis. A practical way to analyze content validity is to examine the suitability between the arrangement of test items and the previously designed grid. Thus, this study examines the assessment instrument for empirical validity, reliability, and objectivity. This is done to provide new scientific development and contribute to development in conducting appropriate interventions to support the gross motor development of early childhood, especially in Indonesia.

The characteristics of the research sample used in this study are early childhood with an age range of 4-5 years in Indonesia. The population in this study involved early childhood aged 4-5 years at Dharma Wanita 1 Ketawang Kindergarten, totaling 17 children, Dharma Wanita Wonotengah Kindergarten, totaling 18 children, and Dharma Wanita Jantok Kindergarten, totaling 10 children. So, the total population obtained from the three schools amounted to 45 children.

The research instrument used in this study was the Test of Gross Motor Development-3 (TGMD-3). The TGMD-3 is an internationally standardized instrument for assessing gross motor skills in children. This instrument consists of a series of tests that measure various motor abilities such as walking, jumping, and throwing. Through a series of structured tests, the examiner can observe and assess the child's performance

in performing a series of tests. The data obtained were then analyzed to determine whether the TGMD-3 is valid and reliable in measuring gross motor skills in children in Indonesia. Data collection is done through the implementation of the TGMD-3 test individually on each research sample. Before conducting the test, both examiners will study and understand the test items to be assessed to ensure consistency in assessment. The test was conducted in a safe open space and equipped with appropriate equipment. Each test item was administered in accordance with the procedures set out in the TGMD-3 manual. The child's performance or process assessment is assessed based on several criteria, such as fluency of movement, coordination, and balance. The results of the assessment are recorded in detail on the observation sheet, both observations made by the first examiner and observations made by the second examiner. Furthermore, a recapitulation of the assessment results carried out by two testers was carried out so that the scores obtained could be analyzed. The validity test results using the Pearson simple formula obtained a result of 0.504. If r count> r table then the instrument can be said to be valid, so this is evident from the validity test results where the results of r count are 0.504> r table of 0.294 obtained from a total sample of 45 with degree of freedom (N-2).

The reliability test results obtained the Cronbach's Alpha reliability coefficient of 0.954. So that according to the [illegible] criteria expressed by Kirkendall, Gruber, & Johnson (1980), the reliability of the process assessment can be categorized as "Excellent" criteria. This shows that the reliability of the instrument is very good, so that the process assessment on the instrument is said to be reliable and feasible to use to measure the truth of motion from the gross motor skills of 4-5 year old children. Reliability tests were also carried out to assess the constancy of the product assessment which obtained a Cronbach's Alpha reliability coefficient of 0.895 which was categorized as "High" criteria. This shows the reliability of the instrument is very high, so 7 that the product assessment on the instrument is said to be reliable and feasible to use to measure the results of gross motor skills of 4-5 year old children.

The objectivity test results between the results of the process assessment conducted by the first examiner and the second examiner were 0.812. The objectivity coefficient results will be categorized according to the criteria expressed by Kirkendall, Gruber, & Johnson (1980) and get the criteria "High". the objectivity coefficient between the results of the product assessment conducted by the first examiner and the second examiner is 0.999 getting the criteria "Excellent". This shows that the process assessment instrument and product assessment used are very clear and unbiased, which is indicated by the results of the product assessment of the two examiners whose results are relatively the same. So that this instrument is unbiased in measuring gross motor skills of 4-5 year old children.

The results of the validity, reliability and objectivity analysis of the TGMD-3 instrument showed satisfactory results. The empirical validity test using Pearson correlation resulted in a calculated r value of 0.504, which is greater than r table, indicating a significant relationship between the process assessment score and the product assessment score. The results of this research can be confirmed from previous research studies that examined the TGMD-3. The research conducted by Wagner et al. (2017) investigated the factorial validity, divergent validity, concurrent validity, and predictive validity specifically concerning ball skills. The findings of their study provided evidence supporting the validity of the German version of the TGMD-3 for assessing motor skills in young children.

Meanwhile, the reliability test with Cronbach's Alpha coefficient resulted in a value of 0.954 for process assessment and 0.895 for product assessment, indicating a very high level of consistency. Previous research by Rintala et al. (2017) reported intrarater reliability, as indicated by Kappa coefficients, ranging from 0.69 to 0.77 for the locomotor subtest, ball skills subtest, and gross motor total score. Percentage agreement ranged from 87% to 91%, demonstrating good to excellent intrarater reliability. In the present study, interrater reliability yielded Kappa coefficients ranging from 0.57 to 0.64 for the locomotor subtest, ball skills subtest, and gross motor total score, indicating moderate to substantial reliability, with a percentage agreement of 83%. This study identified specific skills requiring further attention, such as the hop, horizontal jump, and two-hand strike, which necessitate clearer instructions or scoring criteria. In conclusion, this research affirmed that the TGMD-3 is a reliable instrument for measuring and analyzing children's gross motor skills.

The study by Mamani-Ramos et al. (2023) pursued two primary objectives: to translate the TGMD-3 into Peruvian Spanish and to analyze its psychometric properties, encompassing validity and reliability, for Peruvian children aged 6 to 10 years. Content validity, assessed by a Content Validity Index of 94.4% and a Kappa Coefficient > 0.802, demonstrated that the translated assessment criteria adequately represented the intended constructs. Construct validity analyses indicated excellent model fit, supporting the two-factor structure (locomotor and ball skills) of the Peruvian Spanish version of the TGMD-3. Reliability analyses, using Cronbach's Alpha, revealed a coefficient of 0.851 for internal consistency, 0.951 for test-retest reliability, 0.963 for interrater reliability, and 0.983 for intrarater reliability. These findings confirm that the Peruvian Spanish version of the TGMD-3 provides consistent and stable measurements.

In another study, Garn & Webster (2021) aimed to investigate the factor structure of the TGMD-3,

comparing a bifactor model with one-factor and two-factor models. Results concerning the model fit of the bifactor model demonstrated superior fit compared to the one-factor model, providing a more accurate representation of the relationships between the TGMD-3 items. Reliability analysis of the Fundamental Motor Skill (FMS) general factor revealed high internal consistency (0.797), whereas the locomotor (0.168) and ball skills (0.216) factors demonstrated very low reliability. This indicated that, after accounting for the general FMS factor, the unique variance remaining in the locomotor and ball skills subscales was not sufficiently reliable for separate interpretation. The low relative parameter bias values indicated accurate parameter estimation within the bifactor model. Consequently, the study concluded that interpretation of the TGMD-3 should prioritize the overall FMS score rather than the separate locomotor and ball skills subscales.

The results of the objectivity test produced an objectivity coefficient on process assessment of 0.812 and product assessment of 0.999. These results indicate that the TGMD-3 instrument is reliable for measuring gross motor skills in early childhood 4-5 years. Although the validity value only includes empirical validity, overall this instrument has proven to be valid, reliable, and objective so that it can be used in further research.

The research conducted by Temple & Foley (2016) yielded key findings supporting the developmental validity of the TGMD-3. Specifically, their results provided significant support for the TGMD-3's developmental validity. The obtained scores demonstrated improvement from grade 3 to grade 4, gender differences consistent with prior research, and the absence of a ceiling effect. These findings indicate that the TGMD-3 effectively measures children's motor skill development in accordance with developmental principles, thus validating its use for assessing gross motor skill development in children. The study by Mohammadi et al. (2019) reported a reliability coefficient of 0.82, indicating "excellent" reliability and demonstrating that the TGMD-3 is an instrument with consistent and stable measurement. Construct validity was assessed using both a two-factor model, which exhibited good model fit, and a one-factor model, which demonstrated adequate fit indices. The psychometric properties of the Persian version of the TGMD-3, examined in Iranian children, revealed adequate validity and excellent reliability, thereby supporting the use of the Persian TGMD-3 as a measurement tool for assessing gross motor skill development in this population.

Based on the results found in this study, it can be seen that the TGMD-3 has satisfactory validity, reliability, and objectivity coefficient values. This finding is also proven by previous research, which has similarities in the findings stating that the TGMD-3 instrument has high validity, reliability, and objectivity so that it can be used to measure early childhood gross motor skills both in other countries and in Indonesia. With the findings of this study, it can provide new knowledge related to gross motor assessment instruments that can be used in Indonesia to measure early childhood gross motor skills that have been tested for validity, reliability, and objectivity.

Conclusion

The conclusion of this study is to prove that the TGMD-3 has been tested for empirical validity, reliability using the test-retest method, and objectivity in measuring early childhood motor skills aged 4-5 years in Indonesia. The results of this study indicate that the TGMD-3 is able to measure children's gross motor skills accurately, consistently or steadily and unbiased or objectively. With these findings, it can provide a reference for teachers and further researchers to use the TGMD-3 as one of the measuring instruments that has been tested valid, reliable, and objective to measure the motor skills of early childhood aged 4-5 years, so that an evaluation of child development can be carried out so that intervention can be carried out if the child's gross motor development is hampered. The results of this study also contribute to the development of science in the field of child development, especially related to the assessment of gross motor skills.

Acknowledgments

Thanks are due to the State Universitas Negeri Malang for the budget provided for the implementation of this research. Thanks to the lecturers of the Faculty of Sports Science, State University of Malang who have guided and provided direction regarding this research. Thank you to the teachers and principals of the research sites who have given permission and assisted in the implementation of this research so that the research can run smoothly. Thank you to the research team in data collection who have actively contributed to this research. Thank you to all parties involved in this research so that the research can be completed properly. With the help and support provided, this research can produce extraordinary findings. Thank you very much.

References

Alim, A., Tomoliyus, Refiater, U. H., & Gani, I. (2024). Herramienta de medición de agilidad reactiva basada en sensores para deportes grupales de juegos en red: validez del contenido (Sensor-based Reactive Agility Measurement Tool for Net Game Group Sports: Content Validity). *Retos*, *51*, 167–171. Retrieved from https://recyt.fecyt.es/index.php/retos/article/view/100564

Bangun, S. Y. (2016). Peran Pendidikan Jasmani Dan

Olahraga Kajian Pustaka & Pembahasan. *Jurnal Publikasi Pendidikan*, *6*(1), 156–167.

Budiwanto, S. (2017). *Metode Penelitian dalam Keolahragaan*. Malang: Universitas Negeri Malang.

Colio, B. B., & Donato, F. S. (2022). Validation of the Trade Policy Sports (TP-SPORTS) instrument for the analysis of the commercial policy of sports centres and clubs. *Retos*, *46*(2014), 104–113.

Duncan, M. J., Martins, C., Bandeira, P. F. R., Issartel, J., Peers, C., Belton, S., … Behan, S. (2022). TGMD-3 Short Version: Evidence of Validity and Associations With Sex in Irish Children. *Journal of Sports Sciences*, *40*(2), 138–145. https://doi.org/10.1080/02640414.2021.1978161

Estevan, I., Molina-Garcia, J., Queralt, A., Álvarez, O., Castillo, I., & Barnett, L. (2017). Validity and Reliability of the Spanish Version of the Test of Gross Motor Development-3. *Journal of Motor Learning and Development*. https://doi.org/https://doi.org/10.1123/jmld.2016-0045

Gallahue, D. L. (2017). *Developmental Physical Education for All Children: Theory Into Practice* (5th ed.). Canada: Human Kinetics.

Gallahue, D. L., Ozmun, J. C., & Goodway, J. D. (2012). *Understanding Motor Development: Infants, Children, Adolescent, Adults* (7th ed.). New York: McGraw-Hill.

Garn, A. C., & Webster, E. K. (2021). Bifactor structure and model reliability of the Test of Gross Motor Development-3rd edition. *Journal of Science and Medicine in Sport*, *24*(1), 67–73. https://doi.org/10.1016/j.jsams.2020.08.009

Hanief, Nanda, Yulingga, & Sugito. (2017). Membentuk Gerak Dasar Pada Siswa Sekolah Dasar Melalui Permainan Tradisional. *Jurnal Sportif*, *1*(1), 60–73. Retrieved from https://ojs.unpkediri.ac.id/index.php/pjk/issue/view/318

Hurlock, E. B. (1972). Child Development. In *McGraw-Hill Book Company* (Fifth). New York: McGraw-Hill.

Jaenudin, U., & Sahroni, D. (2021). *Psikologi Pendidikan*. Bandung: Lagood's Publishing.

Kirkendall, D. R., Gruber, J. J., & Johnson, R. E. (1980). *Measurement and Evaluation for Physical Educators* (A. Lockhart, R. Kaplan, & D. Gray, eds.). Dubuqua, lowa: Brown Company Publishers.

Maïano, C., Morin, A. J. S., April, J., Webster, E. K., Hue, O., Dugas, C., & Ulrich, D. (2021). Psychometric Properties of a French-Canadian Version of the Test of Gross Motor Development-Third Edition (TGMD-3): A Bifactor Structural Equation Modeling Approach. *Measurement in Physical Education and Exercise Science*, *26*(1), 51–62. https://doi.org/10.1080/1091367X.2021.1946541

Mamani-Ramos, A. A., Damian-Nuñez, E. F., Torres-Cruz, F., Dextre-Mendoza, C. W., Alcarraz-Curi, M., Quisocala-Ramos, J. A., … Montoya-Castillo, P. M. (2023). Psychometric properties of the Peruvian version of the Gross Motor Development Test - Third Edition. *Retos*, *50*, 1180–1187. https://doi.org/10.47197/retos.v50.100031

Marinšek, M., Bedenik, K., & Kovac, M. (2023). Psychometric Proprieties of the Slovenian Version of the Test of Gross Motor Development–3. *Journal of Motor Learning and Development*, *11*(1), 209–223. https://doi.org/10.1123/jmld.2022-0055

Miller, M. D., Linn, R. L., & Gronlund, N. E. (2009). *Measurement and Assessment in Teaching* (Tenth). New Jersey: Pearson.

Mohammadi, F., Bahram, A., Khalaji, H., Ulrich, D. A., & Ghadiri, F. (2019). Evaluation of the psychometric properties of the Persian version of the test of Gross Motor Development-3rd edition. *Journal of Motor Learning and Development*, *7*(1), 106–121. https://doi.org/10.1123/JMLD.2017-0045

Papalia, D. E., Olds, S. W., & Feldman, R. D. (2009). *Human Development* (11th ed.; M. Ryan, ed.). New York: McGraw-Hill.

Reswanto, F., Nasrulriningrum, A., Iftitah, S. L., & Pangastuti, R. (2022). Perkembangan Fisik dan Motorik Anak. In Syofrianisda (Ed.), *CV. Azka Pustaka*. Pasaman Barat: CV. Azka Pustaka.

Rey, E., Carballo-Fazanes, A., Varela-Casal, C., & Abelairas-Gómez, C. (2020). Reliability of the test of gross motor development: A systematic review. *PLoS ONE*, *15*(7), e0236070. https://doi.org/10.1371/journal.pone.0236070

Rintala, P., Sääkslahti, A., & Livonen, S. (2017). Reliability Assessment of Scores from Video Recorded TGMD-3 Performances. *International Journal of Special Education*, *5*(1), 59–68. https://doi.org/https://doi.org/10.1123/jmld.2016-0007

Samsudin. (2008). *Pembelajaran Motorik di Taman Kanak-Kanak*. Jakarta: Prenada Media Group.

Santrock, J. W. (2010). *Child Development* (13th ed.). New York: McGraw-Hill.

Sugiyono. (2019). Metodelogi Penelitian Pendidikan Pendekatan Kualitatif dan R&D. In *Metodelogi Penelitian Pendidikan Pendektan Kualitatif dan R&D*. Alfabeta.

Suyanto, S. (2005). *Konsep Dasar Pendidikan Anak Usia Dini*. Jakarta: Direktorat Pembinaan Pendidikan Tenaga Kependidikan dan Ketenagaan Perguruan Tinggi.

Syafi'i, A., Marfiyanto, T., & Rodiyah, S. K. (2018).

Studi Tentang Prestasi Belajar Siswa Dalam Berbagai Aspek Dan Faktor Yang Mempengaruhi. *Jurnal Komunikasi Pendidikan*, *2*(2), 115. https://doi.org/10.32585/jkp.v2i2.114

Temple, V. A., & Foley, J. T. (2016). A Peek at the Developmental Validity of the Test of Gross Motor Development – 3. *Journal of Motor Learning and Development*, *29*, 622–627. https://doi.org/http://dx.doi.org/10.1123/jmld.2016-0005

Ulrich, D. A. (2019). *Test of Gross Motor Development: Third Edition* (3rd ed.). Dallas: Pro.ed.

Valentini, N. C., Zanella, L. W., & Webster, E. K. (2017). Test of Gross Motor Development- Third Edition: Establishing Content and Construct Validity for Brazilian Children. *Journal of Motor Learning and Development*, *5*(1), 15–28. https://doi.org/10.1123/jmld.2016-0002

Wagner, M. O., Kipling Webster, E., & Ulrich, D. A. (2017). Psychometric properties of the Test of Gross Motor Development, third edition (German Translation): Results of a pilot study. *Journal of Motor Learning and Development*, *5*(1), 29–44. https://doi.org/10.1123/jmld.2016-0006

Webster, E. K., & Ulrich, D. A. (2017). *Evaluation of the Psychometric Properties of the Test of Gross Motor Development — Third Edition*. (August). https://doi.org/10.1123/jmld.2016-0003

Winarno, M. E. (2006). *Evaluasi Dalam Pendidikan Jasmani dan Olahraga*. Jakarta: Center For Human Capacity Development.

Winarno, M. E. (2013). *Metodologi Penelitian Dalam Pendidikan Jasmani*. Malang: Universitas Negeri Malang (UM PRESS).

Winarno, M. E. (2014). Evaluasi Hasil Belajar Pendidikan Jasmani Olahraga dan Kesehatan. In *UM Press*. Malang: UM Press.