

Standardisation and validation of the TGMD-3 Malay version for assessing fundamental motor skills

Kathiresan Moorthy¹, Siti Hartini Azmi^{2*}, Nohazira Abd Rahim³

^{1,2,3}*Faculty of Sports Science and Coaching, Sultan Idris Education University Malaysia*

ARTICLE INFO

Article history:

Received 2 March 2026

Revised 3 September 2026

Accepted 9 September 2026

Online first 15 September 2026

Published 15 September

Keywords:

TGMD-3

Cross-cultural adaptation

Rasch analysis

Motor competence

Measurement invariance

Validation

ABSTRACT

This study examined the validity and measurement functioning of a Malay adaptation of the Test of Gross Motor Development–Third Edition (TGMD-3) for Malaysian primary-school settings. A structured cross-cultural adaptation process used independent forward translations, synthesis, backward translation and multidisciplinary expert review to preserve semantic, conceptual and contextual equivalence. Content validity was evaluated by seven experts, while instructional clarity and practical usability were appraised by 30 practising Physical Education teachers. Construct validity was subsequently examined using Rasch measurement with 64 Year Four pupils. Item-level content validity indices ranged from 0.86 to 1.00 and the scale-level average was 0.93. Mean clarity ratings ranged from 4.40 to 4.80 on a five-point scale, with Impact Index values from 4.05 to 4.75. Rasch analysis showed acceptable item fit, with Infit Mean Square values of 0.74–1.29 and Outfit Mean Square values of 0.71–1.34. Person reliability was 0.88 with a separation index of 2.65, whereas item reliability was 0.95 with a separation index of 4.37. No item displayed substantial gender-based differential item functioning, and the person–item map indicated satisfactory targeting across the assessed ability continuum. These converging findings support the Malay TGMD-3 as a culturally appropriate and psychometrically defensible instrument for systematic assessment of fundamental motor competence. Its use may strengthen evidence-informed instructional planning, developmental monitoring and more standardised Physical Education assessment in Malaysian primary schools.

^{2*}Corresponding author. *E-mail address* : siti.hartini@fsskj.upsi.edu.my

1. INTRODUCTION

Fundamental movement skills (FMS) provide the movement foundation from which children develop more specialised patterns required for physical activity, games and sport. Their importance extends beyond the execution of isolated motor tasks. A 2025 structural-equation-modelling meta-analysis showed that actual motor competence is meaningfully embedded in the relationship between perceived physical competence and physical activity across childhood and adolescence, reinforcing the developmental relevance of accurate motor assessment (Bourke *et al.*, 2025). Recent intervention syntheses likewise indicate that FMS can be improved through appropriately structured school and physical-activity programmes rather than being assumed to emerge automatically with maturation (Yin *et al.*, 2025; Zhang *et al.*, 2025). In particular, a 2026 network meta-analysis identified FMS-targeted programmes and general structured physical activity as the most effective approaches for improving total FMS in early childhood (Jiang *et al.*, 2026). These developments increase the need for assessment instruments that are sufficiently valid, interpretable and culturally appropriate to detect meaningful differences in motor competence.

The Test of Gross Motor Development–Third Edition (TGMD-3) is a process-oriented assessment designed to evaluate six locomotor and seven ball-skill performances in children (Ulrich, 2019). Its continued relevance is reflected in recent work using the TGMD-3 as a criterion measure for validating a brief preschool motor-competence check (Komura *et al.*, 2025) and in the development of large age-specific normative percentile values from more than 17,000 Italian children (Walters *et al.*, 2025). At the same time, contemporary research has highlighted that the apparent precision of a motor score depends not only on the child's performance, but also on the operational definition of performance criteria, rater judgement and the scoring structure of the instrument. Lorenzo-Martínez *et al.* (2025), for example, found significant score discrepancies when the same movement skills were evaluated with three process-oriented assessment tools. Ng *et al.* (2025) similarly identified subjectivity in rater judgements and inconsistencies in score aggregation as continuing challenges to TGMD comparability and standardisation. These concerns make rigorous adaptation and validation especially important when an established instrument is transferred across languages and educational systems.

Cross-cultural adaptation cannot be reduced to literal translation. The aim is to preserve the meaning of the construct while ensuring that terminology, instructions and performance criteria are understood consistently within the target setting (Beaton *et al.*, 2000; Guillemin *et al.*, 1993). This requirement is particularly salient for motor assessment because movement terminology may be interpreted differently by teachers and assessors whose sporting exposure, professional training and classroom practices vary. International TGMD-3 studies have demonstrated strong psychometric potential after careful adaptation and validation, including Spanish, Italian and French-Canadian applications (Estevan *et al.*, 2017; Magistro *et al.*, 2020; Mañano *et al.*, 2022). Reliability research further shows that rater agreement can improve when assessors explicitly establish shared interpretations of performance criteria (Carballo-Fazanes *et al.*, 2023), supporting the value of linguistically precise and operationally clear instructions.

In Malaysia, the Primary School Standard Curriculum positions Physical Education as part of holistic pupil development (Ministry of Education Malaysia, 2017). However, routine school assessment has traditionally given greater visibility to fitness monitoring than to standardised, process-oriented profiling of locomotor and manipulative competence. This creates a practical assessment gap: teachers may teach movement skills without access to a validated Malay instrument that supports consistent observation,

scoring and developmental monitoring. The gap is consequential because contemporary evidence shows that structured school-based programmes can produce meaningful improvements in FMS (Yin et al., 2025), while structured active play and skill-oriented Physical Education both outperform the assumption that unstructured activity alone will reliably develop proficiency (Liu et al., 2025). A valid Malay TGMD-3 would therefore provide a stronger measurement basis for identifying movement strengths and needs, monitoring instructional response and aligning classroom practice with the broader goal of balanced physical development advocated internationally (UNESCO, 2015).

Accordingly, this study evaluated the Malay adaptation of the TGMD-3 through three complementary sources of validity evidence. First, content relevance and conceptual fidelity were examined through multidisciplinary expert judgement. Second, linguistic clarity and practical usability were evaluated by practising Physical Education teachers who represent the intended school-based users of the instrument. Third, construct functioning was examined using Rasch measurement, including item fit, person and item separation, gender-based differential item functioning and person–item targeting. By combining cross-cultural adaptation, practitioner appraisal and modern psychometric analysis, the study responds directly to current calls for greater comparability and standardisation in motor competence assessment (Ng et al., 2025) and to emerging evidence that motor competence is multidimensional and should be evaluated with construct-sensitive measurement approaches (Bandeira et al., 2025).

2. LITERATURE REVIEW

Motor competence is increasingly understood as a multidimensional construct rather than a single undifferentiated score. Stodden et al. (2008) proposed a developmental relationship in which motor competence and physical activity become increasingly reciprocal across childhood. More recent evidence strengthens this conceptual position: Bourke et al. (2025) found that actual and perceived competence contribute to pathways associated with physical activity, while Bandeira et al. (2025), using exploratory graph analysis, questioned whether commonly used motor assessments always represent the same underlying construct. For school-based research, this means that assessment selection is not a neutral technical decision. The chosen instrument determines which performance features are observed, how proficiency is represented and whether change can be interpreted as true development rather than measurement artefact.

The TGMD-3 is among the most established process-oriented measures of gross motor competence. Evidence from cross-national studies supports its use while also illustrating the need for contextual validation. The Spanish version demonstrated suitable reliability and construct validity (Estevan et al., 2017), a large Italian study supported factorial validity and measurement invariance (Magistro et al., 2020), and a French-Canadian validation supported a bifactor representation with no meaningful differential item functioning across several participant characteristics (Maïano et al., 2022). Rasch analysis has also been used to examine the hierarchy and differential functioning of TGMD-3 performance criteria; Valentini et al. (2022) reported meaningful item-difficulty variation while finding no gender DIF for locomotor and ball-skill items. In 2025, Walters et al. extended the evidence base by establishing normative TGMD-3 percentiles for Italian children, illustrating the value of population-sensitive calibration when scores are used for developmental interpretation.

Measurement quality also depends on implementation. Rintala *et al.* (2017) showed that some TGMD-3 skills are more challenging than others to score consistently from video, and Maeng *et al.* (2017) reported excellent reliability for total and subscale scores among experienced raters but greater variability for individual skills. Carballo-Fazanes *et al.* (2023) subsequently demonstrated that interrater reliability improved after raters reached explicit agreement about performance criteria. Recent research reinforces this concern. Ng *et al.* (2025) highlighted subjectivity, data precision and score aggregation as barriers to comparability, whereas Lorenzo-Martínez *et al.* (2025) showed that three process-oriented tools can produce materially different proficiency estimates for the same skills. Komura *et al.* (2025), in validating the SMC-Kids, used the TGMD-3 as a criterion measure, further demonstrating its role as a reference instrument but also underscoring the importance of establishing validity in the population and language in which it is used.

The need for rigorous measurement is amplified by the growing intervention literature. Zhang *et al.* (2025) found generally positive effects of active-play interventions on children's FMS, although the evidence varied across programme characteristics. Liu *et al.* (2025) reported that structured active play and skill-oriented Physical Education were comparably effective, whereas unstructured active play was less effective for some outcomes. Yin *et al.* (2025) synthesised 33 school-based studies and found an overall positive effect on FMS, with programme duration, frequency and intervention period influencing outcomes. Jiang *et al.* (2026) further reported that FMS-targeted and structured programmes ranked highly for total FMS improvement. Collectively, these findings show why outcome measurement must be linguistically clear, psychometrically stable and sufficiently invariant to support fair comparison across pupils. In Malaysia, a Malay TGMD-3 that demonstrates content relevance, practitioner comprehensibility and Rasch-based construct functioning would address this methodological requirement while supporting more systematic Physical Education assessment.

3. MATERIALS AND METHODS

This study used a structured cross-cultural adaptation and psychometric validation design to evaluate whether the Malay version of the TGMD-3 retained conceptual meaning while remaining linguistically and operationally appropriate for Malaysian primary-school contexts. The adaptation sequence was guided by established principles of forward translation, synthesis, backward translation and expert review (Beaton *et al.*, 2000; Guillemin *et al.*, 1993). Validation was organised across complementary evidence sources: expert content judgement, practitioner clarity appraisal and Rasch analysis of pupil performance.

3.1 Participants

Three participant groups contributed to the validation process. The expert panel comprised seven specialists representing motor development, Physical Education, linguistics, language and curriculum-related expertise. Thirty practising Physical Education teachers from urban and rural primary-school settings evaluated instructional clarity and practical usability as intended end-users of the Malay instrument. Construct validation using Rasch measurement involved 64 Year Four primary-school pupils, comprising 32 boys and 32 girls, who completed all 13 TGMD-3 skills. Participation was voluntary and the study was conducted under institutional ethical approval from Sultan Idris Education University (UPSI/PPPI/PYK/ETIKA (M)/Vol.21(78)).

3.2 Instrument

The Test of Gross Motor Development–Third Edition (TGMD-3) is a process-oriented measure of fundamental gross motor competence (Ulrich, 2019). It comprises two domains: six locomotor skills (run, gallop, hop, skip, horizontal jump and slide) and seven manipulative or ball skills (two-handed strike, one-handed strike, stationary dribble, catch, kick, overhand throw and underhand roll). Each skill is evaluated against observable performance criteria rather than product outcomes such as distance or speed. The present study adapted the instructions and scoring language into Malay while maintaining the performance intent and structure of the original instrument. Formal written permission to adapt and validate the TGMD-3 for use in Malaysia was obtained from Professor Emeritus Dale A. Ulrich before the study commenced.

3.3 Cross-cultural adaptation procedure

The adaptation began with two independent forward translations from English into Malay. One translator had Physical Education and motor-development expertise to protect technical meaning, while a second translator had language expertise to support semantic accuracy. The two versions were compared and reconciled into a synthesised draft. Two independent backward translations were then completed to identify unintended shifts in meaning. Differences between the source and back-translated versions were reviewed systematically, with revisions focused on semantic equivalence, conceptual fidelity, age-appropriate wording and feasibility for Malaysian school use. This staged process followed the principle that cross-cultural equivalence should preserve the intended construct rather than reproduce wording literally (Beaton *et al.*, 2000; Guillemin *et al.*, 1993).

3.4 Data collection

The synthesised Malay TGMD-3 was first reviewed by the seven-member expert panel. Experts evaluated the relevance and appropriateness of each skill item and its instructional wording. The revised version was then appraised by 30 practising Physical Education teachers using a five-point clarity scale ranging from 1 (very unclear) to 5 (very clear), with attention to instructional comprehensibility, terminology and classroom usability. Following these content and language stages, the adapted instrument was administered to 64 Year Four pupils for Rasch analysis. Pupil performance scores were used to examine item functioning, the relationship between item difficulty and pupil ability, and whether items functioned differently for boys and girls.

3.5 Statistical analysis

Content validity was quantified using the Item-Level Content Validity Index (I-CVI), defined as the proportion of experts rating an item as relevant on the four-point relevance scale. An I-CVI of at least 0.78 was treated as acceptable for a panel of this size, and the Scale-Level Content Validity Index based on the average approach (S-CVI/Ave) was calculated across items (Polit & Beck, 2006). Teacher ratings were summarised using mean clarity scores, percentage agreement and the Impact Index reported for each item. Construct validity was examined with Rasch measurement. Item fit was assessed using Infit and Outfit Mean Square statistics, with values between 0.50 and 1.50 treated as productive for measurement. Person and item reliability and separation indices were examined to evaluate discrimination and calibration stability. Differential Item Functioning (DIF) contrasts were examined by gender, and a person–item (Wright) map was used to inspect targeting between pupil ability and item difficulty. These procedures are

consistent with the use of Rasch analysis to evaluate TGMD-3 item hierarchy and invariance (Valentini et al., 2022).

4. RESULTS

The Malay TGMD-3 demonstrated converging evidence of content relevance, instructional clarity and satisfactory Rasch measurement functioning. Results are presented sequentially for expert content validity, teacher-based language clarity and construct validation.

4.1 Content validity

Expert appraisal indicated strong agreement regarding the relevance and representativeness of the adapted skill items. I-CVI values ranged from 0.86 to 1.00, and the S-CVI/Ave across the 13 skills was 0.93. Thus, every item exceeded the minimum criterion used for acceptable content validity. Table 1 presents item-level results.

Table 1. Content Validity Index for the TGMD-3 Malay version

No.	Skill item	Domain	Experts agreeing	I-CVI
1	Run	Locomotor	7/7	1.00
2	Gallop	Locomotor	6/7	0.86
3	Hop	Locomotor	7/7	1.00
4	Skip	Locomotor	6/7	0.86
5	Horizontal jump	Locomotor	7/7	1.00
6	Slide	Locomotor	7/7	1.00
7	Two-handed strike	Manipulative	7/7	1.00
8	One-handed strike	Manipulative	6/7	0.86
9	Stationary dribble	Manipulative	7/7	1.00
10	Catch	Manipulative	7/7	1.00
11	Kick	Manipulative	6/7	0.86
12	Overhand throw	Manipulative	7/7	1.00
13	Underhand roll	Manipulative	7/7	1.00

Source: Authors' own data

4.2 Language validity and clarity of instructions

Teacher appraisal showed consistently high comprehensibility across all 13 skills. Mean clarity ratings ranged from 4.40 to 4.80 on the five-point scale and agreement ranged from 92% to 99%. Impact Index values ranged from 4.05 to 4.75. Catch produced the highest Impact Index, whereas kick produced the lowest value; both remained within the study's 'very clear' interpretation. These results indicate that the adapted terminology and instructions were readily interpretable by practising Physical Education teachers. Table 2 summarises the findings.

Table 2. Average score and Impact Index for the TGMD-3 Malay version

No.	Skill item	Average score	Agreement (%)	Impact Index	Interpretation
1	Run	4.7	98	4.60	Very clear
2	Gallop	4.5	95	4.28	Very clear
3	Hop	4.6	97	4.46	Very clear
4	Skip	4.4	94	4.14	Very clear
5	Horizontal jump	4.7	96	4.51	Very clear

No.	Skill item	Average score	Agreement (%)	Impact Index	Interpretation
6	Slide	4.8	98	4.70	Very clear
7	Two-handed strike	4.6	97	4.46	Very clear
8	One-handed strike	4.5	94	4.23	Very clear
9	Stationary dribble	4.7	98	4.60	Very clear
10	Catch	4.8	99	4.75	Very clear
11	Kick	4.4	92	4.05	Very clear
12	Overhand throw	4.6	96	4.41	Very clear
13	Underhand roll	4.7	97	4.56	Very clear

Source: Authors' own data

4.3 Rasch model analysis for construct validity

Rasch measurement was used to evaluate item fit, reliability and separation, gender-based measurement invariance and targeting of item difficulty to pupil ability.

4.3.1 Item fit statistics

All 13 skills demonstrated acceptable fit to the Rasch model. Infit Mean Square values ranged from 0.74 to 1.29 and Outfit Mean Square values ranged from 0.71 to 1.34, remaining within the pre-specified 0.50–1.50 interval. No skill required removal on the basis of these fit statistics.

4.3.2 Reliability and separation indices

Person reliability was 0.88 and the Person Separation Index was 2.65, indicating that the instrument differentiated approximately three strata of pupil motor ability. Item reliability was 0.95 and the Item Separation Index was 4.37, indicating stable calibration of the observed item-difficulty hierarchy.

4.3.3 Differential item functioning (DIF)

Gender-based DIF analysis did not identify substantial item bias. All observed DIF contrasts remained below ± 0.50 logits, supporting comparable measurement functioning for boys and girls within this sample.

4.3.4 Person–item map (Wright Map)

The Wright Map showed satisfactory targeting between the distribution of pupil ability and the hierarchy of skill difficulty. Locomotor skills were generally positioned at lower logit levels, whereas more coordination-demanding manipulative skills, including the one-handed strike and overhand throw, occupied higher positions. No major unmeasured gap was evident across the central ability continuum.

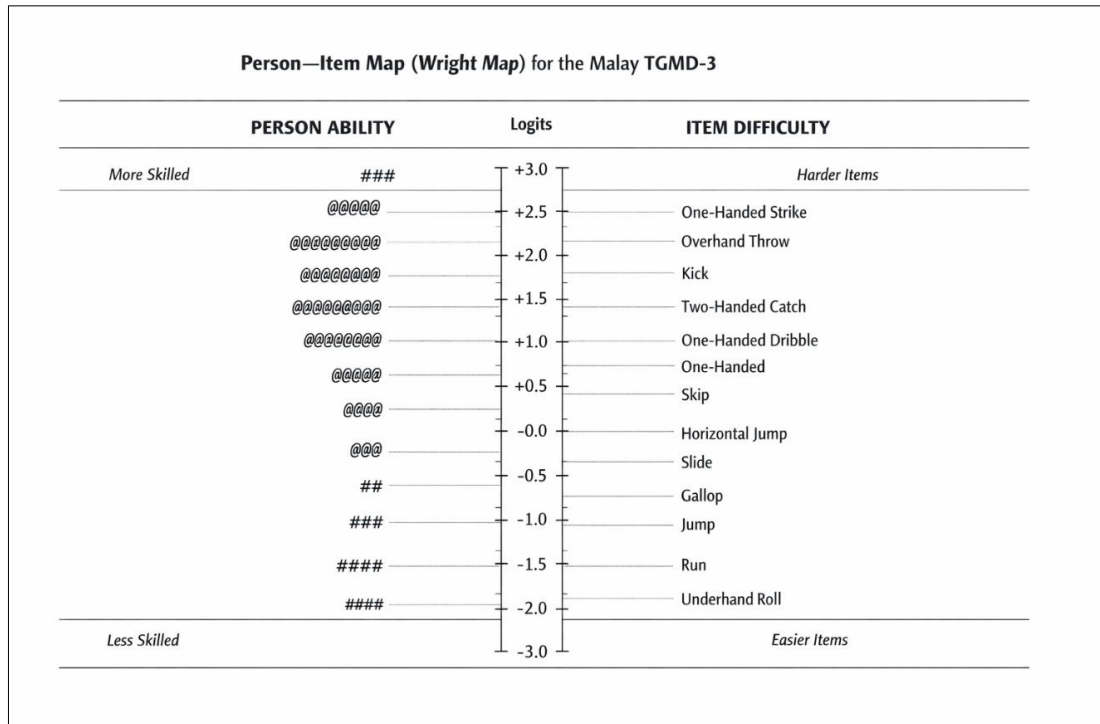


Fig. 1. Person–item map (Wright Map) for the Malay TGMD-3

Source: Authors' own data

The person–item map positions pupil motor ability on the left and item difficulty on the right using a common logit metric. Easier foundational locomotor skills are located lower on the continuum, while more complex object-control demands occur at higher logit levels, indicating an interpretable progression of difficulty across the adapted instrument.

5. DISCUSSION

The present study provides a multi-source validation of the Malay TGMD-3 by integrating expert content judgement, practitioner appraisal and Rasch measurement. The I-CVI range of 0.86–1.00 and S-CVI/Ave of 0.93 indicate that the adapted content retained strong representational relevance, while teacher clarity ratings and Impact Index values show that the instructions were also usable in the intended school context. This combination is important because cross-cultural validity requires both conceptual fidelity and comprehensibility; an instrument may preserve literal terminology yet still fail if end-users interpret performance criteria inconsistently (Beaton et al., 2000; Guillemin et al., 1993). The findings therefore support the adaptation process as more than a linguistic exercise: it produced a Malay version that experts considered relevant and practising teachers considered clear.

The practitioner findings are especially pertinent to current concerns about rater-mediated variation in FMS assessment. Rintala et al. (2017) and Maeng et al. (2017) showed that reliability can vary across

individual TGMD-3 skills even when broader subscale scores are stable. Carballo-Fazanes *et al.* (2023) later demonstrated that reliability improves when raters establish common interpretations of performance criteria. The high clarity scores in the present study suggest that the Malay wording reduces one important source of avoidable ambiguity before scoring occurs. This does not eliminate the need for assessor training, but it provides a more standardised linguistic starting point. That contribution directly addresses the challenges identified by Ng *et al.* (2025), who argued that subjective judgement and inconsistent scoring practices continue to constrain TGMD comparability.

Rasch findings provided complementary structural evidence. All 13 skills fell within the adopted fit interval, person reliability was 0.88 and item reliability was 0.95. The Person Separation Index of 2.65 indicates meaningful discrimination among different ability levels, whereas the Item Separation Index of 4.37 indicates a stable ordering of skill difficulty in this sample. These results are compatible with previous Rasch work showing that TGMD-3 performance criteria occupy distinct difficulty levels and can be organised along an interpretable motor continuum (Valentini *et al.*, 2022). They also complement large-sample psychometric evidence from Italy and French-speaking Canada, where the TGMD-3 demonstrated robust structural validity and measurement functioning after contextual validation (Magistro *et al.*, 2020; Mañano *et al.*, 2022).

The absence of substantial gender-based DIF is another important result. A fair educational assessment should not make a particular skill systematically easier or harder to score solely because a pupil belongs to one gender group after underlying ability is considered. In the present data, all gender DIF contrasts were below ± 0.50 logits. This pattern is consistent with Valentini *et al.* (2022), who also reported no gender DIF for TGMD-3 locomotor and ball-skill items, and it supports the use of a common measurement structure for boys and girls in the current Malaysian sample. The finding does not imply that mean motor performance must be identical between groups; rather, it indicates that the items themselves did not show substantial differential functioning by gender under the adopted criterion.

The Wright Map further suggests that the adapted instrument captured an interpretable span of motor competence, with foundational locomotor tasks at lower difficulty levels and coordination-demanding manipulative tasks at higher levels. This hierarchy is meaningful in light of recent evidence questioning whether different FMS instruments generate directly interchangeable estimates. Lorenzo-Martínez *et al.* (2025) found significant differences in proficiency estimates across three process-oriented tools despite assessing common skills. Bandeira *et al.* (2025) similarly highlighted the importance of evidence-informed construct definition in motor competence assessment. The present findings therefore support the Malay TGMD-3 within its own validated measurement framework rather than assuming that scores are automatically equivalent to results from other instruments or cultural contexts.

Recent 2025 evidence also helps position the practical value of this validation. Komura *et al.* (2025) used the TGMD-3 as a criterion measure when validating a simpler preschool motor-competence check, showing that TGMD-3 continues to function as an important benchmark in measurement research. Walters *et al.* (2025) established normative percentiles for Italian children, illustrating how validated TGMD-3 data can support population-level interpretation. Malaysia does not yet have comparable national norms for the Malay version; however, establishing content, language and construct validity is a necessary first step before larger normative studies can be justified. Future national sampling could extend the present work by developing age- and sex-referenced distributions and testing invariance across regions, school types and socio-demographic groups.

The educational relevance of valid measurement is reinforced by contemporary intervention evidence. Yin *et al.* (2025) reported a positive overall effect of school-based interventions on FMS and identified programme duration and frequency as important moderators. Zhang *et al.* (2025) found generally favourable effects for active-play interventions, while Liu *et al.* (2025) showed that structured active play and skill-oriented Physical Education can both be effective but that unstructured play may be less reliable for developing some skill domains. Jiang *et al.* (2026) further ranked FMS-targeted and structured physical-activity programmes among the strongest approaches for total FMS development. These findings imply that teachers and researchers need outcome measures capable of detecting genuine skill change. A validated Malay TGMD-3 can contribute to that need by supporting consistent pre-intervention profiling, targeted instruction and post-intervention evaluation in Malaysian schools.

The study also has limitations. Rasch validation was conducted with 64 Year Four pupils, so the results should not be interpreted as national norms or as definitive evidence of invariance across all Malaysian age groups and cultural settings. The teacher clarity sample, although drawn from practising Physical Education teachers, was limited to 30 respondents. Longitudinal stability, test–retest reliability and interrater reliability of the Malay version were not examined in the present analysis. Future research should therefore evaluate rater training and agreement directly, test measurement invariance across age, region and school context, and establish larger normative datasets. Technology-assisted scoring may also be explored, particularly given recent calls to improve data precision and reduce subjectivity in TGMD assessment (Ng *et al.*, 2025).

Taken together, the findings advance both local practice and cross-cultural measurement methodology. The Malay TGMD-3 retained strong content relevance, achieved high practitioner comprehensibility and demonstrated satisfactory Rasch fit, reliability, separation, targeting and gender invariance. This convergence of evidence is stronger than relying on translation accuracy or a single reliability coefficient alone. It provides a defensible basis for systematic assessment of fundamental motor competence in Malaysian primary schools while recognising that further large-scale calibration is required before national normative interpretation.

6. CONCLUSION

The Malay adaptation of the TGMD-3 demonstrated strong content validity, high instructional clarity and satisfactory Rasch measurement functioning in the present Malaysian primary-school sample. Expert agreement, teacher appraisal and item-response evidence converged to support conceptual fidelity, practical usability and structural coherence. The absence of substantial gender-based DIF further supports fair item functioning for boys and girls in this sample. The instrument therefore provides a stronger basis for standardised motor-skill assessment, evidence-informed instructional planning and future intervention evaluation in Malaysian Physical Education. Larger multi-site studies should now extend the evidence through age-based invariance testing, interrater and test–retest reliability, and the development of Malaysian normative data.

7. ACKNOWLEDGEMENTS/FUNDING

The authors acknowledge the Faculty of Sports Science and Coaching, Sultan Idris Education University, the multidisciplinary expert panel, participating Physical Education teachers, pupils and cooperating primary schools for their support. Formal written permission to adapt and validate the TGMD-3 for use in Malaysia was obtained from Professor Emeritus Dale A. Ulrich, University of Michigan, before the study commenced. The study was conducted under ethics approval UPSI/PPPI/PYK/ETIKA (M)/Vol.21(78) and the work is protected under Intellectual Property Registration CRLY2025W10963. Funding: None declared.

8. CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of commercial, financial or other relationships that could be construed as a potential conflict of interest.

9. AUTHORS' CONTRIBUTIONS

Author 1: Conceptualisation, methodology, investigation, formal analysis, data curation and writing—original draft. Author 2: Conceptualisation, methodology, supervision, validation and writing - review and editing. Author 3: Methodology, supervision, validation and writing - review and editing.

10. DECLARATION OF GENERATIVE AI IN THE WRITING PROCESS

During the preparation of this work, the authors used ChatGPT (OpenAI) to support language refinement, readability and manuscript formatting. The authors reviewed and verified the scientific content, cited literature and final text, edited the manuscript as needed and take full responsibility for the accuracy and integrity of the publication.

11. DATA AVAILABILITY/SUPPLEMENTARY MATERIALS

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

12. ETHICS STATEMENT

This study was reviewed and approved by the Sultan Idris Education University ethics committee (Approval No. UPSI/PPPI/PYK/ETIKA (M)/Vol.21(78)). Participation was voluntary and all procedures were conducted in accordance with institutional ethical requirements. Participant data were handled anonymously for analysis and reporting.

13. REFERENCES

- Bandeira, P. F. R., Lemos, L. F., Estevan, I., Webster, E. K., Clark, C. T., Duncan, M. J., Mota, J. A., & Martins, C. L. (2025). Are we assessing motor competence? Evidence-informed constructs for motor competence in preschoolers through an Exploratory Graph Analysis. *Journal of Sports Sciences*, 43(1), 109–116. <https://doi.org/10.1080/02640414.2024.2414361>
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186–3191. <https://doi.org/10.1097/00007632-200012150-00014>
- Bourke, M., Wang, H. F. W., Wicks, H., Barnett, L. M., Cairney, J., & Fortnum, K. (2025). Children’s and adolescents’ actual motor competence, perceived physical competence and physical activity: A structural equation modelling meta-analysis. *Sports Medicine*, 55(8), 1923–1936. <https://doi.org/10.1007/s40279-025-02233-2>
- Carballo-Fazanes, A., Rey, E., Valentini, N. C., Varela-Casal, C., & Abelairas-Gómez, C. (2023). Interrater reliability of the Test of Gross Motor Development—Third Edition following raters’ agreement on measurement criteria. *Journal of Motor Learning and Development*, 11(2), 225–244. <https://doi.org/10.1123/jmld.2022-0068>
- Estevan, I., Molina-García, J., Queralt, A., Álvarez, O., Castillo, I., & Barnett, L. M. (2017). Validity and reliability of the Spanish version of the Test of Gross Motor Development-3. *Journal of Motor Learning and Development*, 5(1), 69–81. <https://doi.org/10.1123/jmld.2016-0045>
- Guillemin, F., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. *Journal of Clinical Epidemiology*, 46(12), 1417–1432. [https://doi.org/10.1016/0895-4356\(93\)90142-N](https://doi.org/10.1016/0895-4356(93)90142-N)
- Jiang, S., Zeng, N., Ng, J. Y. Y., Chong, K. H., Zeng, T., Leung, S. K. Y., & Ha, A. S. (2026). Effects of physical activity interventions on fundamental movement skills and cognitive function in early childhood: A systematic review and network meta-analysis. *Journal of Sport and Health Science*, 15, Article 101085. <https://doi.org/10.1016/j.jshs.2025.101085>
- Komura, K., Demura, T., Ogura, Y., Kyan, A., Nawata, R., Takahashi, K., & Matsuura, R. (2025). Validity and reliability of the Simple Motor Competence-check for Kids (SMC-Kids). *Journal of Science and Medicine in Sport*, 28(5), 391–397. <https://doi.org/10.1016/j.jsams.2024.12.010>
- Liu, B., Yan, Y., Jia, J., & Liu, Y. (2025). Can active play replace skill-oriented physical education in enhancing fundamental movement skills among preschool children? A systematic review and meta-analysis. *BMC Public Health*, 25, Article 1399. <https://doi.org/10.1186/s12889-025-22398-9>
- Lorenzo-Martínez, M., Rey, E., & Abelairas-Gómez, C. (2025). Comparative analysis of three process-oriented assessment tools for identifying fundamental motor skills proficiency in children. *Scientific Reports*, 15, Article 10211. <https://doi.org/10.1038/s41598-025-95500-1>
- Maeng, H., Webster, E. K., Pitchford, E. A., & Ulrich, D. A. (2017). Inter- and intrarater reliabilities of the Test of Gross Motor Development—Third Edition among experienced TGMD-2 raters. *Adapted Physical Activity Quarterly*, 34(4), 442–455. <https://doi.org/10.1123/apaq.2016-0026>
- Magistro, D., Piumatti, G., Carlevaro, F., Sherar, L. B., Esliger, D. W., Bardaglio, G., Magno, F., Zecca, M., & Musella, G. (2020). Psychometric proprieties of the Test of Gross Motor Development—Third Edition in a large sample of Italian children. *Journal of Science and Medicine in Sport*, 23(9), 860–865. <https://doi.org/10.1016/j.jsams.2020.02.014>

- Mañano, C., Morin, A. J. S., April, J., Webster, E. K., Hue, O., Dugas, C., & Ulrich, D. A. (2022). 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>
- Ministry of Education Malaysia. (2017). *Primary School Standard Curriculum (KSSR): Physical Education*. Ministry of Education Malaysia.
- Ng, J. Y. Y., Jiang, S., Chan, C. H. S., & Ha, A. S. (2025). Assessing fundamental movement skills using the Test of Gross Motor Development (TGMD): Challenges and solutions to comparability and standardization. *Sports Medicine and Health Science*, 7(4), 299–301. <https://doi.org/10.1016/j.smhs.2025.02.007>
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489–497. <https://doi.org/10.1002/nur.20147>
- Rintala, P. O., Sääkslahti, A. K., & Iivonen, S. (2017). Reliability assessment of scores from video-recorded TGMD-3 performances. *Journal of Motor Learning and Development*, 5(1), 59–68. <https://doi.org/10.1123/jmld.2016-0007>
- Stodden, D. F., Goodway, J. D., Langendorfer, S. J., Robertson, M. A., Rudisill, M. E., Garcia, C., & Garcia, L. E. (2008). A developmental perspective on the role of motor skill competence in physical activity: An emergent relationship. *Quest*, 60(2), 290–306. <https://doi.org/10.1080/00336297.2008.10483582>
- Ulrich, D. A. (2019). *Test of Gross Motor Development–Third Edition (TGMD-3)*. Pro-Ed.
- UNESCO. (2015). *Education 2030: Incheon Declaration and Framework for Action*. UNESCO.
- Valentini, N. C., Duarte, M. G., Zanella, L. W., & Nobre, G. C. (2022). Test of Gross Motor Development-3: Item difficulty and item differential functioning by gender and age with Rasch analysis. *International Journal of Environmental Research and Public Health*, 19(14), 8667. <https://doi.org/10.3390/ijerph19148667>
- Walters, G. W. M., Cooper, S., Carlevaro, F., Magno, F., Boat, R., Vagnetti, R., D'Anna, C., Musella, G., & Magistro, D. (2025). Normative percentile values for the TGMD-3 for Italian children aged 3–11+ years. *Journal of Science and Medicine in Sport*, 28(5), 398–407. <https://doi.org/10.1016/j.jsams.2024.12.013>
- Yin, X., Zhang, D., Shen, Y., Wang, Y., Wang, Z., & Liu, Y. (2025). Effectiveness of school-based interventions on fundamental movement skills in children: A systematic review and meta-analysis. *BMC Public Health*, 25, Article 1522. <https://doi.org/10.1186/s12889-025-22696-2>
- Zhang, X., Tang, C., Geng, M., Li, K., Liu, C., & Cai, Y. (2025). The effects of active play interventions on children's fundamental movement skills: A systematic review. *BMC Pediatrics*, 25, Article 40. <https://doi.org/10.1186/s12887-024-05385-8>

