MALAYSIAN JOURNAL OF LEARNING AND INSTRUCTION
https://e-journal.uum.edu.my/index.php/mjli

How to cite this article
Arsaythamby Veloo, S. Kanageswari Suppiah Shanmugam & Suheysen Revindran. (2024). Examining oral test accommodation in assessing Malaysian *Orang Asli* pupils' mathematics performance for computation and word problem testlets. *Malaysian Journal of Learning and Instruction*, *21*(2), 231-256. https://doi.org/10.32890/mjli2024.21.2.8

EXAMINING ORAL TEST ACCOMMODATION IN ASSESSING MALAYSIAN *ORANG ASLI* PUPILS' MATHEMATICS PERFORMANCE FOR COMPUTATION AND WORD PROBLEM TESTLETS

[1]Arsaythamby Veloo, [2]S. Kanageswari Suppiah Shanmugam & [3]Suheysen Revindran
[1&2]School of Education, Universiti Utara Malaysia, Malaysia
[3]Awang Had Salleh Graduate School,
Universiti Utara Malaysia, Malaysia

[1]Corresponding author: arsay@uum.edu.my

Received: 7/11/2023 Revised: 17/3/2024 Accepted: 27/3/2024 Published: 14/8/2024

ABSTRACT

Purpose - The language load within mathematics test items can lead to inaccuracy in assessing *Orang Asli* pupils' mathematical ability due to their struggle in comprehending the academic language. This study aims to determine the validity of using test accommodations in the form of oral academic language and oral native language when administering mathematics computation and word problem test items among the *Orang Asli* pupils.

Methodology - By employing a quantitative approach, this study utilised random equivalent group design to examine the validity of using oral test accommodations among 334 Grade Four *Orang Asli* pupils. Three testlets were constructed and used in this study: written test in the academic language (WAL), oral test in the academic language (OAL), and oral test in the native language (ONL). The three testlets comprised Paper 1 and Paper 2, each consisting of 20 mathematics computation and 20-word problem test items, respectively. One-way ANOVA was used to analyse and examine if there were differences in the mathematics performance between the three groups of tests administered to the *Orang Asli* pupils.

Findings - The results indicated that the test scores of the pupils on the oral tests were higher compared to the written test administered to the pupils for both mathematics computation and word problem items. For computation items, the pupils performed better in the OAL testlet, followed by the ONL testlet. On the other hand, the pupils obtained the highest test scores in the ONL testlet compared to the OAL testlet for word problem items. Furthermore, the results illustrate that there is a difference in the mathematics performance of the *Orang Asli* pupils between the oral tests in both languages and written tests in the academic language.

Significant - This study clearly shows that language plays a prominent role in affecting the *Orang Asli* pupils' mathematics performance. By utilising the oral test as a promising test accommodation, this measure could help address the language barrier faced by the *Orang Asli* pupils in mathematics testing.

Keywords: *Orang Asli*, oral test accommodation, mathematics, computation, word problem, academic language, native language.

INTRODUCTION

The association between language and culture adds value to recognising cultural validity in mathematics testing as pupils' responses are influenced by their knowledge and cultural experiences (Trumbull & Nelson-Barber, 2019). The cultural responsiveness of Indigenous education enhances the academic growth of the pupils, and failure to recognise some cultural elements, such as language, can be 'detrimental to their attachment needs, their emotional development, their education and their health" (Perso, 2012, p. 12). This holds particularly true among *Orang Asli* communities in the Malaysian context, where the disparity between their home and school cultures is notably different compared to their mainstream counterparts. (Dunbar-Ortiz, 2014).

Similarly, Frigo (1999) asserts the importance of recognising that mathematics classes are essentially a linguistic exercise, which becomes especially complex for Indigenous students, given their diverse backgrounds and language needs (p. 1). The linguistic complexity inherent in test items introduces construct-irrelevant variance, threatening construct and content validity (Lopez, 2023; Roohr & Sireci, 2017) This secondary dimension, unintentional and extraneous, becomes part of the test and is reflected in the pupils' test scores (Messick, 1995). A major concern among the *Orang Asli* pupils is the extent of language load within the mathematics items, which may significantly affect their language comprehension when facing linguistically dense mathematics tests (Ruzlan et al., 2021; S. Kanageswari et al., 2023)

In essence, what is being evaluated may not solely be their ability to solve mathematics problems but their capability to read and understand the language component embedded within the mathematics items. When language becomes a barrier, the genuine mathematics aptitude of *Orang Asli* pupils' cannot be fully determined, particularly for mathematics test items that entail a significant language load (Haag et al., 2013). Therefore, addressing the language barriers confronted by *Orang Asli* pupils is essential to accurately assess their mathematical abilities, ensuring a fair evaluation that reflects their true potential in mathematics regardless of linguistic complexities within test items.

PROBLEM STATEMENT

Despite meticulous efforts and equitable treatment to ensure equal opportunities in mathematics, the pervasive impact of language proficiency remains overwhelming (Attar et al., 2022; Peng et al., 2020; Purpura & Reid, 2016). The influence of language is not only significant during mathematics teaching and learning but also during pupil assessment (Graven & Sibanda, 2018). This is because pupils must read and comprehend mathematics test items before applying the correct mathematical concepts and procedural knowledge for successful problem-solving (Trumbull & Nelson-Barber, 2019). Consequently, the conventional method of assessment used for mainstream pupils may become inadequate when applied to *Orang Asli* pupils. These pupils may require specialized arrangements or adjustments, such as test accommodations, to provide the necessary language support during conventional mathematics tests, as advocated by Zaleha et al. (2020) and S. Kanageswari et al. (2023).

The American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME) (2014) contended that assessing pupils in their dominant language can circumvent unintended language barriers arising from pupils' lack of language comprehension, which have adversely affected the accurate assessment of their mathematical knowledge. The primary objective of mathematical testing is to gauge pupils' mathematical aptitude, skills, or knowledge without the interference of language and culture (Leiss et al., 2019; Mushin et al., 2013). As the *Orang Asli* pupils are evaluated in an academic language that presents considerable linguistic challenges, this study raises a crucial validity question: How accurately do mathematics test scores reflect pupils' true mathematical abilities? There has consistently been a significant correlation between pupils' poor performance on high-stakes mathematics tests and their language skills (Moon, 2016). However, it is nearly impossible to remove the language element from assessments and instructions within them. Hence, this prompts the question: Do the obtained test scores genuinely represent pupils' mathematical abilities?

Current conventional mathematics testing fails to address the language and literacy barriers typically encountered by most *Orang Asli* pupils (Abdul Halim, 2022; Ruzlan et al., 2021; S. Kanageswari et al., 2023). Addressing this gap is crucial, and exploring inclusive testing measures using test accommodations as a means of levelling the playing field is essential (Vidal Rodeiro & Macinska, 2022). Therefore, an important test accommodation to investigate is the use of oral tests that incorporate both the academic language and native language of the pupils (Arsaythamby et al., 2021; S. Kanageswari et al., 2021; Stansfield, 2011). Combining oral tests in both languages with current mathematics testing practices ensures that language does not impede the pupils' mathematics performance (Shohamy, 2011).

On the global front, the context of this research remains relevant to Indigenous pupils. In other parts of the world, similar to *Orang Asli* pupils, they too have to negotiate between their native languages and the academic language taught in schools. Accordingly, this study adds depth to the limited extant literature on a distressing situation that could further impair the development of the Indigenous pupils' educational landscape. With this, the marginalized Indigenous community gains an equal opportunity through their younger generation to showcase their true potential by invoking their latent ability to thrive as key players in the country's development in the 21st century. Therefore, the need to test primary *Orang Asli* pupils in their native language to scaffold their ability to understand non-native languages arises and is further timely justified as the need to master the national and academic language (Malay language) is imminent within the Malaysian context.

LANGUAGE PROFICIENCY ON MATHEMATICS COMPUTATION ITEMS

In general, mathematics computation items developed for mathematics tests are designed to allow pupils to demonstrate mathematical knowledge and problem-solving skills for everyday problems (Tambunan, 2019). Numerical ability, typically assessed in mathematics computation items, has been found to be an important predictor of mathematical achievement in later years of pupils' primary schooling (Barcelos et al., 2018; Wijns et al., 2020). Nonetheless, these test items have been defined in various ways in existing literature (Braeuning, et al., 2021; Geary, 2013).

According to Fuchs et al. (2014), computation item refers to mathematics questions that require arithmetic calculations involving the development of skills in both 'number combination' and 'procedural computation'. The relationship between these two skills needed to solve computation items was highlighted by Brezovszky et al. (2019), who further posited that the interconnectedness of numerical relation and number combination should also consider the characteristics of the numbers placed within the problem to successfully solve the question. Therefore, by this definition, computation items contain numerical and procedural components that enable the pupils to perform the correct calculation to solve the problem.

On the other hand, Geary (2013) defined computation items as items containing basic symbols such as number words or Arabic numerals associated with the quantities or magnitude represented by the numbers.

Moreover, this notion is also present in the definition proposed by Braeuning et al. (2021), whereby computational items measure the pupils' ability to estimate magnitude or perform comparison tasks between numbers. Therefore, pupils are required to determine the position of the two magnitudes within the number line system, usually represented in Arabic symbols. This suggests that pupils should be familiar with the magnitude representation embedded within the test items displayed as Arabic numerals when solving the test items.

Moreover, concerning the linguistic aspects present in mathematics, computation items could also be defined as items that contain minimal language load or demand (Abedi et al., 2003). Several studies have examined the linguistic factors affecting the difficulties of mathematics test items (Boonen et al., 2016; Cho et al., 2020; Daroczy et al., 2015; Pongsakdi et al., 2020). Although computation items typically have a lower linguistic load, the language component may still play an important role in ensuring that pupils can exhibit mathematical computational knowledge during mathematics tests.

Furthermore, the role of language and its impact on mathematics test items requiring computational ability have been examined by various studies, particularly during early formal schooling (Li et al., 2022; Wilkinson, 2018). For instance, Purpura and Reid (2016) found that language proficiency developed during pupils' early childhood may provide the needed language support in the development and acquisition of mathematical ability during early formal schooling. It was further highlighted that the development of language proficiency is important in the development of computational ability when learning mathematics, as suggested by Praet et al. (2013). Consequently, the ability to grasp the language and terms specific to mathematics would subsequently enable pupils to manipulate numbers when dealing with complex mathematical computation items (Prediger et al., 2019; Vukovic & Lesaux, 2013).

Moreover, language proficiency among bilingual pupils has also been found to affect the mathematics performance of pupils on mathematics computation items (Krause et al., 2023; Swanson et al., 2020). For instance, in a study involving 233 bilingual Spanish-speaking dual-language pupils, Li et al. (2022) found that the pupils' language proficiency in their native language significantly predicted mathematics computation skills when items were presented in both the native language and academic language. The study highlighted the impact of native language proficiency when developing basic mathematical skills, which further compounds the pupils' ability to exhibit their computational knowledge during mathematics tests. Concerning the use of the Indigenous native language, Arsaythamby et al. (2021) examined the use of a bilingual test that included both the academic language and native language of the Indigenous *Orang Asli* pupils. The study found that for computation items, pupils taking the bilingual version of the test performed significantly better than those administered with only the academic language. These results may be attributed to the impact of language as a precursor to the cognitive development of computational knowledge needed to carry out basic arithmetic and numeracy skills (Moschkovich, 2019; Swanson, 2020).

LANGUAGE PROFICIENCY ON MATHEMATICS WORD PROBLEM ITEMS

Mathematics word problem items have continued to garner much attention from both educational researchers and test developers in the recent decades due to the complexities they pose to pupils (Clarke & Roche, 2018; Verschaffel et al., 2020). In the existing literature, mathematics word problems remain a key component of mathematics education (Nesher, 2020). However, word problems have been defined in various ways.

In simpler terms, mathematics word problem is defined as mathematics test items with information on a problem embedded in the text rather than in mathematical notation form (Timmermans et al., 2007). However, according to Verschaffel et al. (2020), a mathematics word problem is better defined as a problem situation presented as a verbal description, where one or more questions must be answered by applying mathematical operations derived from the text to obtain the numerical data. This definition highlights the need for text comprehension and discourse processing required by pupils to employ the right solution strategies for word problem items (Fuchs et al., 2018; Pongsakdi et al., 2020; van Dijk, 1993). This definition considers three main corresponding elements of mathematical representation from the text of the word problem (Fuchs et al., 2018).

First, the main context is extracted based on understanding the main ideas embedded within the text. Then, inferences are drawn from this understanding of the context, stemming from the experiences of the pupils. The final element is the construction of a situation or problem model that draws on the relationship between the numerical and mathematical knowledge and concepts, creating the mathematical model necessary to obtain the correct solution to the mathematics word problem item (Leiss et al., 2019). The situation model relies on the linguistic complexity and reading ability of the pupils to accurately create it, subsequently deriving the solution strategies needed to solve the problem (Fuchs et al., 2018; Vershaffel et al., 2020).

Additionally, mathematics word problems can be defined based on linguistic components of the text. Mathematics, as a language with its own grammar, syntax, vocabulary, word order, synonyms, idioms, abbreviations, sentence and paragraph structure (Ledibane et al., 2018; Peng et al., 2020; Riccomini et al., 2015), requires consideration of mathematics word problem from a linguistic perspective. Fuchs et al. (2014) defined a word problem as a linguistic problem statement containing relevant information within its sentences, requiring pupils to derive numerals and mathematical operations from the corresponding words within the text.

Therefore, mathematics word problem items are usually distinguishable from mathematics computation items through their linguistic components and context. Mathematics word problem items typically contain more words or sentences and are presented within a contextualized real-life situation, requiring pupils to use a higher level of language comprehension and cognitive ability to grasp the context and employ the correct mathematical solution (Menaga et al., 2022; Nesher, 2020). In contrast, mathematics computation items are straightforward, requiring pupils to apply their mathematical procedural knowledge directly to obtain the correct solution (Achmetli et al., 2019).

IMPACT OF KINTSCH-GREENO MODEL ON MATHEMATICS PERFORMANCE

The model by Kintsch and Greeno (1985) explores the interaction between comprehension and problem-solving, particularly when text is used to present problem information. This is evident in word problems in mathematics, where the linguistic factor significantly affects conceptual representations of operations such as counting objects, addition, subtraction, and learners' ability to comprehend the requirements of the items.

The model presents three sets of knowledge structures that can be accessed by a proposed set of strategies to build a representation and solve word problems (Chan & Kwan, 2021; Skinner & Cuevas, 2023). These are: 1) Prepositional frames (translating sentences into propositions); 2) Schemata representing properties and relations of sets in general form (constructing macrostructures and problem models) and 3) Schemata representing counting and arithmetic operations in general form (calculating the solutions of problems). Regarding problem representations, information processing begins with the transformation of verbal input into a conceptual representation of its meaning, also known as propositions (Krawec, 2014; Perrenet & Zwaneveld, 2012). These are then organized into a task-specific macrostructure or text base to highlight the general concepts, or sets, and relations among the sets from the text. Next, the problem model, reflecting knowledge of the information needed to solve the problem, is constructed by inferring information not included in the text base for solving the problem. Information unnecessary for the solution of the problem is excluded.

Although the linguistic aspects are not explicitly highlighted in the model, it is evident that language plays a prominent role, especially when mathematical and generic vocabulary are used to contextualize the items, often in the case of word problems (Timmermans et al., 2007). The transfer of key information from descriptions of a problem situation to the application of mathematics operations requires learners to comprehend and process the language of the text before they can arrive at the right solution (Pongsakdi et al., 2020). Expanding on this, the use of unfamiliar language could pose an impediment in the prepositional frames of the model, as pupils may struggle to extract the necessary information required to conceptualize the problem embedded within the mathematical model.

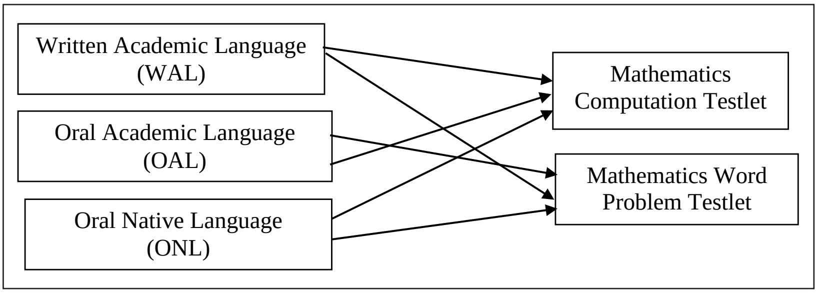
CONCEPTUAL FRAMEWORK FOR ORAL TEST ACCOMMODATION

This study employed three independent variables and two dependent variables to assess the feasibility of utilizing oral tests as potential test accommodation for evaluating the mathematics performance of *Orang*

Asli pupils. The three independent variables comprised a written test in the academic language (WAL), an oral test in the academic language (OAL), and an oral test in the native language (ONL). The two dependent variables were the mathematics computation testlet and mathematics word problem testlet. Drawing from these dependent and independent variables, a conceptual framework was developed to guide the assessment of *Orang Asli* pupils' mathematics performance, as illustrated in Figure 1.

Figure 1

Conceptual Framework for Oral Test Accommodation in Assessing Orang Asli Pupils' Mathematics Performance



The objective of this study is justified by the imperative need to offer language and linguistic support to *Orang Asli* pupils, facilitating the demonstration of their mathematical knowledge and skills. Introducing culturally responsive measure for Indigenous pupils holds promise for inclusive and equitable mathematics testing, eliminating language barriers that may hinder their mathematical performance. Therefore, this study aims to determine the validity of using oral tests in both the academic and native language of *Orang Asli* pupils as test accommodations for mathematics computation and word problem items.

METHODOLOGY

Research Design

This study adopted a quantitative approach to investigate the validity of the oral test in both the academic language (AL) and the native language (NL) of *Orang Asli* pupils. A random equivalent group design was employed to mitigate stress-inducing effects and counterbalance any order effects resulting from the administration of three different tests. This research design ensures that groups of test takers experiencing each test are comparable, thus equalizing their test-taking experience (Livingston & Kim, 2010).

A total of 334 Grade Four *Orang Asli* pupils from two primary schools in the district of Sungai Siput, Perak (160 pupils) and two primary schools in Gua Musang, Kelantan (174 pupils) participated in this study. The sample consisted of pupils from the *Temiar* native tribe. Among the 334 *Orang Asli* pupils, 156 (47%) were boys, while 178 (53%) were girls.

Instrument

In this study, three mathematics tests were constructed: the written test in the academic language (WAL), the oral test in the academic language (OAL) and the oral test in the native language (ONL). Each test comprised 40 mathematics computation and mathematics word problem items, with 20 computation items presented in Paper 1 and 20 words problem items in Paper 2. These test items were carefully selected to assess the mathematics skills and knowledge of Grade 4 primary school *Orang Asli* pupils.

The oral tests were developed based on the written test as a master test. Each of the 40 mathematics test items from the written test was individually audio recorded, with an expert teacher fluent in *Bahasa Melayu* (BM) reading each test item aloud in *Bahasa Melayu* (Malay language). The expert involved in developing the OAL test is from the *Temiar* native tribe in Gua Musang, Kelantan. The OAL test consists of audio recordings for the 20 computation items and 20-word problem items presented in Paper 1 and Paper 2, respectively. All test items in the OAL test were validated by the expert to ensure item equivalence with the WAL test.

For the ONL test, the mathematics test items from the WAL test were translated into the native language (*Bahasa Temiar or Temiar* Language) of the *Orang Asli* pupils and transadapted to ensure linguistic and cultural equivalence suitable for the *Temiar* ethnic group's cultural context. Each test item was then read aloud and simultaneously audio recorded in the *Temiar* language. The audio recordings of the 20 computation items and 20-word problem items were compiled into audio tests for the respective test booklets for Paper 1 and Paper 2.

Data Collection

Before commencing the study, permission was obtained from the Education Policy Planning and Research Division (EPRD) under the Ministry of Higher Education (MoHE), followed by approvals from the State Education Departments of Perak and Kelantan, and the Department of *Orang Asli* Development (*Jabatan Kemajuan Orang Asli* - JAKOA). Briefings were conducted with headmasters and school management teams to introduce the study's background.

Test administration followed a random equivalent group design with spiral administration. *Orang Asli* pupils were randomly assigned to three different groups, each administered one of the tests: WAL, OAL, or ONL. The tests were administered independently to ensure equitable pupil ability and test difficulty. The mathematics computation testlet in Paper 1 was administered first, followed by the mathematics word problem testlet in Paper 2 with a 15-minute break between sessions.

For pupils administered the oral tests (OAL and ONL), a copy of the written test in the academic language (WAL) was provided. The audio recording for each test item was played twice, allowing pupils sufficient time to complete the item before proceeding to the next. The duration to complete both testlets was approximately two hours.

Data Analysis

This study assessed the mathematics performance of *Orang Asli* pupils through their mathematics test scores in mathematics computation and word problem testlets across three administered mathematics tests. The three tests served as independent variables, while mathematics test scores for both testlets served as the dependent variable. Data analysis was conducted using SPSS version 28. One-way ANOVA was used to examine the three independent variables: the written test in the academic language (WAL), oral test in the academic language (OAL), and oral test in the native language (ONL), on both the mathematics computation and mathematics word problem testlets.

One-way ANOVA was utilized to examine the three independent variables, on both the mathematics computation and word problem testlets. One-way ANOVA allowed for the comparison of group means to determine whether statistically significant differences existed between the groups (Mishra et al., 2019). Additionally, the Tukey HSD post hoc test was conducted to identify specific group differences (Kucuk et al., 2016). This post hoc analysis provided a comprehensive understanding of the discrepancies observed among group pairs comprising the WAL, OAL, and ONL tests.

FINDINGS

For the mathematics computation testlet, Levene's test indicated non-significance ($p > 0.05$), indicating homogeneity of variance among the three tests. Subsequently, a one-way ANOVA was conducted to assess if there were significant differences in the mathematics test scores of the *Orang Asli* pupils across the three tests. The analysis revealed a statistically significant difference [$F(2, 331) = 17.72$, $p < 0.05$] suggesting variations in mathematics performance among the *Orang Asli* pupils across the three tests. Table 1 presents the results of one-way ANOVA for the mathematics computation testlets.

Table 1

One-way ANOVA for Mathematics Computation Testlets

Mathematics Test	N	Mean	SD	df1	df2	F	p
WAL	109	11.50	3.33	2	331	17.72	0.00*
OAL	108	13.72	3.11				
ONL	117	13.53	2.73				

The Tukey HSD post hoc test was then conducted to identify differences among the three tests for the mathematics computation testlet. The analysis revealed a significant difference between the WAL and OAL mathematics computation testlets ($p < 0.05$) and between the WAL and ONL mathematics computation testlets ($p < 0.05$). However, there was no significant difference between the OAL and ONL mathematics computation testlets ($p > 0.05$). The results suggest differences in the mathematics performance of *Orang Asli* pupils when administered the oral tests in both languages, compared to the written test in the academic language for mathematics computation items. Moreover, there was no difference in the pupils' mathematics performance between the oral test in the academic language and the oral test in the native language. Table 2 presents the results of the Tukey HSD post hoc test for the mathematics computation testlets.

Table 2

Tukey HSD Post Hoc Test for Mathematics Computation Testlets

Mathematics Test		Mean Difference	p
OAL	WAL	2.22	0.00*
ONL	WAL	2.03	0.00*
OAL	ONL	0.19	0.89

For the mathematics performance of the *Orang Asli* pupils, as indicated by their mean test scores, the study found that the pupils performed the best when administered the OAL test (M = 13.72, SD = 3.11), followed by the ONL test (M = 13.53, SD = 2.73) and finally, the WAL test (M = 11.50, SD = 3.33). The results suggest that the pupils were able to achieve significantly better performance when administered both the oral test in the academic language and the oral test in the native language, compared to the written test in the academic language for mathematics computation items. Table 3 presents the descriptive statistics computed for the mathematics computation testlets.

Table 3

Descriptive Statistics of Mathematics Computation Testlets

Mathematics Test	N	Mean	SD
OAL	108	13.72	3.11
ONL	117	13.53	2.73
WAL	109	11.50	3.33

On the other hand, for the mathematics word problem testlets, the results of the Levene's test were not significant ($p > 0.05$), suggesting homogeneity of variance between the WAL, OAL and ONL tests. Subsequently, a one-way ANOVA was carried out, revealing a statistically significant difference between the three tests [F (2,

331) = 8.98, $p < 0.05$]. This indicates variations in the mathematics performance of the *Orang Asli* pupils across the three mathematics tests. Table 4 presents the results of the one-way ANOVA for mathematics word problem testlets.

Table 4

One-way ANOVA for Mathematics Word Problem Testlets

Mathematics Test	N	Mean	SD	df1	df2	F	p
WAL	109	9.50	4.25	2	331	8.98	0.00*
OAL	108	10.92	3.78				
ONL	117	11.67	3.60				

Subsequently, to determine the difference between the groups of mathematics tests, the Tukey HSD post hoc test was carried out. The analysis revealed a significant difference between the WAL and OAL word problem testlets ($p < 0.05$) and between the WAL and ONL word problem testlets ($p < 0.05$). However, there was no significant difference between the OAL and ONL word problem testlets ($p > 0.05$). Therefore, for mathematics word problem items, the results suggest differences in the mathematics performance of the *Orang Asli* pupils when administered the oral tests in both languages, compared to the written test in the academic language. Table 5 presents the results of the Tukey HSD post hoc test for mathematics word problem testlets.

Table 5

Tukey HSD Post Hoc Test for Mathematics Word Problem Testlets

Mathematics Test		Mean Difference	p
OAL	WAL	1.41	0.02*
ONL	WAL	2.16	0.00*
ONL	OAL	0.75	0.32

*$p < 0.05$, and $p > 0.05$

Regarding the mathematics performance of *Orang Asli* pupils', the study found that the pupils achieved the highest test scores when administered the ONL test (M = 11.67, SD = 3.60), followed by the

OAL test (M = 10.92, SD = 3.78) and finally, the WAL test (M = 9.50, SD = 4.25). Based on these findings, this study concluded that the pupils performed significantly better when administered both the oral test in the academic language and the oral test in the native language, compared to the written test in the academic language for mathematics word problem items. Table 6 presents the descriptive statistics computed for the mathematics word problem testlets.

Table 6

Descriptive Statistics of Mathematics Word Problem Testlets

Mathematics Test	N	Mean	SD
ONL	117	11.67	3.60
OAL	108	10.92	3.78
WAL	109	9.50	4.25

DISCUSSION

The results for the computation testlets indicate that oral tests, whether in academic or native languages, yielded higher test scores compared to the written test in the academic language. This underscores the significant role of language in mathematics testing (Boonen et al., 2016; Cho et al., 2020; Daroczy et al., 2015; Pongsakdi et al., 2020). The findings suggest that using oral tests in both academic and native languages can assist *Orang Asli* pupils overcome language difficulties inherent in mathematics test items. The lower test scores attained by pupils on mathematics computation items conducted in the academic language align with previous research indicating that inadequate proficiency in the academic language used in mathematics tests continues to affect *Orang Asli* pupils (Abdul Halim, 2022). This language barrier may have contributed to their lower performance in mathematics tests (Ruzlan et al., 2021; S. Kanageswari et al., 2023), even on mathematics computation items with minimal language load.

Furthermore, similar results also revealed that the *Orang Asli* pupils performed better in both oral tests in academic and native languages, respectively, with no significant difference between the two forms of oral tests. This finding aligns with studies by Arsaythamby et al. (2021) and Zaleha et al. (2020), where *Orang Asli* pupils tended to perform better when administered oral tests, compared to mathematics written tests. This suggests that the language and linguistic support provided by oral tests may be beneficial in helping these pupils better comprehend test items more effectively than when presented in written form.

As described by Abedi et al. (2003), mathematics computation items contain minimal language load or demand. Hence, it can be inferred that when *Orang Asli* pupils attempt test items with low language demands, they are well-equipped to solve them regardless of the language used. Having predominantly learned the subject in the academic language, they can perform better in the oral test in the academic language as they are more familiar with the mathematical terms used in the tests. Therefore, findings for the mathematics computation testlet support the use of oral tests as a valid test accommodation capable of mitigating the impact of linguistic components embedded within test items when assessing the true mathematical ability of *Orang Asli* pupils.

As for the mathematics word problem testlet, the oral test in the native language proved the most effective among *Orang Asli* pupils compared to both the written test and oral test in the academic language. Similar to the mathematics computation testlet, pupils' performance on the oral tests in both the academic and native languages was significantly better compared than their performance on the written test in the academic language. This outcome underscores the pivotal role of language in mathematics testing, highlighting that *Orang Asli* pupils exhibit improved performance when their native language is incorporated as a supportive element in the mathematics test, as evidenced by Attar et al. (2022). This aligns with previous studies by Trumbull and Nelson-Barber (2019) which emphasize the advantage of employing a language familiar to pupils, resulting in enhanced comprehension of mathematics test items. The utilization of pupils' native language asserts its significance in validating inclusive and equitable mathematics tests for *Orang Asli* pupils. This underscores the importance of embracing linguistic diversity to create fair testing environments that accommodate the linguistic backgrounds of students.

This distinction between the two testlets highlights the considerable challenge posed by mathematics word problem items in problem-solving (Kintsch & Greeno, 1985), especially for *Orang Asli* pupils, where language becomes a determining factor in their textual understanding of the content and context of the test items. This challenge arises due to pupils' unfamiliarity with the core mathematical concepts embedded within the linguistic text, underscoring the importance of language proficiency in understanding complex mathematical test items when administered to pupils (Leiss et al., 2019; Menaga et al., 2022; Nesher, 2020). This obstacle hinders further information processing and hampers pupils' ability to solve word problems effectively.

The findings of this study contribute to the existing evidence regarding linguistic complexity and language barriers that affect Indigenous pupils' performance in content subjects such as mathematics. Therefore, until these pupils attain sufficient proficiency in addressing textual comprehension of mathematics test items, oral test accommodations, either in the academic or native language, may provide crucial support in addressing the challenges faced by *Orang Asli* pupils. Consequently, these findings should be contextualized within the unique circumstances of Indigenous pupils who may be isolated due to their locality and should not be generalized to pupils with limited proficiency in the academic language but have access to educational resources.

Future studies should aim to confirm the efficacy of oral test accommodations in mathematics among Indigenous tribes in various regions worldwide. By conducting such studies, we can establish with certainty, the utility of oral test accommodations as a valid measure of Indigenous pupils' mathematical ability. Furthermore, considering the recent digital advancements in testing, there is a need for further studies to explore the technological mediation of oral test accommodations, particularly in developing countries. This could involve the implementation of digital read-alouds or computer-aided testing.

In conclusion, results obtained for both mathematics computation and word problem testlets suggest that oral tests, whether in academic or native languages, represent a valid and promising testing alternative for assessing *Orang Asli* pupils' mathematical performance. As outlined in Goal 2 (Universal primary education) of UNESCO's (2015) Education for All initiative, countries worldwide are striving to ensure that all children, especially those in challenging circumstances and belonging to ethnic minorities, have access to quality primary education. To achieve this vision, assessments must also adopt innovative approaches that translate testing practices into strategic actions within mathematics classrooms.

Therefore, this study, which investigates oral test accommodations in mathematics for *Orang Asli* pupils, underscores the timely need to promote valid testing practices and transform mathematics assessment methodologies in the 21st century for Indigenous pupils globally. The audacious reality is that oral mathematics tests can serve as a viable and forward-thinking solution for Indigenous pupils to attain equitable access to education. Furthermore, they may provide a valid alternative testing method that accurately measures their mathematical abilities, aligning with the principles of equity in education advocated in both the Malaysian Education Blueprint and UNESCO's Education for All initiative.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

ACKNOWLEDGMENT

This research was supported by the Ministry of Higher Education of Malaysia through the Fundamental Research Grant Scheme (FRGS/1/2021/SSI01/UUM/02/1).

REFERENCES

Abdul Halim Abdullah. (2022). A systematic review of what Malaysia can learn to improve Orang Asli students' mathematics learning from other countries. *Sustainability*, *14*(20), 13201. https://doi.org/10.3390/su142013201

Abedi, J., Leon, S., & Mirocha, J. (2003). Impact of student language background on content-based performance: Analyses of extant data. *Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing*, Graduate School of Education & Information Studies, University of California, Los Angeles.

Achmetli, K., Schukajlow, S., & Rakoczy, K. (2019). Multiple solutions for real-world problems, experience of competence and students' procedural and conceptual knowledge. *International Journal of Science and Mathematics Education*, *17*(8), 1605-1625. https://doi.org/10.1007/s10763-018-9936-5

American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME). (2014). *Standards for Educational and Psychological Testing.* AERA.

Arsaythamby Veloo, S. Kanageswari Suppiah Shanmugam, Ruzlan Md-Ali, Yus'aiman Jusoh Yusoff, & Rosna Awang-Hashim. (2021). Grade five Indigenous (*Orang Asli*) pupil's achievement in bilingual versions of mathematics test. *Journal of Language and Linguistic Studies*, *17*(4), 1863-1872.

Attar, Z., Blom, E., & Le Pichon, E. (2022). Towards more multilingual practices in the mathematics assessment of young refugee students: Effects of testing language and validity of parental assessment. *International Journal of Bilingual Education and Bilingualism*, *25*(4), 1546-1561. https://doi.org/10.1080/13670050.2020.1779648

Barcelos, T. S., Muñoz-Soto, R., Villarroel, R., Merino, E., & Silveira, I. F. (2018). Mathematics learning through computational thinking activities: A systematic literature review. *Journal of Universal Computer Science*, *24*(7), 815-845.

Boonen, A. J., de Koning, B. B., Jolles, J., & Van der Schoot, M. (2016). Word problem solving in contemporary math education: A plea for reading comprehension skills training. *Frontiers in Psychology*, *7*(191), 1-10. https://doi.org/10.3389/fpsyg.2016.00191

Braeuning, D., Hornung, C., Hoffmann, D., Lambert, K., Ugen, S., Fischbach, A., Schiltz, C., Hubner, N., Nagengast, B., & Moeller, K. (2021). Long-term relevance and interrelation of symbolic and non-symbolic abilities in mathematical-numerical development: Evidence from large-scale assessment data. *Cognitive Development*, *58*, 101008. https://doi.org/10.1016/j.cogdev.2021.101008

Brezovszky, B., McMullen, J., Veermans, K., Hannula-Sormunen, M. M., Rodríguez-Aflecht, G., Pongsakdi, N., Laakkonen, E., & Lehtinen, E. (2019). Effects of a mathematics game-based learning environment on primary school students' adaptive number knowledge. *Computers and Education*, *128*, 63-74.

Chan, W. W. L., & Kwan, J. L. Y. (2021). Pathways to word problem solving: The mediating roles of schema construction and mathematical vocabulary. *Contemporary Educational Psychology*, *65,* 101963. https://doi.org/10.1016/j.cedpsych.2021.101963

Cho, E., Fuchs, L. S., Seethaler, P. M., Fuchs, D., & Compton, D. L. (2020). Dynamic assessment for identifying Spanish-speaking English learners' risk for mathematics disabilities: Does language of administration matter? *Journal of Learning Disabilities*, *53*(5), 380-398. https://doi.org/10.1177%2F0022219419898887

Clarke, D., & Roche, A. (2018). Using contextualized tasks to engage students in meaningful and worthwhile mathematics learning. *The Journal of Mathematical Behavior*, *51*, 95-108. https://doi.org/10.1016/j.jmathb.2017.11.006

Daroczy, G., Wolska, M., Meurers, W. D., & Nuerk, H. C. (2015). Word problems: A review of linguistic and numerical factors contributing to their difficulty. *Frontiers in Psychology*, *6,* 348. https://doi.org/10.3389/fpsyg.2015.00348

Dunbar-Ortiz, R. (2014). *An indigenous peoples' history of the United States* (Vol. 3). Beacon Press.

Frigo, T. (1999). *Resources and teaching strategies to support Aboriginal children's numeracy learning*. Australian Council for Educational Research. https://research.acer.edu.au/indigenous_education/11

Fuchs, L. S., Powell, S. R., Cirino, P. T., Schumacher, R. F., Marrin, S., Hamlet, C. L., Fuchs, D., Compton, D. L., & Changas, P. S. (2014). Does calculation or word-problem instruction provide a stronger route to prealgebraic knowledge*? Journal of Educational Psychology*, *106*(4), 990-1006.

Fuchs, L. S., Gilbert, J. K., Fuchs, D., Seethaler, P. M., & N. Martin, B. (2018). Text comprehension and oral language as predictors of word-problem solving: Insights into word-problem solving as a form of text comprehension. *Scientific Studies of Reading*, *22*(2), 152-166. https://doi.org/10.1080/10888438.2017.1398259

Geary, D. S. (2013). Early foundations for mathematics learning and their relations to learning disabilities. *Current Direction in Psychological Science*, *22*(1), 23-27.

Graven, M., & Sibanda, L. (2018). Can mathematics assessments be considered valid if learners fail to access what is asked of them? *South African Journal of Childhood Education*, *8*(1), 1-12. https://hdl.handle.net/10520/EJC-15d538c9bd

Haag, N., Heppt, B., Stanat, P., Kuhl, P., & Pant, H. A. (2013). Second language learners' performance in mathematics: Disentangling the effects of academic language features. *Learning and Instruction*, *28*, 24-34.

Krause, G. H., Vanderberg, M., E. Hung, E., & Skuratowicz, E. (2023). Computational thinking in a bilingual kindergarten classroom: Emergent ideas for teaching across content areas. *Education and Information Technologies*, *28*(8), 9767-9782. https://doi.org/10.1007/s10639-022-11454-1

Krawec, J. L. (2014). Problem representation and mathematical problem solving of students of varying math ability. *Journal of Learning Disabilities*, *47*(2), 103-115. https://doi.org/10.1177/0022219412436976

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review*, *92*(1), 109-129. https://psycnet.apa.org/doi/10.1037/0033-295X.92.1.109

Kucuk, U., Eyuboglu, M., Kucuk, H. O., & Degirmencioglu, G. (2016). Importance of using proper post hoc test with ANOVA. *International Journal of Cardiology*, *209*, 346. https://doi.org/10.1016/j.ijcard.2015.11.061

Ledibane, M., Kaiser, K., & Van der Walt, M. (2018). Acquiring mathematics as a second language: A theoretical model to illustrate similarities in the acquisition of English as a second language and mathematics. *Pythagoras*, *39*(1), 1-12.

Leiss, D., Plath, J., & Schwippert, K. (2019). Language and mathematics-key factors influencing the comprehension process in reality-based tasks. *Mathematical Thinking and Learning*, *21*(2), 131-153. https://doi.org/10.1080/10986065.2019.1570835

Li, J. T., Arizmendi, G. D., & Swanson, H. L. (2022). The role of language comprehension skills and instructional practices in cross-language influence of Spanish-speaking dual language learners' calculation skills. *Early Childhood Research Quarterly*, *61*, 90-105. https://doi.org/10.1016/j.ecresq.2022.05.004

Livingston, S. A., & Kim, S. (2010). Random-groups equating with samples of 50 to 400 test takers. *Journal of Educational Measurement*, *47*(2), 175-185. https://doi.org/10.1111/j.1745-3984.2010.00107.x

Lopez, A. (2023). Examining how Spanish-speaking English language learners use their linguistic resources and language modes in a dual language mathematics assessment task. *Journal of Latinos and Education*, *22*(1), 198-210.

Menaga Suseelan, Chew, C. M., & Chin, H. (2022). Research on mathematics problem solving in elementary education conducted from 1969 to 2021: A bibliometric review. *International Journal of Education in Mathematics, Science and Technology*, *10*(4), 1003-1029.

Messick, S. (1995). Standards of validity and the validity of standards in performance asessment. *Educational Measurement: Issues and Practice*, *14*(4), 5-8.

Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Annals of Cardiac Anaesthesia*, *22*(1), 67-72.

Moon, T. R. (2016). Differentiated instruction and assessment: An approach to classroom assessment in conditions of student diversity. In G. Brown & L. Harris (Eds.), *Handbook of Human and Social Conditions in Assessment* (pp. 300-317). Routledge.

Moschkovich, J. (2019). Codeswitching and mathematics learners: How hybrid language practices provide resources for student participation in mathematical practices. In J. MacSwan & C. J. Faltis (Eds.), *Codeswitching in the Classroom* (pp. 88-113). Routledge.

Mushin, I., Gardner, R., & Munro, J. M. (2013). Language matters in demonstrations of understanding in early years mathematics assessment. *Mathematics Education Research Journal*, *25*, 415-433. https://doi.org/10.1007/s13394-013-0077-4

Nesher, P. (2020). Solving multiplication word problems. In G. Leinhardt, R. Putnam, & R. A. Hattrup (Eds.), *Analysis of arithmetic for mathematics teaching* (pp. 189-219). Routledge.

Peng, P., Lin, X., Ünal, Z. E., Lee, K., Namkung, J., Chow, J., & Sales, A. (2020). Examining the mutual relations between language and mathematics: A meta-analysis. *Psychological Bulletin*, *146*(7), 595-634. https://psycnet.apa.org/doi/10.1037/bul0000231

Perrenet, J. C., & Zwaneveld, B. (2012). The many faces of the mathematical modeling cycle. *Journal of Mathematical Modelling and Application*, *1*(6), 3-21.

Perso, T. (2012). *Cultural responsiveness and school education: With particular focus on Australia's first peoples: A review & synthesis of the literature*. Menzies, Australia.

Pongsakdi, N., Kajamies, A., Veermans, K., Lertola, K., Vauras, M., & Lehtinen, E. (2020). What makes mathematical word problem solving challenging? Exploring the roles of word problem characteristics, text comprehension, and arithmetic skills. *ZDM-Mathematics Education*, *52*(1), 33-44. https://doi.org/10.1007/s11858-019-01118-9

Praet, M., Titeca, D., Ceulemans, A., & Desoete, A. (2013). Language in the prediction of arithmetics in kindergarten and grade 1. *Learning and Individual Differences*, *27*, 90-96. https://doi.org/10.1016/j.lindif.2013.07.003

Prediger, S., Erath, K., & Opitz, E. M. (2019). The language dimension of mathematical difficulties. In A. Fritz, V. G. Haase, & P. Räsänen (Eds.), *International handbook of mathematical learning difficulties: From the laboratory to the classroom* (pp. 437-455). Springer.

Purpura, D. J., & Reid, E. E. (2016). Mathematics and language: Individual and group differences in mathematical language skills in young children. *Early Childhood Research Quarterly*, *36*, 259-268. https://doi.org/10.1016/j.ecresq.2015.12.020

Riccomini, P. J., Smith, G. W., Hughes, E. M., & Fries, K. M. (2015). The language of mathematics: The importance of teaching and learning mathematical vocabulary. *Reading & Writing Quarterly*, *31*(3), 235-252.

Roohr, K. C., & Sireci, S. G. (2017). Evaluating computer-based test accommodations for English learners. *Educational Assessment*, *22*(1), 35-53.

Ruzlan Md-Ali, Arsaythamby Veloo, S. Kanageswari Suppiah Shanmugam, Yus'aiman Jusoh Yusoff, & Rosna Awang Hashim. (2021). The issues and challenges of mathematics teaching and learning in Malaysia "Orang Asli" primary schools from teachers' perspectives. *Malaysian Journal of Learning and Instruction*, *18*(2), 129-160.

S. Kanageswari Suppiah Shanmugam, Arsaythamby Veloo, Ruzlan Md-Ali. (2021). Culturally responsive assessment: Assessing mathematics performance of Indigenous pupils in Malaysia using trilingual test. *Diaspora, Indigenous, and Minority Education*, *15*(2), 113-136.

S. Kanageswari Suppiah Shanmugam, Arsaythamby Veloo, & Yus'aiman Bin Jusoh. (2023). The challenges of virtual learning: An exploratory study among Orang Asli Pupils from the mathematics teachers' perspectives. *Diaspora*, *Indigenous*, *and Minority Education*, *18*(3), 211-229. https://doi.org/10.1080/15595692.2023.2212823

Shohamy, E. (2011). Assessing multilingual competencies: Adopting construct valid assessment policies. *The Modern Language Journal*, *95*(3), 418-429. https://doi.org/10.1111/j.1540-4781.2011.01210.x

Skinner, M. G., & Cuevas, J. A. (2023). The effects of schema-based instruction on word-problems in a third-grade mathematics classroom. *International Journal of Instruction*, *16*(1), 855-880. https://doi.org/10.29333/iji.2023.16148a

Stansfield, C. W. (2011). Oral translation as a test accommodation for ELLs. *Language Testing*, *28*(3), 401-416. https://doi.org/10.1177/0265532211404191

Swanson, H. L. (2020). The relationship between executive processing and computational growth among monolingual and English learners with and without math difficulties: Does it help to be bilingual? *Cognitive Development*, *56*, 100961. https://doi.org/10.1016/j.cogdev.2020.100961.

Swanson, H. L., Kong, J., Petcu, S. D., & Asencio Pimentel, M. F. (2020). Can difficulties in language acquisition and specific learning disabilities be separated among English learners? *Exceptional Children*, *86*(3), 293-309. https://doi.org/10.1016/j.cogdev.2020.100961

Tambunan, H. (2019). The effectiveness of the problem solving strategy and the scientific approach to students' mathematical capabilities in high order thinking skills. *International Electronic Journal of Mathematics Education*, *14*(2), 293-302.

Timmermans, R. E., Van Lieshout, E. D. C. M., & Verhoevan, L. (2007). Gender related effects of contemporary math instruction for low performers on problem-solving behaviour. *Learning and Instruction*, *17*(1), 42-54.

Trumbull, E., & Nelson-Barber, S. (2019). The ongoing quest for culturally-responsive assessment for Indigenous students in the US. *Frontiers in Education*, *4*(40), 1-11. https://doi.org/10.3389/feduc.2019.00040

UNESCO (2015). *Education for All 2000–2015: Achievements and challenges*. EFA Global Monitoring Report. Paris: UNESCO.

van Dijk, T. A. (1993). Principles of critical discourse analysis. *Discourse & Society*, *4*(2), 249-283. https://doi.org/10.1177%2F0957926593004002006

Verschaffel, L., Schukajlow, S., Star, J., & Van Dooren, W. (2020). Word problem in mathematics education: A survey. *ZDM Mathematics Education*, 52, 1-16.

Vidal Rodeiro, C., & Macinska, S. (2022). Equal opportunity or unfair advantage? The impact of test accommodations on performance in high-stakes assessments. *Assessment in Education: Principles, Policy & Practice*, *29*(4), 462-481.

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

Wijns, N., De Smedt, B., Verschaffel, L., & Torbeyns, J. (2020). Are preschoolers who spontaneously create patterns better in mathematics? *British Journal of Educational Psychology*, *90*(3), 753-769. https://doi.org/10.1111/bjep.12329

Wilkinson, L. C. (2018). Teaching the language of mathematics: What the research tells us teachers need to know and do. *The Journal of Mathematical Behavior*, *51*, 167-174. https://doi.org/10.1016/j.jmathb.2018.05.001

Zaleha Ismail, Tan, Y. C., & Nur Amira Muda. (2020). Numeracy competency of Year 5 aboriginal students using written and oral tests. *The Mathematics Enthusiast*, *17*(1), 32-62. https://doi.org/10.54870/1551-3440.1479