



Higher-order thinking word problem-solving errors made by low-performing pupils: Comparative case study across school types in Malaysia

Menaga Suseelan¹ · Cheng Meng Chew¹ · Huan Chin¹

Accepted: 23 May 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

The achievement gaps existed among the three types of Malaysian elementary schools with different instruction mediums. Thus, this study sought to compare the types of errors made by the low-performing students from different school types in solving word problems involving higher-order thinking skills by conducting a multiple case study. The study involved 18 low-performing students from a National Primary School (NPS), National-Type Chinese Primary School (NTCPS), and National-Type Tamil Primary School (NTTPS) in Malaysia. The students' errors in solving eight word problems were identified qualitatively through conducting four sessions of diagnostic interviews, followed by the Newman's Error Analysis. Then, a cross-case analysis was conducted. The findings indicated that the participants of the three school types made reading, comprehension, transformation, process skills and encoding errors. However, the NPS and NTTPS low-performing students made more process skills errors and encoding errors, as compared to the NTCPS low-performing students. The findings imply the need to enhance the procedural fluency of NPS and NTTPS low-performing students for reducing the achievement gaps among the school types.

Keywords Higher-order thinking skills · Low-performing · Newman's error analysis · School type · Word problem

Introduction

Geometric measurements such as perimeter, area, and volume are the fundamental knowledge required in most fields such as science, technology, and engineering (Smith et al., 2011). These important mathematical topics are usually included in either the geometry strand or the measurement strand in the elementary mathematics curriculum of various countries (Huang & Witz, 2011; National Council of Teachers of Mathematics [NCTM], 2000). After learning the mathematical concepts, students are engaged in solving routine word problems involving geometric measurements so that they can relate the concepts learnt with the real-world applications (Verschaffel et al., 2010).

While real-world problems are often presented in an unfamiliar context and have no definite solution, solving this

type of problem involves manipulating the learnt knowledge and synthesizing a new strategy, instead of applying the known strategy. In other words, higher-order thinking skills (HOTS) are highly in demand for solving such problems. To improve students' problem-solving competency, the students are also engaged in solving the non-routine word problems that involve deriving the numerical information which is not explicitly given based on the properties of the geometric shape (Jones, 2002) and the relationship between the formulae (Yeo, 2008). These non-routine word problems are regarded as word problems involving HOTS in this study.

While there is a greater emphasis on problem-solving and HOTS, the result of the Trend in International Mathematics and Science Study (TIMSS) 2019 indicated that the performance gaps between high-performing students and low-performing students are large (Mullis et al., 2020). The low-performing students often fall far behind their counterparts in solving geometric measurement word problems involving HOTS (Satsangi & Bouck, 2015). To support the low-performing students in problem-solving skill acquisition, there is a need to identify the difficulties faced by the students in solving the word problems involving HOTS. In

✉ Cheng Meng Chew
cmchew@usm.my

¹ School of Educational Studies, Universiti Sains Malaysia,
11800 USM, Penang, Malaysia

this regard, Newman's Error Analysis could be a promising method to deepen educators' understanding of student difficulties because it can provide a highly specific diagnosis of types of problem-solving errors made (Watson, 1980).

As a multi-ethnic country, Malaysia practises a vernacular education system to accommodate the needs of the three major ethnic groups, namely Malays (69.6%), Chinese (22.6%) and Indians (6.8%) (Department of Statistics Malaysia, 2021). The elementary schools in Malaysia are streamed into three school types based on the medium of instruction: Malay-medium National Primary School (NPS); Mandarin-medium National-Type Chinese Primary School (NTCPS), and Tamil-medium National-Type Tamil Primary School (NTTPS). Thus, the three main ethnic groups have equal opportunities to receive education in their mother tongue (L1).

Despite the same curriculum being adopted in the three types of schools, researchers (i.e., Chew et al., 2021; Chin & Chew, 2022b; Sia & Lim, 2020) constantly found that students from NTCPS performed better than students from NPS and NTTPS in solving word problems. This indicates the existence of mathematics achievement gaps in Malaysia due to school type difference besides the learners' ability. To narrow down the gaps, this study sought to compare the errors made by the low-performing students from NPS, NTCPS, and NTTPS in solving word problems involving HOTS. The findings of the study would provide insight into the hindrance which obstructs students to solve word problems from a cultural perspective. This would eventually deepen teachers' understanding of the problem-solving difficulties faced by their students with different cultural backgrounds while the education system in many countries is increasingly diversified due to globalisation (Zhang, 2019).

Newman's Error Analysis (NEA)

Newman's Error Analysis (NEA) is a simple diagnostic procedure introduced by Newman (1977, 1983) for identifying errors in solving mathematical word problems. In order to solve the word problems correctly, Newman (1977, 1983) claimed that the student must pass through five successive hurdles: (i) Reading (ii) Comprehension, (iii) Transformation, (iv) Process Skills, and (v) Encoding. In the first stage, the students would have to read the word problems. This was followed by comprehending the numerical data, the quantitative relationship of the objects, and the numerical tasks in the second stage. Then, the students would have to transform the numerical data into mathematical sentences by applying the appropriate arithmetic operation to solve the numerical tasks. After that, the students would have to perform the calculation based on the mathematical sentences formulated. Lastly, the students would have to express the final answer in the unit as requested in the word problem. For example,

rather than expressing the final answer for the area of the square in *cm*, the students should express the final answer in the unit of cm^2 .

Anchoring on the claim made by Newman (1977, 1983), an error would arise if the students failed to pass through any of the five successive stages in solving word problems. To identify the error made by the students in solving word problems, Newman (1977, 1983) proposed a diagnostic procedure that involves conducting an individual interview with the students while they are solving the word problems. Since each Newman's interview prompt corresponds with a problem-solving hurdle (White, 2010), the type of error made by the student could be specified systematically (Watson, 1980).

Numerous past studies had applied NEA in their studies to get insights into students' errors while solving various problems. The study conducted by Abdullah et al. (2015) showed that Malaysian secondary students made all the five errors with almost equal proportions when solving mathematics problems involving HOTS. Rather than conducting studies in the secondary school context, the researchers such as Chin and Chew (2022a, b), Singh et al. (2010), as well as Sibanda (2017) also determined elementary students' errors in solving word problems in Malaysia and South Africa. The findings indicate that majority of the students failed to solve the word problems due to comprehension errors (Sibanda, 2017; Singh et al., 2010). On the other hand, Chin and Chew (2022a, b) found that the transformation error contributed the most to the student's failure in solving word problems. Despite the numerous studies on NEA, the past studies only diagnosed the errors made by students with all ranges of abilities in different contexts.

Higher-Order Thinking Skills (HOTS)

Higher-order thinking skills (HOTS) refer to creative, critical, and analytical thinking skills used to solve non-routine problems through manipulating existing knowledge and known algorithms (Puspitasari et al., 2018; Yeung, 2012). In other words, solving the non-routine word problems would require the top three levels of the cognitive domain in Bloom's Taxonomy, namely: analysing, evaluating, and creating (Anderson et al., 2001). The students need to extract scattered information in the non-routine word problems and link them accordingly (analyzing), make suitable inferences based on prior conceptual and procedural knowledge (evaluating), and eventually manipulate the knowledge and devise a procedure to solve the word problem (creating). Thus, non-routine word problems are considered word problems involving HOTS in this study.

Past studies related to HOTS were mostly related to large-scale assessments such as TIMSS and Programme for International Student Assessment (PISA). Even though Mullis

et al. (2020) found that majority of Grade Four and Eight students could not solve complex problems in TIMSS 2019, Puspitasari et al. (2018) indicated that non-routine word problems could be solved by the Indonesian Grade Eight high-achieving students without any difficulty. In Malaysia, most studies on HOTS were conducted in the secondary school context. Based on the study conducted by Abdullah et al. (2017), majority of Grade 10 students could not solve the non-routine word problems. Likewise, Suseelan et al. (2022) also found that most Malaysian students in elementary schools performed poorly in solving word problems involving higher-order thinking skills. Although various studies related to HOTS have been conducted in the past, the findings were inconsistent.

School-Type Differences in Mathematics Teaching and Learning in Malaysia

The diversity in the Malaysian education system sparked the interest of researchers to study the school-type differences in mathematics teaching and learning. Despite the same mathematics curriculum used (Lim, 2003) and the same teachers' qualifications among the school types (Hamid et al., 2012), the existence of an achievement gap was reported in several studies (i.e., Ghazali & Sinnakaudan, 2014; Lim, 2003; Sia & Lim, 2020). This is because mathematics teaching and learning could be affected by the instruction medium used (Vukovic & Lesaux, 2013). While the Mandarin language has a simpler and consistent number system, the use of the language decreased students' memory load and hence contributed to the effective retrieval of procedural facts learned (Zhang et al., 2019). Thus, the use of Mandarin language as an instruction medium benefits L1 students in mathematical concept acquisition (Lim, 2003).

Rather than focussing on achievement differences and the impact of instruction medium, several researchers (e.g., Chia & Lim, 2020; Ghazali & Sinnakaudan, 2014; Roscoe & Sriraman, 2011) compared the teachers' beliefs and instructional practice across school types. The NTCPS teachers held stronger beliefs in constructivist teaching in mathematics learning compared to the NPS and NTTPS teachers (Ghazali & Sinnakaudan, 2014). Thus, the mathematics lessons in NTCPS focused on developing students' conceptual understanding followed by procedural fluency. The NTCPS teachers allocated more time to explain the mathematical concepts (Chia & Lim, 2020; Ghazali & Sinnakaudan, 2014; Lim, 2003). After that, the students were given drills and practices to reinforce the concepts they had learned (Ghazali & Sinnakaudan, 2014; Lim, 2003). Besides, the NTCPS teachers actively engaged the students in mathematics by developing problem-solving activities to enhance their problem-solving competency (Ghazali & Sinnakaudan, 2014).

On the contrary, the NPS and NTTPS teachers believed that a teacher-centred learning environment with adequate reinforcement practices was sufficient to support students' learning (Roscoe & Sriraman, 2011). With this perception, more time was allocated for desk instruction in the NPS mathematics classroom (Chia & Lim, 2020). Instead of explaining the mathematical concepts explicitly, the NPS teachers spent more time walking around the class to check their students' work while they were working out the exercise individually in the classroom. The NPS and NTTPS teachers also had low confidence in students' capability in solving word problems (Ghazali & Sinnakaudan, 2014). Since they perceived that the students could not solve mathematics problems independently, they taught them to solve the word problems by applying the learnt algorithm (Ghazali & Sinnakaudan, 2014; Lim, 2003), rather than engaging them to solve the problems in the lessons.

Purpose of the Study

Various studies on assessing students' competency in solving problems involving HOTS and diagnosing students' errors made in solving word problems have been conducted in the past. However, the reviewed studies involved students with different ranges of abilities, and studies that solely focused on low-performing students are considered scarce. Whilst Malaysia has a diverse school system, the previous studies mainly focused on comparing the teachers' beliefs, teaching approaches and the impact of language used on students' mathematics learning. To fill the research gaps, this study sought to compare the errors made by the Malaysian low-performing pupils from different types of schools in solving word problems involving measurement formulae and higher-order thinking skills based on NEA. The research questions addressed in this study are:

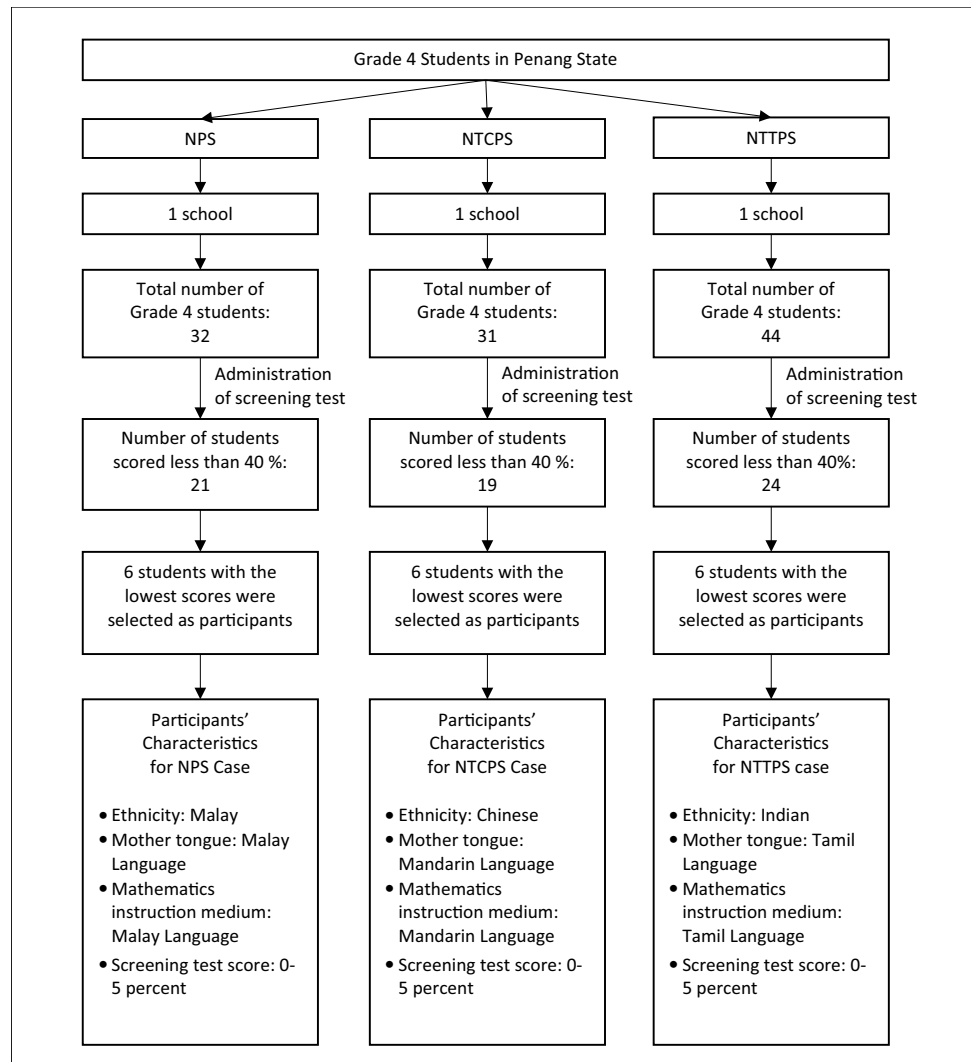
- (1) What are the types of errors made by low-performing pupils in solving word problems involving HOTS and measurement formulae based on NEA?
- (2) What are the differences in the types of errors made by low-performing NPS, NTCPS and NTTPS pupils in solving word problems involving HOTS and measurement formulae?

Methodology

Research Design

The study was conducted by employing a multiple case study research design. As advocated by Creswell and Poth (2016), multiple case study involves performing an in-depth analysis of the data collected from multiple sources

Fig. 1 Sample Selection Process



to explore the phenomenon in the multiple bounded systems explicitly. In other words, multiple case studies are appropriate to be used when the cases involve different contexts (Ridder, 2017) because the robustness of the study could be increased through cross-case comparison (Yin, 2018; Miles & Huberman, 1994). In this study, school type served as the boundaries of each case. With school type as the unit of analysis, school-type comparisons of errors made by the low-performing students in solving higher-order thinking word problems could be made. While each case bounded by the school type could be used for confirming or disconfirming the conclusion drawn from each other, the use of multiple case studies would enhance the external validation of the findings.

Population and Sampling

The population of the study was Grade Four students from NPS, NTCPS, and NTTPS because measurement

formulae are introduced to the students in Grade Four. Due to the practical constraint, the sampling frame was limited to Penang state, Malaysia. Since the study aimed to compare the errors made by the low-performing students in solving word problems, the sample of the study was selected using stratified purposive sampling (Radhakrishnan, 2014).

The sample selection process is shown in Fig. 1. The process began by stratifying the population into three strata based on school type, namely NPS, NTCPS, and NTTPS. Then, one school was selected to represent each stratum. This was followed by selecting the low-performing students as the sample of study through the administration of a problem-solving screening test involving HOTS to all Grade Four students in the three selected schools (NPS: 32; NTCPS: 31; NTTPS: 44). In this study, the low-performing students were operationalized as the students who scored below 40% based on the guidelines given by the Malaysian Examination Syndicate (2016).

There were 21 NPS students, 19 NTCPS students and 24 NTTPS students who were categorized as low-performing students. Due to the time constraint, only six students who scored the lowest in the problem-solving screening test from each school type were selected as the participants of the study. This was supported by the previous studies (i.e., Baxter et al., 2005; McAuley & McLaughlin, 1992; Methe et al., 2012) that involved a small sample size ($n \leq 6$). Since the students received mathematics instruction in their mother tongue and scored at most five marks on the problem-solving test, the participants from each school type were equivalent to each other.

Instruments of the Study

Problem-Solving Screening Test

In this study, the problem-solving screening test served two purposes: (i) identifying the low-performing students, and (ii) eliciting participants' responses to determine their errors. The problem-solving screening test consisted of eight open-ended non-routine word problems involving geometric measurements and HOTS. They were adopted from Electronic Supplementary Material 1¹ published by Menaga et al. (2022) and covered all aspects of geometric measurement included in the Malaysian Grade Four Mathematics Curriculum: (i) perimeter of squares, rectangles, triangles, and regular polygons; (ii) area of squares, rectangles, and triangles, and (iii) volume of cubes and cuboids.

All eight problems required the students to analyse the given pieces of information in the word problems (analyse), identify the appropriate numerical information in the problems (evaluate) and manipulate the necessary measurement formula (create). Since they involved the top three levels of the cognitive domain in the Revised Bloom's Taxonomy (Anderson et al., 2001), they were categorized as HOTS problems. The content and cognitive domains involved and the mark allocation for each item in the test are summarized in Table 1. The scores allocated for each problem ranged from four to six depending on the complexity of the solution. The more complex the solution was, the higher the marks allocated. The scores were given based on the methods used, the calculations performed, and the final answer written. The marking scheme for Item Q1 is shown in Electronic Supplementary Material 2² prepared by Menaga et al. (2022).

Table 1 Cognitive Domain Involved and Content Covered in the Problem-Solving Test

Items	Content Covered	Cognitive Domain	Mark Allocated
Q1	• Perimeter of rectangle	Analysing, Evaluating, Creating	4
Q2	• Perimeter of polygon	Analysing, Evaluating, Creating	4
Q3	• Area of rectangle • Area of triangle	Analysing, Evaluating, Creating	6
Q4	• Perimeter of rectangle • Area of rectangle	Analysing, Evaluating, Creating	6
Q5	• Perimeter of square • Area of square	Analysing, Evaluating, Creating	5
Q6	• Perimeter of square • Perimeter of triangle • Area of triangle	Analysing, Evaluating, Creating	6
Q7	• Volume of cube	Analysing, Evaluating, Creating	6
Q8	• Volume of cuboid	Analysing, Evaluating, Creating	4

Initially, the problem-solving screening test was formed in English. Then, it was translated to Malay, Mandarin, and Tamil languages to ensure its usability for the NPS, NTCPS, and NTTPS students respectively. Upon the construction of the problem-solving test, it was sent to two subject matter experts from NPS, NTCPS, and NTTPS, each with at least 10 years of teaching experience for the evaluation of content coverage and item relevance. Together with the problem-solving test, the test specification, the content coverage judgemental form and the item relevance judgemental form were emailed to the experts. The experts were requested to rate the content coverage (the extent to which each item was covered in the syllabus) and the item relevance (the extent to which the items were relevant to the constructs and cognitive domains measured) on the judgemental form with a five-point Likert Scale.

To evaluate the consensus of the experts, the Scale-level Content Validity Index (S-CVI) was calculated for both content coverage and item relevance of the problem-solving test based on the ratings given by the experts (Polit & Beck, 2006). The S-CVI for the content coverage of the problem-solving screening test was 1.00 for all the three types of schools. This indicated that all the items in the problem-solving screening test were covered in the Grade Four syllabus (Polit & Beck, 2006). The item relevance S-CVI for the Malay, Mandarin and Tamil versions of the problem-solving test were .94, 1.00, and 1.00 respectively. All the S-CVI values surpassed .80, indicating that the problem-solving test items were relevant to the construct and cognitive domain measured (Polit & Beck, 2006).

¹ Please refer to the Electronic Supplementary Material 1 (Menaga et al., 2022) at <https://link.springer.com/article/10.1007/s10763-021-10245-3>

² Please refer to the Electronic Supplementary Material 2 (Menaga et al., 2022) at <https://link.springer.com/article/10.1007/s10763-021-10245-3>

Table 2 Newman's Error Analysis Coding Framework (Adopted from Watson, 1980)

Code	Error categories	Descriptors
R	Reading error	<ul style="list-style-type: none"> • The student fails to recognize the words or symbols in the problem.
C	Comprehension error	<ul style="list-style-type: none"> • The student doesn't know what was asked by the problem. • The student does not understand certain terms and symbols in the word problems.
T	Transformation error	<ul style="list-style-type: none"> • The students fail to transform the word problems into mathematics sentences using correct arithmetic operations.
P	Process skills error	<ul style="list-style-type: none"> • The students do not know how to perform the calculation. • The students make mistakes in performing the calculation.
E	Encoding error	<ul style="list-style-type: none"> • The students fail to express the answer in the correct unit. • The students cannot conclude the final answer based on their calculations.

After the instrument validation, the Malay, Mandarin, and Tamil versions were piloted at the NPS, NTCPS and NTPPS schools respectively to evaluate their reliability. The estimate of reliability was calculated by obtaining Cronbach's alpha value using SPSS Version 24. With the Cronbach's alpha coefficient (Malay version: .83; Mandarin version: .81; Tamil version: .78) surpassing the reliability common rule of thumb of .70, the three versions of the problem-solving test were reliable (Pallant, 2016).

NEA Interview Protocol

The NEA interview protocol was used to guide the semi-structured interview for collecting the qualitative data on students' errors made. The NEA interview protocol consisted of five interview prompts adopted from White (2010). Each interview prompt was used to identify a specific error made in solving the word problems:

- (1) Reading Error: Please read the question to me. If you don't know a word, leave it out.
- (2) Comprehension Error: Tell me what the question is asking you to do.
- (3) Transformation Error: Tell me how you are going to find the answer.
- (4) Procedural Skill Error: Show me what to do to get the answer. "Talk aloud" as you do it so that I can understand how you are thinking.
- (5) Encoding Error: Now, write down your answer to the question.

Research Procedure

Before the data collection, permission to conduct the study was obtained from the Malaysian Educational Planning and Research Division (EPRD) and the Penang State Education Department. With the permission of the headmaster of each participating school and parents' consent, the research participants identified through the problem-solving screening test were invited to attend four one-to-one NEA interview sessions

which were conducted using the students' mother tongue. The four interview sessions with the duration of 30 minutes each were conducted on four consecutive days during school hours. Each interview session began by engaging the student to solve two word problems in the problem-solving test [i.e., Session 1: Q1 and Q2; Session 2: Q3 and Q4; Session 3: Q5 and Q6; Session 4: Q7 and Q8], followed by identifying the student's errors made in each word problem guided by the NEA interview protocol. All sessions were video-recorded with the student's consent and transcribed verbatim into text for data analysis purposes by three researchers who are fluent in Malay, Mandarin and Tamil respectively. After transcription, the transcripts were rechecked with the original audio to ensure the accuracy of the data (Braun & Clarke, 2006).

Data Analysis

The data analysis was conducted by three researchers who are fluent in Malay, Mandarin and Tamil respectively. To address Research Question One, Newman's Error Analysis was conducted. The analysis process began by coding the transcripts based on the coding framework proposed by Watson (1980) as shown in Table 2. The reading, comprehension, transformation, process skills and encoding errors were identified and coded as 'R', 'C', 'T', 'P' and 'E' respectively. Then, the frequency of each type of error made was tabulated. Besides, the mean frequency was calculated for each error type to indicate the average number of errors made by each student. To address Research Question Two, a cross-case analysis was conducted. For each case (school type), the mean frequency of each type of error made was calculated. Then, the cross-case comparison was conducted to determine the differences in terms of errors made.

Findings

Types of Errors Made by Low-Performing Students

Based on the coded transcripts of the NEA interview, it was found that the students from all three types of schools made

Table 3 Descriptive Analysis of Errors Made by the Low-Performing Students from the Three School Types

Types of Error	Total Frequency (Percentage)	Mean Frequency
Reading	37 (25.69%)	2.06
Comprehension	100 (69.44%)	5.56
Transformation	144 (100.00%)	8.00
Process Skills	89 (61.81%)	4.94
Encoding	51 (35.42%)	2.83

Mean frequency indicates the average errors made by each student

all the five errors, namely reading, comprehension, transformation, process skills and encoding errors. Since each error could be made by the participants once for each item, the total possible errors made by the students from the three types of schools was 144 (3 types of school \times 6 students per school type \times 8 items \times 1 time for each error type = 144 times for each type of error). The percentages of all types of errors based on NEA made by the low-performing Grade Four students in solving word problems involving measurement formulae and HOTS identified in this study are shown in Table 3.

In general, transformation errors are the most common errors made by the low-performing students from all three types of schools. All students from NPS, NTCPS, and NTTPS made transformation errors (100%) in solving the eight word problems involving geometric measurement and HOTS. With the guidance given by the researcher, all low-performing students have rectified the reading and comprehension errors made. Yet, they failed to transform the eight word problems into mathematical sentences using correct sentences.

The comprehension error ranked second in terms of its occurrence. A total of 100 comprehension errors have been made by the 18 low-performing students from the three types of schools. Even though their reading errors have been corrected, they failed to understand the terminology used and the diagram shown in the word problems. Thus, they could not explain the problem situation using their own words even with the researcher's prompts. Besides, they also failed to state the requirements of the word problems. With a mean frequency of 5.56, each low-performing student made a comprehension error in about six word problems on average.

The total number of processing skills errors made by the low-performing students was slightly lower than comprehension errors. The 18 low-performing students made 89 process skill errors in solving the eight word problems. In other words, each student made a process skills error in nearly five word problems. Even though they had formulated the correct mathematical sentences for solving the word problems, they made errors in performing arithmetic operations. Consequently, they obtained the wrong answer.

In general, the students made the least errors in reading the word problems. There were only 37 reading errors that had been made by the 18 low-performing students from the three types of schools. In other words, most of the students could recognise and pronounce the words and symbols shown in the word problems correctly. Out of the eight word problems, each student only made a reading error for two word problems on average.

Comparison of Types of Errors Made across School Types

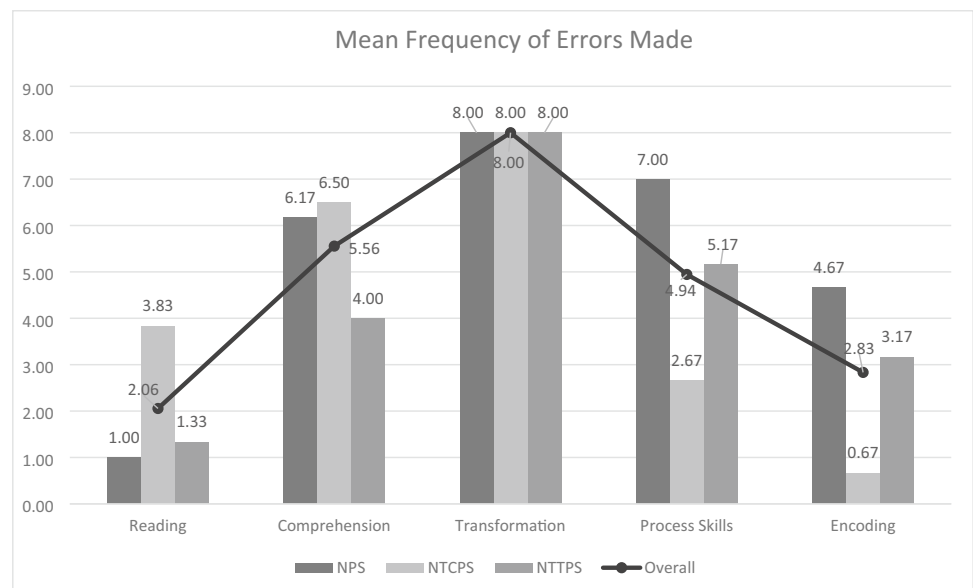
To compare the errors made by the low-performing students from NPS, NCTPS, and NTTPS, the mean frequency of each type of error made was calculated. Then, the combo chart was plotted to illustrate the comparison of the mean frequency of errors made. As shown in Fig. 2, the mean frequency was not equal across school types for most of the types of problem-solving errors, except transformation errors.

The mean frequency of reading and comprehension errors made by the low-performing NTCPS students was higher than that of the low-performing students from NPS and NTTPS. However, there was only a slight difference between the mean frequency of comprehension errors made by the low-performing students from NPS and NCTPS. On average, the low-performing students from NPS ($M_{Comprehension, NPS} = 6.17$) and NTCPS ($M_{Comprehension, NTCPS} = 6.50$) did not comprehend about six word problems. Meanwhile, the low-performing students from NTTPS ($M_{Comprehension, NTTPS} = 4.00$) made comprehension errors for four word problems on average.

With a mean frequency of 3.83, the reading errors made by the low-performing NTCPS students were at least three-fold of the reading errors made by the low-performing students from NPS ($M_{Reading, NPS} = 1.00$) and NTTPS ($M_{Reading, NTTPS} = 1.33$). In the interview sessions, the low-performing NTCPS students were asked to read aloud the word problems presented to them. However, the read-aloud session was paused several times because the low-performing NTCPS students did not know how to pronounce the words. They only could continue reading aloud after the researcher guided them to pronounce the unfamiliar words. Besides, the low-performing NTCPS students also pronounced the words wrongly. For example, SJKC11 pronounced the mandarin character '该' [correct pronunciation: gāi] wrongly as 'kè' [the pronunciation for the mandarin character '刻']. In contrast, the low-performing students from NPS and NTTPS were fluent in reading the word problems. The long pauses and incorrect pronunciation of words were seldom found in their read-aloud sessions.

The mean frequency of process skills and encoding errors made by the low-performing NTCPS students was

Fig. 2 Mean Frequency of Each Type of Errors Made across School Type



lower than that of the low-performing students from NPS and NTTPS. A large difference was observed in the mean frequency for the two types of errors among the three school types. The mean frequency of process skills errors made by the low-performing NTCPS ($M_{Process\ skills, NTCPS} = 2.67$) was half of the mean frequency of process skills errors made by the low-performing NTTPS students ($M_{Process\ skills, NTTPS} = 5.17$). Meanwhile, the mean frequency of process skills errors made by the low-performing NCTPS ($M_{Process\ skills, NCTPS} = 2.67$) was less than half of the mean frequency of process skills errors made by the low-performing NPS students ($M_{Process\ skills, NPS} = 7.00$). Most of the low-performing NCTPS students only made computation errors in performing division. However, the process skills errors were made by the low-performing students from NPS and NTTPS in other arithmetic operations besides division. For example, the low-performing NPS student, SK32 obtained 66 by adding six to six in Interview Session 3. On the other hand, the low-performing NTTPS student, SJKT32 answered six when the researcher asked, ‘5 times what is 35?’ during Interview Session 1.

With a mean frequency of 0.67, the encoding errors made by the low-performing NTCPS students were less than the low-performing students from NPS ($M_{Encoding, NPS} = 4.67$) and NTTPS ($M_{Encoding, NPS} = 3.17$). Specifically, the mean frequency of encoding errors made by the low-performing NPS students was nearly seven times the mean frequency of reading errors made by the low-performing students from NTCPS. On the other hand, the mean frequency of encoding errors made by the low-performing NTTPS students was nearly five times the mean frequency of reading errors made by low-performing students from NTCPS. Most of the low-performing students from NPS and NTTPS failed to express

the final answer with the correct unit. For example, the low-performing NPS and NTTPS students (i.e., SK28 and SJKT32) gave the final answer for the area in the unit of *cm*, rather than *cm²* during Interview Session 2 conducted in the respective school. However, the encoding errors were rarely made by the low-performing students in NTCPS. In the four interview sessions, most of them expressed the final answers with the correct unit as requested in the word problems.

Discussion

What are the Types of Errors Made by Low-Performing Pupils in Solving Word Problems Involving HOTS and Geometric Measurement Based on NEA?

This study revealed that the low-performing students from NPS, NTCPS, and NTTPS made all five types of errors, namely reading, comprehension, transformation, process skills and encoding in solving word problems involving geometric measurement and HOTS. These five types of errors corresponded with the four-step problem-solving process model introduced by Polya (2004): (i) understand the problem; (ii) devise a plan; (iii) carry out the plan; and (iv) look back. The reading and understanding errors were made by the low-performing students when they were understanding the word problem. The transformation errors were made by the students when they were formulating the mathematical sentences to solve the word problem (devise a plan). The procedural errors were made by the students when they were performing the calculation to solve the word problem (carry out the plan). The encoding errors

were made by the students when they checked whether their solution made sense or not (look back). In other words, the low-performing students from NPS, NTCPS and NTTPS made errors in all the problem-solving steps proposed by Polya (2004).

The most frequent errors were the transformation errors, followed by the comprehension errors, process skills errors, encoding errors and finally reading errors. In fact, the low-performing students from all three types of schools made transformation errors in all the word problems. This is in line with the study conducted by Abdullah et al. (2015), Chin and Chew (2022a, b) and Newman (1977). The findings are also in accordance with the claim made by Jiang et al. (2020), whereby many learners committed systematic errors at the stage of transformation due to the inappropriate application of the learnt heuristic. However, the findings contradicted the findings of the study by Raduan (2010) which indicated that comprehension errors were the most common errors made by the students in solving word problems. This might be due to the difference in the nature of word problems as well as the sample used in the study conducted by Raduan (2010). Unlike the study conducted by Raduan (2010) which involved students with different ranges of abilities in solving routine word problems in the real-world context, this study only involved low-performing students in solving non-routine word problems with high cognitive demands. These types of word problems might contain extraneous information. Besides, the mathematical relationship of the numerical information was not presented explicitly in the word problems. While the low-performing students are commonly characterized by poor mathematical representation ability (Montague, 2003), transforming the non-routine word problems into the mathematical sentence using correct operations might be a notoriously difficult task for them (Xin et al., 2005).

Reading error was the error least encountered in this study. This is in line with the study by Abdullah et al. (2015) as well as Chin and Chew (2022a, b). This might be because most of the Year Four pupils were able to read all the words or symbols in the word problems. However, they failed to understand the problems and hence could not proceed with the succeeding stages to solve the problem. Also, the word problems in the test were presented in their mother tongue, thus might have made it less difficult for them to read the problems (Ganuza & Hedman, 2017). In addition, the word problems have been validated by the subject matter experts to ensure the language used was on par with the level of Year Four students. The terminologies which the students were unfamiliar with had been replaced using simpler terms. In other words, the word problems were presented using the terms that were appropriate to the student's reading level. Thus, reading errors recorded the lowest percentage as compared to other types of errors.

In general, the students conducted more errors in transformation, process skills and encoding, as compared to reading and comprehension errors. According to Chan and Kwan (2021), students' word problem-solving proficiency was affected by their content knowledge and reading ability. Pivoting on this claim, the reading and comprehension errors were associated with language factors (Fuchs et al., 2018) while the transformation, process skills and encoding errors were associated with the content-knowledge factor (Lin, 2021; Singh et al., 2010). In other words, the proportion of errors made by the low-performing students related to content knowledge in this study was much higher compared to errors related to language factor. This was in line with the study conducted by Chin and Chew (2022a, b) as well as Clements and Ellerton (1996). The high proportion of errors made by the low-performing students on transformation, process skills, and encoding might be due to their low proficiency in mathematics content knowledge (Lin, 2021; Singh et al., 2010).

What are the Differences in the Type of Errors Made by Low-Performing NPS, NTCPS and NTTPS Pupils in Solving Word Problems Involving HOTS and Measurement Formulae?

The findings indicated that there was no difference in transformation errors made by the low-performing students from NPS, NTCPS, and NTTPS. Despite the mean frequency of comprehension errors made by low-performing students from NTCPS being higher than those from NPS and NTTPS, the difference in the mean frequency was not large. This might be due to the nature of word problems solved by the students. While the mathematical relationship among the numerical information was not explicitly presented in the word problems included in the screening test, it could be difficult to be comprehended by the low-performing students regardless of school type. Rather than visualising the mathematical relationship underlying the word problem, they only focused on the numbers and occasionally, the keywords in the problems (Montague, 2003). With insufficient mathematical vocabulary, they might also fail to recognise the keywords or misinterpret the keywords (Peng & Lin, 2019). Consequently, the low-performing students might simply formulate the mathematical sentence with the numerical information given in the word problems by using any arithmetic operations. Due to limited work memory capacity, the low-performing students might also fail to retrieve the schema which associated the keywords with the appropriate arithmetic operation (Namkung et al., 2019). Thus, the students tended to make transformation errors when solving the non-routine word problems regardless of school type.

The magnitude of differences in the reading errors made by the low-performing students from the three school types

was considerably large. The NTCPS low-performing students made more reading errors compared to the low-performing students from NPS and NTCPS. This could be due to the linguistic factor. According to Sung and Wu (2011), the Mandarin language is notably more difficult than the Malay and Tamil languages due to the complexity of its writing system. While Malay words consist of letter strings (Yap et al., 2010) and the Tamil scripts consist of vowels and consonants (Nag & Narayanan, 2019), reading the Malay words and Tamil scripts involves pronouncing the letter morpheme or vowels and consonants presented in the script. Unlike the Malay words and Tamil scripts, Chinese words are composed of characters with several strokes occupied in a two-dimensional box-shaped spatial layout (Yu & Reichle, 2017). To read the Chinese words, the NTCPS low-performing students had to recognize the character and the corresponding morphemic syllables called *pīnyīn*, instead of reading the word based on letter morpheme or vowels and consonants presented in the script (Chua & Tan, 2015; Yu & Reichle, 2017). While low-performing students are commonly associated with poor working memory (Xin et al., 2005), they tended to make reading errors when solving the word problems in this study.

The findings also indicated that NTCPS low-performing students performed significantly better in terms of process skills than NPS and NTPPS students. This is in line with the study by Chin et al. (2022), Lim and Chan (1993), as well as Sia and Lim (2020). The strong computational skills of NTCPS low-performing students might result from the teachers' practice in the mathematics classroom. As reported by the study conducted by Chia and Lim (2020), Ghazali and Sinnakaudan (2014), as well as Lim (2003), the mathematics lessons in NTCPS commonly began with the explanation of the concepts. Since this activity had the longest time allocation (Chia & Lim, 2020), the teachers had sufficient time to promote students' conceptual understanding. After explaining the concepts, the students would be given drills and practice to reinforce the concepts learned, and hence procedural fluency was built (Ghazali & Sinnakaudan, 2014; Lim, 2003). Thus, students from the NCTPS had better computational skills compared to those from the NPS and NTPPS.

Like the process skills error, the NTCPS low-performing students made significantly lower encoding errors than those from NPS and NTPPS. The findings are parallel with the findings reported by Sia and Lim (2020). This might be due to the different teaching methods used in each type of school. As reported in the study conducted by Chia and Lim (2020), the mathematical concepts were explained explicitly in the mathematics classroom. For example, various concrete examples were used by the teachers in NTCPS to support students' understanding of measurement units such as centimetre, and millimetre (Chia & Lim, 2020). With a

strong conceptual understanding, the students were able to identify the errors made when they looked back at the questions and their solutions. Moreover, the teachers in NTCPS gave plenty of additional word problems to the students at the end of every topic and guided them till the stage of writing the final answer correctly (Lim, 2002). This might have helped the low-performing pupils to learn to identify what should be written in the final answer space.

Conclusion

This study provides an insight into the comparison errors made by the low-performing students from the NPS, NTCPS and NTPPS in solving word problems involving geometric measurement and HOTS. Regardless of school type, the low-performing students made reading, comprehension, transformation, process skills and encoding errors. However, the tendency of conducting problem-solving errors varies among the school types. The low-performing students from NTCPS made more reading errors, while the low-performing students from NPS and NTPPS made more process skills errors and encoding errors. In short, the findings of this study provide case-based evidence on the differences in the problem-solving errors made by the low-performing students and hence explain their weaknesses in solving higher-order thinking word problems. While the study was conducted following the qualitative research paradigm, the findings of this study still need to be confirmed with a quantitative nationwide study for ensuring the generalisability of the findings.

Practical Implications

Despite all the types of errors that had been made by the low-performing students from NPS, NTCPS, and NTPPS, transformation errors were the most common errors which hindered them from solving the word problems involving geometric measurement and HOTS correctly. Even though they fully understood the numerical information and the problem situations as described in the higher-order thinking word problems, they failed to formulate the mathematical sentence for solving the word problems correctly. Thus, it is suggested that mathematics educators should plan for appropriated intervention that may help the low-performing students overcome the transformation errors in solving the word problems. This may eventually enhance their problem-solving competency.

The findings of this study highlighted that more process skills errors and encoding errors were made by the NPS and NTPPS low-performing students compared to their peers in NTCPS. According to Singh et al. (2010), process skills errors and encoding errors are content-knowledge related

errors. The low-performing students made mistakes in performing the computation based on the mathematical sentence formulated and failed to write the final answer using the correct units. This indicates the lack of procedural fluency among the low-performing students from NPS and NTPPS. Thus, procedural fluency should be emphasised in NPS and NTPPS mathematics classrooms for enhancing the low-performing students' computation skills. As such, the performance gap among the three types of schools could be reduced.

In addition, the low-performing students from NTCPS were reported for making more reading errors in solving higher-order thinking word problems, compared to the low-performing students from NPS and NTPPS. The NTCPS students could not recognise the Chinese characters, and hence failed to pronounce the words correctly. This would eventually obstruct their understanding of the problem situation described in the text. Consequently, they failed to solve the word problems. While reading error is a language-related error (Singh et al., 2010), the findings of this study call upon the collaboration among the Mandarin and Mathematics teachers to plan for appropriate remediation that could support low-performing students in rectifying the reading errors made during higher-order thinking word problem-solving.

Limitations and Recommendations

There are several limitations to this study. Firstly, the findings of this study could not be generalised to the entire population due to the use of multiple case studies that follow a qualitative research paradigm. Moreover, the findings of this study might be subjected to errors due to the small sample size of the study in view of the practical constraints such as sampling accessibility arising from the Covid-19 pandemic. To ensure the generalisability of the study, it is suggested to conduct a nationwide quantitative study involving a larger sample size in the future. Besides, it is suggested to use inferential statistics for making statistical inferences on school-type differences in problem-solving errors made so that more meaningful findings could be obtained.

Acknowledgements Acknowledgement to “Ministry of Higher Education for Fundamental Research Grant Scheme with Project Code: FRGS/1/2019/SS109/USM/02/13”.

Authors Contributions All authors have contributed to the study conception and design. Material preparation, data collection and analysis were performed by Menaga Suseelan, Chew Cheng Meng and Chin Huan. The first draft of the manuscript was written by Menaga Suseelan. Chew Cheng Meng and Chin Huan have commented on previous versions of the manuscript. All authors have read and approved the final manuscript.

Funding This work was supported by Fundamental Research Grant Scheme (FRGS) with Project Code: FRGS/1/2019/SS109/USM/02/13 of the Malaysian Ministry of Higher Education (MOHE).

Data Availability The datasets generated during and/or analysed during the current study are not publicly available due to privacy and ethical concern but are available from the corresponding author on reasonable request.

Code Availability Not applicable.

Declarations

Ethics Approval This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Human Research Ethics Committee of Universiti Sains Malaysia (Ethics Approval No: USM/JePEM/19090510).

Consent to Participate Written informed consent was obtained from the parents.

Consent to Publish Additional informed consent was obtained from all individual participants for whom identifying information is included in this article.

Conflicts of Interest/Competing Interests The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

- Abdullah, A. H., Abidin, N. L. Z., & Ali, M. (2015). Analysis of pupils' errors in solving higher order thinking skills (HOTS) problems for the topic of fraction. *Asian Social Science*, *11*(21), 133–144.
- Abdullah, A. H., Mokhtar, M., Halim, N. D. A., Ali, D. F., Tahir, L. M., & Kohar, U. H. A. (2017). Mathematics teachers' level of knowledge and practice on the implementation of higher order thinking skills (HOTS). *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(1), 3–17.
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Baxter, J. A., Woodward, J., & Olson, D. (2005). Writing in mathematics: An alternative form of communication for academically low-achieving students. *Learning Disabilities Research & Practice*, *20*(2), 119–135.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101.
- 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>
- Chew, C. M., Chin, H., & Suseelan, M. (2021). Year four pupils' performance in solving word problems and school-type difference. *Turkish Journal of Computer and Mathematics Education*, *12*(14), 4871–4878.
- Chia, H. M., & Lim, C. S. (2020). Characterising the pedagogical practices in mathematics lessons among selected Malaysian primary schools. *The Mathematics Enthusiast*, *17*(1), 307–323.
- Chin, H., & Chew, C. M. (2022a). Cognitive diagnostic assessment with ordered multiple-choice items for word problems involving 'time'. *Current Psychology*. [Advance online publication]. <https://doi.org/10.1007/s12144-022-02965-8>.

- Chin, H., & Chew, C.M. (2022b). Online cognitive diagnostic assessment with ordered multiple-choice items for word problems involving 'time'. *Education and Information Technologies*. Advance online publication. <https://doi.org/10.1007/s10639-022-10956-2>.
- Chin, H., Chew, C. M., Lim, H. L., & Thien, L. M. (2022). Development and validation of a cognitive diagnostic assessment with ordered multiple-choice items for addition of time. *International Journal of Science and Mathematics Education*, 20(4), 817–837. <https://doi.org/10.1007/s10763-021-10170-5>
- Chua, H. W., & Tan, T. G. (2015). A review of challenges in learning Chinese characters among non-native learners in Malaysia. *History*, 5(16), 93–100.
- Clements, M. A., & Ellerton, N. (1996). *The Newman procedure for analysing errors on written mathematical tasks*. Compasstech. <https://www.compasstech.com.au/ARNOLD/PAGES/newman.htm>
- Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.
- Department of Statistics Malaysia (2021). *Current population estimates, Malaysia, 2020*. Department of Statistics Malaysia.
- Fuchs, L. S., Gilbert, J. K., Fuchs, D., Seethaler, P. M., Martin, N., & 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.
- Ganuza, N., & Hedman, C. (2017). The impact of mother tongue instruction on the development of biliteracy: Evidence from Somali–Swedish bilinguals. *Applied Linguistics*, 40(1), 108–131.
- Ghazali, M., & Sinnakaudan, S. (2014). A research on teachers' beliefs about mathematics teaching and learning between Sekolah Kebangsaan (SK), Sekolah Jenis Kebangsaan Cina (SJKC) and Sekolah Jenis Kebangsaan Tamil (SJKT). *Journal of Education and Practice*, 5(31), 10–20.
- Hamid, S. R., Hassan, S., & Ismail, N. A. H. (2012). Teaching quality and performance among experienced teachers in Malaysia. *Australian Journal of Teacher Education*, 37(11), 85–103.
- Huang, H. M. E., & Witz, K. G. (2011). Developing children's conceptual understanding of area measurement: A curriculum and teaching experiment. *Learning and Instruction*, 21(1), 1–13.
- Jiang, R., Li, X., Xu, P., & Lei, Y. (2020). Do teachers need to inhibit heuristic bias in mathematics problem-solving? Evidence from a negative-priming study. *Current Psychology*. [Advance online publication]. <https://doi.org/10.1007/s12144-020-01209-x>.
- Jones, K. (2002). Issues in the teaching and learning of geometry. In L. Haggarty (Ed.), *Aspects of teaching secondary mathematics: Perspectives on practice* (pp. 121–139). Routledge.
- Lim, C. S. (2002). *The culture of mathematics learning in two chinese primary schools: Drill and practice*. [Paper presentation]. National Conference of Mathematics Education.
- Lim, C. S. (2003). Cultural differences and mathematics learning in Malaysia. *The Mathematics Educator*, 7(1), 110–122.
- Lim, S. K., & Chan, T. B. (1993). A case study comparing the learning of mathematics among Malay pupils in primary national school and primary national type Chinese schools. *Journal of Science and Mathematics Education in South East Asia*, 16(2), 49–53.
- Lin, X. (2021). Investigating the unique predictors of word-problem solving using meta-analytic structural equation modeling. *Educational Psychology Review*, 33(3), 1097–1124.
- Malaysian Examination Syndicate (2016). *Gred markah peperiksaan UPSR, PT3 dan SPM*. Kementerian Pendidikan Malaysia. Retrieved August 19, 2020, from <https://www.ipendidikan.my/gred-markah-peperiksaan-upsr-pt3-dan-spm.html>
- McAuley, S. M., & McLaughlin, T. F. (1992). Comparison of add-A-word and Compu spell programs with low-achieving students. *The Journal of Educational Research*, 85(6), 362–369. <https://doi.org/10.1080/00220671.1992.9941139>
- Menaga, S., Chew, C. M., & Chin, H. (2022). School-type difference among rural grade four Malaysian students' performance in solving mathematics word problems involving higher order thinking skills. *International Journal of Science and Mathematics Education*. <https://doi.org/10.1007/s10763-021-10245-3>.
- Methe, S. A., Kilgus, S. P., Neiman, C., & Riley-Tillman, T. C. (2012). Meta-analysis of interventions for basic mathematics computation in single-case research. *Journal of Behavioural Education*, 21, 230–253.
- Miles, M. B., & Huberman, A. M. (1994). *An expanded sourcebook qualitative data analysis* (2nd ed.). Sage publications.
- Montague, M. (2003). Teaching division to students with learning disabilities: A constructivist approach. *Exceptionality*, 11(3), 165–175.
- Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 international results in mathematics and science*. Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2019/international-results/>
- Nag, S., & Narayanan, B. (2019). *Orthographic knowledge, reading and spelling development in Tamil: The first three years*. In *Handbook of literacy in akshara orthography* (pp. 55–83). Springer.
- Namkung, J. M., Peng, P., & Lin, X. (2019). The relation between mathematics anxiety and mathematics performance among school-aged students: A meta-analysis. *Review of Educational Research*, 89(3), 459–496.
- National Council of Teachers of Mathematics [NCTM] (2000). *Principles and standards for school mathematics*. NCTM.
- Newman, M. A. (1977). An analysis of sixth-grade pupils' errors on written mathematical tasks. *Victorian Institute for Educational Research Bulletin*, 39, 31–43.
- Newman, M. A. (1983). *Strategies for diagnosis and remediation*. Harcourt Brace Jovanovich.
- Pallant, J. (2016). *SPSS survival manual: A step by step guide to data analysis using SPSS*. Open University Press/McGraw-Hill.
- Peng, P., & Lin, X. (2019). The relation between mathematics vocabulary and mathematics performance among fourth graders. *Learning and Individual Differences*, 69, 11–21.
- 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.
- Polya, G. (2004). *How to solve it: A new aspect of mathematical method* (No. 246). Princeton university press.
- Puspitasari, L., In'am, A., & Syaifuddin, M. (2018). Analysis of students' creative thinking in solving arithmetic problems. *International Electronic Journal of Mathematics Education*, 14(1), 49–60.
- Radhakrishnan, G. (2014). Sampling in mixed methods research. *International Journal of Advances in Nursing Management*, 2(1), 24–27.
- Raduan, I. (2010). Error analysis and the corresponding cognitive activities committed by year five primary pupils in solving mathematical word problems. *Procedia Social and Behavioral Sciences*, 2(2), 3836–3838.
- Ridder, H. G. (2017). The theory contribution of case study research designs. *Business Research*, 10(2), 281–305.
- Roscoe, M., & Sriraman, B. (2011). A quantitative study of the effects of informal mathematics activities on the beliefs of preservice elementary school teachers. *ZDM*, 1(21), 967–974.
- Satsangi, R., & Bouck, E. C. (2015). Using virtual manipulative instruction to teach the concepts of area and perimeter to secondary students with learning disabilities. *Learning Disability Quarterly*, 38(3), 174–186.
- Sia, C., & Lim, C. S. (2020). Knowledge states in the learning of "time": Comparison of primary six pupils from different types of schools in Penang. *The Mathematics Enthusiast*, 17(1), 6–31.

- Sibanda, L. (2017). Grade 4 learners' linguistic difficulties in solving mathematical assessments. *African Journal of Research in Mathematics, Science and Technology Education*, 21(1), 86–96.
- Singh, P., Rahman, A. A., & Teoh, S. H. (2010). The Newman procedure for analyzing primary four pupils errors on written mathematical tasks: A Malaysian perspective. *Procedia Social and Behavioral Sciences*, 8, 264–271.
- Smith, J. P., Heuvel-Panhuizen, M., & Teppo, A. R. (2011). Learning, teaching, and using measurement: Introduction to the issue. *ZDM*, 43(5), 617–620.
- Sung, K., & Wu, H. (2011). Factors influencing the learning of Chinese characters. *International Journal of Bilingual Education and Bilingualism*, 14(6), 683–700.
- Suseelan, M., Chew, C. M., & Chin, H. (2022). School-type difference among rural grade four Malaysian students' performance in solving mathematics word problems involving higher order thinking skills. *International Journal of Science and Mathematics Education* [Advance online publication]. <https://doi.org/10.1007/s10763-021-10245-3>.
- Verschaffel, L., Van Dooren, W., Greer, B., & Mukhopadhyay, S. (2010). Reconceptualising word problems as exercises in mathematical modelling. *Journal für Mathematik-Didaktik*, 31(1), 9–29.
- 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.
- Watson, I. (1980). Investigating errors of beginning mathematics. *Educational Studies in Mathematics*, 11(3), 319–329.
- White, A. L. (2010). Numeracy, literacy and Newman's error analysis. *Journal of Science and Mathematics Education in Southeast Asia*, 33(2), 129–148.
- Xin, Y. P., Jitendra, A. K., & Deatline-Buchman, A. (2005). Effects of mathematical word problem—Solving instruction on middle school students with learning problems. *The Journal of Special Education*, 39(3), 181–192.
- Yap, M. J., Liow, S. J. R., Jalil, S. B., & Faizal, S. S. B. (2010). The Malay lexicon project: A database of lexical statistics for 9,592 words. *Behavior Research Methods*, 42(4), 992–1003.
- Yeo, J. K. K. (2008). Teaching area and perimeter: Mathematics-pedagogical-content knowledge-in-action. In M. Goos, R. Brown, & K. Makar (Eds.), *Proceedings of the 31st Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 621–627). MERGA.
- Yeung, S. Y. (2012). Conceptualizing higher-order thinking for reforming school curriculum and teaching. *Hong Kong Teachers' Centre Journal*, 11, 131–144.
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage publications.
- Yu, L., & Reichle, E. D. (2017). Chinese versus English: Insights on cognition during reading. *Trends in Cognitive Sciences*, 21(10), 721–724.
- Zhang, J. (2019). Educational diversity and ethnic cultural heritage in the process of globalization. *International Journal of Anthropology and Ethnology*, 3(1), 7. <https://doi.org/10.1186/s41257-019-0022-x>
- Zhang, S., Cao, Y., Wang, L., & Li, X. (2019). Characteristics of teaching and learning single-digit whole number multiplication in China: The case of the nine-times table. *ZDM*, 51(1), 81–94.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.