

FROM MANIPULATION TO MEASUREMENT REASONING: RETHINKING VIRTUAL MANIPULATIVES FOR PRIMARY MATHEMATICS

Saifulnizan Che Ismail ^{1*}
Mohamad Ikram Zakaria ²

¹ Faculty of Educational Sciences and Technology, Universiti Teknologi Malaysia
(E-mail: saifulnizan@graduate.utm.my)

² Faculty of Educational Sciences and Technology, Universiti Teknologi Malaysia
(E-mail: mohamad.ikram@utm.my)

*Corresponding author: saifulnizan@graduate.utm.my

Article history

Received date : 15-8-2026

Revised date : 16-8-2026

Accepted date : 15-10-2026

Published date : 20-10-2026

To cite this document:

Che Ismail, S., & Zakaria, M. I. (2026). From manipulation to measurement reasoning: Rethinking virtual manipulatives for primary mathematics. *Journal of Islamic, Social, Economics and Development (JISED)*, 11 (87), 170 – 185.

Abstract: *Understanding relationships among measurable quantities is an important aspect of learning primary mathematics; however, measurement instruction may emphasise procedures rather than how quantities are related and change together. This study developed and conducted a preliminary classroom-based evaluation of the Virtual Manipulative Model of Measurement Relationships (VM-MR Model), implemented through the VM Measurement Y6 digital environment. The model was designed around three measurement relationships—length–mass, length–volume of liquid, and mass–volume of liquid—and three design principles: dynamic representation, guided exploration, and relational explanation. A quasi-experimental pre-test–post-test comparison-group design was employed with 60 Year 6 pupils from two intact classes in a Malaysian primary school. Thirty pupils received instruction using the VM-MR Model, while 30 pupils received textbook-based instruction covering the same measurement content. Both groups received 12 hours of instruction over six weeks. Pupils' Understanding of Relationships Among Measurable Quantities (URMQ) was assessed using a 15-item test, and the data were analysed using a two-way mixed ANOVA. The mean URMQ score of the VM-MR group increased from 31.33% at pre-test to 74.53% at post-test, while the comparison group increased from 29.33% to 61.07%. A significant Time × Group interaction was found, $F(1, 58) = 17.67$, $p < .001$, $\eta^2 = .233$, indicating different patterns of change between the two instructional groups. Significant Time × Group interactions were also found across all three measurement-relationship domains. These findings provide preliminary classroom-based evidence that virtual manipulation can be purposefully structured around explicit opportunities for measurement reasoning rather than digital interactivity alone.*

Keywords: *measurement reasoning, measurement relationships, primary mathematics, quantitative reasoning, virtual manipulatives*

Introduction

Measurement is an essential element of primary mathematics insofar as it involves having a learner make sense of quantities, units and relationships of measurable qualities. Park and Yeo (2023) highlight the importance of quantitative reasoning about units in elementary pupils' measurement activity. Learning measurement is, then, more than applying formulas, converting units, and finding correct numerical solutions. Thompson (1993) emphasises the conceptualisation of quantities and the development of relationships between them, while Skemp (1976) distinguishes relational understanding from more procedural learning. Pupils must thus have opportunities not just to measure quantities, but to identify and describe the relationships between quantities that can be measured.

The capability in this study is conceptualised as Understanding of Relationships Among Measurable Quantities (URMQ). Based on the description of the relational understanding provided by Skemp (1976), URMQ encompasses the consideration of meaningful relationships not procedures alone. It also shows the attention of Thompson (1993) towards conceptualisation of quantities and relationships between quantities. In cases where students are required to focus on changes in related quantities that occur simultaneously, Carlson et al. (2002) offer another theoretical line of reasoning in covariational reasoning. To this end, the operationalisation of URMQ in this study is the capability of the pupils to recognise measurable quantities, interpret relationships between them, coordinate changes between related quantities, and explain measurement relationships across three contexts: length–mass, length–volume of liquid, and mass–volume of liquid.

One of the ways mathematical objects and relationships can be brought to life in order to be explored interactively is virtual manipulatives. Although the role of digital representations in education cannot be reduced to interactivity, Moyer-Packenham and Suh (2012) demonstrated the potential of the virtual manipulatives with regard to learning mathematics. Similarly, in a critical review of the use of virtual manipulatives in elementary mathematics, Rich (2021) also paid attention to how the tools are integrated into the learning of mathematics. Weigand et al. (2024) in greater detail argue that the technological affordances are to be bridged in terms of their connection with the mathematical teaching and learning goals. With respect to mathematical visualisation, there is additional evidence to indicate that the representations may aid the learning process when they facilitate working with the mathematical concepts under investigation (Schoenherr et al., 2024).

This is a specific problem with measurement, in which the pupils might be required to observe a change in more than a single quantity. Panorkou (2020) demonstrated how changing quantities of measurement can be available to reasoning in the context of area and volume by dynamically changing environments. Building on this, Panorkou (2021) studied dynamic measurement reasoning about right prisms and cylinders in pupils, demonstrating how dynamic representations have the potential to organize changing quantities. Similar evidence provided by Yeo and Webel (2024) shows how a lively measurement environment could help elementary students in their reasoning about the relations among quantities. Despite these developments, relatively limited attention has been paid to how virtual manipulatives can be purposefully structured in terms of relationships among various measurable properties in primary mathematics. This raises an essential design challenge: how can virtual manipulation be designed in such a way that pupils can go beyond manipulation of digital representations to observe, coordinate, and explain relationships between measurable quantities?

The current study addresses this question through the Virtual Manipulative Model of Measurement Relationships (VM-MR Model) that is applied with the help of the VM Measurement Y6 digital environment. On the theoretical level, Skemp (1976) informs the model's focus on relational rather than procedural understanding. The quantitative reasoning view of Thompson (1993) focuses attention on quantities and the relationships between them, whereas Carlson et al. (2002) inform opportunities for pupils to coordinate changes between the related quantities. Such views are converted into three design principles: dynamic representation, guided exploration and relational explanation. By so doing, the model also responds to the broader argument that the organisation of digital affordances must be based on mathematical purposes (Weigand et al., 2024). The theoretical viewpoints thus serve as foundations for designing reasoning opportunities rather than as independent outcomes measured by the URMQ Test. The VM-MR Model thus frames virtual manipulation as a mediational space to think about measurement and not as a pedagogical end in itself.

To this end, the study carried out a preliminary classroom-based evaluation of whether patterns of change in Year 6 pupils' URMQ differed between pupils who underwent the VM-MR Model with VM Measurement Y6 and those who learned the same measurement content using the textbook. It also tested the hypothesis of different patterns of change in the three measurement-relationship domains. The research questions that were discussed in the study are as follows:

1. Does the pattern of change in Year 6 pupils' URMQ performance from pre-test to post-test differ between the VM-MR and textbook-based comparison groups?
1. Do the two instructional groups vary in the pattern of change across the length–mass, length–volume of liquid, and mass–volume of liquid relationship domains?

Literature Review

This literature review examines the theoretical and empirical foundations underlying the development of the VM-MR Model. It focuses on measurement reasoning and the understanding of relationships among measurable quantities, the role of virtual manipulatives and dynamic representations in mathematics learning, and the theoretical foundations and design principles underpinning the VM-MR Model.

Measurement Reasoning and Understanding of Relationships Among Measurable Quantities

Measurement learning requires learners not only to know how quantities are measured but also to understand the relationships among quantities, units, and measurable properties. As demonstrated by Park and Yeo (2023), the measurement activity in elementary pupils does not mean that they have to provide numerical responses but quantitative reasoning about units. Thompson (1993) argues that learners must be able to conceptualise quantities and develop relationships between quantities instead of working with numeric values. This relational orientation is especially critical in the case of two quantities that change simultaneously as learners should be concerned about how variations in one quantity are related to variations in another, which is a pivotal aspect of covariational reasoning (Carlson et al., 2002). A dynamic measurement situation also provides evidence that elementary students are able to interact with relationships between quantities when the relationships are represented in ways that allow students to explore them (Yeo & Webel, 2024).

This broader understanding of measurement reasoning is conceptualised in the current study as Understanding of Relationships Among Measurable Quantities (URMQ). One reason for focusing on meaningful mathematical relations rather than procedural performance alone is the

distinction between relational and instrumental understanding, made by Skemp (1976). URMQ is operationalised through four related skills, namely, identifying measurable quantities, interpreting relationships between quantities, coordinating changes between related quantities, and explaining measurement relationships. The emphasis on conceptualising quantities and building relationships between them is guided by Thompson's (1993) account of quantitative reasoning and the coordination of changes between related quantities is informed by the view of covariational reasoning by Carlson et al. (2002). These skills are analysed in three contexts: length–mass, length–volume of liquid, and mass–volume of liquid where measurement is not only discussed in the context of the attributes of any given amount but also the relations between the measurable amounts.

Collectively, these views give supplementary grounds to structure measurement reasoning within the VM-MR Model. Relational understanding moves the focus to elucidating and signifying mathematical connections (Skemp, 1976), whereas quantitative reasoning draws attention to quantities and the relationships between them (Thompson, 1993). This orientation is expanded by covariational reasoning which draws attention to the coordination of changes in related quantities (Carlson et al., 2002). Notably, dynamic settings can offer learners the opportunity to see and investigate such relations as quantities vary, rather than encountering them only through static numerical representations (Panorkou, 2020, 2021; Yeo & Webel, 2024). In the current research, however, these theoretical perspectives serve as design foundations for reasoning opportunities in the VM-MR Model; they are not treated as distinct constructs independently measured by the URMQ Test.

Virtual Manipulatives and Dynamic Representation

Virtual manipulatives are interactive interfaces through which learners can manipulate mathematical objects and test the effects of their actions. The study by Reimer and Moyer (2005) conducted in an early classroom setting demonstrated how virtual manipulatives might be integrated into learning elementary mathematics, whereas the study provided by Moyer-Packenham and Suh (2012) showed that learners can also learn mathematics with the help of technology-based manipulatives. We cannot, however, attribute to digital interactivity alone their educational potential. Rich (2021) emphasised the importance of considering how virtual manipulatives are incorporated into mathematical learning, while more recent reviews have similarly highlighted the role of instructional design and implementation in shaping learning with manipulatives (Byrne et al., 2023; Ochugboju & Díez-Palomar, 2025).

Recent studies have focused more on the relationship between technological affordances and mathematical learning objectives instead of the existence of technology. As Weigand et al. (2024) stress, digital technologies should be thought of in terms of mathematical activities that they make possible and facilitate. This viewpoint is especially applicable to virtual manipulatives since representations can enable learners to interact with mathematical objects, yet then such actions can only become educational when they are related to mathematical concepts and arguments. The findings of visualisation studies support this distinction: Schoenherr et al. (2024) found that visual representations can support students' mathematics learning, whereas employing technology does not necessarily mean more successful learning. In the same way, the study of elementary students found that virtual manipulatives must have a purposeful role in instruction instead of the role of a virtual replacement of other teaching tools (Siller & Ahmad, 2024).

Measurement reasoning especially concerns dynamic representations as they allow changes in related quantities to be observed simultaneously. Panorkou (2020) showed the ability of dynamic environments to facilitate exploration with varying quantities in measurement scenarios, enabling learners to focus on the relationships that might be challenging to see in static representations. Panorkou (2021) later demonstrated the ability of pupils to work with dynamic measurement reasoning with right prisms and cylinders. Through such environments, pupils can manipulate one quantity to see a related change in another, compare successive states, and think about how quantities change in relation to each other. Other related studies that apply a dynamic measurement scheme also demonstrate how elementary students are able to reason about relationships between quantities through the use of interactive representations (Yeo & Webel, 2024). Collectively, this work implies that the pedagogic promise of dynamic representations does not just consist of getting quantities to move onscreen, but rather in how relationships between changing quantities can be explored and explained mathematically.

The VM-MR Model

Studies of virtual manipulatives have shifted from merely questioning the general effectiveness of digital manipulatives towards investigating the effects of their characteristics and didactic application on mathematical activity. Rich (2021) emphasises the need to focus on the role of integrated virtual manipulatives in the learning process in elementary mathematics, as opposed to addressing the digital tool as the learning itself. Even more recent research involving elementary students also suggests that the learning value of manipulatives hinges on their incorporation into teaching (Siller & Ahmad, 2024). On a higher level, Weigand et al. (2024) highlight the capacity of technological affordances to align with mathematical teaching and learning purposes. This is of particular interest to the field of measurement, in which dynamic environments may make changing quantities available for exploration and reasoning (Panorkou, 2020, 2021). Despite this literature, relatively little theoretical focus has been placed on theory-informed virtual manipulative models that purposefully structure the activity of pupils around relationships among different measurable attributes.

To address this gap, the Virtual Manipulative Model of Measurement Relationships (VM-MR Model) joins three theoretical facets of measurement relationships in a complementary way. The relational conception by Skemp (1976) informs the sense-making focus of the mathematical relationships used in the model rather than using procedures as the main focus of learning. The quantitative reasoning described by Thompson (1993) focuses on the conceptualisation of the quantities and the creation of the relations between them. Another view of changing quantities is offered by Carlson et al. (2002) through covariational reasoning, which focuses on coordinating changes in quantities that vary in relation to one another. In the VM-MR Model, these three views are combined to frame measurement learning in terms of relationships between quantities, and not in terms of individual measurements or arithmetic operations. They thus serve as design foundations of reasoning opportunities in the model and not as independent results to be measured by URMQ Test.

The theoretical principles are transformed into three design principles, dynamic representation, guided exploration and relational explanation. Dynamic representation allows pupils to see how changes in measurable quantities occur along with changes in related quantities, which are based on findings that dynamic measurement situations should make changing quantities accessible to mathematical reasoning (Panorkou, 2020, 2021). Guided exploration helps focus the attention of pupils on relevant quantities, comparisons, and emerging patterns, and this approach is aligned with the importance of linking technological affordances with mathematical

uses and purposes (Rich, 2021; Weigand et al., 2024). Relational explanation takes manipulation a step further towards interpretation by asking pupils to explain the relationships they see, which is in keeping with the focus on meaningful mathematical relationships on which relational understanding is based (Skemp, 1976). Collectively, these principles arrange virtual manipulation in terms of opportunities to observe, compare, coordinate, and explain relationships between measurable quantities. They are implemented through VM Measurement Y6 which offers learning tasks based on the length–mass, length–volume of liquid, and mass–volume of liquid relationships with the possibility of virtual experimentation.

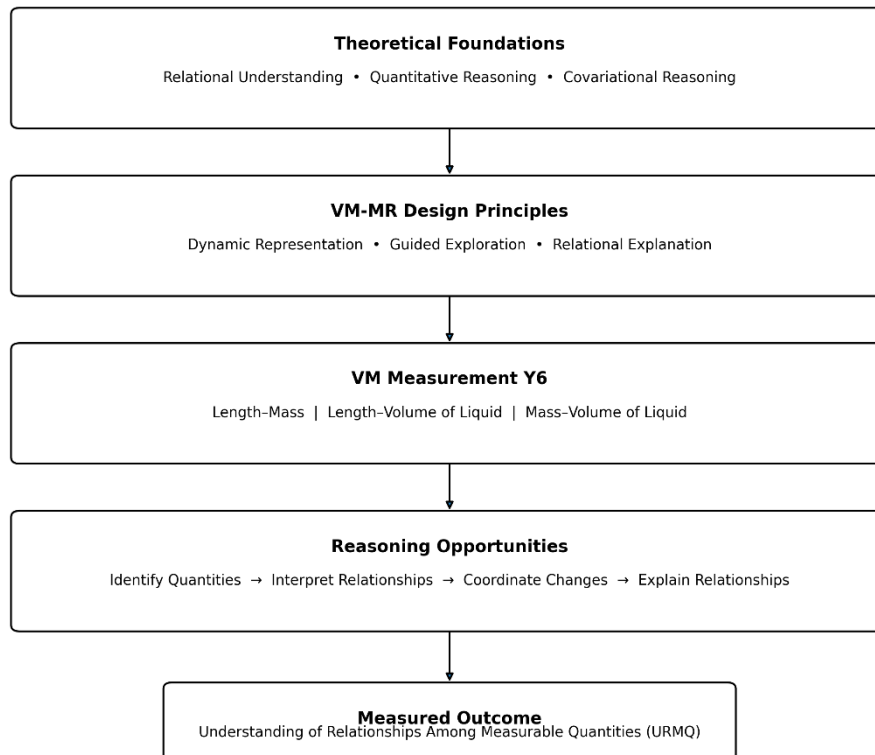


Figure 1: Conceptual Framework of the VM-MR Model

The VM-MR Model conceptualises virtual manipulation as a means of supporting measurement reasoning rather than as digital interactivity alone. Its role is to make a purposeful linkage between what is being manipulated by the pupils and what they are supposed to observe, compare, coordinate and explain. This conceptualisation provides a rationale for exploring the proposition that pupils who go through the VM-MR Model show a different URMQ change pattern from pupils learning the same measurement content through the textbook.

Method

Research Design

In this study, a quasi-experimental comparison-group design was used, which included a pre-test and post-test in order to carry out a preliminary classroom-based evaluation of the VM-MR Model. The participants of the study were two intact Year 6 classes. One group served as the intervention group and studied using the VM-MR Model with VM Measurement Y6, while the other served as the comparison group and studied the same measurement content using the textbook. Pupils were not randomly assigned because the study was conducted within existing classroom settings.

Since instructional condition was also associated with intact class membership, the possible influence of classes could not be completely separated from the instructional condition. The study was thus positioned as a preliminary classroom-based evaluation rather than a randomised efficacy trial. The analyses were aimed at determining whether the patterns of change were different in the two instructional conditions and the results were interpreted within the context of this quasi-experimental design.

The ADDIE framework was also used in the study to inform the structured design and evaluation of the VM-MR Model and digital implementation, VM Measurement Y6. According to Branch (2009), ADDIE is an instructional design process that consists of five stages, which are interrelated: Analysis, Design, Development, Implementation, and Evaluation. The Analysis stage in the current research involved the identification of learning requirements in terms of relationships between measurable quantities. The Design phase formed the theoretical background and the design principles of the VM-MR Model, whereas the Development phase consisted of the creation of VM Measurement Y6. During the Implementation phase, the model and the digital environment were used with the intervention group over the six-week instructional period. Lastly, the Evaluation phase looked at the performance of the pupils in terms of URMQ through pre-test and post-test assessment. The five phases offered a methodological framework for development, implementation and evaluation of the instructional intervention (Branch, 2009). The way the ADDIE phases were operationalised in the current study is summarised in Figure 2.

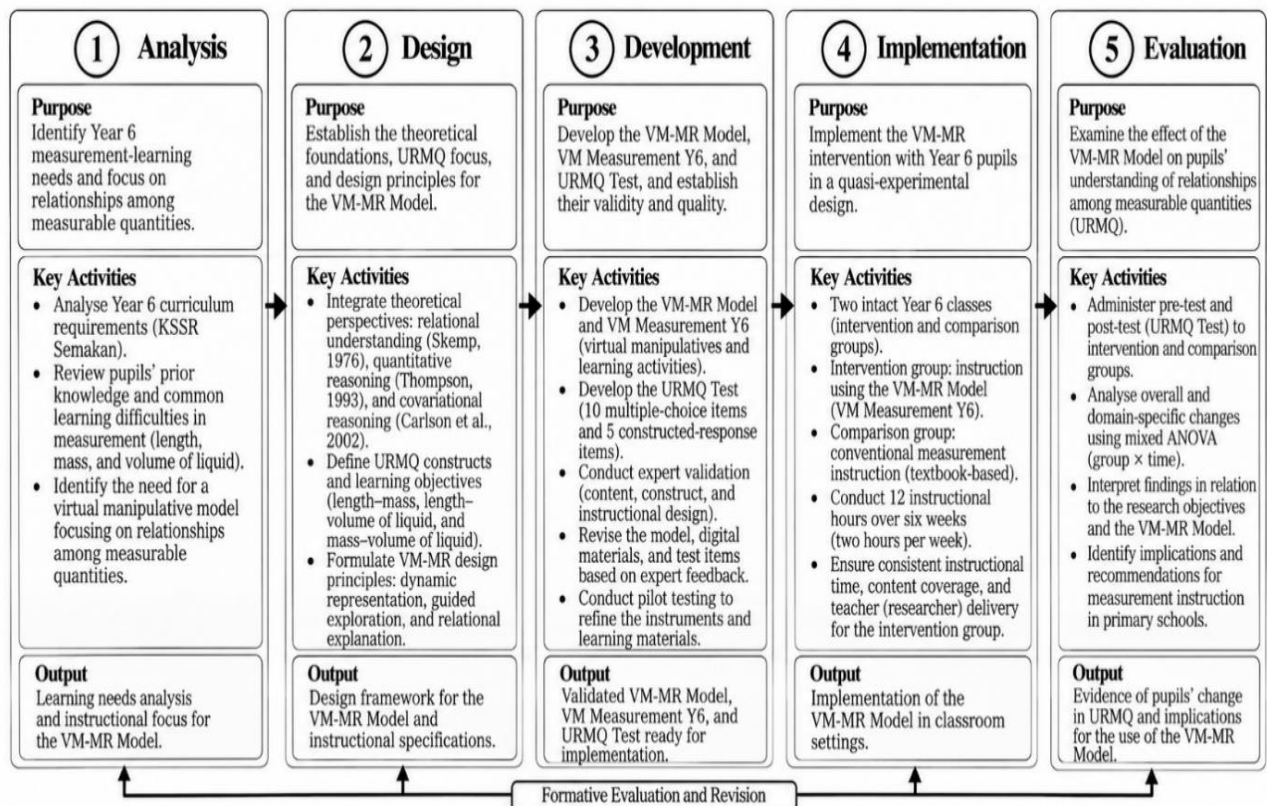


Figure 2: ADDIE-Based Development and Evaluation Process of the VM-MR Model

Participants and Instructional Context

The sample consisted of 60 pupils (aged 12 years) in two intact classes of a Malaysian primary school. The number of pupils in each group was 30. The intervention group ($n = 30$) received instruction using the VM-MR Model with VM Measurement Y6 and the comparison group ($n = 30$) studied the same measurement content using the textbook. The mathematics teacher taught the two groups during the period of instruction.

The instructional period lasted six weeks, with a total of 12 hours of teaching. The measurement content of the two groups was identical with three measurement relationships: length–mass, length–volume of liquid and mass–volume of liquid. The main difference between the instructional conditions was that the intervention group interacted with the content through VM Measurement Y6 as part of the VM-MR Model, while the comparison group worked with the textbook.

VM-MR Intervention

The VM-MR Model was implemented through VM Measurement Y6, a digital environment that was created to be used in the study. The intervention applied the model's three design principles—dynamic representation, guided exploration, and relational explanation—to the classroom learning activities. Instead of virtual manipulation being used to simply model the concepts of measurement, the pupils were able to manipulate measurable quantities, record changes that they observed, compare measurement data, identify patterns of change, and explain how quantities may be related to one another.

VM Measurement Y6 was composed of domain-specific learning modules that covered length–mass, length–volume of liquid, and mass–volume of liquid relationships as well as a Virtual Experiment Module. To illustrate, in the length–volume of liquid module, pupils were able to control the volume of liquid and observe changes in the water level. These activities were expanded in the Virtual Experiment Module, which gave pupils opportunities to make predictions, manipulate quantities, record and compare measurement data, determine patterns of change, and explain the relationships they discovered. Thus, the digital environment realised the VM-MR design principles through activities that linked virtual manipulation to recognising, interpreting, coordinating, and explaining relationships among measurable quantities.

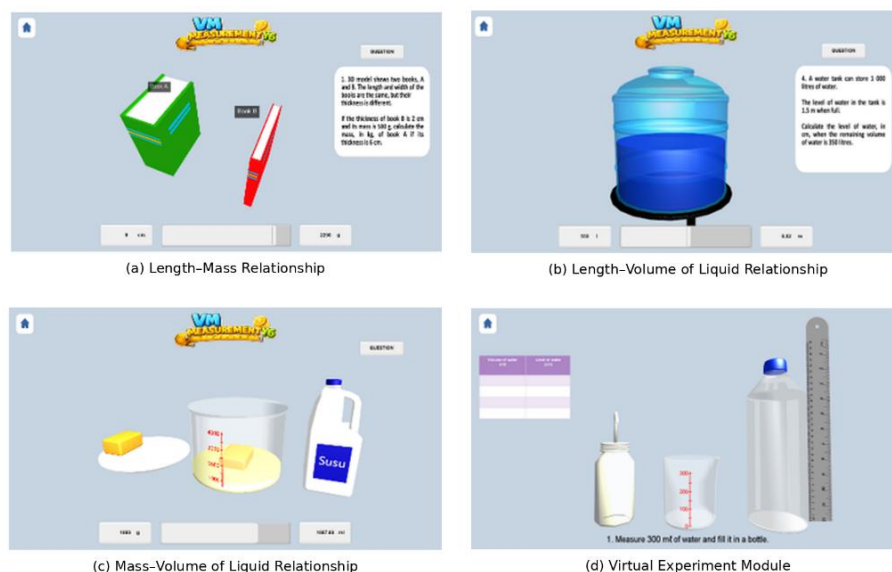
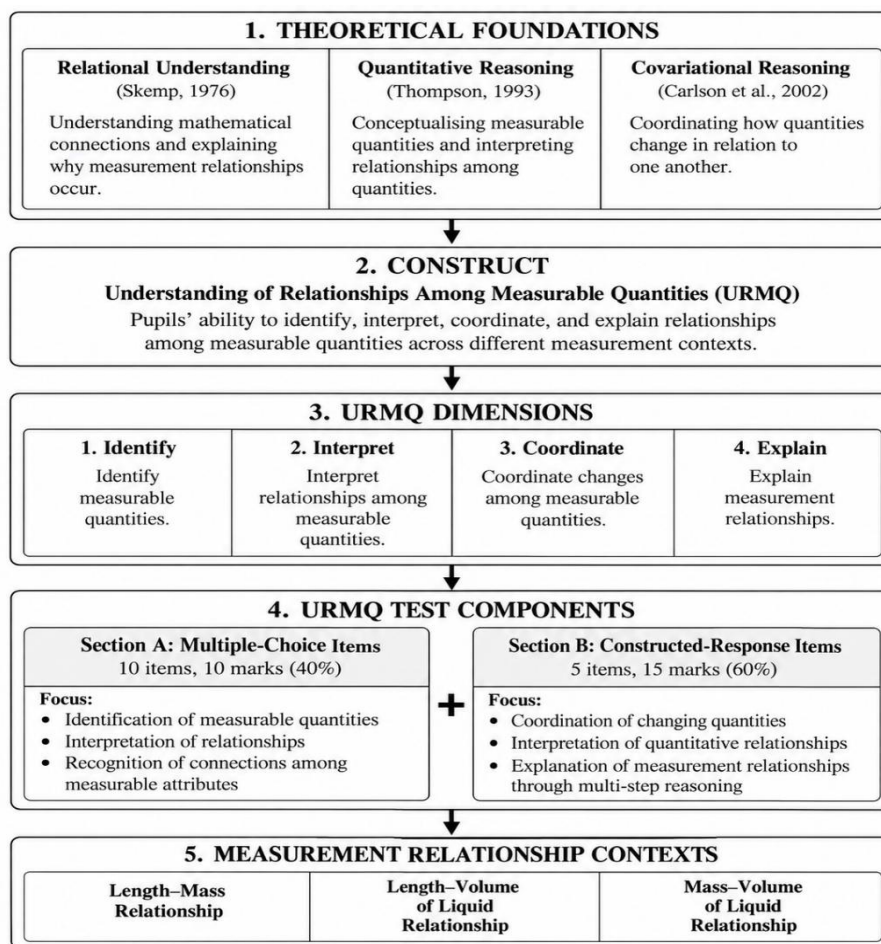


Figure 3: Implementation of the VM-MR Model through VM Measurement Y6

Research Instrument

The URMQ Test was designed to measure pupils' Understanding of Relationships Among Measurable Quantities (URMQ) and to operationalise the construct examined in this study. In line with the conceptualisation of URMQ, the test assessed four related abilities: identifying measurable quantities, interpreting relationships between quantities, coordinating changes between related quantities, and explaining measurement relationships. These skills were explored under the three measurement-relationship situations that the VM-MR Model tackles: length–mass, length–volume of liquid, and mass–volume of liquid.

The test consisted of 15 questions with a total of 25 marks and was administered within a 60-minute period. The items in Section A were 10 multiple-choice questions worth 10 marks (40%), which focused on the identification and interpretation of measurable quantities and their relationships. Section B comprised five constructed-response items with 15 marks (60%) which allowed the pupils to coordinate changes between related quantities, interpret quantitative relationships, and explain measurement relationships by way of multi-step arguments. In this way, both types of items were used in combination to reflect the variety of abilities represented by the operational definition of URMQ instead of considering URMQ as performance on a single type of measurement task. The alignment between the theoretical foundations, URMQ construct, the assessed dimensions, the test components, and measurement-relationship contexts is summarised in Figure 4.



Notes: MCQ = Multiple-choice items; CR = Constructed-response items.

Figure 4: Operationalisation of Understanding of Relationships Among Measurable Quantities (URMQ) and Alignment of the URMQ Test

Three mathematics education experts assessed the content validity, which led to the overall Content Validity Ratio (CVR) of .92. The instrument was then piloted using a separate sample of 30 Year 6 pupils. The pilot study produced a Cronbach's alpha coefficient of .86, indicating satisfactory internal consistency of the URMQ Test items in the pilot sample.

Data Collection and Analysis

Both groups went through the URMQ Test prior to the instructional period and after the six-week intervention. The intervention and comparison groups were administered the URMQ Test using the same procedures. The URMQ Test scoring system was used to score pupils' responses, and the scores were converted into percentages for analysis.

To analyse the pre-test and post-test performance of both groups, descriptive statistics were first calculated. Then, a two-way mixed analysis of variance (mixed ANOVA) was conducted. Time (pre-test and post-test) was considered as the within-subject factor and Group (VM-MR and comparison) as the between-subject factor. The Time $\times$ Group interaction was used to determine whether the pattern of change differed between the two instructional groups. The three measurement-relationship domains were also analysed using separate mixed ANOVAs to help answer the second research question. Statistical significance was set at $p < .05$, and partial eta squared (η^2) was reported as a measure of effect size.

Interpretation of the mixed ANOVA results was preceded by preliminary assumption checks. The results of Shapiro-Wilk tests showed that the intervention group had no significant deviation at pre-test ($W = .944, p = .115$) but the comparison group had a significant deviation ($W = .908, p = .013$). Significant deviations from normality at post-test were seen in both the intervention ($W = .902, p = .009$) and comparison ($W = .881, p = .003$) groups. Levene's test indicated no significant evidence of unequal variances at either pre-test, $F(1, 58) = 0.65, p = .423$, or post-test, $F(1, 58) = 3.66, p = .061$. The mixed ANOVA was retained as the main analysis, because the design was balanced ($n = 30$ per group), and the results could be interpreted cautiously as there were some violations of normality. Because the within-subject factor (Time) had only two levels, the sphericity assumption was not applicable.

As a sensitivity analysis, between-group differences in gain scores were also analysed with Welch's t-test and Mann-Whitney U test to determine whether the overall trend was parallel to the main Time $\times$ Group analysis.

Results and Discussion

Overall URMQ Performance

Table 1 shows the descriptive statistics of the overall Understanding of Relationships Among Measurable Quantities (URMQ) performance of pupils at pre-test and post-test. The mean scores of both groups were higher after instruction. The intervention group increased from 31.33 (SD = 7.58) at pre-test to 74.53 (SD = 6.77) at post-test, which is a descriptive gain of 43.20 percentage points. The comparison group increased from 29.33 (SD = 8.29) at pre-test to 61.07 (SD = 9.27) at post-test, with a descriptive increase of 31.73 percentage points.

Table 1: Descriptive Statistics of URMQ Test Percentage Scores

Group	n	Pre-test M (%)	Pre-test SD	Post-test M (%)	Post-test SD	Gain
Intervention (VM-MR)	30	31.33	7.58	74.53	6.77	43.20
Comparison (Textbook)	30	29.33	8.29	61.07	9.27	31.73

The overall URMQ performance of intervention and comparison groups did not differ significantly at pre-test, $t(58) = 0.97$, $p = .334$. Therefore, the overall URMQ performance did not show any statistically significant difference between the groups at baseline. Nonetheless, since the researchers did not randomly assign individual pupils, but instead used intact classes, group equivalence cannot be established based on this finding alone.

Figure 5 shows the changes in the mean URMQ scores across the two measurement occasions. Both groups experienced improvements between the pre-test and post-test, although the descriptive increase was greater in the VM-MR intervention group. The divergent lines offer a visual depiction of the different patterns of change which are then analysed according to the Time $\times$ Group interaction.

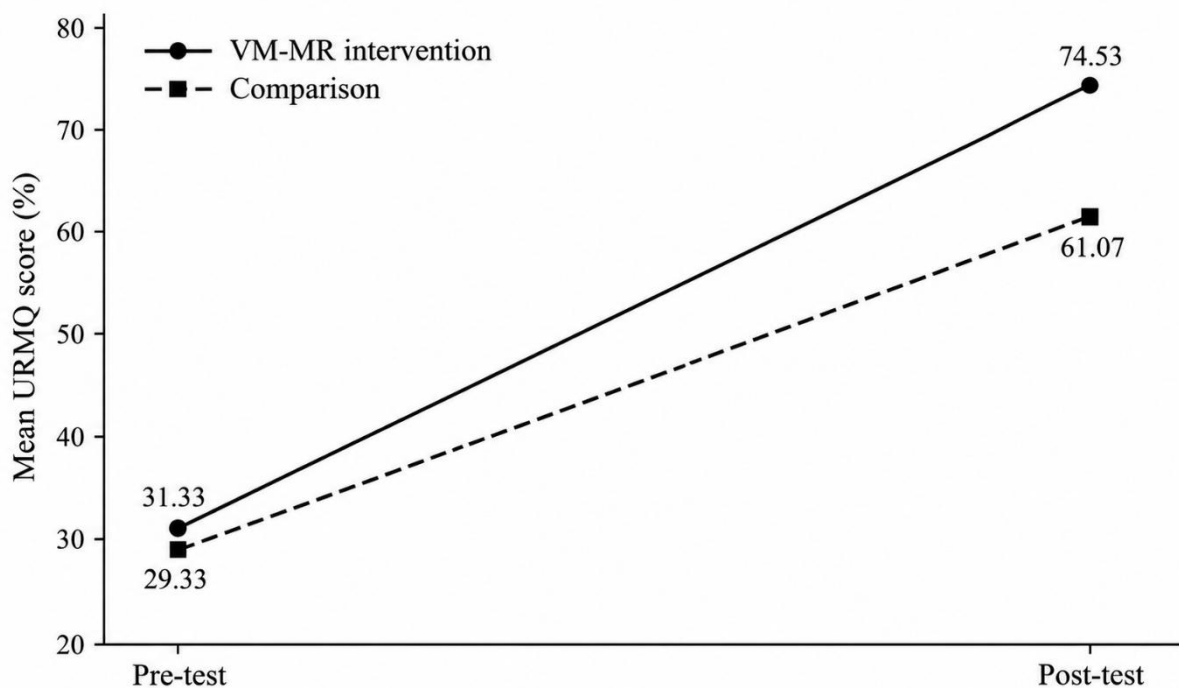


Figure 5: Mean URMQ Percentage Scores at Pre-test and Post-test by Instructional Group

The mixed ANOVA indicated a significant main effect of Time, $F(1, 58) = 754.49$, $p < .001$, $\eta^2 = .929$, indicating an overall change in URMQ performance between the two measurement occasions. A significant main effect of Group was also observed, $F(1, 58) = 4.10$, $p = .047$, $\eta^2 = .066$. In relation to the first research question, the Time $\times$ Group interaction was also significant, $F(1, 58) = 17.67$, $p < .001$, $\eta^2 = .233$. This interaction indicates that the pattern of change in URMQ performance between the pre-test and post-test was significantly different between the two instructional groups.

Table 2: Mixed ANOVA Results for Overall URMQ Performance

Effect	F	df	p	η^2
Time	754.49	1, 58	< .001	.929
Group	4.10	1, 58	.047	.066
Time $\times$ Group	17.67	1, 58	< .001	.233

The descriptive results should be understood along with the significant interaction. Pupils in both groups showed improvement from pre-test to post-test; hence, the results do not indicate that improvement occurred only in the VM-MR group. Instead, the descriptive increase of 43.20 percentage points in the intervention group was higher than the descriptive increase of 31.73 percentage points in the comparison group and the significant Time $\times$ Group interaction shows the difference in patterns of improvement between the two instructional conditions.

Since the preliminary assumption checks indicated departures from normality, sensitivity analyses were performed on the gain scores. Welch's t-test showed a significant between-group difference in gain scores, $t(57.78) = 4.20$, $p < .001$. The Mann-Whitney U test yielded the same inferential conclusion, $U = 700.50$, $p < .001$. The consistency of these sensitivity analyses with the substantial Time $\times$ Group interaction offers further evidence of the overall pattern observed in the initial analysis.

The difference in the pattern of improvement can be observed in relation to the instructional design of the VM-MR Model. VM Measurement Y6 was not created only to provide measurement content digitally but to involve pupils in manipulating measurable quantities, observing changes, comparing quantities, identifying patterns, and explaining relationships. Such an orientation reflects Skemp's (1976) distinction between procedural knowledge and the understanding of mathematical relationships, with the VM-MR activities emphasising the latter.

The tasks also gave pupils opportunities to treat measurable quantities as attributes rather than as isolated numbers. This sort of activity is in line with the description of quantitative reasoning proposed by Thompson (1993), which focuses on the conceptualisation of quantities and relationships among them. Similarly, activities in which pupils viewed and compared concurrent variations in measurable quantities offered opportunities for reasoning consistent with covariational perspectives (Carlson et al., 2002). These conceptual relationships explain the reasoning opportunities inherent in the VM-MR design; the URMQ outcome information does not in itself indicate that pupils acquired each of these types of reasoning as individual constructs.

Performance Across Measurement-Relationship Domains

The second research question examined whether there was a difference in the pattern of change between the instructional groups across the three measurement-relationship domains measured by the URMQ Test. Separate mixed ANOVAs showed significant Time $\times$ Group interactions for length-mass, $F(1, 58) = 6.72$, $p = .012$, $\eta^2 = .104$; length-volume of liquid, $F(1, 58) = 9.68$, $p = .003$, $\eta^2 = .143$; and mass-volume of liquid, $F(1, 58) = 5.01$, $p = .029$, $\eta^2 = .079$. The differences in the patterns of change were therefore evident between the instructional groups across all three measurement-relationship domains.

Table 3: Time × Group Interaction Effects Across URMQ Domains

Measurement Relationship Domain	F	df	p	η^2
Length–Mass	6.72	1, 58	.012	.104
Length–Volume of Liquid	9.68	1, 58	.003	.143
Mass–Volume of Liquid	5.01	1, 58	.029	.079

The domain-level results indicate that the differences in patterns of improvement noted were not limited to a single measurement context. This is consistent with the design of the VM-MR Model, as VM Measurement Y6 structured activities around the relationships among quantities in all three domains as opposed to a single isolated measurement relationship. Pupils were given a chance to identify measurable quantities, manipulate or observe changes, record and compare measurement information, find patterns, and develop explanations of the relationships they observed.

Numerically, the largest Time × Group interaction among the three domains was observed for the length–volume of liquid relationship ($\eta^2 = .143$). A potential reason is related to the representational nature of this VM-MR module. The length–volume of liquid activities allowed the pupils to notice variations in the volume of liquid and relate them to changes in the related length measure, particularly the water-level height. As such, two connected quantities could be displayed simultaneously in the digital space.

Dynamic forms of these representations can be especially relevant when pupils need to observe how quantities change in relation to each other, a meaning that aligns with previous studies of dynamic representations of measurement and reasoning about changing quantities (Panorkou, 2020, 2021). Nevertheless, the current research did not gather process information that would explain the numerically larger interaction in this domain. There was not a statistical comparison of the interaction effects in the three domains. The difference observed should instead be viewed as a descriptive trend in the domain-level findings as opposed to the idea that the VM-MR Model was statistically more successful in one area than the other or that a specific cognitive process was responsible for the observed pattern.

Combined, the significant interactions within the three domains suggest that variations in patterns of change between the instructional conditions were apparent through a variety of types of measurable relationships. This observation matches the idea behind creating virtual manipulatives based on clear mathematical connections, as opposed to using digital manipulation as a goal by itself. Other studies have also stressed the same idea that the learning value of virtual manipulatives is not necessarily based on their digital or interactive characteristics but is related to the relation between their representational affordances and mathematical practice and learning (Moyer-Packenham & Suh, 2012; Weigand et al., 2024).

Implications for Virtual Manipulative Design and Measurement Learning

Combined, the results can be used to re-evaluate the concept of virtual manipulatives in primary measurement learning. The value of the VM-MR Model is not merely to substitute physical or static resources with virtual ones but rather to structure the virtual manipulation by making explicit opportunities to engage in reasoning about measurements. In the three domains of measurement-relationship, pupils interacted with quantities by observing or manipulating quantities, attuning to the attendant changes, comparing, detecting relationships, and giving explanations. In this regard, manipulation was used as a tool for making relationships between measurable quantities accessible through mathematical action rather than as an educational goal in itself.

This is a crucial difference since interactivity in itself is not mathematical reasoning. The VM-MR Model relates specific types of digital activity to what learners should observe, compare, coordinate, and explain. The observed differences in patterns of URMQ change are consistent with the rationale for structuring virtual manipulation around these reasoning opportunities. Nevertheless, the data do not determine which particular design feature or reasoning process contributed to those differences. This contribution should therefore be interpreted as a preliminary classroom-based evaluation of a theory-based virtual manipulative model rather than as validation of a particular cognitive process.

In the case of measurement teaching, this view expands the teaching emphasis to include more than measuring, calculating, and converting units. These processes are still significant, but learning to measure also entails the ability to identify how measurable quantities relate to each other, and how a change in one quantity relates to a change in another. The three domains considered in this study illustrate how virtual environments can be intentionally designed to allow such relationships to be explored, compared and explained. This indicates a design principle that is not as specific to the digital context being studied here: the relevance of technological features is to be organised in terms of the mathematical relationships and reasoning opportunities that pupils should engage with.

The results must, however, be interpreted within the confines of the research design. Sixty pupils in two intact Year 6 classes participated in one Malaysian public primary school and individual pupils were not randomly assigned to conditions of instruction. Further, because instructional condition corresponded with intact class membership, potential class-level influences could not be completely separated from the instructional condition. Though the two groups had the same mathematics instructor, the same measurement content and the same amount of time (12 hours) of instruction over six weeks, the potential influence of classroom context cannot be ruled out. The research thus presents evidence of different patterns of URMQ change between the two instructional conditions instead of a conclusive causal estimate of the impact of the VM-MR Model.

The model can be further tested in future research using larger and more varied samples across several schools, classrooms and teachers. Such process-oriented evidence as pupils' verbal descriptions, written arguments, interaction logs and screen-recorded activities might also explain how pupils interact and organise measurable quantities when working in the virtual environment. This evidence would enable future research to study more directly the relationship between specific VM-MR design characteristics and reasoning processes in pupils and whether the reasoning opportunities inherent in the model are manifested in observable changes in quantitative and covariational reasoning.

Conclusion

This research formulated and implemented a preliminary classroom-based evaluation of the VM-MR Model, which is a theory-based program aimed at helping Year 6 pupils with their Understanding of Relationships Among Measurable Quantities (URMQ). Both instructional groups showed improvement in URMQ performance following 12 hours of instruction over six weeks. Nonetheless, the Time $\times$ Group interaction was significant, indicating that the pattern of improvement differed between the two groups, with a larger descriptive increase in the VM-MR group. Significant Time $\times$ Group interactions were also observed across all three measurement-relationship domains: length–mass, length–volume of liquid, and mass–volume of liquid.

The primary contribution of the research is that virtual manipulation is carefully structured around the concepts of measurement reasoning. Using VM Measurement Y6, pupils had a chance to manipulate measurable quantities, observe changes, compare measurement information, identify relationships, and provide mathematical explanations. The results thus point to the importance of designing virtual manipulative environments around explicit mathematical reasoning objectives rather than interactivity alone.

However, the results should be interpreted within the context of a quasi-experimental study involving one school and two intact classes. The research provides evidence of different patterns of URMQ improvement between the two instructional conditions but does not identify the specific reasoning processes responsible for these differences. Further studies using larger and more heterogeneous samples across multiple schools and teachers, together with process-related information such as pupils' explanations, interviews, and screen-recorded interactions, may explore how specific aspects of virtual manipulation relate to pupils' reasoning about relationships between measurable quantities.

Acknowledgements

The authors are thankful to the school, mathematics teacher, pupils, parents or guardians, mathematics education experts, and Universiti Teknologi Malaysia for their support in the development, validation, implementation, and preparation of this study.

References

- Branch, R. M. (2009). *Instructional design: The ADDIE approach*. Springer. <https://doi.org/10.1007/978-0-387-09506-6>
- Byrne, E. M., Jensen, H., Thomsen, B. S., & Ramchandani, P. G. (2023). Educational interventions involving physical manipulatives for improving children's learning and development: A scoping review. *Review of Education*, 11(2), Article e3400. <https://doi.org/10.1002/rev3.3400>
- Carlson, M., Jacobs, S., Coe, E., Larsen, S., & Hsu, E. (2002). Applying covariational reasoning while modeling dynamic events: A framework and a study. *Journal for Research in Mathematics Education*, 33(5), 352–378. <https://doi.org/10.2307/4149958>
- Moyer-Packenham, P. S., & Suh, J. M. (2012). Learning mathematics with technology: The influence of virtual manipulatives on different achievement groups. *Journal of Computers in Mathematics and Science Teaching*, 31(1), 39–59.
- Ochugboju, A. O., & Díez-Palomar, J. (2025). Systematic review on the use of virtual and physical manipulatives by primary school teachers. *Social Sciences & Humanities Open*, 12, Article 102188. <https://doi.org/10.1016/j.ssaho.2025.102188>
- Panorkou, N. (2020). Dynamic measurement reasoning for area and volume. *For the Learning of Mathematics*, 40(3), 9–13.
- Panorkou, N. (2021). Exploring students' dynamic measurement reasoning about right prisms and cylinders. *Cognition and Instruction*, 39(4), 477–511. <https://doi.org/10.1080/07370008.2021.1958218>
- Park, J., & Yeo, S. (2023). Elementary school students' levels of quantitative reasoning of units: Using open number line tasks. *The Mathematical Education*, 62(4), 457–471. <https://doi.org/10.7468/mathedu.2023.62.4.457>
- Reimer, K., & Moyer, P. S. (2005). Third graders learn about fractions using virtual manipulatives: A classroom study. *Journal of Computers in Mathematics and Science Teaching*, 24(1), 5–25.

- Rich, K. M. (2021). Virtual manipulatives in elementary mathematics: A critical integrative review of research. *Journal of Computers in Mathematics and Science Teaching*, 40(2), 163–177.
- Schoenherr, J., Strohmaier, A. R., & Schukajlow, S. (2024). Learning with visualizations helps: A meta-analysis of visualization interventions in mathematics education. *Educational Research Review*, 45, Article 100639. <https://doi.org/10.1016/j.edurev.2024.100639>
- Siller, H.-S., & Ahmad, S. (2024). The effect of concrete and virtual manipulative blended instruction on mathematical achievement for elementary school students. *Canadian Journal of Science, Mathematics and Technology Education*, 24, 229–266. <https://doi.org/10.1007/s42330-024-00336-y>
- Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching*, 77, 20–26.
- Thompson, P. W. (1993). Quantitative reasoning, complexity, and additive structures. *Educational Studies in Mathematics*, 25(3), 165–208. <https://doi.org/10.1007/BF01273861>
- Weigand, H.-G., Trgalova, J., & Tabach, M. (2024). Mathematics teaching, learning, and assessment in the digital age. *ZDM–Mathematics Education*, 56, 525–541. <https://doi.org/10.1007/s11858-024-01612-9>
- Yeo, S., & Webel, C. (2024). Elementary students' fraction reasoning: A measurement approach to fractions in a dynamic environment. *Mathematical Thinking and Learning*, 26(1), 20–46. <https://doi.org/10.1080/10986065.2022.2025639>