MODESTUM

EURASIA Journal of Mathematics, Science and Technology Education, 2022, 18(6), em2119
ISSN:1305-8223 (online)

OPEN ACCESS

Review Article

https://doi.org/10.29333/ejmste/12087

Structure of the Observed Learning Outcomes (SOLO) model: A mixed-method systematic review of research in mathematics education

Saidat Morenike Adeniji [1,2*], Penelope Baker [1], Martin Schmude [1]

[1] University of New England, Armidale, AUSTRALIA
[2] University of Ilorin, Ilorin, NIGERIA

Received 4 March 2022 ▪ Accepted 10 May 2022

Abstract

The review followed the preferred reporting items for systematic review and meta-analysis (PRISMA) standard to search and report relevant articles on the use of SOLO model in mathematics. A systematic search was conducted in Education Source, ERIC, JSTOR, and PsycINFO databases and yielded 198 studies. After screening and appraisal, 62 papers (37 qualitative, 17 quantitative, and eight mixed-method studies) published in English between 1990 and 2020 were reported using a narrative synthesis. The findings indicated that SOLO model appropriately reflects students' learning outcomes; there is a direct relationship between students' performances and their SOLO levels; and SOLO model could explain several other developmental theories and contribute to the development of mathematics curricula. These findings highlight the gaps between theory and practice of this model in mathematics education, and informs education professionals about the diverse applications of the SOLO model for improving mathematics teaching and learning, fair assessment, and curriculum development.

Keywords: mathematics assessment, mathematics teaching & learning, SOLO model, systematic review

INTRODUCTION

Assessment is an essential part of the teaching and learning processes because it provides information on students' current understanding of a concept and identifies whether the goals of teaching and learning are being achieved (Tosuncuoglu, 2018). Mathematics assessment aims to be rigorous and has undergone several reforms. For example, in the United Kingdom, the Assessment Reform Group (2002) considers that assessment should not only gather information about students' achievements but also inform ongoing teaching and learning. This position emphasizes 'assessment for learning' and 'assessment as learning' against the traditional and common dimension of 'assessment of learning' (Serow et al., 2019). Under this reform approach, it is necessary to incorporate assessing the quality of learning ('how well') into the traditional assessment practices of quantity of learning ('how much'). With this approach, students' learning assessment would provide information on what students know and how well they know it in order to guide future learning (Pegg & Panizzon, 2007).

Biggs and Collis (1982) argued that traditional quantitative assessment which consists of marks or scores does not provide sufficient information on the learning growth of students to inform current and future learning. They suggested that the process of learning moves from quantitative change–superficial learning, to qualitative change–deep learning. To address this view, they developed the structure of the observed learning outcome (SOLO) model, which provides a taxonomy for assessing students' cognitive growth both quantitatively and qualitatively. With this approach, rather than simply testing students' capacity to memorize information, teachers could assess students' intellectual ability which involves reasoning, rate of learning, abstract thinking, and making connections between ideas (Lucander et al., 2010). This taxonomy analyzes the structure of students' responses to a task to assess students' learning from the perspective of levels of understanding. The SOLO model has been reported to not only improve assessment of educational outcomes but also promote students'

✉ sadeniji@myune.edu.au **(*Correspondence)** ✉ pep.baker@une.edu.au ✉ mschmud2@une.edu.au

Contribution to the literature

- In a quest for a fair assessment of mathematics learning, this study followed the PRISMA guidelines to examine the contributions of SOLO model to mathematics education.
- This review is significant for its methodological approach and the coverage of articles published within 30 years (1990-2020) out of the 38 years of existence of SOLO model.
- The findings from this review highlighted SOLO model as a sound cognitive model that has contributed immensely, not only to fair mathematics assessment around the world, but also to mathematics pedagogy and mathematics curriculum.

Table 1. SOLO modes

SOLO modes	Description
Sensorimotor (from birth)	This mode of functioning helps in acquisition & development of fine motor skills (tacit knowledge)–knowing how to complete a physical task such as counting, sorting, making shapes, & building numbers. Sensorimotor mode extends beyond childhood, & an individual's fine motor skills continue to develop throughout lifespan.
Ikonic (from 18 months)	In this mode, children acquire intuitive knowledge, where they can use images or words to represent events, objects, or things. They link images to language, which helps in developing mathematical vocabulary. Form of knowledge they acquire within this mode is subjective because they have no reason, inference, or experience to justify their representations. Examples are naming & defining different shapes, & comparing, ordering, & matching objects by length.
Concrete symbolic (from 6 years)	Usually, from age of six years, a child applies a symbol system to express his or her thoughts. Use of symbols such as written language & number system increases their experiences of empirical world, which helps to build abstraction such as demonstrating rules for mathematical equations & facts. School curriculum focuses mainly on this mode & it is often referred to as target mode. Knowledge in this mode is declarative.
Formal (from 16 years)	Response in this mode involves abstract concepts such as principles & theories. There is no restriction or strong tie to refer to empirical or concrete world, thereby creating a range of possibilities to explore such as manipulating a theoretical construct.
Post formal (from 20 years)	Abstract concepts identified in previous mode are questioned, queried, & challenged. Individuals within this mode perceive real & likely possible things, which helps them to operate using theoretical knowledge.

learning (Hattie & Brown, 2004). Despite this, the SOLO taxonomy has not been widely used in the classrooms of many countries around the world after nearly four decades of its existence (Hook, 2016). Also, there is no study, either empirical or theoretical, to indicate the impact of the SOLO model on mathematics education. Hence, this study critically synthesizes evidence from the literature to determine the impact of the SOLO model on mathematics education and understand the state of the research evidence relating to the model. It is expected that the evidence from this review will incorporate diverse study types to identify knowledge gaps in the research. The results of this review have the potential to inform current and future research and practice in mathematics education. A mixed-method systematic review was undertaken and a description of the SOLO model is discussed next.

DESCRIPTION OF THE SOLO MODEL

The SOLO cognitive and developmental model was developed by Biggs and Collis (1982) as a response to the inadequacies identified in a comprehensive analysis of Piaget's theory. They realized that despite what Piaget's stage theory claims, students' responses are not consistent within each stage and that Piaget's assumption of staged theory does not hold for several subject areas. Rather than focusing on individual cognitive developmental stages as Piaget's model does, the SOLO model modified the four developmental stages (sensorimotor, preoperational, concrete operational, and formal operational) into modes and added one new mode (post-formal), as a way to overcome the inadequacies of Piaget's theory (Biggs & Collis, 1991; Pegg & Tall, 2005). Also, the SOLO model advocated for multi-modal functioning against the complete exit of each mode suggested by Piaget. This means that an individual could function in any of the previously attained modes if required. Hence, the SOLO model further developed Piaget's cognitive developmental model. Like other developmental theorists, Biggs and Collis (1991) also focused on how students' understanding grows (Haynes, 2009), and described students' understanding of concepts in terms of observable learning outcomes (Hattie & Brown, 2004). Thus, SOLO model explains the increasing complexity of understanding and provides a taxonomy for assessing students' cognitive learning outcomes. The SOLO model comprises two elements: the mode of functioning and levels of attainment (Serow et al., 2019). The five levels of SOLO attainment are replicated in each mode of functioning. These two components are discussed next.

Modes of Functioning

There are five modes of functioning in the SOLO taxonomy: sensorimotor, ikonic, concrete symbolic, formal, and post formal. These modes are identified by

Table 2. SOLO modes

SOLO levels	Description
Pre-structural	This level is below intended mode required to complete a task & response requires no or a minimal amount of working memory. Operations at this level are demonstrated by irrelevant responses to task such as tautology, declining to answer, insufficient understanding, or possibly distracted by irrelevant aspects of task. There is no consistency at this stage because learner might not understand problem.
Uni-structural	This level is the first of aspired level, where response indicates a piece of relevant idea, fact, or information about task. Demand for working memory is minimal & there is no meaningful conclusion between single aspect identified in response & whole picture. Thus, responses at this level are inconsistent. But, amount of used working memory has increased over pre-structural level.
Multi-structural	In this second target level, students' responses focus on several relevant but unrelated & fragmented ideas that indicate a lack of full picture. Each idea is considered independently because connections between ideas are not given, which makes response at this level prone to inconsistencies. Students' understanding at uni-structural level has improved & each idea acquired at uni-structural level is over-learned, aiding automatic retrieval of information or ideas, which facilitates ability to focus on multiple ideas. Amount of working memory used for this level of response is medium.
Relational	Response at this level is identified by focusing on how several relevant ideas are connected. This third level of response reflects generalizing ideas within experienced or given contexts using related aspects. At this level, students demonstrate understanding by providing multiple ideas & integrating task components into a coherent whole. Working memory considers individual ideas as well as interrelationship between ideas, which prompts students to investigate reasons for connections between ideas. Hence, amount of working memory used is high because this level witnesses a qualitative shift from concrete & superficial learning to abstract & deeper understanding. Since there is a coherent connection between ideas, response is consistent within specified system. Response at this level represents higher-order thinking.
Extended abstract	This level is above target mode. At this level, students can generalize their experiences to new situations. Responses are characterized by concise & relevant data, multiple & coherent interconnectedness, predicting situations, & testing possible hypotheses. As a result, amount of working memory used at this level is significant because it may sometimes involve clear construction of knowledge. At this level, students' responses are beyond expectation because they use related prior knowledge & experience to logically conclude outside scope of what was learned. They are proficient at hypothesizing, theorizing, criticizing, & generalizing a specific circumstance to a higher level of abstraction.

the levels of abstraction in the responses of individuals (Bhagwat, 2017; Haynes, 2009). The five learning modes proposed by Biggs and Collis (1982) are shown in **Table 1**. While the ages shown for each mode are a guide to when the modes function, individuals may reach these modes of functioning at different ages.

The SOLO modes of functioning are not fixed or invariant. This means a newly acquired mode does not subsume the previous modes, as they can come into play when needed. Individuals can therefore use several modes at a time for a specific task and each mode continues to develop (Pegg, 1992). The SOLO model explains how students' responses to tasks increase in structure, complexity, and abstraction, showing the progress of understanding from being an incompetent learner to an expert. When students understand the SOLO model, they have ownership of their learning and feel more confident in their learning journey, which helps them to become masters of their knowledge (Hook, 2016). Rather than relying on the teachers' feedback, they can assess their own progress in learning and determine what to learn next.

Levels of the SOLO Model

The SOLO model describes students' responses using five hierarchical levels of increasing complexity. The levels are pre-structural, uni-structural, multi-structural, relational, and extended abstract (Biggs & Collis, 1982). The level of response is identified by analyzing the components and structure of the response. These levels are reported based on the amount of working memory used, relevant operations, consistency, and closure of learning outcomes (Biggs & Collis, 1982; Hattie & Brown, 2004). Pegg (1992) and Serow et al. (2019) have described the levels as presented in **Table 2**.

The prototypical competencies for individual levels of learning outcomes are shown in the **Figure 1**.

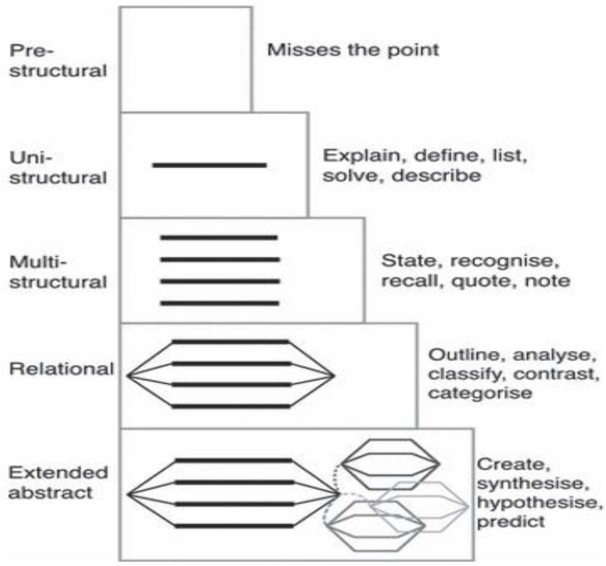

Figure 1. Levels of SOLO model (Adapted from Lucander et al., 2010)

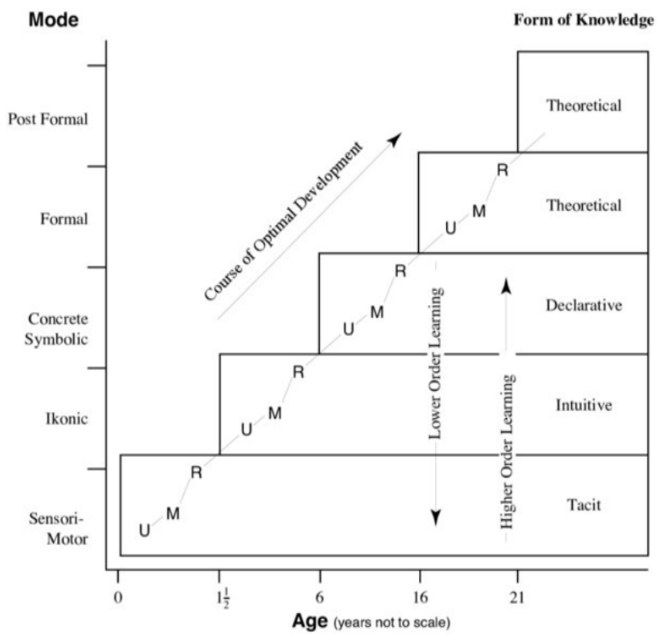

Figure 2. SOLO learning cycle (Adapted from Biggs & Collis, 1991)

While the SOLO modes offer a sequential abstraction to students' responses, the SOLO levels provide an understanding of the hierarchical complexity in the structure of students' responses. However, this does not imply that the SOLO model commonly follows a linear progression; it often results in iterative and spiral cycles because the levels are repeated within each mode of functioning. Moreover, students' SOLO level in each topic differs–while some topics may be easy to understand, others require additional efforts to reach the multi-structural and relational levels. It is essential to note that only uni-structural, multi-structural, and relational levels fall within the target mode (Lucander et al., 2010). The SOLO modes and their learning cycles are illustrated in **Figure 2**.

The quantitative change (lower-order learning) in understanding is observed within the pre-structural, uni-structural, and multi-structural levels, while the qualitative change (higher-order learning) exists at the relational and extended abstract levels. In the same vein, relational responses beyond the step-by-step algorithms of mathematical procedure commence at the relational level (Pegg, 2010). A simple analogy of the series of SOLO levels is that the pre-structural level misses the point, the uni-structural and multi-structural levels achieve the aim of learning, the relational level achieves the learning with merit, and the extended abstract level achieves learning with excellence.

When compared to other educational taxonomies (such as Bloom's (1956) taxonomy), this potentially useful metacognitive tool has the advantage of generalization: it does not depend on the content and applies to different abilities of students. Of key importance is the contribution of the SOLO model to assessment of students' mathematical understanding. In line with this, Chick (1998) stated that the SOLO framework is comprehensive and has easy assessment criteria, and the SOLO model has been used across a range of research studies to classify and measure students' cognitive abilities (Chan et al., 2002). The following subsection presents a sample of a task analyzed using SOLO levels.

An example of a task analyzed using the SOLO framework

Students' levels of understanding can be evaluated by considering the structural complexities of their responses to task. Below is a problem relating to simultaneous linear equation and the characteristics of students' responses are presented in **Table 3**.

Question: Find the values of x and y that satisfy these equations: $2x + y = 3$ and $x - y = 0$.

Upon placement of students' responses into the SOLO levels, a value can be assigned to each level: pre-structural=0, uni-structural=1, multi-structural=2, relational=3, and extended abstract=4. These values can then be used to calculate the aggregate score for a test.

METHODOLOGY

The review was conducted to identify, synthesize, and critically appraise available evidence on the SOLO

Table 3. Description of students' responses according to the SOLO levels

SOLO category description	Response
Pre-structural (misses the point)	-No response. -x & y are not known.
Uni-structural (one relevant aspect)	-Identifies this problem as simultaneous equations i.e., values of x & y in two equations must be equal.
Multi-structural (more than one relevant independent aspect)	-Identifies equations as simultaneous equations, names each of equations, & selects a method of solving equations but could not arrive at correct answer due to disjointed ideas.
Relational (several connected aspects integrated into a structure)	-Identifies equations as simultaneous equations, names each of equations, selects an appropriate method for solving equations, finds first unknown & inserts it in one of equations to find second unknown, & could check if answers are correct (x=1, y=1).
Extended abstract (generalized into a new domain)	-Points out that two simultaneous equations with alternate signs, just as above, are to be added together when solving using elimination method; compares method (elimination, substitution, & graphical methods) to determine fastest for this type of equations; and/or proposes a shorter method of finding solutions to problem outside what was taught.

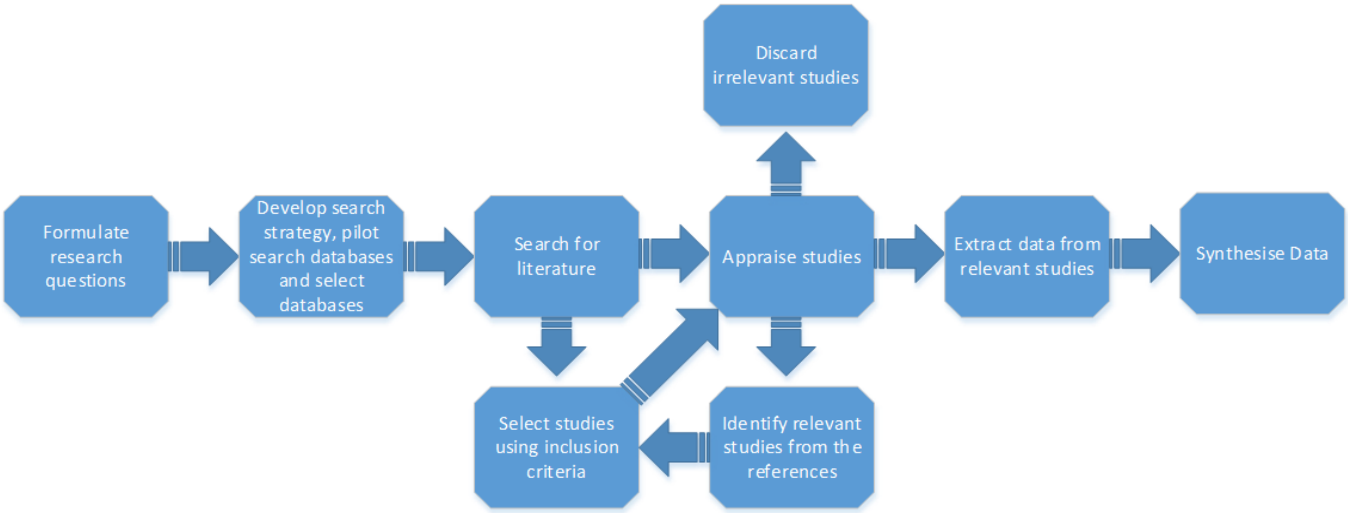

Figure 3. Methodological design

model and mathematics education. A mixed-method systematic approach was utilized to present existing knowledge and evidence on the SOLO model. This review was conducted according to the preferred reporting items for systematic review and meta-analysis (PRISMA) methodology (Moher et al., 2009), which is a leading evidence-based minimum standard for improving the methodology and reporting of systematic reviews (Page et al., 2021). Systematic review is known for its transparency and comprehensiveness and for resolving conflicting evidence. This method was selected because of the researchers' interest in understanding the extent of the use of the SOLO model in mathematics education. In order to avoid duplication of studies, the protocol of the review was preregistered on open science framework (OSF) registries. To guide the study, the following questions were formulated:

1. Do the reviewed studies suggest that the SOLO model is adequate for assessing students' mathematical understanding?
2. What relationship exists between students' mathematics performances and their assessment based on the SOLO model?
3. What observable similarities and differences were evident from the reporting of studies relating to the SOLO model?
4. What generalizations were observed from the studies relating to the SOLO model that could inform future studies?

After the research questions were set, the inclusion and exclusion criteria for studies were determined, the search strategy was developed, the quality appraisal of the literature was undertaken, data were extracted from the selected studies, and narrative synthesis of the data was carried out (as shown in Figure 3). Details about the procedures are discussed next.

Eligibility Criteria

Five criteria were formulated to guide the selection of studies for this review. First, as the focus of the systematic review was on the SOLO model and mathematics, all the studies included in this review needed to be relevant to at least one aspect of the SOLO taxonomy (SOLO levels or modes) and mathematics. It could be related to any topic in mathematics. For example, studies focusing on algebra and the SOLO levels would meet this criterion and would be included in the review, while studies on the SOLO model but not in any way related to mathematics would be excluded. The second criterion was that the study has used a qualitative, quantitative, or mixed-method methodology. Therefore, position papers or theoretical articles that addressed the SOLO model in mathematics were excluded. Third, studies were included if they were carried out between January 1990 and December 2020. Setting the time frame to 30 of the 38 years of the existence of the SOLO model was considered appropriate for identifying its impact on mathematics education. Fourth, to reduce the threat of publication bias, Thornton and Lee (2000) recommended the use of peer-reviewed articles and other grey literature (unpublished research such as thesis, conference proceedings, government reports) in systematic reviews. In line with this recommendation, this study included both peer-reviewed articles and grey literature that met the inclusion criteria. Lastly, the included studies must be written in the English language. There was no restriction on the type of participants involved in the study, which meant included studies could have focused on both students at various levels and teachers.

Development of a Search Strategy & Study Selection

To ensure a comprehensive search of relevant databases, a search strategy was developed. An education librarian at the University of New England, Australia, provided support in developing a search strategy that was appropriate for the objectives of this review. The search terms entered into each of the database search queries were "solo-model" OR "solo model" OR "solo taxonomy" AND Mathematic* (mathematic, mathematics, mathematical, mathematically). A pilot search was undertaken to identify and select relevant databases. The researchers agreed to select four databases that covered a wide range of material: Education Source, Education Resources Information Centre (ERIC), Journal Storage (JSTOR), and PsycINFO.

The literature search was conducted on 16 November 2020 and again on 5 January 2021. These searches resulted in 32 hits in Education Source, 40 hits in ERIC, 108 hits in JSTOR, and 18 hits in PsycINFO, making 198 articles. The eligibility criteria were used to screen out 145 articles: 31 articles did not meet the timespan, eight articles were identified to have duplicates, two articles were not written in English, 17 articles did not refer to the SOLO model, 63 articles did not focus on mathematics, 20 articles were not empirical investigations, and one article was excluded because the full text was not available. The remaining 56 articles met the eligibility criteria.

Three pairs of multiple reports of the same study were identified and each pair was treated as one study. For example, Watson (2002, 2007) had the titles "inferential reasoning and the influence of cognitive conflict" and "the role of cognitive conflict in developing students' understanding of average", respectively. These two sources were linked as one study after identifying that the two reports had a similar number and type of participants, focused on cognitive conflict, and utilized a similar methodology. This step reduced the number of included articles by three. After two independent screening was carried out, the researchers compared the results, resolved disagreements by discussion, and agreed to accept 53 articles for the study.

Lastly, the references of the accepted articles were searched and snowballing was used to search Google Scholar to identify other relevant articles that could meet the inclusion criteria. This step resulted in nine additional articles. Quality appraisal of the selected articles were carried out twice by the researchers to ensure appropriateness of the study sample, ethical considerations, suitability of methodology, adequacy of the data analysis tool and inferences from the results. Thus, a total of 62 articles were reviewed for this study. The process of searching and screening articles is demonstrated in Figure 4.

Data Extraction

The researchers developed a data extraction form. This form was sectioned into three parts. The first section contained the features of the author, the topic, and the year of publication. The second section extracted data on

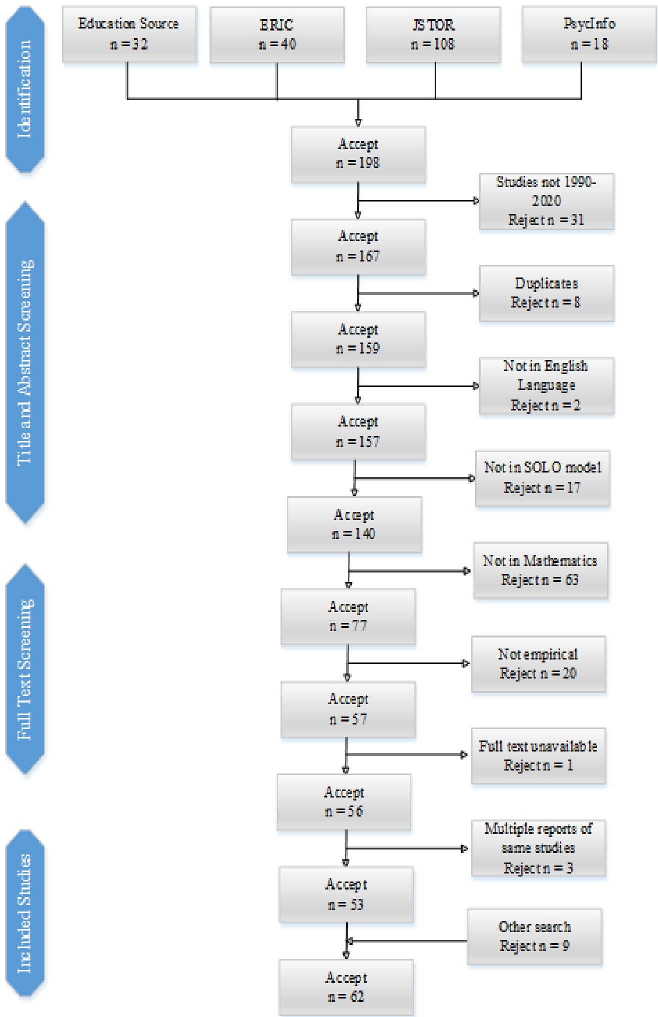

Figure 4. Flow chart illustrating the systematic search and screening processes

the objectives of the studies, the methodology used, what the SOLO model was used for, the perception of the authors of the SOLO model, how SOLO model was used, the type of participants, and the mathematics content. The last section extracted information on the findings related to the use of the SOLO model. Hence, the extracted data were then synthesized.

Data Synthesis

This review was synthesized using a narrative synthesis approach that summarized the results from the included studies. Narrative synthesis, which depends on text and the use of words to describe and summarize the findings of the review, is appropriate for organizing, describing, exploring, and interpreting results of statistically and methodologically diverse studies (Popay et al., 2006). In line with this, the articles reviewed were coded as A, B, or C according to their relevance to the research questions. A, B, and C represented research questions one, two, and three respectively. The SOLO related findings from the categories of the relevant articles were then textually

Table 4. Description of included studies

No Description	Number (%)
1. Focus	
Assessment	48 (67%)
Teaching and learning	17 (24%)
To explain other theory	4 (6%)
Curriculum	3 (4%)
2. Methodology	
Qualitative research	37 (60%)
Quantitative research	17 (27%)
Mixed methods	8 (13%)
3. Data collection instruments	
Tests	37 (45%)
Interviews	29 (35%)
Questionnaires	11 (13%)
Observations	4 (5%)
Document analysis	2 (2%)
4. Type of participants	
High school students	37 (52%)
Pre-service/college/university students	15 (21%)
Primary school students	9 (13%)
Servicing teachers	8 (11%)
Pre-schoolers/kindergarten children	2 (3%)
5. Mathematics content	
Algebra	19 (32%)
Statistics	18 (30%)
Geometry	14 (23%)
Probability	4 (7%)
Word problem	4 (7%)
Algorithm	1 (2%)
Graph	1 (2%)

synthesized to provide answers to the research questions. Hence, the findings of this review were based on the volume of SOLO model results from the 62 articles reviewed.

RESULTS

Characteristics of Included Studies

The 62 articles that were reviewed in this study were described in terms of what the SOLO model was used for, the methodology, the research instruments, the type of participants, and the content of mathematics that was addressed. **Table 4** provides a summary of the data extracted from the 62 included studies.

This review identified that while some articles had multiple foci (i.e., for assessment and to inform mathematics instruction (Lian et al., 2010)), a larger percentage of the reviewed studies focused on utilizing the SOLO model as an assessment tool to measure students' cognitive attainment (Chaphalkar, 2016; Huey, 2011; Kamol & Ban Har, 2010; Peters, 2010). Also, most of the studies used a qualitative research approach (Aoyama & Stephens, 2003; Holmes, 2005), which may be because the SOLO framework emphasizes measuring the quality of students' understanding from the structure and abstraction of the responses. This approach was apparent through the use of interview, observation, and survey materials. Moreover, several studies combined two or more instruments to provide evidence for the SOLO model (Ozdemir & Goktepe Yildiz, 2015; Pfannkuch, 2005; Reading, 2004). Overall, the reviewed studies appear to provide evidence that the SOLO model (either by itself or together with other theories) had been used for mathematics assessment (Canigla & Meadows, 2018; Rider, 2004; Sun, 2013), to inform mathematics teaching and learning (Rooney, 2012; Sudihartinih, 2019; Watson et al., 2014), and to improve mathematics curricula across different school levels in several mathematics topics (Alsaadi, 2001; Brabrand & Dahl, 2009; Callingham & Pegg, 2010).

Adequacy of SOLO Model for Assessing Students' Mathematical Understanding

In this review, understanding is conceived to be an increment in the quantity and quality of students' responses as demonstrated through the number and complexity of connections in ideas. The SOLO model was originally developed to evaluate the quality of students' learning through the structural complexity of their responses to a task (Biggs & Collis, 1982). In line with this intention, 50 (34 qualitative, 10 quantitative, and six mixed-method studies) of the 62 reviewed studies used the SOLO model as a framework for measuring students' mathematical knowledge and cognitive skills. For example, the SOLO model was used to determine the growth pattern or developmental trends of students' thinking (Chaphalkar, 2016; Elassabi & Kacar, 2020; Gagani & Misa, 2017; Huey, 2011; Kamol & Ban Har, 2010; Peters, 2010; Putri et al., 2017; Ronda, 2015), to identify conceptions, misconceptions, and difficulties of mathematical ideas (Campbell, 2006; Fonger, 2019; Mhlolo & Schafer, 2013; Peters, 2010), to distinguish surface learning from deep learning (Afriyani et al., 2018; Atasoy & Konyalihatipoglu, 2019; Easdown et al., 2019; Holmes, 2005; Ozdemir & Goktepe Yildiz, 2015), and to clarify the procedural and conceptual mathematical knowledge demonstrated in students' responses (Caniglia & Meadows, 2018; Niemela et al., 2018; Taplin, 1998).

While most of the 50 articles that assessed students' understanding exclusively used the SOLO model for assessment, a few other articles used the SOLO model with other theories to explain students' mathematical understanding (Aoyama & Stephens, 2003; Caniglia & Meadows, 2018; Sun, 2013). The combination of the SOLO model and other models or theories seems to provide a broader understanding of students' attainment beyond the limitations of the SOLO model. For instance, Caniglia and Meadows (2018) reported how pre-service teachers' level of geometric thinking related to their capacities to pose problems during teaching. The SOLO model was used to assess their responses to a geometry test, and the Crespo (2003) problem-posing framework was used to categorize students' problem-posing practices. The SOLO analysis indicated that the majority of the pre-service teachers lacked conceptual understanding. Similarly, Crespo's (2003) categorization revealed that a substantial number of problems posed during teaching were procedural questions. Hence, the SOLO model was combined with the Crespo (2003) framework to explore connections between pre-service teachers' pedagogical and mathematical knowledge. The authors concluded that the combination is a useful tool for achieving a holistic understanding of students' thinking.

From the reviewed studies, it appears that the SOLO model could sufficiently measure students' mathematical understanding across different school levels. None of the 50 articles that assessed students' responses using the SOLO model challenged the effectiveness of the model in understanding students' cognitive attainment. However, a few studies indicated that the SOLO model could not fully depict students' understanding when multiple-choice questions were posed, due to limited information in the responses (Aoyama & Stephen, 2003; Lam & Foong, 1996). Aoyama and Stephen (2003), however, advocated that multiple-choice questions are suitable for SOLO-based assessment when students are requested to state the reasons for their answers. The SOLO model relies on comprehensive responses from students to appropriately assess students' cognitive attainment. The model was also used to monitor students' progression from informal to formal mathematical reasoning and to classify students' abilities based on the structural complexities in their responses to a task (Goss, 2015). Furthermore, the sampled studies for this review indicated that the SOLO model is useful for diagnosing students' learning and informing future teaching (Caniglia & Meadows, 2018). Because different strands of mathematics were involved in these studies, the SOLO model is not limited to particular topics in mathematics. In addition, from a Rasch analysis of raters' assessment of students' responses, Ilhan and Cetin (2016) reported that the rubrics of the SOLO taxonomy have high consistency among raters and no significant difference was found between raters' leniency and severity. These results suggest that SOLO assessment has clear and fair criteria for assessing students' learning and can potentially provide comprehensive feedback on students' learning.

Relationship Between Students' Performances and Their SOLO-Based Performances

Students' performance, as conceived in this study relates to the results of non-SOLO assessment (usually from marking the quantity of response) on a variety of dimensions that could contribute to their holistic growth in the cognitive, affective, and psychomotor domains of learning. In the studies reviewed, 10 studies provided evidence of correspondence between a range of student performances and their SOLO levels. The dimensions of performance available in the reviewed studies were: learning outcomes (Atasoy & Konyalihatipoglu, 2019; Easdown et al., 2019), students' ability levels (Lian et al., 2010), spatial visualization skills (Ozdemir & Goktepe Yildiz, 2015), activeness in learning activities (Kaharuddin & Hajeniati, 2020), cognitive styles (Mulbar et al., 2017), levels of self-efficacy (Putri et al., 2017), problem-solving abilities (Lian & Idris, 2006), attitude (Sudihartinih, 2019), and students' beliefs (Watson & Moritz, 2001). The available literature related the affective, behavioral, and cognitive several dimensions of students' performances to the SOLO model.

These dimensions have the potential to contribute positively to students' learning, and have been reported to have a direct correlation with students' SOLO levels. For example, Kaharuddin and Hajeniati (2020) compared students' learning outcomes, activeness in learning, and their SOLO level of reasoning. This mixed-method and explanatory sequential research design used observational techniques, test, questionnaire, and interviews to measure students' activeness, learning outcomes and to place students into SOLO levels, respectively. The quantitative data, which were obtained from observations of students' activeness in six meetings and their responses to the test and questionnaire, were analyzed using descriptive and inferential statistics, while the qualitative data from the interview were narratively analyzed. The results of the analysis indicated that students who have high learning outcomes were very active and exhibited extended abstract and relational SOLO levels, students with moderate learning outcomes were active and operated at the multi-structural and uni-structural SOLO levels, and students with low learning outcomes were less active and operated at pre-structural levels of attainment. The authors concluded that students' SOLO levels could inform not only cognitive learning outcomes but also participation in class activities.

In addition, Putri et al. (2017) investigated grade 8 students' self-efficacy and the quality of their thinking in circle geometry based on the SOLO taxonomy. Questionnaires and interviews were used to place students into low, medium, and high self-efficacy groups. Each self-efficacy group was then tested and their responses were analyzed using the SOLO taxonomy. The results of the analysis revealed that high self-efficacy students demonstrated uni-structural, multi-structural, and relational thinking levels, medium self-efficacy students displayed only uni-structural and multi-structural levels, and low self-efficacy students met only the pre-structural and uni-structural levels. This result indicated that students' self-efficacy corresponds with their SOLO thinking levels. Moreover, Ozdemir and Goktepe Yildiz (2015) analyzed the spatial orientation skills of 81 preservice teachers in Turkey. Using the Purdue spatial visualization test (PSVT), the pre-service teachers were classified into low, medium, and high spatial levels. The teachers' responses to a geometric achievement test were then analyzed based on the SOLO taxonomy by two independent researchers to ensure consistencies. A descriptive analysis of the data indicated that preservice teachers with high spatial visualization skills were at relational SOLO levels, while the middle and low spatial teachers were mostly at the multi-structural and uni-structural levels. This implies that the pre-service teachers' spatial visualization abilities correlated with the development levels of the SOLO model. This result indicates that the SOLO framework appears to be a cognitive theory that aligns with other developmental dimensions of individual performance. Hence, SOLO-based assessment has a direct link to student performance.

Similarities Observed from Reviewed Studies

In terms of the similarities in the studies, many of the studies only reported their findings in relation to the levels of students' responses without mentioning the modes in which the students were functioning. As stated earlier, the SOLO model provides both the modes in which students could operate and the levels that exist within each mode. However, only 20.97% of the studies referred to the SOLO modes of functioning. The reviewed studies were all consistent in referring to a learning cycle as a combination of three SOLO levels–**U**ni-structural, **M**ulti-structural, and **R**elational (U-M-R)–and many studies identified two U-M-R cycles within the ikonic, concrete symbolic, or formal modes (Callingham & Pegg, 2010; Groth & Bergner, 2006; Mulligan & Watson, 1998). Goss (2015) called for the inclusion of a third U-M-R cycle (necessitated by students' performance) within the concrete symbolic mode, which is the mode targeted by the school curriculum (Pegg & Panizzon, 2007), suggesting that including a third cycle would allow students to consider both data and context to draw informal inferences. Thus, there is consistent evidence of multiple learning cycles within the modes of functioning (Goss, 2015; Jurdak & El Mouhayar, 2018; Reading, 2004). The emerging multiple cycles of learning (i.e., U_1-M_1-R_1 and U_2-M_2-R_2) observed in the literature could provide explicit explanation for the increasingly diverse structure of students' responses and may provide more in-depth analysis of how individual student understanding progresses rather than a broad view of students' understanding gained from a single U-M-R cycle.

Observable Differences from the Reviewed Studies

In terms of the differences observed in the reviewed studies, most of the studies used the SOLO model as an alternative for measuring students' cognitive attainment (Ilhan & Cetin, 2016; Lam & Foong, 1996), while other studies relied on this framework to inform ongoing teaching and learning (Rider, 2004; Rooney, 2012; Sun, 2013). Some studies used the SOLO model to analyze mathematics curriculum documents (Alsaadi, 2001; Callingham & Pegg, 2010). The studies appear to indicate that the SOLO model could help teachers to meet the expectations of the curriculum - standard learning outcomes. Also, while the majority of the studies did not report how they avoided bias in classifying students' responses to levels (Afriyani et al., 2018; Mulbar et al., 2017; Putri et al., 2017), a few studies reported that as part of the data analysis, two or more independent individuals assigned the students responses into levels, and conflicts about the classification were resolved (Cetin & Ilhan, 2017; Wongyai & Kamol, 2004).

Most of the studies used open-ended questions or tests to obtain SOLO responses from the students. Only two studies mentioned the use of multiple-choice questions to elicit responses that could be assigned to the SOLO levels (Lam & Foong, 1996; Aoyama & Stephen, 2003). While Aoyama and Stephens (2003) combined the multiple-choice questions with an open-ended questionnaire to identify the detailed differences in students' thinking, Lam and Foong (1996) provided evidence that multiple-choice questions (developed through the SOLO levels) are adequate for assessing students' responses using the SOLO model. The latter constructed 10 testlets of three items each (arranged in hierarchical SOLO levels) using the criterion-referenced approach. The three items were developed to assess various levels of thinking (U-M-R) and are associated with mathematical concepts to be tested. The analysis of students' responses to the testlets provided a viable basis for multiple-choice questions in SOLO assessment. Lastly, the majority of the studies acknowledged the usefulness of the SOLO model for assessing mathematics learning outcomes. In contrast, a few other studies claimed that the SOLO model is best used in combination with other theories, as presented later (Aoyama & Stephens, 2003; Caniglia & Meadows, 2018; Sun, 2013).

Generalizations from the Reviewed Studies

The reviewed literature suggests that the SOLO model appears to be useful framework that has contributed substantially to the quality assessment of students' learning outcomes in mathematics. The model seems to inform future teaching and learning processes by identifying students' current stage of learning and how they could progress. Interestingly, the SOLO model was shown to be useful in the planning and implementing of curriculum documents (Brabrand & Dahl, 2009; Alsaadi, 2001). Evidence from the reviewed studies indicated that the SOLO model could also provide a means for understanding students' thinking (Huey, 2011; Peters, 2010). The SOLO model emphasizes the need for conceptual and deep understanding rather than relying on superficial procedural learning (Afriyani et al., 2018; Holmes, 2005). The conceptual understanding of mathematical concepts allows knowledge to be acquired consciously and fully, which helps in knowledge automaticity and the start of learning higher-order skills. The use of the SOLO framework in teaching and assessment seems to distinguish between meaningful learning and 'learning without understanding', which may rely on rote. Learning with understanding is attributed to use of the SOLO model because students often need to present reasons for their responses. Having explicit reasons for the responses is likely to help in the acquisition of schema and the reduction of cognitive load (Pegg, 2010). Many of the reviewed studies required students to provide reasons for their responses when using the SOLO model to assess their understanding. This position of meaningful learning aligns with the standards of mathematics curricula around the world.

There were limitations observed in the reviewed studies. While most of the studies provide consistent evidence to explain students' cognitive abilities in relation to the SOLO framework, there is little evidence about any non-cognitive attributes of students that could contribute to learning (such as motivation, self-efficacy, and belief) with the SOLO model. Only three of the reviewed studies addressed the affective domains of learning (Putri et al., 2017; Sudihartinih, 2019; Watson & Moritz, 2003). Hence, unlike the cognitive attributes, it may not be possible to conclude that students' SOLO levels relate to their non-cognitive attributes. The review of the sampled studies also suggested that young children are under-researched, as only two articles investigated young children (Drefs, 2006; Thouless & Gifford, 2019). Therefore, there is likely to be limited evidence on the ikonic mode of SOLO functioning or the transition to the concrete-symbolic mode. Also, only three of the 62 studies were carried out in Africa (Apawu, 2018; Mukuka, 2020; Wessels, 2007), which indicates that the SOLO model has not been adequately investigated in every cultural context. Finally, while it has been established that the SOLO model could be used to improve the mathematics curriculum, more studies are required to further validate this claim. Future research on the SOLO model could develop cycles of student learning for different topics in mathematics, which may help in the building of pathways for teaching and assessment that are closely linked with the curriculum content.

DISCUSSION

The focus of this review was to determine the adequacy of the SOLO taxonomy for assessing students' mathematical understanding, indicating the relationship between students' performances and their SOLO levels, and addressing the consistencies and disagreements in the research findings in the literature in order to inform future research. From the perspective of the reviewed articles, it appears to a large extent that the SOLO model could sufficiently identify students' mathematical understanding, especially when the responses are open-ended. This result corresponds with Jimoyiannis (2011), whose findings indicated significant empirical evidence to support the efficacy of the SOLO model for explaining students' mental models in computer programming. The results of SOLO-based assessments of students correspond with their performances (cognitive and non-cognitive) using other assessments. The limited evidence on corresponding non-cognitive outcomes may be because the SOLO model mainly focuses on cognitive attainments. These results have implications for mathematics teaching and learning, assessment, and curriculum development.

First, in terms of mathematics teaching, the SOLO framework may help teachers in the planning and development of learning outcomes, and may also inform them of how best to sequence their teaching activities in order to reflect the links between conceptual ideas. This approach could help teachers to teach from simple to complex, concrete to abstract, and how a single idea progresses to multiple, then to create relationships between single and multiple ideas, and then to integrate and extend the relationships to new situations. In contrast, SOLO may become difficult for teachers to use when students are placed at different SOLO levels. For the students, an understanding of the SOLO model could help them to identify and reflect on how much and how well they have learned a concept and what to do to progress in their learning, which means that learners could have control over their learning progression. However, focusing on teaching students the explicit metacognitive tools to understand their learning progression may deviate students' attention from the core subject content and may cause unnecessary cognitive load to the working memory as part of the working memory, resources are allocated to understanding their learning. Further research is needed to clarify some of these ideas.

Second, it appears from the reviewed studies that the SOLO model effectively measures the quality and quantity of students' learning. This framework seems to provide an assessment tool that deviates from traditional evaluation, which mainly depends on the quantity of response. It could also be inferred that SOLO-based assessments measure both the procedural and conceptual knowledge of students, as demonstrated by requesting students to state the reasons for their mathematical procedures. Hence, SOLO-based assessment may encourage students to learn individual mathematical concepts more deeply and link them together meaningfully. Moreover, a SOLO model assessment provides extensive and comprehensive feedback to the teachers on the mathematical competencies of the students, which helps them to decide on the next step in the instructional process.

Hence, the use of the SOLO model to guide teaching and assessment may lead to more meaningful learning of mathematical concepts.

Despite the substantial evidence in the reviewed literature of using the SOLO model for assessment, there are some limitations to its effective use. First, some studies identified the ambiguity of classifying responses into SOLO levels, especially when students' cognitive attainment is concentrated in one or two levels. Therefore, Serow (2007) advocated for inclusion of sublevels to ease the fair placement of students' responses. As most of the studies do not report the reliability of assessing students based on the SOLO model, we encourage SOLO assessors to consider intra-rater (agreement between two assessments by the same person) and/or inter-rater (agreement between two independent assessors) reliabilities, to ensure consistency of assessment. Lastly, as identified earlier, the SOLO model could not sufficiently assess students' responses to multiple-choice questions unless these were carefully designed against the SOLO model. Hence, multiple-choice questions should include a statement of the reasons for the responses.

Third, the SOLO model could help in organizing the curriculum content, defining the levels of educational outcomes, and catering for different possibilities of levels of students' thinking. The incorporation of each of these features may improve mathematics curricula around the world and provide a greater chance of achieving the expected standard specified in these curricula. This review informs mathematics educators and researchers about the extent of use of the SOLO model as evidenced across the 30 years of investigations, but was limited to the inclusion criteria. Further review on the use of the SOLO model could use broader search items to consider articles from selected mathematics education journals that follow a similar methodology and statistical approach, conduct an in-depth longitudinal analysis on the trend of use of SOLO model in mathematics education and take a meta-analysis approach. Further studies on the SOLO model can build on the findings of this study in the domain of mathematics education.

CONCLUSION

This systematic review, which was conducted following the PRISMA guidelines, critically examined studies on the SOLO model and mathematics education. The sampled studies highlighted the SOLO model's strengths in developing mathematics teaching and learning, mathematics assessment, and mathematics curriculum. Although, the few limitations identified on the effectiveness of the SOLO model may discourage education practitioners from relying solely on this model, solutions to the highlighted limitations were presented. Hence, it seems that the strengths of the SOLO model outweigh its weaknesses. The results of this systematic review indicated that:

i. the SOLO model is sufficient to measure and represent learning outcomes of different categories of students in several mathematics topics,
ii. there are direct relationships between dimensions of students' performances and their SOLO levels of attainment,
iii. there are substantially more consistent results than contradictory findings related to using the SOLO framework for assessing and informing instruction, and
iv. more robust research methods and statistical approaches are required to provide information on the mediating and moderating roles of the SOLO model in mathematics education.

The review showed that the SOLO model has been generally demonstrated to be a sound cognitive and developmental theory. Therefore, countries around the world could incorporate it into mathematics teaching, learning, and assessment. This integration may have the potential benefit of promoting students' meaningful learning and awareness of "learning to learn".

Author notes: This review article is part of the PhD thesis of the first author.

Author contributions: All authors have sufficiently contributed to the study, and agreed with the results and conclusions.

Funding: The first author is supported by a University of New England International Postgraduate Research Award.

Protocol registration: Details of protocol registration is available at: https://doi.org/10.17605/OSF.IO/5Y3HF

Declaration of interest: No conflict of interest is declared by authors.

REFERENCES

Afriyani, D., Sa'dijah, C., Subanji, & Muksar, M. (2018). Characteristics of students' mathematical understanding in solving multiple representation task based on solo taxonomy. *International Electronic Journal of Mathematics Education, 13*(3), 281-287. https://doi.org/10.12973/iejme/3920

Alsaadi, A. (2001). A comparison of primary mathematics curriculum in England and Qatar: The SOLO taxonomy. *Research into Learning Mathematics*, *21*(3), 1.

Aoyama, K., & Stephens, M. (2003). Graph interpretation aspects of statistical literacy: A Japanese perspective. *Mathematics Education Research Journal, 15*(3), 207-225. https://doi.org/10.1007/BF03217380

Apawu, J., Owusu-Ansah, N. A., & Akayuure, P. (2018). A study on the algebraic working processes of senior high school students in Ghana. *European Journal of Science and Mathematics Education, 6*(2), 62-68. https://doi.org/10.30935/scimath/9523

Assessment Reform Group. (2002). *Assessment for learning: 10 principles*. https://pdf4pro.com/cdn/assessment-for-learning-10-principles-hkeaa-2b14b0.pdf

Atasoy, E., & Konyalihatipoglu, M. E. (2019). Investigation of students' holistic and analytical thinking styles in learning environments assisted with dynamic geometry software. *Eğitim ve Bilim* [*Education and Science*], *44*(199), 49-74. http://doi.org/10.15390/eb.2019.8003

Bhagwat, M. S. (2017). *Developing and implementing instructional strategy on the structure of observed learning outcomes (SOLO) taxonomy for mathematics of class-ix*. [Unpublished doctoral thesis]. Maharaja Sayajirao University of Baroda, India.

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy (structure of the observed learning outcome)*. Academic Press.

Biggs, J. B., & Collis, K. F. (1991). *Multimodal learning and the quality of intelligent behavior*. Lawrence Erlbaum.

Brabrand, C., & Dahl, B. (2009). Using the SOLO taxonomy to analyze competence progression of university science curricula. *Higher Education, 58*(4), 531-549. https://doi.org/10.1007/s10734-009-9210-4

Callingham, R., & Pegg, J. (2010). Using developmental frameworks to support curriculum outcomes. In L. Sparrow, B. Kissane, & C. Hurst (Eds.), *Shaping the future of mathematics education: Proceedings of the 33rd Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 125-132). MERGA.

Campbell, K. A. (2006). *The SOLO taxonomy and probability: Development of a set of superitems* [Unpublished doctoral dissertation]. State University of New York, Albany.

Caniglia, J. C., & Meadows, M. (2018). An application of the SOLO taxonomy to classify strategies used by pre-service teachers to solve "one question problems". *Australian Journal of Teacher Education, 43*(9), 75-89. https://doi.org/10.14221/ajte.2018v43n9.5

Cetin, B., & Ilhan, M. (2017). An analysis of rater severity and leniency in open-ended mathematic questions rated through standard rubrics and rubrics based on the SOLO taxonomy. *Eğitim ve Bilim* [*Education and Science*], *42*, 189. https://doi.org/10.15390/EB.2017.5082

Chan, C. C., Tsui, M., Chan, M. Y., & Hong, J. H. (2002). Applying the structure of the observed learning outcomes (SOLO) taxonomy on student's learning outcomes: An empirical study. *Assessment & Evaluation in Higher Education, 27*(6), 511-527. https://doi.org/10.1080/0260293022000020282

Chaphalkar, R. M. (2016). *A longitudinal study of students' reasoning about variation in distributions in an introductory college statistics course* [Unpublished PhD thesis]. University of Montana.

Chick, H. (1998). Cognition in the formal modes: Research mathematics and the SOLO taxonomy. *Mathematics Education Research Journal, 10*(2), 4-26. https://doi.org/10.1007/BF03217340

Drefs, M. A. (2006). *Assessment of number concepts in kindergarten and grade one students* [Unpublished doctoral thesis]. University of Calgary, Calgary, AB. https://doi.org/10.11575/PRISM/1601

Easdown, D., Papadopoulos, G., & Zheng, C. (2019). Summer school versus term-time for fundamental mathematics at the tertiary level. *International Journal of Innovation in Science and Mathematics Education, 27*(5), 13-26. https://doi.org/10.30722/IJISME.27.05.002

Elassabi, A., & Kacar, A. (2020). Investigation of Libyan and Turkish students' thinking levels in solving quadratic word problems based on SOLO taxonomy. *PEGEM Journal of Education and Instruction, 10*(1), 283-316. https://doi.org/10.14527/pegegog.2020.010

Fonger, N. L. (2019). Meaningfulness in representational fluency: An analytic lens for students' creations, interpretations, and connections. *Journal of Mathematical Behavior, (54)*, 100678. https://doi.org/10.1016/j.jmathb.2018.10.003

Gagani, R. F. M., & Misa, R. O. (2017). SOLO based-cognition levels of inductive reasoning in geometry. *Alberta Journal of Educational Research, 63*(4), 344-356. https://doi.org/10.11575/ajer.v63i4.56331

Goss, J. (2015). *A method for describing the informal inferential reasoning of middle school students.* [Unpublished doctoral dissertation]. Western Michigan University.

Groth, R. E., & Bergner, J. A. (2006). Preservice elementary teachers' conceptual and procedural knowledge of mean, median, and mode. *Mathematical Thinking & Learning: An International Journal, 8*(1), 37-63. https://doi.org/10.1207/s15327833mtl0801_3

Hattie, J. A. C., & Brown, G. T. L. (2004). Cognitive processes in AsTTle: The SOLO taxonomy. *Ministry of Education (New Zealand)*. https://www.yumpu.com/en/document/view/31644453/43-the-solo-taxonomy-2004pdf-e-asttle

Haynes, J. E. (2009). *Qualitative analyses of a fundamental motor skill across the lifespan; linking practice and theory* [Unpublished doctoral thesis]. University of New England, Australia, AU. https://hdl.handle.net/1959.11/3015

Holmes, K. (2005). Analysis of asynchronous online discussion using the SOLO taxonomy. *Australian Journal of Educational & Developmental Psychology, 5*, 117-127. https://files.eric.ed.gov/fulltext/EJ815606.pdf

Hook, P. (2016). *First steps with SOLO taxonomy: Applying the model in your classroom*. Essential Resources Educational Publishers Limited.

Huey, M. E. (2011). *Characterizing middle and secondary preservice teachers' change in inferential reasoning* [Unpublished doctoral dissertation]. University of Missouri-Columbia.

Ilhan, M., & Cetin, B. (2016). The identification of the views of raters on standard rubrics and rubrics based on the SOLO taxonomy. *Eğitimde Kuram ve Uygulama* [*Theory and Practice in Education*], *12*(1), 1-16.

Jimoyiannis, A. (2011). Using SOLO taxonomy to explore students' mental models of the programming variable and the assignment statement. *Themes in Science and Technology Education*, *4*(2), 53-74. http://earthlab.uoi.gr/theste/index.php/theste/article/view/99/69

Jurdak, M. E., & El Mouhayar, R. R. (2014). Trends in the development of student level of reasoning in pattern generalization tasks across grade level. *Educational Studies in Mathematics*, *85*(1), 75-92. https://doi.org/10.1007/s10649-013-9494-2

Kaharuddin, A., & Hajeniati, N. (2020). An identification of students' responses based on SOLO taxonomy in mathematics learning toward learning activities and learning outcomes. *Al-Jabar: Jurnal Pendidikan Matematika* [*Al-Jabar: Journal of Mathematics Education*], *11*(2), 191-200. https://doi.org/10.24042/ajpm.v11i2.6270

Kamol, N., & Ban Har, Y. (2010). Upper primary school students' algebraic thinking. In L. Sparrow, B. Kissane, & C. Hurst (Eds.), *Shaping the future of mathematics education: Proceedings of the 33rd Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 289-296). MERGA.

Lam, P., & Foong, Y.-Y. (1996). Rasch analysis of math SOLO taxonomy levels using hierarchical items in testlets [Report]. *ERIC Document Reproduction Service No. ED 398271*. https://files.eric.ed.gov/fulltext/ED398271.pdf

Lian, L. H., & Idris, N. (2006). Assessing algebraic solving ability of form four students. *International Electronic Journal of Mathematics Education*, *1*(1), 55-76. https://doi.org/10.29333/iejme/171

Lian, L. H., Yew, W. T., & Idris, N. (2010). Superitem test: An alternative assessment tool to assess students' algebraic solving ability. *International Journal for Mathematics Teaching and Learning*, 1-15. https://www.cimt.org.uk/journal/lian.pdf

Lucander, H., Bondemark, L., Brown, G., & Knutsson, K. (2010). The structure of observed learning outcome (SOLO) taxonomy: A model to promote dental students' learning. *European Journal of Dental Education*, *14*(3), 145-150. https://doi.org/10.1111/j.1600-0579.2009.00607.x

Mhlolo, M. K., & Schafer, M. (2013). Consistencies far beyond chance: An analysis of learner preconceptions of reflective symmetry. *South African Journal of Education*, *33*(2), 1-17. https://doi.org/10.15700/saje.v33n2a686

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Mukuka, A., Balimuttajjo, S., & Mutarutinya, V. (2020). *Applying the SOLO taxonomy in assessing and fostering students' mathematical problem-solving abilities* [Paper presentation]. The 28th Annual Conference of the Southern African Association for Research in Mathematics, Science and Technology Education.

Mulbar, U., Rahman, A., & Ahmar, A. (2017). Analysis of the ability in mathematical problem-solving based on SOLO taxonomy and cognitive style. *World Transactions on Engineering and Technology Education*, *15*(1). https://doi.org/10.26858/WTETEV15I1Y2017P6873

Mulligan, J., & Watson, J. (1998). A developmental multimodal model for multiplication and division. *Mathematics Education Research Journal*, *10*(2), 61-86. https://doi.org/10.1007/BF03217343

Niemela, P., Mikkolainen, V., & Vuorinen, J. (2018). Compute mindlessly. Not! Map consciously. *Universal Journal of Educational Research*, *6*(11), 2669-2678. https://doi.org/10.13189/ujer.2018.061133

Ozdemir, A. S., & Goktepe Yildiz, S. (2015). The analysis of elementary mathematics preservice teachers' spatial orientation skills with SOLO model. *Eurasian Journal of Educational Research*, *61*, 217-236. https://doi.org/10.1016/j.sbspro.2015.01.588

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hrobjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-ilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, *372*, 71. https://doi.org/10.1136/bmj.n71

Pegg, J. (1992). Students' understanding of geometry: Theoretical perspectives. In B. Southwell (Ed.), *Space: The first and final frontier: Proceedings of the Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 18-36). MERGA.

Pegg, J. (2010). *Promoting the acquisition of higher order skills and understandings in primary and secondary mathematics* [Paper presentation]. Australian Council for Education Research Conference, Melbourne.

Pegg, J., & Panizzon, D. (2007). Addressing changing assessment agendas: Impact of professional development on secondary mathematics teachers in NSW. *Mathematics Teacher Education & Development, 9*, 66-80. http://doi.org/10.1.1.496.4499

Pegg, J., & Tall, D. (2005). Using theory to advance our understandings of student cognitive development. *International Group for the Psychology of Mathematics Education*, 187.

Peters, S. A. (2010). *Developing an understanding of variation: Ap statistics teachers' perceptions and recollections of critical moments* [Unpublished doctoral dissertation]. Pennsylvania State University, University Park, PA.

Pfannkuch, M. (2005). Characterizing year 11 students' evaluation of a statistical process. *Statistics Education Research Journal, 4*(2), 5-26. https://doi.org/10.52041/serj.v4i2.512

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006). Guidance on the conduct of narrative synthesis in systematic reviews: A product from the ESRC methods programme. *Lancaster University*. https://doi.org/10.13140/2.1.1018.4643

Putri, U., Mardiyana, M., & Saputro, D. (2017). How to analyze the students' thinking levels based on SOLO taxonomy? *Journal of Physics: Conference Series, 895*, 012031. https://doi.org/10.1088/1742-6596/895/1/012031

Reading, C. (2004). Student description of variation while working with weather data. *Statistics Education Research Journal, 3*(2), 84-105. https://doi.org/10.52041/serj.v3i2.548

Rider, R. L. (2004). Using the SOLO taxonomy to evaluate student learning of function concepts in developmental algebra. *North American Chapter of the International Group for the Psychology of Mathematics Education.*

Ronda, E. (2015). Growth points in linking representations of function: A research-based framework. *Educational Studies in Mathematics, 90*(3), 303-319. https://doi.org/10.1007/s10649-015-9631-1

Rooney, C. (2012). How am I using inquiry-based learning to improve my practice and to encourage higher order thinking among my students of mathematics? *Educational Journal of Living Theories, 5*(2), 99-127.

Serow, P. (2007). Utilising the Rasch model to gain insight into students' understandings of class inclusion concepts in geometry. *Mathematics: Essential Research, Essential Practice, 2*, 651-666.

Serow, P., Callingham, R., & Muir, T. (2019). *Primary mathematics: Integrating theory with practice*. Cambridge University Press. https://doi.org/10.1017/9781108665377

Sudihartinih, E. (2019). Facilitating mathematical understanding in three-dimensional geometry using the SOLO taxonomy. *Erudio Journal of Educational Innovation, 6*(1), 11-18. https://doi.org/10.18551/erudio.6-1.2

Sun, X. (2013). Assessment of prospective teachers' multiple proof construction of a trapezoid area formula. *New Waves, 16*(1), 124-145.

Taplin, M. (1998). Preservice teachers' problem-solving processes. *Mathematics Education Research Journal, 10*(3), 59-75. https://doi.org/10.1007/BF03217058

Thornton, A., & Lee, P. (2000). Publication bias in meta-analysis: Its causes and consequences. *Journal of Clinical Epidemiology, 53*(2), 207-216. https://doi.org/10.1016/s0895-4356(99)00161-4

Thouless, H., & Gifford, S. (2019). Dotty triangles. *For the Learning of Mathematics, 39*(2), 13-18.

Tosuncuoglu, I. (2018). Importance of assessment in ELT. *Journal of Education and Training Studies, 6*(9), 163-167. https://doi.org/10.11114/jets.v6i9.3443

Watson, J. M. (2002). Inferential reasoning and the influence of cognitive conflict. *Educational Studies in Mathematics, 51*(3), 225-256. https://doi.org/10.1023/A:1023622017006

Watson, J. M. (2007). The role of cognitive conflict in developing students' understanding of average. *Educational Studies in Mathematics, 65*(1), 21-47. https://doi.org/10.1007/s10649-006-9043-3

Watson, J. M., & Moritz, J. B. (2001). Development of reasoning associated with pictographs: Representing, interpreting, and predicting. *Educational Studies in Mathematics, 48*(1), 47-81. https://doi.org/10.1023/A:1015594414565

Watson, J. M., & Moritz, J. B. (2003). Fairness of dice: A longitudinal study of students' beliefs and strategies for making judgements. *Journal for Research in Mathematics Education, 34*(4), 270-304. https://doi.org/10.2307/30034785

Watson, J., Chick, H., & Callingham, R. (2014). Average: The juxtaposition of procedure and context. *Mathematics Education Research Journal, 26*(3), 477-502. https://doi.org/10.1007/s13394-013-0113-4

Wessels, H. M. (2007). *Types and levels of data arrangement and representation in statistics as modeled by grade 4 to 7 learners* [Unpublished doctoral dissertation]. University of South Africa.

Wongyai, P., & Kamol, N. (2004). *A framework in characterizing lower secondary school students' algebraic thinking*. https://www.icme-organisers.dk/tsg09/PiyavadeeWongyai.pdf

APPENDIX A

Articles Included in the Systematic Review

Afriyani, D., Sa'dijah, C., Subanji, & Muksar, M. (2018). Characteristics of students' mathematical understanding in solving multiple representation task based on solo taxonomy. *International Electronic Journal of Mathematics Education, 13*(3), 281-287. https://doi.org/10.12973/iejme/3920

Alsaadi, A. (2001). A comparison of primary mathematics curriculum in England and Qatar: The SOLO taxonomy. *Research into Learning Mathematics*, *21*(3), 1.

Aoyama, K., & Stephens, M. (2003). Graph interpretation aspects of statistical literacy: A Japanese perspective. *Mathematics Education Research Journal, 15*(3), 207-225. https://doi.org/10.1007/BF03217380

Apawu, J., Owusu-Ansah, N. A., & Akayuure, P. (2018). A study on the algebraic working processes of senior high school students in Ghana. *European Journal of Science and Mathematics Education, 6*(2), 62-68. https://doi.org/10.30935/scimath/9523

Atasoy, E., & Konyalihatipoglu, M. E. (2019). Investigation of students' holistic and analytical thinking styles in learning environments assisted with dynamic geometry software. *Eğitim ve Bilim* [*Education and Science*], *44*(199), 49-74. http://doi.org/10.15390/eb.2019.8003

Biber, A. C., & Incikabi, L. (2016). Problems posed by prospective elementary mathematics teachers in the concept of functions: An analysis based on SOLO taxonomy. *Mersin Üniversitesi Eğitim Fakültesi Dergisi* [*Journal of Mersin University Faculty of Education*], *12*(3), 796-809. https://doi.org/10.17860/mersinefd.282381

Brabrand, C., & Dahl, B. (2009). Using the SOLO taxonomy to analyze competence progression of university science curricula. *Higher Education, 58*(4), 531-549. https://doi.org/10.1007/s10734-009-9210-4

Callingham, R., & Pegg, J. (2010). Using developmental frameworks to support curriculum outcomes. In L. Sparrow, B. Kissane, & C. Hurst (Eds.), *Shaping the future of mathematics education: Proceedings of the 33rd Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 125-132). MERGA.

Campbell, K. A. (2006). *The SOLO taxonomy and probability: Development of a set of superitems* [Unpublished doctoral dissertation]. State University of New York, Albany.

Caniglia, J. C., & Meadows, M. (2018). An application of the SOLO taxonomy to classify strategies used by pre-service teachers to solve "one question problems". *Australian Journal of Teacher Education, 43*(9), 75-89. https://doi.org/10.14221/ajte.2018v43n9.5

Cetin, B., & Ilhan, M. (2017). An analysis of rater severity and leniency in open-ended mathematic questions rated through standard rubrics and rubrics based on the SOLO taxonomy. *Eğitim ve Bilim* [*Education and Science*], *42,* 189. https://doi.org/10.15390/EB.2017.5082

Chaphalkar, R. M. (2016). *A longitudinal study of students' reasoning about variation in distributions in an introductory college statistics course* [Unpublished PhD thesis]. University of Montana.

Chaphalkar, R., & Wu, K. (2020). Students' reasoning about variability in graphs during an introductory statistics course. *International Electronic Journal of Mathematics Education, 15*(2), 1-22.https://doi.org/10.29333/iejme/7602

Drefs, M. A. (2006). *Assessment of number concepts in kindergarten and grade one students* [Unpublished doctoral thesis]. University of Calgary, Calgary, AB. https://doi.org/10.11575/PRISM/1601

Easdown, D., Papadopoulos, G., & Zheng, C. (2019). Summer school versus term-time for fundamental mathematics at the tertiary level. *International Journal of Innovation in Science and Mathematics Education, 27*(5), 13-26. https://doi.org/10.30722/IJISME.27.05.002

El Mouhayar, R. (2018). Trends of progression of student level of reasoning and generalization in numerical and figural reasoning approaches in pattern generalization. *Educational Studies in Mathematics, 99*(1), 89-107. https://doi.org/10.1007/s10649-018-9821-8

Elassabi, A., & Kacar, A. (2020). Investigation of Libyan and Turkish students' thinking levels in solving quadratic word problems based on SOLO taxonomy. *PEGEM Journal of Education and Instruction, 10*(1), 283-316. https://doi.org/10.14527/pegegog.2020.010

English, L. D., & Watson, J. M. (2015). Exploring variation in measurement as a foundation for statistical thinking in the elementary school. *International Journal of STEM Education, 2*(1), 1-20. https://doi.org/10.1186/s40594-015-0016-x

Fonger, N. L. (2019). Meaningfulness in representational fluency: An analytic lens for students' creations, interpretations, and connections. *Journal of Mathematical Behavior, (54)*, 100678. https://doi.org/10.1016/j.jmathb.2018.10.003

Fujita, T., & Yamamoto, S. (2011). The development of children's understanding of mathematical patterns through mathematical activities. *Research in Mathematics Education, 13*(3), 249-267. https://doi.org/10.1080/14794802.2011.624730

Gagani, R. F. M., & Misa, R. O. (2017). SOLO based-cognition levels of inductive reasoning in geometry. *Alberta Journal of Educational Research, 63*(4), 344-356.

Goss, J. (2015). *A method for describing the informal inferential reasoning of middle school students.* [Unpublished doctoral dissertation]. Western Michigan University.

Groth, R. E., & Bergner, J. A. (2006). Preservice elementary teachers' conceptual and procedural knowledge of mean, median, and mode. *Mathematical Thinking & Learning: An International Journal, 8*(1), 37-63. https://doi.org/10.1207/s15327833mtl0801_3

Holmes, K. (2005). Analysis of asynchronous online discussion using the SOLO taxonomy. *Australian Journal of Educational & Developmental Psychology, 5*, 117-127.

Huey, M. E. (2011). *Characterizing middle and secondary preservice teachers' change in inferential reasoning* [Unpublished doctoral dissertation]. University of Missouri-Columbia.

Ilhan, M., & Cetin, B. (2016). The identification of the views of raters on standard rubrics and rubrics based on the SOLO taxonomy. *Eğitimde Kuram ve Uygulama* [*Theory and Practice in Education*], *12*(1), 1-16.

Irwin, K. C., & Britt, M. S. (2005). The algebraic nature of students' numerical manipulation in the New Zealand numeracy project. *Educational Studies in Mathematics, 58*(2), 169-188.https://doi.org/10.1007/s10649-005-2755-y

Jurdak, M. E., & El Mouhayar, R. R. (2014). Trends in the development of student level of reasoning in pattern generalization tasks across grade level. *Educational Studies in Mathematics, 85*(1), 75-92. https://doi.org/10.1007/s10649-013-9494-2

Kaharuddin, A., & Hajeniati, N. (2020). An identification of students' responses based on SOLO taxonomy in mathematics learning toward learning activities and learning outcomes. *Al-Jabar: Jurnal Pendidikan Matematika* [*Al-Jabar: Journal of Mathematics Education*], *11*(2), 191-200. https://doi.org/10.24042/ajpm.v11i2.6270

Lam, P., & Foong, Y.-Y. (1996). Rasch analysis of math SOLO taxonomy levels using hierarchical items in testlets [Report]. *ERIC Document Reproduction Service No. ED 398271*. https://files.eric.ed.gov/fulltext/ED398271.pdf

Lian, L. H., & Idris, N. (2006). Assessing algebraic solving ability of form four students. *International Electronic Journal of Mathematics Education, 1*(1), 55-76. https://doi.org/10.29333/iejme/171

Lian, L. H., Yew, W. T., & Idris, N. (2010). Superitem test: An alternative assessment tool to assess students' algebraic solving ability. *International Journal for Mathematics Teaching and Learning*, 1-15.

Mhlolo, M. K., & Schafer, M. (2013). Consistencies far beyond chance: An analysis of learner preconceptions of reflective symmetry. *South African Journal of Education, 33*(2), 1-17. https://doi.org/10.15700/saje.v33n2a686

Mukuka, A., Balimuttajjo, S., & Mutarutinya, V. (2020). *Applying the SOLO taxonomy in assessing and fostering students' mathematical problem-solving abilities* [Paper presentation]. The 28th Annual Conference of the Southern African Association for Research in Mathematics, Science and Technology Education.

Mulbar, U., Rahman, A., & Ahmar, A. (2017). Analysis of the ability in mathematical problem-solving based on SOLO taxonomy and cognitive style. *World Transactions on Engineering and Technology Education, 15*(1). https://doi.org/10.26858/WTETEV15I1Y2017P6873

Mulligan, J., & Watson, J. (1998). A developmental multimodal model for multiplication and division. *Mathematics Education Research Journal, 10*(2), 61-86. https://doi.org/10.1007/BF03217343

Niemela, P., Mikkolainen, V., & Vuorinen, J. (2018). Compute mindlessly. Not! Map consciously. *Universal Journal of Educational Research, 6*(11), 2669-2678. https://doi.org/10.13189/ujer.2018.061133

Olive, J. (1991). Logo programming and geometric understanding: An in-depth study. *Journal for Research in Mathematics Education, 22*(2), 90-111. https://doi.org/10.2307/749587

Ozdemir, A. S., & Goktepe Yildiz, S. (2015). The analysis of elementary mathematics preservice teachers' spatial orientation skills with SOLO model. *Eurasian Journal of Educational Research*, *61*, 217-236. https://doi.org/10.1016/j.sbspro.2015.01.588

Pegg, J., & Panizzon, D. (2007). Addressing changing assessment agendas: Impact of professional development on secondary mathematics teachers in NSW. *Mathematics Teacher Education & Development, 9*, 66-80.

Peters, S. A. (2010). *Developing an understanding of variation: Ap statistics teachers' perceptions and recollections of critical moments* [Unpublished doctoral dissertation]. Pennsylvania State University, University Park, PA.

Pfannkuch, M. (2005). Characterizing year 11 students' evaluation of a statistical process. *Statistics Education Research Journal, 4*(2), 5-26. https://doi.org/10.52041/serj.v4i2.512

Putri, U., Mardiyana, M., & Saputro, D. (2017). How to analyze the students' thinking levels based on SOLO taxonomy? *Journal of Physics: Conference Series, 895,* 012031. https://doi.org/10.1088/1742-6596/895/1/012031

Reading, C. (2002). *Profile for statistical understanding* [Paper presentation]. The 6th International Conference on Teaching Statistics, Cape Town, South Africa.

Reading, C. (2004). Student description of variation while working with weather data. *Statistics Education Research Journal, 3*(2), 84-105. https://doi.org/10.52041/serj.v3i2.548

Rider, R. L. (2004). Using the SOLO taxonomy to evaluate student learning of function concepts in developmental algebra. *North American Chapter of the International Group for the Psychology of Mathematics Education*.

Ronda, E. (2015). Growth points in linking representations of function: A research-based framework. *Educational Studies in Mathematics, 90*(3), 303-319. https://doi.org/10.1007/s10649-015-9631-1

Rooney, C. (2012). How am I using inquiry-based learning to improve my practice and to encourage higher order thinking among my students of mathematics? *Educational Journal of Living Theories, 5*(2), 99-127.

Salcedo, A. (2014). Statistics test questions: content and trends. *Statistics Education Research Journal, 13*(2), 202-217. https://doi.org/10.52041/serj.v13i2.291

Serow, P. (2007). Utilising the Rasch model to gain insight into students' understandings of class inclusion concepts in geometry. *Mathematics: Essential Research, Essential Practice, 2*, 651-666.

Sudihartinih, E. (2019). Facilitating mathematical understanding in three-dimensional geometry using the SOLO taxonomy. *Erudio Journal of Educational Innovation, 6*(1), 11-18. https://doi.org/10.18551/erudio.6-1.2

Sun, X. (2013). Assessment of prospective teachers' multiple proof construction of a trapezoid area formula. *New Waves, 16*(1), 124-145.

Taplin, M. (1998). Preservice teachers' problem-solving processes. *Mathematics Education Research Journal, 10*(3), 59-75. https://doi.org/10.1007/BF03217058

Thouless, H., & Gifford, S. (2019). Dotty triangles. *For the Learning of Mathematics, 39*(2), 13-18.

Watson, J. M. (2002). Inferential reasoning and the influence of cognitive conflict. *Educational Studies in Mathematics, 51*(3), 225-256. https://doi.org/10.1023/A:1023622017006

Watson, J. M., & Moritz, J. B. (2001). Development of reasoning associated with pictographs: Representing, interpreting, and predicting. *Educational Studies in Mathematics, 48*(1), 47-81. https://doi.org/10.1023/A:1015594414565

Watson, J. M., & Moritz, J. B. (2003). Fairness of dice: A longitudinal study of students' beliefs and strategies for making judgements. *Journal for Research in Mathematics Education, 34*(4), 270-304. https://doi.org/10.2307/30034785

Watson, J., Chick, H., & Callingham, R. (2014). Average: The juxtaposition of procedure and context. *Mathematics Education Research Journal, 26*(3), 477-502. https://doi.org/10.1007/s13394-013-0113-4

Wessels, H. M. (2007). *Types and levels of data arrangement and representation in statistics as modeled by grade 4 to 7 learners* [Unpublished doctoral dissertation]. University of South Africa.

Whitehead, A. N., & Walkowiak, T. A. (2015). Pre-service teachers' understanding of fraction operations: Providing justification for common algorithms. *North American Chapter of the International Group for the Psychology of Mathematics Education*.

Wongyai, P., & Kamol, N. (2004). *A framework in characterizing lower secondary school students' algebraic thinking*. https://www.icme-organisers.dk/tsg09/PiyavadeeWongyai.pdf