Assessment of Prospective Teachers' Multiple Proof Construction of a Trapezoid Area Formula

Xuhua Sun
University of Macau, Macao

Abstract

The concept of multiple proofs is important, but measuring the multiple-proofs ability of teachers and developing a systematic assessment method are relatively new endeavors. Examining the structures of the participants' proof spaces, this study extends the perspective for understanding proving by combining the structure of observed learning outcomes (SOLO) framework and tools of validity. A multiple-proof task involving the trapezoid area formula is used as an investigative tool. A total of 19 elementary and 23 secondary prospective teachers were asked to provide multiple proofs. As mainstream proof research frequently takes a validity perspective, in this paper, we added a conceptual development perspective of multiple proofs, which enabled us to see which parts of proof assessment are complementary and interrelated in a more integrative perspective.

Keywords: proof, proof spaces, conceptual development, validity, teacher

Background

Multiple-solution Tasks (MST)

Ideally, students should develop a variety of methods to solve problems and the flexibility to select the most appropriate solutions for a given problem. A number of studies (e.g., Star & Bethany, 2009) have observed that students develop inert strategies which are not applied to solving novel problems. As Cai and Nie (2007) noted, "All too often students hold the misconception that there is only one 'right' way to approach and solve a problem and, therefore, they fail to develop flexibility in inventing and selecting appropriate strategies and finding solutions. This misconception might be largely due to their lack of experience of using multiple ways to approach a problem." Creating multiple solutions for a problem (i.e., problem solving in different ways) is recommended as an important way to integrate mathematical knowledge (National Council of Teachers of Mathematics, 2000), develop expertise in problem solving (Silver, Ghousseini, Gosen, Charalambous, & Strawhun, 2005), and improve flexibility in the transfer of knowledge (Star & Bethany, 2009) and mathematical thinking (Krutetskii, 1976; Sun, 2009).

In teacher education, the ability to develop multiple solutions for a problem is a key characteristic of an expert teacher (e.g., Ball, 1993), as well as an important characteristic of teaching performance in high-performing countries (Stigler and Hiebert, 1999). In Japan, the ability to generate multiple solutions for a problem is regarded as a characteristic of professional teaching because during their lesson preparation, professional teachers are required to anticipate strategies that students may likely employ (Shimizu, 2009). In China, generating multiple solutions for a problem [一题多解, yiti duojie] is a key reflection of the substance of a teacher's subject knowledge (Ma, 1999).

Much has been written on the cognitive benefit of multiple solutions for student learning and the significance of teacher preparation. However, little is known about the performance of

Correspondence should be addressed to Xuhua Sun, Faculty of Education, University of Macau, Macao, China. E-mail: sunxuhua@gmail.com

prospective teachers on developing multiple solutions and the approaches to assessing that performance.

Assessments of Multiple-Solution Tasks

Although the notion of multiple-solutions as a thinking development tool is well-known, its in-depth assessment is rarely studied. The assessment of multiple solution tasks (called process-open-ended questions in general) (e.g., Cai & Silver, 1995; Leikin & Lev, 2007a) has mainly used a quantitative approach or has simply listed the solutions, and has involved measures of multiple representations, multiple responses, and multiple strategies in cross-national comparative studies. One distinctive study series done by Cai and his colleagues (e.g. Cai, 2007) found that, when faced with a process-open task that could be solved in multiple ways, such as verbal (primarily written words), pictorial (a picture or drawing), arithmetic (arithmetic expressions), and algebraic (algebraic expressions), Chinese students used symbolic solutions—arithmetic or algebraic—significantly more often than U.S. students, whereas U.S. students used verbal and pictorial solutions significantly more often than Chinese students. In some studies of teachers, U.S. teachers have been much more likely than Chinese teachers to use drawing and guess-and-check strategies whereas Chinese teachers were much more likely than U.S. teachers to use algebraic strategies (Ma, 1999). However, these results raise the question of whether there are any differences between guess-and-check strategies and algebraic ones in terms of conceptual development? Another study (Silver, Leung, & Cai, 1995) used a marble arrangement problem in a cross-cultural comparison to examine the mathematical responses of American and Japanese students. The results simply pointed out that the Japanese students used "multiplication solutions," whereas American students primarily utilized "addition (counting) solutions." Again, this leaves open the question of whether there are any differences between addition (counting) solutions and multiplicative ones in terms of number concept development.

SOLO Framework

A number of studies on learning have claimed that various learning outcomes at a given stage always follow the principle that complexities are accumulated structurally (e.g., Marton, 1981). That is, pre-structural responses represent the use of no relevant aspect; uni-structural represents the use of only one relevant aspect; multi-structural represents the use of several disjoint aspects, usually in a sequence; relational represents several aspects related into an integrated whole; and extended abstract represents taking the whole process into a higher mode of functioning. Biggs and Collis (1982) argued that the traditional quantitative assessment of learning based on aggregating units fails to chart longitudinal growth of conceptual knowledge and proposed the structure of observed learning outcomes (SOLO) taxonomy using qualitative methods. That is, they proposed identifying the different levels of conceptual knowledge based on their use of available information and the complexity with which they are put together, namely, extended abstract, rational, multi-structural, uni-structural, and pre-structural levels. Biggs and Collis claimed that learning processes go from simple to complex, from the lower level to a higher level, from quantitative change to qualitative change. The SOLO framework, which builds upon Piagetian theory or neo-Piagetian theory, is broadly applied to assess structures of any learning result that may occur within a Piagetian stage (Sensorimotor; Intuitive/Preoperational; Concrete Operational; Formal Operational). These include diverse learning results such as mathematics, English, history, geography, economics, and specifically, reasoning on school-related tasks and mathematical problem solving (Collis, Romberg, & Jurdak, 1986). The SOLO framework may also be applied to examine multiple solution tasks as a specific form of problem solving, specifically in the context of proof.

Proving and Proof Assessment

A growing consensus exists that proving should be the core of mathematical education worldwide (Stylianides & Stylianides, 2009a). "Proving for all" is an important direction for current curriculum reforms. In addition, the development of knowledge on the teaching and learning of proof should be highly prioritized to provide a basis for curriculum development and instruction. In the past three decades, the focus in proof assessment has been on the validity of the proof alone. For example, the levels of geometrical reasoning were enumerated as recognition, analysis, ordering, deduction, and rigor, as identified by the van Hieles (Hoffer, 1981). Blum and Kirsch (1991) identified three levels of proof evidence, namely, experimental verifications (verification of a finite number of examples), pre-formal proofs (substantial argumentation on a non-formal basis), and formal proofs (argument at a professional level). Lin (2005) grouped the proof construction performance of students in geometry into four types, namely, acceptable, incomplete (reasoning missing one or two steps), improper (non-deductive reasoning, reasoning based on incorrect properties or reasoning based on correct properties but that do not satisfy the given premise), and intuitive (reasoning based on visual judgment or authority). Most other assessment studies have followed this trend of proof validation (e.g., Stylianides, Stylianides, & Philippou, 2007; Goetting, 2005).

Another characteristic of proof assessment in existing research is that it has generally focused on a single proof at a time. For example, most studies (e.g., Selden and Selden, 2003; Stylianides, Stylianides, & Philippou, 2007; Goetting, 2005; Martin & Harel, 1989) investigated various constructs involved in producing and validating arguments, such as the "proof frame" of the given arguments or the "proof scheme" of a single proof (e.g., Martin & Harel, 1989). A number of approaches to analyzing proof activity, including analysis of oral justifications and identification of a single proof (e.g. Lin, 2005), analysis of the given arguments of a single proof (e.g., Selden & Selden, 2003), and analysis of the "construction-evaluation" activity of a single proof (Stylianides & Stylianides, 2009b) fail to focus on the relations among different proofs, namely, the perspective from multiple-proofs. Would it be possible to examine the structures of multiple-proofs, in a manner similar to the analysis of multiple-solutions, by using the SOLO assessment framework described above?

Multiple-proof Tasks

Although multiple-proof tasks combine multiple-solutions with proving, these tasks have rarely received much attention due to the one-proof-only tradition. Dhombres (1993) argued there is a certain cultural habit of one-proof-only despite the varied ways of dealing with the subject, and that this tradition is attributable to the fact that mathematical development is presented axiomatically with simple demonstration, to the epistemological basis of rational intuitionism of the most economical method, and to variations on a theme that do not generate different theories. The one-proof-only tradition is also easily traced to teacher education due to teachers' weak knowledge (Leikin & Lev, 2007b) and the limited conceptions held by many teachers that their role is not to produce multiple proofs but to deliver the curriculum-prescribed proofs to the students. Besides the one-proof-only tradition and the rich knowledge required for implementation of multiple proofs, an important issue is how to assess multiple-proof tasks. Focusing on the number of proofs or quantitative assessment fails to provide a direction for conceptual development in multiple-proof tasks (e.g. Leikin & Lev, 2007a). Developing a systematic assessment method for examining conceptual development in proof understanding is a relatively new endeavor in mathematics education.

A Multiple-proof Task: Area Formula of a Trapezoid

The separating–combining concept is an important idea in geometry, which was called the “Out-in Complementary Principle” (出入相补原理) by the ancient Chinese mathematician Liuhui (刘徽). This concept has been regarded as the foundation of ancient Chinese mathematics (Siu, 1993). For example, Huang Hengfan (华蘅芳), an ancient Chinese mathematician, produced more than 20 proofs of the Pythagorean Theorem by using this separating - combining concept. This concept has also been broadly applied to develop mathematics education. For example, Zhang Jingzhong (张景中), a famous modern Chinese mathematician, not only proved all geometrical properties and theorems in the secondary curriculum, but also proposed an innovative notion of educational mathematics and associated curriculum reform plan by using this separating - combining concept (Zhang, 2005).

The separating-combining concept is very useful for developing in-depth understanding for teacher education, and is broadly applied to solve problems of area and volume in the modern curriculum. For example, when addressing the topic of the area formula for a trapezoid, the traditional Chinese curriculum materials generally present three types of separating - combining explanations (Figure 2). The concept of separating is used in two of these explanations, namely, a trapezoid is divided into two triangles or into a parallelogram and a triangle. The concept of combining is used in the third explanation, in which a trapezoid is reorganized into a parallelogram by using a copy of the trapezoid. Clearly, the combining approach requires more a complex cognitive load, extending beyond the given trapezoid figure in a way not required in the separating approach. However, all these proofs reflect a relatively basic use of the separating - combining concept. The proof represented in Figure 3 indicates a much deeper application of the separating - combining concept, which integrates both separating and combining. Thus, this type of separating – combining conceptual knowledge is needed to characterize the differences in proofs when assessing multiple proofs of the area formula of a trapezoid.

Figure 1: Example introducing three explanations on the trapezoid area formula in a Chinese textbook (Mathematics Textbook Developer Group for Elementary School, 2003, Vol. 1, Grade 4, p.88)

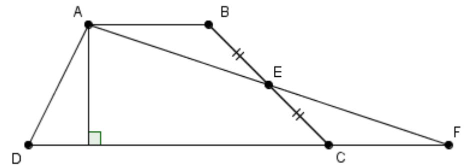


Figure 2: Example of combining separating - combining concept of the trapezoid area formula

However, the separating-combining concept is not the only way to identify and analyze multiple proofs of the area formula of a trapezoid. For example, one could consider generality with respect to the different classes of trapezoids. For example, an isosceles trapezoid and a right-angle trapezoid are less general classes of trapezoids than acute angle trapezoids (where both bottom angles are acute angles). Similarly, an acute angle trapezoid is less general than a trapezoid. When analyzing multiple proofs of the trapezoid area formula, one might consider how general the proofs are (i.e., what classes of trapezoids the proof works for).

Research Bases, Questions, and Methods

Research Bases

Multiple-proof tasks are a combination of multiple-solutions and proving, and thus naturally comprise logical knowledge and conceptual knowledge related to a statement. Logical knowledge refers to the components of the building of various logical forms (producing conjectures and justifications) and their logical relations (classifying justifications as either experiment or deductive justifications). Conceptual knowledge related to a statement refers to the meaning of the statement and its conditions. For example, the proof of the statement "an odd plus an odd is an even," comprises logical knowledge of building formal derivations or using examples from the premise. Conceptual knowledge related to such a proof would include knowledge of number names, properties, and the number system.

Logical knowledge therefore denotes the methods of verifying and justifying in a rigorous way, thinking accurately to a point of a high degree of completeness using formal derivations from the axioms and premises to the conclusions. This is common knowledge for all proof problems. Conceptual knowledge denotes the understanding of the subject associated with a statement, which requires different subject knowledge for different proof problems. Accordingly, the assessment of proof tasks should attend to both types of knowledge. However, the current assessment of multiple-proof tasks actually focuses on either examining logical knowledge or examining conceptual knowledge. This captures only part of the relevant information and fails to provide a whole picture.

In fact, the conceptual development levels between different correct proofs in the current study would not be differentiated if one were to use existing assessment tools for proving. Thus, the SOLO framework, as a general assessment of conceptual development levels, was used to examine multiple proof tasks to understand their structural differences. However, the logical knowledge of a single proof cannot be identified using the SOLO framework alone, so a validity tool was used to assess the multiple proof tasks to analyze the validity differences for the proofs. Thus the analysis combined two tools to offer a more detailed description of the structure of multiple-proofs and their validity.

This combined strategy appears reasonable, but a key issue was how to apply this in the assessment of proving tasks generated by prospective teachers, the research orientation for this study. From a logical and conceptual knowledge perspective, we chose to concentrate on proof spaces, which are the collections of proofs of a statement that individuals or groups can produce. Our research orientation tries to expand the literature on assessment of multiple-proofs, the literature on assessment of multiple-solutions, and the literature on understanding proof from a conceptual perspective.

Research Questions

In this research, we assessed multiple proofs using both the SOLO framework and tools of validity to address the following research questions:

1. Using the SOLO tool, what are the structural differences of multiple proofs generated by prospective teachers for the area formula of a trapezoid within an individual or group proof space?
2. Using a validity assessment tool, what are the validity differences of multiple proofs generated by prospective teachers for the area formula of a trapezoid?

"Individual proof space" or "group proof space" denotes the collections of proofs of a statement that individuals or groups can produce (see elaborated definitions in Leikin, 2009; cf. example spaces defined by Watson & Mason, 2005).

Data Source

The current study was set in the context of a four-year program to obtain a bachelor's degree in mathematics or elementary education. The experiment was conducted in a semester course of mathematics education involving 23 third year prospective secondary teachers and 19 third year prospective elementary teachers (the two populations were not combined). A secondary group (mathematics majors) was recruited using an entrance test of mathematical knowledge at the University of Macau. This group was required to take more advanced mathematics courses than the primary group.

The task presented to the participants was as follows: *Generate multiple proofs for the area formula of a trapezoid and then write down the proofs in your worksheet.* Worksheets were collected after three hours. Participants were invited to present one of their proofs on the whiteboard at the end of the class as a means of sharing and learning from others. A key point in this study was that three hours is an extraordinary length of time to work on a multiple-solution proof task, and this allowed the prospective teachers to fully explore their different solutions.

Analysis methods

The multiple proofs generated by the prospective teachers were analyzed in terms of the validity of a single proof and on the extent of the use of the separating – combining concept in multiple-proofs. Below, we introduce the criteria and explain the data analysis procedure.

The structure of the proof space refers to the longitude or depth of the separating – combining conceptual development for a particular proof task. The analytical method was adopted and revised according to the SOLO framework (Biggs & Collis, 1982): pre-structural responses represent the use of no relevant aspect of the separating – combining concept; uni-structural, only one relevant aspect; multi-structural, several disjoint aspects, usually in a sequence; relational, several aspects related into an integrated whole; and extended abstract takes the whole process into a higher mode of functioning. However, this description of the SOLO levels is rather general. In the specific case of the trapezoid formula multiple-proof task, the

SOLO categories were ascribed based on the extent of completeness of the application of the separating –combining concept in a collection of multiple-proofs.

Within the context of the trapezoid formula task, a proof space was considered pre-structural if it showed no relevant aspect of separating or combining methods in the collection of multiple-proofs. A uni-structural proof space included at least one proof using separating or combining methods (e.g., a trapezoid may be separated into a triangle and a parallelogram, a trapezoid may be separated into two triangles, or two same trapezoids may be combined into a parallelogram using a combining method) in the collection of multiple-proofs. Proofs with combining methods might require higher cognition than those with separating methods because the latter requires reorganizing a quadrangle or triangle within a trapezoid, whereas the former requires reorganizing a polygon outside of the trapezoid. Proofs with separating or combining methods were differentiated between simple and complex reasoning. Therefore, a uni-structural proof space could be classified as one of four levels, namely, a proof space with simple separating methods, a proof space with complex separating methods, a proof space with simple combining methods, and a proof space with complex combining methods.

A multi-structural proof space included at least two proofs using both separating and combining methods in the collection of multiple-proofs. A relational proof space was characterized by at least one proof combining separating and combining methods in the collection of multiple-proofs. Finally, an extended abstract proof space denoted at least one proof with the characteristics of a relational proof space, but which also synthesized various methods (both separating and combining) in the collection of multiple-proofs. Clearly, a systematic examination of the use of the separating-combining concept requires examination of more than one proof. Note that the use of the SOLO-based tool in this study was limited to the analysis of the use of the separating-combining concept. We did not examine proofs using other approaches, such as using a definite integral. The method for analyzing the validity of proofs in this study was adapted from previous studies (Zazkis & Leikin, 2007, 2008; Stylianides & Stylianides, 2009b). Some proofs were found to make use of different classes of trapezoids, whereas others reflected deficiencies in understanding the classes of a trapezoid. Some proofs, for example, referred to general trapezoids, whereas other proofs referred to acute angle trapezoids or isosceles trapezoids only. Still others specified trapezoids with specific characteristics, such as a=2 h or b=4 cm. Thus, the generality of a proof was considered an indication of the participants' mathematical understanding of proofs.

The following categories for proof validity were adopted in the study. A complete proof for the truth of a claim denoted a sequence of reasoning that referred to all cases involved in the claim. An incomplete proof for the truth of a claim denoted a sequence of reasoning that referred to some cases, specifically a proper subset of all cases involved in the proof construction. An empirical proof implied invalid reasoning that provided inconclusive evidence for the truth of a claim. Finally, an invalid proof was a non-genuine or irrelevant response. Each proof construction within a proof space was classified into one of these four categories.

One might ask whether the two analytic tools are interrelated. We categorized a proof space by identifying the most complex use of the separating-combining concept in a complete proof in the space. Therefore, in order to be considered a uni-structural proof space, at least one valid proof must be included in the space. Accordingly, the proof with the most complex reasoning in a uni-structural proof space should be a compete proof. However, the SOLO tool fails to describe the state of a pre-structural proof space. Instead, the validity tool has the potential to further classify the multiple-proofs collection in a pre-structural proof space. The proofs in a pre-structural proof space could be characterized as incomplete general, empirical, and invalid proofs in terms of their validity and logical rigor. Conversely, the validity tool fails to further to differentiate the conceptual development of complete proofs. In this case, the SOLO tool has the capacity to further classify the multiple-proofs collection into uni-structural, multi-structural, relational, and extended abstract categories.

Results

Structural Differences of Individual Proof Spaces

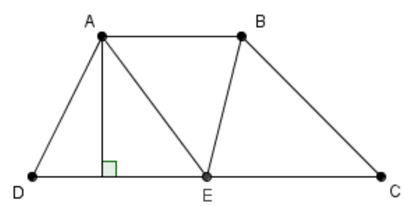
The results indicated some qualitative differences in the application of the separating – combining concept in individual proof spaces. Using the SOLO framework described above, the proof spaces were classified according to the different criteria of application of separating – combining concept. Here, we present illustrative examples from each category.

Uni-structural proof space. The structures of the proof spaces of nine elementary teachers and five secondary teachers were categorized into the uni-structural (U) level. A uni-structural proof space outlines at least one proof with separating or combining methods. It includes four categories, namely, proof space with simple separating methods, proof space with complex separating methods, proof space with simple combining methods, and proof space with complex combining methods. Illustrative examples are detailed below for the four levels of uni-structural proof space.

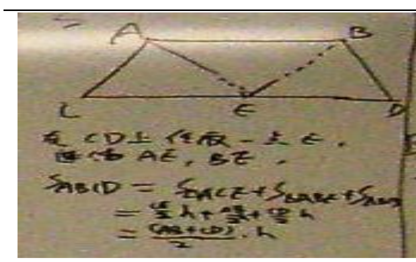
Level 1: Proof Space with Simple Separating Methods

E is the midpoint of CD. Connect AE and BE (Note: AB=a; DC=b). Therefore,

$$S_{ABCD} = S_{\Delta ADE} + S_{\Delta ABE} + S_{\Delta BCE} = \frac{1}{2}\cdot\frac{b}{2}\cdot h + \frac{ah}{2} + \frac{1}{2}\cdot\frac{b}{2}\cdot h = \frac{(a+b)h}{2}.$$

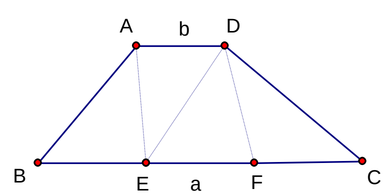


COMMENT: The trapezoid is divided into three triangles. The key goal of the method is finding the midpoint, which simplifies the task of proof construction. In addition, any point on the line DC is also available.



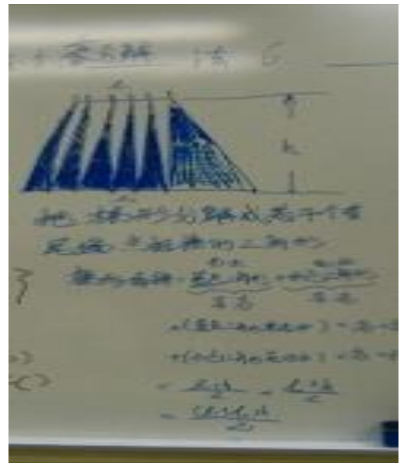
Level 2: Proof Space with Complex Separating Methods

E is the 1/m point of BC (divide BC into m equal sections). Connect AE and DE (Note: AD=b; BC=a). Therefore,



SABCD =1/2 1/m a h m+1/2bh= $\frac{(a+b)h}{2}$.

COMMENT: This approach is an extension of the method above from two sections to m sections.



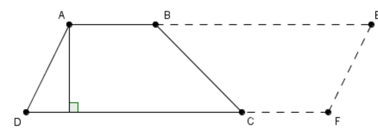
Level 3: Proof Space with Simple Combining Methods

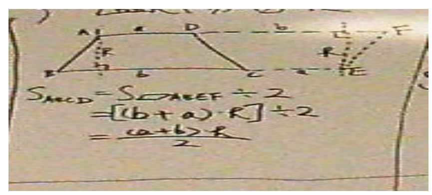
Extend AB to E as $BE = CD$. Extend DC to F as $CF = A$

(Note: AB=a; DC=b). Then, $AE = FD$ and $AE // FD$.

Therefore, AEFD is a parallelogram and is expressed as:

$$S_{ABCD} = \frac{1}{2} S_{AEFD} = \frac{(a+b)h}{2}.$$





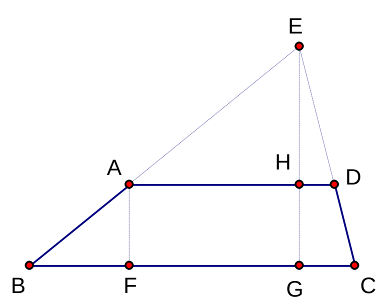
COMMENT: A trapezoid is reorganized into a parallelogram by making a copy of the trapezoid.

Level 4: Proof Space with Complex Combining Methods

Extend BA and DC. E is the intersection of BA and CD. Draw height EG and height AF. G is the intersection of EG and BC. F is the intersection of AF and BC (Note: AD=a; BC=b).

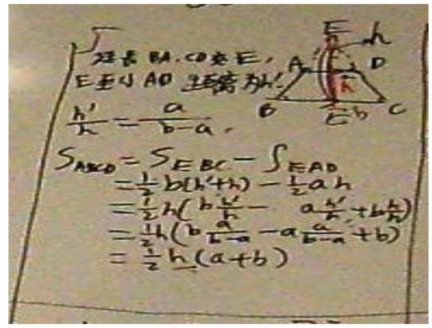
Given that $AD // BC$, the triangle EAD is similar to the triangle EBC,

$$\frac{EH}{EG} = \frac{EH}{h+EH} = \frac{AD}{BC} = \frac{a}{b}.$$



Then, EH= $\frac{ah}{b-a}$ and:

$$S_{ABCD} = S_{\Delta EBC} - S_{\Delta EAD}$$
$$= \frac{b(h + \mathrm{EH})}{2} - \frac{a\mathrm{EH}}{2} = \frac{(a + b)h}{2}\cdot$$

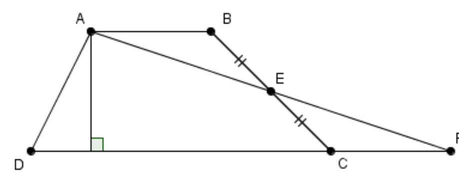


COMMENT: A trapezoid is formed into a triangle by extending its two sides. EH was eliminated according to the property of similar triangles.

Multi-structural proof space. The proof spaces of six elementary teachers and 11 secondary teachers were classified into the multi-structural (M) level. A multi-structural proof space might include at least two proofs collectively employing both separating and combining methods like those shown above in the collection of proofs, but it does not include an individual proof that brings both methods together. Examples of this proof space would be collections of proofs similar to those above (including examples of both separating and combining proofs), and thus are not repeated here.

Relational proof space. A relational (R) proof space is characterized by containing at least one single proof in the collection of proofs that combines the separating and combining methods. This is distinct from multi-structural proof spaces in which the two methods are never combined in a single individual proof. Five secondary teachers, but no elementary teachers, were classified into the R level.

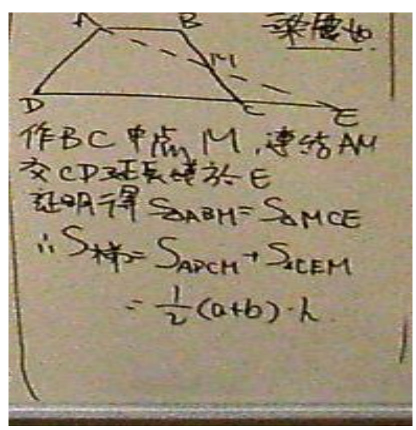
Example 1:
E is the midpoint of BC. Connect AE. F is the intersection of extended line DC and extended line AE (Note: AB=a; DC=b).



$$\begin{cases} \angle ABE = \angle FCE \\ BE = CE \\ \angle BEA = \angle CEF \end{cases} \Rightarrow \Delta ABE \cong \Delta FCE \Rightarrow \begin{cases} AB = FC \\ S_{\Delta ABE} = S_{\Delta FCE} \end{cases}$$

Then,

$$S_{ABCD} = S_{\Delta ADF} = \frac{(a+b)h}{2}.$$



COMMENT: A trapezoid is skillfully transformed into a triangle with the same area by replacing ΔABE by ΔFCE. This example is more advanced compared with those in the multi-structural proof space because it integrates the separating and combining methods.

Extended abstract proof space. An extended abstract (E) proof space contains at least one proof with the characteristics of a relational proof, but then generalizes both separating and combining to untaught applications in the collection of multiple-proofs. Only two secondary teachers had proof spaces whose structures were categorized into the E level.

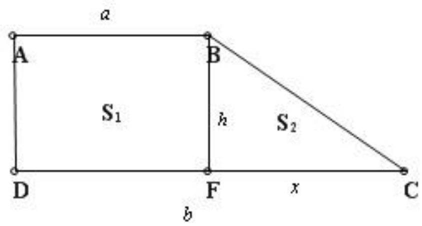
The following example is a general solution from a specific proof to a general proof, which reflects thinking about proving in broad generalizations and is, therefore, categorized into the extended abstract proof spaces. This thinking is more advanced than the thinking reflected in the proofs of a rational proof space.

With a right angle, the general trapezoid can be divided into a triangle and a rectangle.

Suppose $AB = a, \quad CD = b, \quad AD = h$

$$S_2 = \frac{(b-a)h}{2}$$

$$S = S_1 + S_2 = ah + \frac{(b-a)h}{2} = \frac{2ah + bh - ah}{2} = \frac{(a+b)}{2}h$$



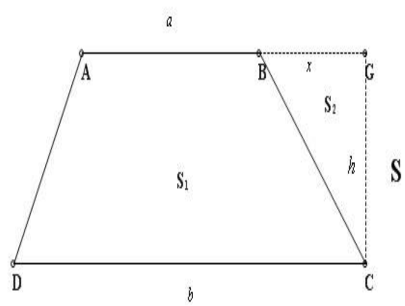
Without a right angle, the trapezoid can be transformed into one with a right angle (i.e., by drawing its height). Suppose BG=x. Then, according to the conclusion on the right angle trapezoid mentioned earlier, the following is derived:

$$S_{right-angle-trapzoid} = \frac{(a+x+b)}{2}h$$

$$S_{\Delta} = \frac{x}{2}h$$

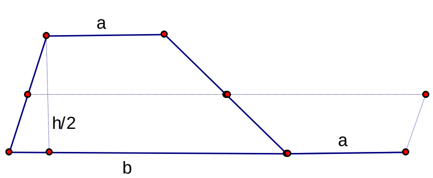
$$S_{trapzoid} = S_{right-angle-trapzoid} - S_{\Delta}$$

$$= \frac{(a+x+b)}{2}h - \frac{x}{2}h - = \frac{(a+b)}{2}h$$



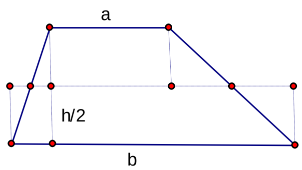
The following proofs indicate creative applications of separating and combining by reverse thinking or by constructing a figure from an algebraic formula, such as 1/2h (a+b). These examples theorize or generalize the concept (both separating and combining), which results in a more advanced proof than those in the rational proof space.

$$S = \frac{1}{2}(a+b)h$$



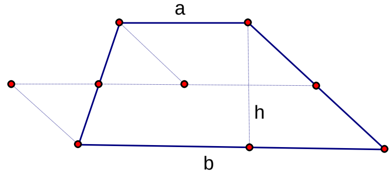
,

$$S = a \cdot \frac{h}{2} + b \cdot \frac{h}{2} = \frac{1}{2}(a+b)h$$



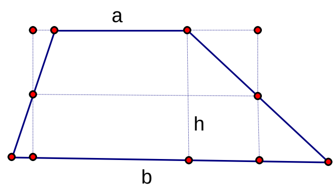
,

$$S = a \cdot \frac{h}{2} + b \cdot \frac{h}{2} = \frac{1}{2}(a+b)h$$



,

$$S = \frac{1}{2}(a+b)h$$

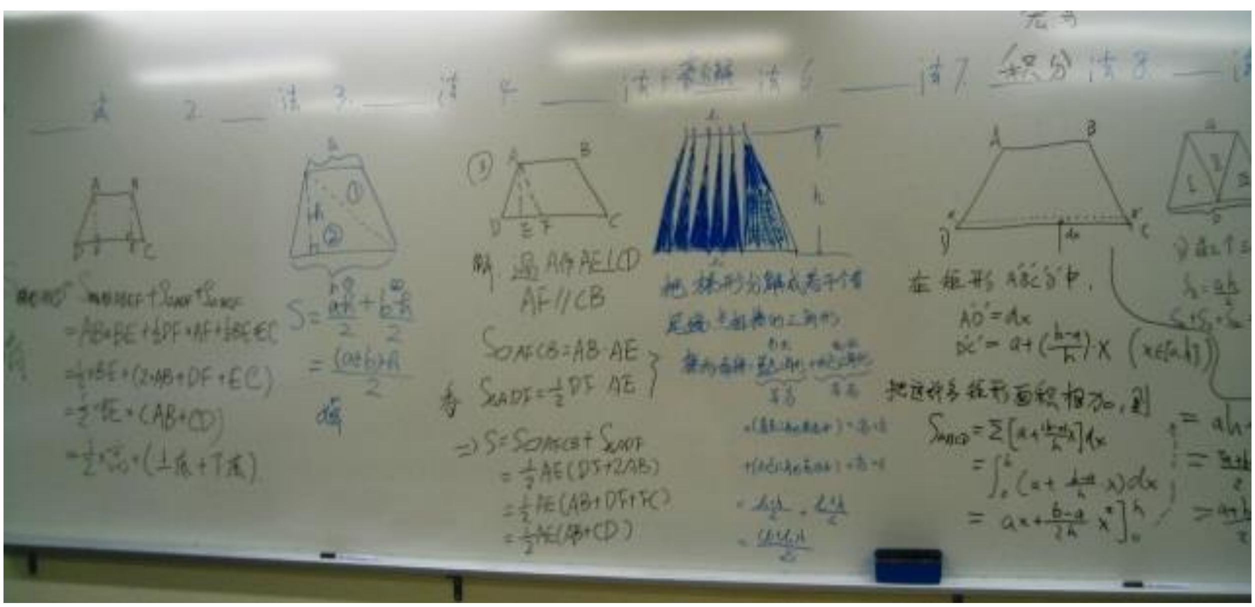


.

The approaches above clearly reflect the advanced use of the separating-combining concept, using deep "geometrical eyes" in proof construction from multiple perspectives, which prompts learners to prove in a more reflective way. The proofs are more advanced than those in relational proof space, and are instead categorized as extended abstract proof spaces.

Structural Differences of Group Proof Spaces

The prospective teachers were given an opportunity to present a portion of their multiple proofs in the classroom (Figure 3). The distribution of structures within individual/group proof spaces are presented below.



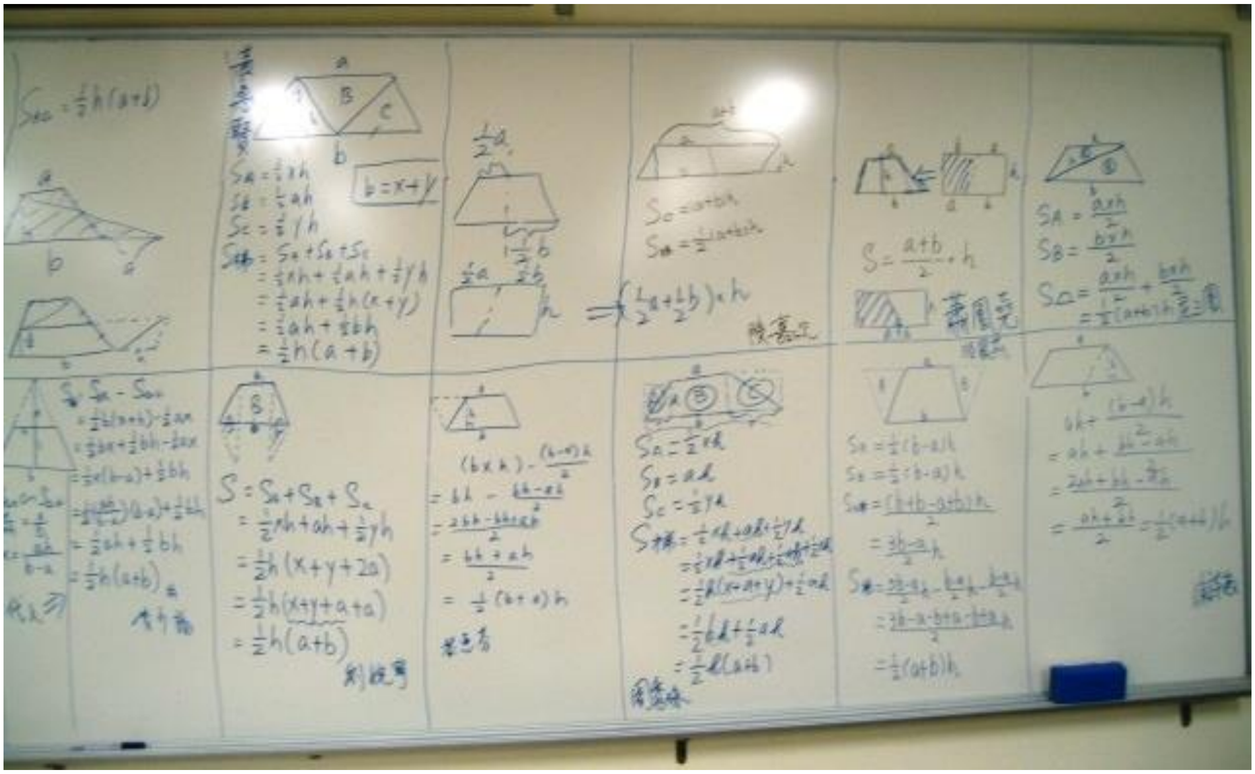


Figure 3: Photographs of multiple proofs presented by students in the classroom

The most complex group proof spaces generated by the secondary and elementary teachers fell into the extended abstract and multi-structural categories. Although the elementary teachers generated a greater number of proofs, the levels at which they applied the separating-combining concept in multiple-proofs (structures of proof space) were obviously lower than those of the secondary group. The proof space structures most often observed in the secondary and elementary groups were multi-structural and uni-structural (see Table 1).

Table 1
Distribution of the SOLO Structural Levels of Proof Spaces among the Groups of Secondary and Elementary Teachers

	Most Complex Structure Level of Proof Space at the Group Level	Most Complex Structure Level of Proof Space at the Individual Level
Prospective Secondary Teachers	E	M
Prospective Elementary Teachers	M	U

For the elementary group (see Table 2), only six of the 19 teachers (approximately 32 %), were in the M category, nine teachers (47 %), fell into in the U category, and four teachers (21%), were in the pre-structural (P) category. The classification of the average SOLO level of functioning for the elementary group was above U. The elementary group was in transition from level P to M and none were in the R category. The data suggest that the majority of the elementary teachers could not produce the R and E levels of structure of the trapezoid area formula proofs.

For the secondary group, 18 of the 23 teachers (about 78%) were classified into the M to E categories. The data suggest that the majority of secondary teachers could present deep structure in this specific proof task. Only five secondary teachers (about 22%) were classified in the U category, and none were classified in the P category. The classification of the average SOLO level of functioning for the secondary group was near M. The majority of the teachers in the secondary group ranged from levels M to R.

The data indicate that across both groups, 18 teachers (approximately 30.9%) were classified into the U and P categories and were unable to present adequately deep proof space structures.

Table 2

Distribution of levels of structure of individual proof spaces

	Structure of Proof Space				
	Pre-structural	Uni-structural	Multi-structural	Relational	Extended Abstract
Number of Prospective Secondary Teachers	0	5	11	5	2
Number of Prospective Elementary Teachers	4	9	6	0	0

The structures of proof spaces provided by prospective primary teachers were unsurprisingly weaker than those of the secondary teachers. The primary teachers were recruited with weak mathematical requirements for this study and their programs were designed to prepare them for all subjects (English, Chinese, Music, etc.) without specialization.

In summary, the data suggest that the majority of secondary teachers (78%) and a minority of elementary teachers (32%) could provide a multi-structural proof space (i.e., could present proving in multiple ways). Across the two groups, 30% of the teachers could provide only uni-structural and pre-structural proof spaces, which is inadequate for presenting mathematics in multiple ways in this specific proving task.

Overall Validity Differences

Although four of the elementary teachers were categorized as having pre-structural proof spaces, 15 of the 42 elementary and secondary teachers (35.7%) produced proofs belonging to the pre-structural proof space. The elementary teachers produced seven invalid proofs, eight empirical proofs, and one incomplete general proof, all of which indicated inadequate proof knowledge. The secondary teachers produced no invalid proofs, seven empirical proofs, and two incomplete general proofs. The validity distribution provided by both groups of teachers is presented in Table 3.

Table 3
Validity Distribution Provided by Elementary and Secondary Prospective Teachers

	Secondary	Elementary
Valid (n=36)	Complete general proof (n=12)	Complete general proof (n=6)
	Incomplete general proof (n=2)	Incomplete general proof (n=1)
	Empirical proof (n=7)	Empirical proof (n=8)
Invalid (n=7)	Non-genuine or irrelevant response (n=0)	Non-genuine or irrelevant response (n=7)

Proof Validity Performance

In addition to the complete general proofs, which successfully established the area formula for all trapezoids generally, the proofs produced within the pre-structural level were examined. In the following, these proofs are classified and some illustrative examples according to the given framework of the validity of a proof are presented.

Incomplete general proofs. Incomplete proofs of a formula did not apply to a trapezoid in general, but made reference to some special figures or a proper subset of all cases. The prospective teachers generated three incomplete general proofs (Figure 4). The primary incomplete general proofs in this case referred to the acute angle trapezoid (both bottom angles were acute angles), and did not consider the situation wherein one of the bottom angles is an obtuse angle. It is interesting to note that these proofs were initially believed by majority of the participants to be correct.

$$S = ah + \frac{1}{2}xh + \frac{1}{2}yh = \frac{1}{2}(a+b)h$$

$$S = bh - \frac{1}{2}xh - \frac{1}{2}yh = \frac{1}{2}(2b - x - y)h = \frac{1}{2}(a+b)h$$

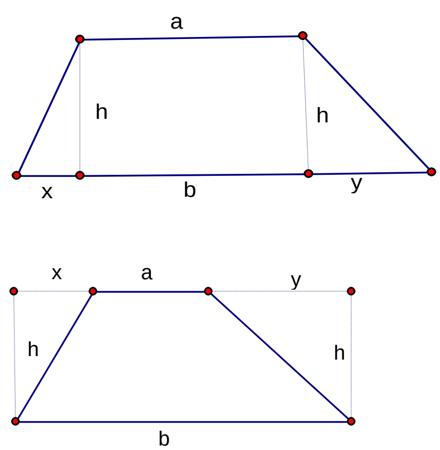


Figure 4: Proofs for an acute angle trapezoid

The proof for an obtuse angle trapezoid (i.e., an obtuse angle exists at the bottom) was neglected (Figure 5).

$$a + y = b + x,\ x - y = a - b$$
$$S = (b + x)h - 1/2xh - 1/2yh = bh + 1/2xh - 1/2yh$$
$$= bh + 1/2(x - y)h = bh + 1/2(a - b)\ h$$
$$= 1/2(a + b)h$$

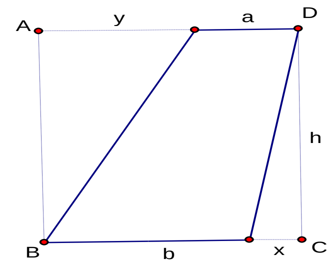
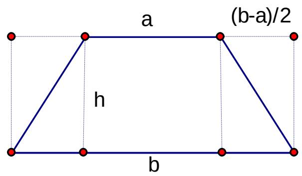


Figure 5: Proof of an obtuse angle trapezoid

Empirical proofs. The empirical proofs included some special cases, such as isosceles or right-angle trapezoids. The prospective secondary and elementary teachers generated seven and eight empirical proofs, respectively. Two examples of these responses are provided below, one for the isosceles trapezoid (Figure 6) and one for the right-angle trapezoid (Figure 7). Other general mistakes found in the empirical proofs included proving with a conditional premise (e.g., a=b, a=h; a=2b, h=3 cm, or b=4 cm) in the justification.



$$bh - \left(\frac{b-a}{4}h + \frac{b-a}{4}h\right)$$
$$= \frac{2bh}{2} - \frac{b-a}{2}h$$
$$= \frac{(a+b)}{2}h$$

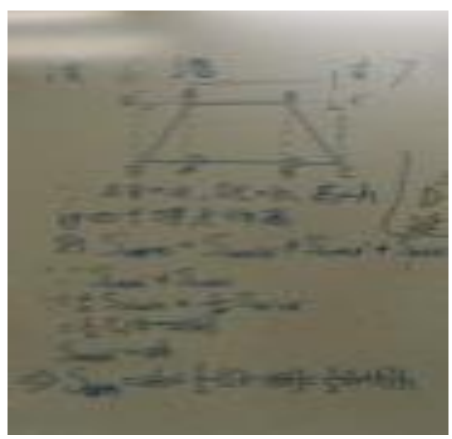
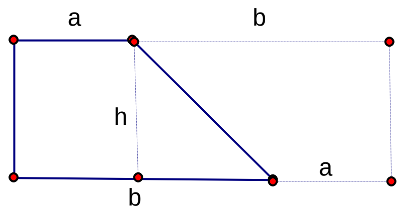


Figure 6: An empirical proof in reference to an isosceles trapezoid as all cases

S=1/2ah+$\frac{1}{2}(a-b)h=\frac{1}{2}(a+b)h$



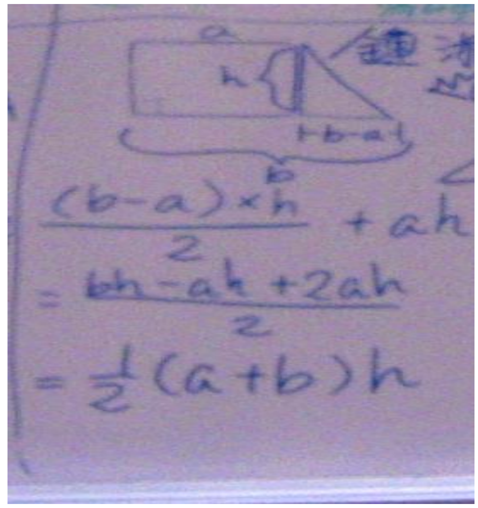


Figure 7: An empirical proof in reference to a right-angle isosceles trapezoid as all cases

Invalid proofs. An invalid proof comprises flawed or irrelevant answers in proof construction. The prospective elementary teachers generated seven invalid proofs, but the secondary teachers generated none. An invalid proof showed logical flaws in the justification (Figure 8). One of the primary mistakes in justification was cyclical reasoning, in which the trapezoid area formula was used in the proof.

Original student's diagram of proof:

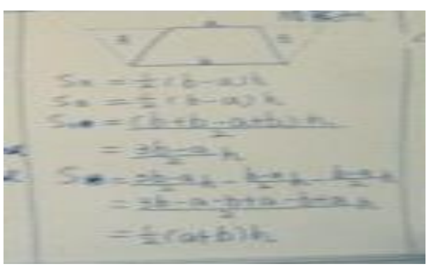


Figure 8: Invalid proof with wrong justifications

In summary, these data reveal the performance of prospective teachers with respect to proof validity when proving the area formula of a trapezoid. Across the two groups, 35.7% of the prospective teachers produced deficient proofs (seven invalid proofs, 15 empirical proofs, and three incomplete general proofs), reflecting difficulties with generality and their proof knowledge with regard to the trapezoid area formula. Consistent with previous research, these students demonstrated poor competencies in justifying and proving (e.g., Bell, 1976; Senk, 1989), especially in their reliance on the use of examples as a means of demonstrating and verifying the truth of a statement.

Discussion

The existing trend in the assessment of proofs does not focus on conceptual development, but rather on the validity of the proof alone. This is because examining a single proof rarely provides sufficient information about conceptual structure. From a conceptual knowledge and a logical knowledge perspective, this study provides two tools for examining proof spaces, specifically, the structures associated with the separating-combining concept and the validity of a proof. Obviously, these are complementary, and together they describe proof understanding better than either tool would individually when assessing multiple-proof tasks. From a separating-combining concept perspective, multiple-proofs can be classified using the following levels: extended abstract, rational, multi-structural, uni-structural, and pre-structural. Note that a proof space, as a holistic structure, falls into only one of these levels; it does not span more than one level. From a logical knowledge perspective, the validity of a single proof can be identified as either complete general, incomplete general, empirical, or invalid. In contrast, a complete general proof may belong to one of several proof space levels, namely, the extended abstract, rational, multi-structural, and uni-structural levels.

Implications, Limitations, and Future Directions

Although the concept of multiple-proofs is important for teachers and students, the assessment of teachers' multiple-proofs ability and the development of a systematic assessment method are relatively new endeavors. By combining the SOLO framework and tools of validity, this study extends the perspective of proof understanding. The proof spaces presented above reveal structured and unstructured perspectives at several different conceptual levels: extended abstract, rational, multi-structural, uni-structural, and pre-structural. Also, with respect to validity, the single proofs in this study were identified as complete general, incomplete general, empirical, and invalid. The data suggest that 78% of the secondary teachers and 32% of the elementary teachers could present multiple proofs of the trapezoid area formula. However, 35.7% of the prospective teachers produced some number of invalid, empirical, or incomplete general proofs, reflecting deficiencies in their proving knowledge. These findings have two significant implications for multiple-solution assessment and proof assessment.

Multiple-solution Assessment

Mathematics educators have long recognized the importance of being able to solve a problem in multiple ways, and that this capacity is a hallmark of expertise in a given domain. The findings of this study indicate that the assessment of multiple solutions would benefit from a focus on qualitative characteristics rather than only on quantity of solutions or on listing solutions without further analysis of underlying conceptual understanding. For example, Silver, et al. (1995) used a marble arrangement problem to show that Japanese students used multiplication solutions, whereas American students primarily utilized addition (counting) solutions. These different responses indicate different levels of understanding which can be classified in the SOLO framework. A range of cross-cultural comparisons that have examined multiple solutions follow the same trend of simply listing different solutions without discerning different conceptual development levels. These findings are rarely identified in existing studies.

Proof Assessment

The significance of this study also lies in its effort to capture a more holistic view of assessment of the written products of proving. The conceptual (SOLO) and validity frameworks extended our perspective for analyzing learners' understanding of proving. In addition to analyzing the validity of single proofs, this study develops an assessment tool to understand multiple proofs from a proof space perspective, which adds a new dimension to proof evaluation and points to a new trend of analyzing the conceptual understanding represented in the products of the activity of proving. As mainstream proof research is frequently focused on a validity perspective, we seek to integrate a conceptual development perspective related to a pedagogical practice favored in Chinese teaching—OPMS (one problem multiple solutions) (Sun, 2011). This may enable us to see which parts of proof assessment are complementary and interrelated in a more integrative perspective. Similar results to those presented here have been found in work with the area formula of a triangle and the area formula of a parallelogram. Further work will be required to identify the extent to which the conclusions presented here hold more broadly.

It is also appropriate to acknowledge the limitations of this study. The findings of the study should not be overgeneralized. Empirical studies with other multiple solution tasks or with the use of secondary students as samples are also recommended. In addition, teachers' proof spaces might look very different across different proving items. The work presented here may be strengthened by examining teachers' performance across several different proving items. Finally, interviews with the participants could provide additional useful information for classifying students' proving understandings.

References

Ball, D. L. (1993).With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *The Elementary School Journal, 93* (4), 373-397.

Bell, A. (1976). A study of pupils' proof-explanations in mathematical situations. *Educational Studies in Mathematics*, *7,* 23-40.

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

Blum, W., & Kirsch, A. (1991). Pre-formal proving: Examples and reflections. *Educational Studies in Mathematics, 22*(2), 183-203.

Cai, J., & Nie, B. (2007). Problem solving in Chinese mathematics education: Research and practice, *ZDM Mathematics Education, 39,* 459-473.

Cai, J., & Silver, E. A. (1995). Solution processes and interpretations of solutions in solving a division-with-remainder story problem: Do Chinese and U.S. students have similar difficulties? *Journal for Research in Mathematics Education, 26*(5), 491-497.

Cai, J. (2007). *A series of empirical studies of U.S. and Chinese students' learning of mathematics*. Beijing, China: Educational Sciences Publishing House.

Collis, K. F., Romberg, T. A., & Jurdak, M. E. (1986). A technique for assessing mathematical problem solving ability. *Journal for Research in Mathematics Education*, *17,* 206-221.

Dhombres, J. (1993). Is one proof enough? Travels with a mathematician of the Baroque period, *Educational Studies in Mathematics, 24*, 401-419.

Goetting, M. (1995). *The college students' understanding of mathematical proof,* Unpublished doctoral dissertation, University of Maryland.

Hoffer, A. (1981). Geometry is more than proof. *Mathematics Teacher, 74*, 11-18.

Krutetskii, V. A. (1976). *The psychology of mathematical abilities in school children.* Chicago and London: University of Chicago Press.

Leikin, R., & Lev, A. (2007a). Multiple solution tasks as a magnifying glass for observation of mathematical creativity. In J-H Woo, H-C Lew, K-S Park & D-Y Seo (Eds.), *Proceedings of the 31st International Conference for the Psychology of Mathematics Education. (vol. 3, pp. 161-168).* Korea: The Korea Society of Educational Studies in Mathematics.

Leikin, R., & Lev, A. (2007b). Exploring mathematics teacher knowledge to explain the gap between theory-based recommendations and school practice in the use of connecting tasks. *Educational Studies in Mathematics, 66*, 349-371.

Leikin, R. (2009). Multiple proof tasks: Teacher practice and teacher education. Proceedings of *International Commission on Mathematical Instruction, ICMI STUDY 19, Proof and proving in mathematics education,* Taipei.

Lin, F. (2005). Modeling students' learning on mathematical proof and refutation. In Chick, H. L. & Vincent, J. L. (Eds.), *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 3-18). Melbourne: PME.

Ma, L. (1999). *Knowing and teaching elementary mathematics: teachers' understanding of fundamental mathematics in China and the United States.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Martin, W. G., & Harel, G. (1989). Proof frames of pre-service elementary teachers, *Journal for Research in Mathematics Education, 20*, 41-51.

Marton, F. (1981). Phenomenography – Describing conceptions of the world around us. *Instructional Science, 10,* 177-200.

Mathematics Textbook Developer Group for Elementary School. (2003). *Mathematics* (in Chinese). Beijing: People's Education Press.

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

Selden, A., & Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem. *Journal for Research in Mathematics Education, 34,* 4-36.

Senk, S. L. (1989). van Hiele levels and achievement in writing geometry proofs. *Journal for Research in Mathematics Education, 20*, 309-321.

Shimizu, Y. (2009).Characterizing exemplary mathematics instruction in Japanese classrooms from the learner's perspective, *ZDM Mathematics Education, 41,* 311-318.

Silver, E. A., Ghousseini, H., Gosen, D., Charalambous, C., & Strawhun, B.T.F. (2005). Moving from rhetoric to praxis: Issues faced by teachers in having students consider multiple solutions for problems in the mathematics classroom. *Journal of Mathematical Behavior, 24*, 287-301.

Silver, E. A., Leung, S. S., & Cai, J. (1995). Generating multiple solutions for a problem: A comparison of the responses of U.S. and Japanese students. *Educational Studies in Mathematics, 28*, 35-54.

Siu, M. (1993). Proof and pedagogy in ancient China: Examples from Liu Hui's Commentary on Jiu Zhang Suan Shu , *Educational Studies in Mathematics, 24,* 345-357.

Star, J., & Bethany, R. (2009). It pays to compare: An experimental study on computational estimation. *Journal of Experimental Child Psychology, 102,* 408-426.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom.* New York: Free Press.

Stylianides, G. J., & Stylianides, A. J. (2009a). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education, 40*, 314-352.

Stylianides, G. J., & Stylianides, A. J. (2009b). Ability to construct proofs and evaluate one's constructions. *Proceedings of International Commission on Mathematical Instruction, ICMI STUDY 19, Proof and Proving in Mathematics Education*, Taipei.

Stylianides, G. J., Stylianides, A. J., & Philippou, G. N. (2007). Preservice teachers' knowledge of proof by mathematical induction. *Journal of Mathematics Teacher Education, 10,* 145-166.

Sun, X, (2009). An experiment for enhancement of a trapezoid area formula proof constructions of student teachers by 'One Problem, Multiple Solutions,' *Proceedings of the International Commission on Mathematical Instruction, ICMI STUDY 19, Proof and Proving in Mathematics Education,* Taipei.

Sun, X, (2011). An insider's perspective variation problems and their cultural grounds in Chinese curriculum practice, *Journal of Mathematics Education, 4*(1), 101-114.

Watson, A. & Mason, J. (2005). *Mathematics as a constructive activity: Learners generating examples*. Mahwah USA: Erlbaum.

Zazkis, R., & Leikin, R. (2007). Generating examples: From pedagogical tool to a research tool. *For the Learning of Mathematics, 27*, 11-17.

Zazkis, R., & Leikin, R. (2008). Exemplifying definitions: A case of a square. *Educational Studies in Mathematics, 27,* 11-17.

Zhang J.Z. (2005). *From mathematical education to educational mathematics*, Beijing: Juvenile and Children Publishing House.

Acknowledgement

This study was funded by the Research Committee, University of Macau, Macao, China [MYRG092(Y1-L2)-FED11-SXH].The opinions expressed in the article are those of the author.