2016 NCTM RESEARCH CONFERENCE
April 11-13 • San Francisco

Paper Title: “It won’t work every time”: The refutations of pre-service elementary teachers

Author(s): Michael Jarry-Shore & Marta Kobiela

Session Title: “It won’t work every time”: The refutations of pre-service elementary teachers

Session Type: Brief Research Report

Presentation Date: April 13, 2016

Presentation Location: San Francisco, California

# “It won’t work every time”: The refutations of pre-service elementary teachers

## Introduction

In today’s reform mathematics classroom, teachers are to support students in both generating mathematical conjectures and refuting those that are false through the use of counterexamples (Common Core State Standards Initiative, 2010; Lampert, 1990; Stylianides & Ball, 2008). Rather than follow a series of rules provided to them by their teachers, students today are to devise their own conjectures that will guide them through their mathematical work. Many student conjectures, however, while sound when applied to particular cases, are certain not to generalize (Stafylidou & Vosniadou, 2004). As a result, teachers today must be capable of demonstrating for students why their conjectures might fail to generalize. One response to students’ false conjectures is a thoughtful counterexample.

Although studies show that experienced teachers have greater skill in generating counterexamples than their less experienced pre-service counterparts (Peled & Zaslavsky, 1997; Zaslavsky & Peled, 1996), not all experienced teachers display proficiency in this realm. As such, it is far from guaranteed that teachers will develop the skill in generating counterexamples that they now require merely through on-the-job experience. One approach to ensuring teachers possess such skill is through teacher education. However, while several studies have examined the refutations of pre-service teachers at the secondary level (Peled & Zaslavsky, 1997; Zaslavsky & Peled, 1996), few have been conducted at the elementary level, leaving the elementary pre-service teacher educator with little guidance in this area.

Accordingly, this study examined pre-service elementary teachers' refutations of students' false conjectures for the process of comparing fractions. The comparison of fractions was chosen as, in the past, students have typically been told to convert fractions into equivalents with common denominators in order to compare them, whereas today, students are encouraged to devise their own conjectures when engaged in this process. By consequence, teachers are likely to not only encounter a range of student conjectures when teaching the comparison of fractions, but must be able to generate effective counterexamples when teaching such a unit, as well.

## Theoretical Perspectives

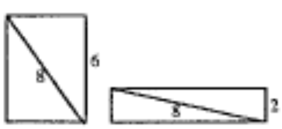
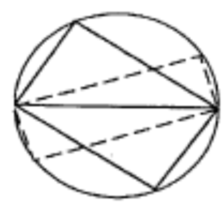
### Pedagogical Power of Counterexamples

In the math education literature, scholars have argued that two counterexamples that are both mathematically correct may differ in terms of their pedagogical power. While a mathematician may regard two counterexamples as roughly equivalent in a mathematical sense, an educator may identify differences between two such counterexamples that render one more helpful in supporting student understanding.

Almost two decades ago, Peled and Zaslavsky (1997) examined the knowledge of refutation displayed by both in-service and pre-service secondary teachers. In their study, teachers were presented with the task of refuting false student conjectures through the use of counterexamples (Peled & Zaslavksy, 1997). For instance, one false conjecture presented to participants in this study read as follows: two rectangles having congruent diagonals are congruent. The counterexamples generated by participants to refute this and other conjectures were classified as either adequate (i.e., successful in refuting the conjecture) or inadequate. Adequate counterexamples were further classified as specific, semi-general, or general.

Specific counterexamples succeeded in refuting the incorrect student conjecture, but failed to hint at some procedure to follow in generating many more such counterexamples (see Table 1). Semi-general counterexamples, on the other hand, not only refuted the false student conjecture, but hinted at some procedure for generating other similar such counterexamples, as well. Finally, general counterexamples succeeded not only in refuting the false student conjecture, but also suggested a procedure to follow in generating an infinite number of counterexamples to refute the conjecture.

Table 1
*Specific, semi-general, and general counterexamples for the rectangles conjecture*

| Rectangles conjecture<br>"Two rectangles having congruent diagonals are congruent" | | |
|---|---|---|
| Specific counterexample | Semi-general counterexample | General counterexample |
|  |  |  |

Peled and Zaslavsky argued that counterexamples should serve two purposes: 1) to demonstrate why a claim is false and 2) to suggest a procedure for generating many more such counterexamples. Counterexamples that hint at a procedure to follow in generating many more counterexamples, they argue, are better able to explain why a conjecture is false, and therefore, have greater pedagogical power than those that fail to do so. In failing to suggest a procedure for generating many counterexamples, a specific counterexample, while correct mathematically, has less pedagogical power than either semi-general or general counterexamples. While a specific counterexample can successfully refute a false conjecture, it may leave a student with the impression that such a counterexample represents a single "pathological case" (p. 51) rather than an instance of a group of cases. If presented, however, with a general counterexample, a student is faced with a wealth of disconfirming evidence refuting his/her conjecture and is, therefore, more likely to abandon his/her flawed reasoning.

## Accessibility of Counterexamples

**Mathematical complexity.** The framework of Peled and Zaslavsky has done much to further the field's understanding of how two counterexamples equivalent mathematically possess different pedagogical power. However, this framework, while incredibly useful, fails to capture a number of additional pedagogical distinctions between two given counterexamples.

In recent years, scholars in another area have devoted significant attention to exploring and unpacking a unique form of mathematical knowledge required by those who teach the subject to students, what is commonly referred to as mathematical knowledge for teaching (Hill, Rowan, & Ball, 2005). A key component of this unique mathematical knowledge is skill with exemplification, specifically, the selection of appropriate examples for the teaching of a given mathematical concept. According to Ball, Thames, and Phelps (2008), the teaching of mathematics involves, among other things, "considering what numbers are strategic to use in an example" (p. 398). To illustrate their point, Ball et. al provide the example of the subtraction problem, 307-168, and describe how teachers skilled with exemplification would not select numbers like those in this particular problem haphazardly when teaching subtraction to students. Such teachers would acknowledge that this particular subtraction problem requires two regroupings, thereby making the problem less than ideal for an initial discussion of multi-digit subtraction. Slight modifications to this particular subtraction problem, however, could preclude the need for regrouping and thereby render the problem both less mathematically complex and more accessible to children. Perhaps a teacher might choose to save the problem 307-168 until many days into a unit on multi-digit subtraction, starting such a unit instead with simpler problems like 368-107, which, unlike 307-168, requires no regrouping.

To a layperson, the problems 307-168 and 368-107 may appear roughly equivalent, yet to a teacher skilled in the realm of exemplification, these problems present vastly different demands on a young learner of mathematics. The intentional and thoughtful selection of numbers is likely relevant not only in the domain of arithmetic, however, but when generating counterexamples, as well.

**Mirroring of logic employed in comprehending both confirming- and counter-examples.** In addition to mathematical complexity, counterexamples are likely to vary along other dimensions, as well. Studies of students' responses to refutations of their false conjectures demonstrate that, when faced with conflicting evidence, students may choose to amend, rather than abandon, their faulty conjectures. Zazkis and Chernoff (2008) shared a revealing classroom vignette from their work with a group of pre-service elementary teachers revolving around the comparison of fractions. In this vignette, one pre-service teacher named Tanya described how, when comparing two fractions, one need only find the difference between each fraction's numerator and denominator, as the fraction with the larger such difference will always be smaller. Tanya accompanied her description of this procedure with the confirming case of 2/7 & 5/7. As Tanya's teacher educator recognized that this procedure would not result in the correct identification of the larger of two fractions in every fraction comparison, she proceeded to present Tanya with a series of counterexamples. While correct mathematically, several of these counterexamples failed to support Tanya in abandoning her faulty conjecture.

Tanya was first presented with the counterexample 1/2 & 2/4, to which she responded, "they never give you fractions that are the same to compare" (p. 204), likely referring to the fact that 1/2 & 2/4 are equivalent fractions. Realizing that this counterexample had failed to lead Tanya to relinquish her faulty conjecture, Tanya's teacher educator next presented Tanya with another counterexample, 5/6 & 6/7. Again, Tanya displayed resistance when faced with this disconfirming evidence, noting that her conjecture wouldn't apply "if the difference is the same" (p. 204).

The counterexamples provided by Tanya's teacher educator appeared not to mirror Tanya's confirming case well enough to lead her to abandon her faulty conjecture. Unlike the teacher educator's counterexample of 1/2 & 2/4, Tanya's confirming case, 2/7 & 5/7, involved two fractions that represented different, not equivalent, quantities. Furthermore, unlike the teacher educator's other counterexample, 5/6 & 6/7, Tanya's confirming case involved two fractions missing a different number of pieces, not the same number. The logic employed by Tanya in comparing fractions was likely that, "in any fraction comparison, the fraction missing more pieces is the smaller of the two fractions." The first counterexample presented by Tanya's teacher educator, 1/2 and 2/4, did not mirror Tanya's logic very well, as the counterexample consisted of two fractions that, unlike her confirming case, were not different in size, but were the same size. The second counterexample of 5/6 and 6/7, on the other hand, also failed to mirror Tanya's logic, as, unlike her confirming case, 2/7 & 5/7, this counterexample consisted of a pair of fractions missing the same number of pieces.

**Mapping of confirming examples and counterexamples.** As the vignette involving Tanya continues, Zazkis and Chernoff describe how Tanya was next presented with a third correct counterexample, 9/10 & 91/100. When faced with this particular counterexample, Tanya again resisted, choosing to amend her conjecture and arguing that it still worked, simply not with numbers that were "ridiculously large" (p. 204). According to Zazkis and Chernoff, the "relative size of numbers" in counterexamples like 9/10 & 91/100 compared to those comprising the fractions in Tanya's confirming case, 2/7 & 5/7, likely played little role in uprooting Tanya's faulty reasoning. And yet, Tanya's emphatic disregard for the counterexample involving "ridiculously large numbers" suggests that the disparity between the numbers comprising the confirming and disconfirming cases could very well have played a role in her initial dismissal of this third counterexample. Presumably, what Tanya was referring to with the phrase "ridiculously large" were the double-digit numerator and triple-digit denominator of the fraction 91/100, which bared little resemblance to either of the fractions comprising Tanya's confirming case, 2/7 & 5/7. Similar to Tanya's confirming example, on the other hand, were still more counterexamples presented to her after 9/10 & 91/100 like 2/3 & 5/7 and 3/4 & 8/11, which "appeared to the students as more convincing than the originally suggested fractions of 9/10 and 91/100" (p. 205). While it is not entirely clear why these particular fractions were more convincing, it would not be implausible to contend that the convincing power of these counterexamples stemmed from them being comprised of fractions that, like Tanya's confirming example, consisted of reasonable, rather than ridiculous, numbers.

This study addressed three research questions:

1. What is the pedagogical power of counterexamples generated by pre-service elementary teachers in refuting students' false conjectures?
2. How do these counterexamples vary in their potential accessibility to a young learner of mathematics?

## Method

### Setting and Participants

For this study, 17 pre-service teachers from an elementary math-methods course in number and operation were recruited to take part in semi-structured interviews (Ginsburg, Jacobs, & Lopez, 1998). This methods course was the first of two taken by these pre-service teachers and was taken during the second year of participants' education program, prior to any experience teaching in a field placement. In this course, pre-service teachers spent considerable time exploring both operations on fractions and fraction equivalency concepts. Additionally, pre-service teachers engaged in in-depth explorations of pictorial models for the multiplication of fractions, as well as placed fractional numbers on number lines.

Participants were recruited from multiple sections of the course, each taught by a different instructor, in order to increase the likelihood of obtaining a diverse sample. Teachers recruited from one section of the course (n=4) participated in pilot interviews, which resulted in several revisions to the interview protocol. Students from a second (n=10) and third section (n=7) were later interviewed using a revised version of the interview protocol; the interviews of these 17 pre-service teachers were the focus of this particular paper. Typical of pre-service elementary teachers, the teachers in our sample were about 20-21 years of age and predominantly female (16/17=94%).

**Data Collection**

The 17 pre-service elementary teachers in this study took part in hour-long structured interviews. Interviews were conducted after the course from which participants were recruited had completed and grades for the course had been submitted, as it was feared that conducting the interviews while the course was still in session may have led participants to worry that their responses might adversely affect their grades. In the interviews, each pre-service teacher was presented with three measures, the first being the focus of this paper. The first measure (see Figure 1), henceforth referred to as "the fractions measure," was borrowed for use from the *Content Knowledge for Teaching Mathematics (CKT-M)*, an assessment tool designed as part of the Measures for Effective Teaching (MET) initiative. For the fractions measure, pre-service teachers were asked to determine if each of five student explanations for why 7/8 is greater than 6/9 were mathematically valid or not. The mathematical explanations provided by students A and C in this measure were the only ones that lacked mathematical validity and were thus ideal for investigating pre-service teachers' refutations.

During each interview, the interviewer first read through the measure then provided the pre-service teacher 4 minutes to think about and respond to the measure independently. After 4 minutes had elapsed, pre-service teachers were given 14 minutes to share their thinking in response to the measure. During this 14-minute time period, pre-service teachers were encouraged to discuss each of the student solutions presented in the measure in whatever order they wanted and were posed a series of probing questions intended to uncover more about their reasoning. All pre-service teachers who indicated that student C's explanation lacked mathematical validity were asked to generate a counterexample to refute the student's conjecture. In some cases, pre-service teachers' discussion of student A's explanation involved mention of counterexamples, although, due to time limitations, teachers were not explicitly asked to refute this faulty explanation.

Pre-service teachers' mathematical work and discussion of the fractions measure were captured by a video camera placed directly above the teachers' designated workspace. In addition, the interviewer took detailed field notes as pre-service teachers shared their responses and related reasoning.

**Data Analysis**

A first pass through the video-recorded interviews was conducted to gauge the pedagogical power of the counterexamples provided by pre-service teachers. While this first pass revealed distinctions in the pedagogical power of pre-service teachers' counterexamples, as will be discussed below, additional distinctions not captured using the framework borrowed from Peled and Zaslavksy (1997) were identified, resulting in additional analysis that sought to capture further distinctions regarding the accessibility of counterexamples.

**Coding-scheme development.** The framework of Peled and Zaslavsky (1997) was operationalized for use in coding the counterexamples provided by pre-service teachers in the

Mr. Lee asked his students to compare $\frac{7}{8}$ and $\frac{6}{9}$. All of his students correctly answered that $\frac{7}{8}$ is greater than $\frac{6}{9}$, but they offered a variety of responses when asked to explain their reasoning. Of the following, which student responses provide mathematically valid explanations for why $\frac{7}{8}$ is greater than $\frac{6}{9}$? For each student response, indicate whether or not it provides a mathematically valid explanation.

| | | **Provides a Mathematically Valid Explanation** | **Does Not Provide a Mathematically Valid Explanation** |
|---|---|---|---|
| A) | When you compare them, $\frac{7}{8}$ is greater than $\frac{6}{9}$ because 7 is greater than 6. | | |
| B) | You can see that $\frac{7}{8}$ is greater than $\frac{6}{9}$ because ninths are smaller than eighths, which means that $\frac{6}{9}$ is less than $\frac{6}{8}$ which is less than $\frac{7}{8}$. | | |
| C) | You just need to look at how many pieces are missing. $\frac{7}{8}$ is greater than $\frac{6}{9}$ because $\frac{7}{8}$ is only missing one piece from the whole, but $\frac{6}{9}$ is missing three pieces from the whole. | | |
| D) | I think $\frac{7}{8}$ is greater than $\frac{6}{9}$ because $\frac{7}{8}$ has more pieces than $\frac{6}{9}$ and those pieces are larger. | | |
| E) | $\frac{7}{8}$ is greater than $\frac{6}{9}$ because $\frac{6}{9}$ is equal to $\frac{2}{3}$, and because $\frac{1}{3}$ is greater than $\frac{1}{8}$, $\frac{2}{3}$ is farther away from 1 than $\frac{7}{8}$ is. | | |

*Figure 1*. Fractions measure. Reprinted from "Content knowledge for teaching: Mathematics grades 4-5 assessment," G. Phelps & D. Gitomer, 2012, *Educational Testing Service*, 11. 

present study. However, as initial viewings of the interviews of a subsample of pre-service teachers revealed that not all counterexamples provided by participants could be classified as specific, semi-general, or general, additional categories were added to our operationalized framework. As such, this operationalized framework, developed via constant comparison (Glaser & Strauss, 1965), contained not three, but six levels (Table 2): R - no counterexample provided; $R_0$ - inadequate counterexample; $R_1$ - hinted at, but didn't provide, counterexample; $R_2$ - specific counterexample; $R_3$ - semi-general counterexample; and $R_4$ - general counterexample.

Instances assigned a code of R, no counterexample provided, received such a code either because the pre-service teacher was not asked to provide a counterexample – relevant for Solution A – or because the pre-service teacher believed the conjecture to be refuted was actually valid. Instances assigned a code of $R_0$, inadequate counterexample, received this code as the counterexample was actually consistent with the false student conjecture. Instances assigned a code of $R_1$ - hints at, but doesn't provide a counterexample - were coded as such, as the pre-service teacher suggested that he/she or one of his/her students could create a counterexample to refute the conjecture under consideration, but no such counterexample was actually generated. Instances assigned a code of $R_2$, specific counterexample, received such a code as the pre-service teacher provided a counterexample that did refute the false student conjecture, yet did not describe some procedure to follow in generating more such counterexamples. Instances that received a code of $R_3$, semi-general counterexample, received such a code as the pre-service teacher not only provided a counterexample that refuted the false student conjecture, but described some procedure to be used in generating other similar such counterexamples. Finally, instances assigned a code of $R_4$, general counterexample, were coded as such, as the pre-service teacher both provided a counterexample that refuted the false student conjecture and described some procedure that could be followed in generating an infinite number of similar such counterexamples.

Table 2
*Refutation coding scheme*

| Level of Refutation | Description |
|---|---|
| R: No counterexample provided because not asked or conjecture viewed as valid | Pre-service teacher does not provide a counterexample because they weren't asked to or they believe the conjecture to be refuted is actually valid. |
| $R_0$: Counterexample does not refute, but is consistent with conjecture | Pre-service teacher provides an example that, rather than refute the conjecture under consideration, is consistent with this conjecture. |
| $R_1$: Hints at, but doesn't provide, counterexample | Pre-service teacher suggests that he/she or one of his/her students could create a counterexample to refute the conjecture under consideration, but no such counterexample is actually generated. |
| $R_2$: Specific counterexample | Pre-service teacher provides a counterexample that refutes the conjecture under consideration, but in doing so, does not hint at a procedure to follow in generating many similar such counterexamples. |
| $R_3$: Semi-general counterexample | Pre-service teacher provides counterexample that refutes the conjecture under consideration and hints at a procedure for generating many similar such counterexamples. |
| $R_4$: General counterexample | Pre-service teacher provides counterexample that refutes the conjecture under consideration and makes evident a procedure for generating an infinite number of similar such counterexamples. |

**Determining instances and the coding process.** Once our operationalized framework was ready for use, pre-service teachers' discussion of student A's and C's explanations were parsed into instances, which began when one of these two explanations was first addressed and ended when discussion of that particular explanation was over. Next, a diverse subsample of pre-service teacher interviews were coded by two raters and agreement between the codes of each rater was assessed. Wherever disagreement existed, discussion ensued until consensus was reached.

As an example of such disagreement, consider the example of a pre-service teacher providing a pair of equivalent fractions to refute student A's conjecture that, in any fraction comparison, the fraction with the larger numerator is the larger of two fractions. One rater coded this particular instance at the "general" level, arguing that, in providing a pair of equivalent fractions, the pre-service teacher suggested that any pair of equivalent fractions, of which there are an infinite number, could be used to refute student A's conjecture. However, the second rater disagreed, arguing that the pre-service teacher would need to make explicit mention of this in their response, rather than just provide a pair of equivalent fractions, in order for their refutation to be coded as "general."

Wherever disagreement arose, revisions to the operationalized framework, including examples used to exemplify each level of the framework, were made. Revised frameworks were then used to code additional subsamples of pre-service teachers' interviews. This process of coding, calculating agreement between raters, and revising the coding scheme, continued until agreement between raters' codes exceeded 80%, at which point, the final version of the operationalized framework was used by one rater to code every instance.

**Accessibility of Counterexamples**

As stated previously, after coding the pedagogical power of pre-service teachers' counterexamples, it was felt that additional distinctions in these counterexamples existed that were not captured using our operationalized framework. Efforts were thus made to capture additional distinctions in the accessibility of counterexamples. To accomplish this objective, instances coded at the same level in terms of pedagogical power were subjected to further scrutiny. To gauge accessibility, counterexamples were quantified along three dimensions: a) mapping, b) logic, and c) complexity. Mapping was defined as the degree to which the numbers comprising the fractions in a given counterexample mirrored those in the student's confirming example, 7/8 & 6/9. Logic, on the other hand, gauged how closely the reasoning employed in understanding a counterexample matched that employed by the student who authored the false claim being refuted. Finally, complexity was a measure of the number of mathematical concepts conceivably involved in understanding a given counterexample.

**Mapping.** As stated above, mapping was defined as the degree to which the numbers comprising the fractions in a given counterexample mirrored those in the student's confirming example, 7/8 & 6/9. Mapping of counterexamples to student A's conjecture was quantified by first calculating the difference between the numerators in both the first fractions of the confirming example and counterexample, then calculating the difference between the numerators in both the second fractions of the confirming example and counterexample, and finally, summing these two values (Table 3). For example, for the counterexample 6/8 & 3/4, mapping would have been quantified as follows: $7 - 6 = 1$, $6 - 3 = 3 \Rightarrow 1 + 3 = 4$. In quantifying counterexamples to student A's conjecture, the focus was on the numerators alone, as this is what student A was focused on in making his/her conjecture, and as such, what pre-service teachers were themselves focused on in designing counterexamples to refute this particular conjecture.

In quantifying counterexamples to student C's conjecture, first, the difference was found between the number of pieces missing from the first fractions of both the confirming example and counterexample, then the difference was found between the number of pieces missing from the second fractions of both the confirming example and the counterexample, and finally, these values were added. Following this procedure, the counterexample 1/2 & 2/4, for example, would have received a mapping value of 1, calculated as follows: 7/8 & 6/9 are missing 1 and 3 pieces, respectively, whereas 1/2 & 2/4 are missing 1 and 2 pieces, respectively, therefore, 1 - 1 = 0, 3 - 2 = 1 => 0 + 1 = 1. When quantifying counterexamples to student C's conjecture, the focus was on the number of pieces missing, as this is what student C was focused on in making his/her conjecture, and by consequence, what pre-service teachers were themselves focused on in designing counterexamples for this particular student conjecture.

Table 3

*Quantifying "mapping" of counterexamples*

| Student Explanation | Confirming example | Counter-example | Procedure for quantifying "mapping" | Mapping score |
|---|---|---|---|---|
| A | 7/8 & 6/9 | 6/8 & 3/4 | Difference between numerators in 1st fractions: 7 - 6 = 1<br><br>Difference between numerators in 2nd fractions: 6 - 3 = 3<br><br>Sum of differences: 1 + 3 = 4 | 4 |
| C | 7/8 & 6/9 | 1/2 & 2/4 | Difference between # of pieces missing in 1st fractions: 1 - 1 = 0<br><br>Difference between # of pieces missing in 2nd fractions: 3 - 2 = 1<br><br>Sum of differences: 0 + 1 = 1 | 1 |

**Logic.** As described above, logic was a measure of how closely the reasoning employed in understanding a pre-service teacher's counterexample matched that employed by the student who authored the false claim the counterexample refuted. Student A's logic was that, when comparing two fractions, *the fraction with the larger numerator is larger*. As such, counterexamples that followed a similar logic received stronger logic scores (Table 4). Any counterexample that successfully refutes student A's conjecture, however, will necessarily diverge from student A's logic somewhat, otherwise, the counterexample would confirm, rather than refute, the student's conjecture. That being said, when quantifying, it was determined that the less a counterexample diverged from student A's logic, the stronger logic score it would receive. If, for example, a counterexample followed the logic that, in a fraction comparison, the fraction with the larger numerator is not the larger, but the smaller fraction, it was assigned a value of 0, the best value obtainable, for logic. A counterexample like 7/20 & 1/2 follows student A's logic closely, as the fraction 7/20 has a larger numerator, yet is smaller. If, on the other hand, a counterexample followed the logic that the fraction with the smaller numerator is larger, the counterexample was assigned a value of 1. A counterexample like 6/9 & 7/100, for example, follows this particular logic. However, a student like student A may be thrown by such a counterexample, for he/she was talking about fractions with larger, not smaller, numerators, when making his/her conjecture. As such, a counterexample like 6/9 & 7/100 received a weaker logic score. Finally, if a counterexample followed the logic that, in comparing two fractions, the fraction with the larger numerator is the same size as the fraction with the smaller denominator, the counterexample was assigned a value of 2. A counterexample like 9/12 & 3/4, for example, is similar to the confirming example, 7/8 & 6/9, as it consists of one fraction with a larger numerator than the other. However, as the fraction pair 9/12 & 3/4 consists of fractions that, unlike the confirming example of 7/8 & 6/9, are the same size, a counterexample like this received the weakest logic score.

Table 4

*Quantifying "logic" of counterexamples*

| Student Explanation | Confirming example | Counter-example | Mirroring of student's and pre-service teacher's "logic" | Logic score |
|---|---|---|---|---|
| A | 7/8 & 6/9 | 7/20 & 1/2 | • Student's logic: The fraction with the *larger* numerator is the *larger* fraction.<br>• Pre-service teacher's logic: The fraction with the *larger* numerator is the *smaller* fraction | 0 |
| | 7/8 & 6/9 | 6/9 & 7/100 | • Student's logic: The fraction with the *larger* numerator is the *larger* fraction.<br>• Pre-service teacher's logic: The fraction with the *smaller* numerator is the *larger* fraction. | 1 |
| | 7/8 & 6/9 | 9/12 & 3/4 | • Student's logic: The fraction with the *larger* numerator is the *larger* fraction.<br>• Pre-service teacher's logic: The fraction with the *larger* numerator is the *same size* as the other fraction. | 2 |
| C | 7/8 & 6/9 | 1/2 & 7/10 | • Student's logic: The fraction missing *fewer* pieces is the *larger* fraction.<br>• Pre-service teacher's logic: The fraction missing *fewer* pieces is the *smaller* fraction. | 0 |
| | 7/8 & 6/9 | 3/4 & 9/12 | • Student's logic: The fraction missing *fewer* pieces is the *larger* fraction.<br>• Pre-service teacher's logic: The fraction missing *fewer* pieces is *equivalent* to the fraction missing more pieces | 1 |
| | 7/8 & 6/9 | 5/8 & 1/2 | • Student's logic: The fraction missing *fewer* pieces is the *larger* fraction.<br>• Pre-service teacher's logic: The fraction missing *more* pieces is the *larger* fraction. | 2 |
| | 7/8 & 6/9 | 1/2 & 3/4 | • Student's logic: The fraction missing *fewer* pieces is the *larger* fraction.<br>• Pre-service teacher's logic: The fractions are missing the *same number* of pieces, but are *different sizes*. | 3 |

Student C's logic posited that, when comparing two fractions, *the fraction missing fewer pieces is the larger fraction*. As such, counterexamples that followed a similar such logic received better logic scores. However, as noted above, the logic of any such counterexample would, by virtue of being a counterexample, necessarily diverge from that underlying student C's conjecture. As an example, a counterexample like 1/2 & 7/10, which follows the logic that the fraction missing fewer pieces is smaller, was assigned a value of 0, the best logic score attainable. Up until the final conclusion (i.e., "smaller" instead of "larger"), such a counterexample follows student C's logic exactly, thereby earning the counterexample a strong logic score. If, on the other hand, a counterexample followed the logic that the fraction missing fewer pieces is equivalent to the fraction missing more pieces, it was assigned a value of 1. Such counterexamples were assigned a value of 1 because they followed student C's logic initially, yet, as they consisted of equivalent fractions, not fractions that differed in size like 7/8 & 6/9 in the confirming case, they received a worse logic score. If, on the other hand, a counterexample followed the logic that the fraction missing more pieces is larger, the counterexample was assigned a value of 2. Such counterexamples received a weaker logic score as, unlike student C, they focused on the fraction missing more, not fewer, pieces, which we anticipate might confuse a student like student C. Finally, if a counterexample consisted of two fractions each missing an equal number of pieces, the counterexample was assigned a value of 3. Such counterexamples received the weakest possible logic score, as, unlike student C's confirming example, they consisted of fractions both missing the same number of pieces.

When quantifying logic, the order of the fractions provided in the counterexample was maintained, as it was determined that this order may not have been chosen arbitrarily by the pre-service teacher. Furthermore, we believed that two counterexamples comprised of the same fractions, but in reverse order, may mirror the student's logic quite differently. As an example, while reversing the order of the fractions in the counterexample 7/20 & 6/12 to 6/12 & 7/20 may appear inconsequential, the latter counterexample, 6/12 & 7/20, could potentially throw a student like student A, as his/her confirming example, 7/8 & 6/9, is comprised of two fractions in which the first fraction, not the second one, consists of the larger numerator. Similarly, reversing the order of the fractions in the counterexample 1/2 & 4/6 to 4/6 & 1/2 would diminish the degree to which the counterexample mirrors student C's logic, as, unlike the confirming example 7/8 & 6/9, in the latter counterexample, the first fraction, not the second, is missing more pieces.

**Mathematical complexity.** Mathematical complexity was quantified in order to capture further how two counterexamples, despite being correct mathematically or equally "pedagogically powerful," may not be equally accessible to a student. Complexity was equated with the number of mathematical concepts a student would conceivably need to understand in order to make sense of a given counterexample. As the purpose in providing a counterexample is to support a student in coming to see why their conjecture is false, not to engage them in a series of lessons on a number of related mathematical concepts, those counterexamples that precluded the need for such lessons received the best scores. In quantifying complexity, each concept involved in comprehending a given counterexample was assigned a value of 1 (Table 5). Hence, to refute student A's conjecture, a counterexample like 6/8 & 3/4 would receive a complexity score of 1, as this counterexample would require one to comprehend only one concept, that of equivalent fractions. A counterexample like 6/8 & 5/6, on the other hand, would require a student to comprehend several concepts, not one, and as such, received a worse score. As it is not immediately evident which of the two fractions in the counterexample, 6/8 & 5/6 is larger, a student would conceivably need to understand several concepts before they could comprehend the counterexample, specifically: (a) having a common denominator makes fractions easier to compare, (b) a common denominator can be found by converting fractions into equivalent fractions (e.g., 6/8 = 36/48 and 5/6 = 40/48), and (c) equivalent fractions consist of different numbers, but represent the same quantity. As such, a counterexample like 6/8 & 5/6 is far more complex mathematically than the equally correct counterexample of 6/8 & 3/4.

The complexity of counterexamples to student C's conjecture was calculated by following the same procedure as that described above for counterexamples to student A's conjecture.

Table 5
*Quantifying "complexity" of counterexamples*

| Student Explanation | Confirming example | Counter-example | Concepts a student would conceivably need to know to comprehend the counterexample | Complexity score |
|---|---|---|---|---|
| A | 7/8 & 6/9 | 6/8 & 3/4 | • Equivalent fractions | 1 |
| | 7/8 & 6/9 | 6/8 & 5/6 | • Having common denominators facilitates the comparison of fractions<br>• Common denominators<br>• Equivalent fractions | 3 |

## Results

### Pedagogical Power of Counterexamples

The participants in this study refuted student conjectures using almost exclusively counterexamples coded as "specific." Only once did a pre-service teacher provide a semi-general counterexample and not once was a "general" counterexample provided (see Table 6). While pre-service teachers were not asked explicitly to provide a counterexample to refute student A's false conjecture, 6 pre-service teachers did provide adequate counterexamples when discussing student A; of these 6 counterexamples, all but one were specific. On the other hand, when asked explicitly to refute student C’s conjecture, 15 of 17 pre-service teachers provided adequate counterexamples, all 15 of which were specific.

Table 6
*Specific, semi-general, and general counterexamples*

| Pre-Service Teacher* | Counterexample for Student A ($R_0$, $R_1$, $R_2$, $R_3$, or $R_4$) | Counterexample for Student C ($R_0$, $R_1$, $R_2$, $R_3$, or $R_4$) |
|---|---|---|
| Emilia | R | $R_2$ |
| Ella | $R_3$ | $R_0$, $R_1$, & ($R_2$) |
| Claudia | $R_0$ | $R_0$ |
| Nina | R | $R_2$ |
| Delaney | R | $R_2$ |
| Olga | $R_0$ | R & $R_2$ |
| | $R_2$ | |
| Cynthia | $R_2$ | $R_2$ |
| Maria | R | $R_2$ |
| Katherine | R | $R_2$ & $R_1$ |
| Melanie | $R_2$ | $R_2$ |
| Ashley | R | $R_2$ |
| Annabelle | R | $R_0$ & $R_2$ |
| | $R_2$ | |
| Brian | R | $R_1$ & $R_2$ |
| Caroline | R | ($R_2$) |
| Esther | $R_2$ | $R_2$ |
| | $R_2$ | |
| Lane | $R_0$ | $R_0$ |
| Veronica | R | $R_2$ |

R: No counterexample provided, $R_0$: Counterexample does not refute, but is consistent with conjecture, $R_1$: Hints at, but doesn’t provide, counterexample, $R_2$: Specific counterexample, $R_3$: Semi-general counterexample, $R_4$: General counterexample

*All names are pseudonyms.

*Note.* If a pre-service teacher’s thinking was supported by the interviewer in such a way as to render the validity of this thinking questionable, the refutation code assigned was placed in parentheses.

*Note.* If a pre-service teacher provided evidence of reasoning at multiple levels in a given instance, all such levels were indicated, not merely the highest of such levels.

*Note.* In discussing student A's false conjecture, three pre-service teachers, Olga, Annabelle, and Esther, provided different counterexamples at two separate stages in their discussion of student A.

As pre-service teachers were not asked to refute student A's false conjecture, a preponderance of "R" codes - no counterexample provided - are found under the column "Counterexample for Student A" in Table 6. On the other hand, given that every pre-service teacher was asked to refute student C's false conjecture, only one such "R" code is found under the column in Table 6 labeled, "Counterexample for Student C."

As an example of a specific counterexample, consider the counterexample provided by pre-service teacher Annabelle in her discussion of student A's false conjecture that, "7/8 is greater than 6/9 because 7 is greater than 6":

> Yeah, um, because, in fractions, it doesn't, the numerator being higher than the next one doesn't mean that's the larger fraction. In, for instance, what else, what else can I do that? Okay, 4/8 or ½, they would be, say the 4 is bigger than the 1, but it's actually the same fraction.

Here, Annabelle refuted student A's conjecture using a specific case, that of 4/8 & 1/2. While it could be argued that Annabelle may have been recommending here that any pair of equivalent fractions would constitute an effective counterexample to student A's conjecture, given that such a procedure was not stated explicitly, the counterexample could not be coded as anything more than specific.

As another example, consider the following specific counterexample provided by pre-service teacher Maria in refuting student C's false conjecture that, "7/8 is greater than 6/9 because 7/8 is only missing one piece from the whole, but 6/9 is missing three pieces from the whole":

> Um, I think that, that's wrong because say you have 1/16, and you have, or you have 15/16 and you have 7/8, they're both missing one, but the size of the 1/16 that's missing from the 15/16 is much smaller.

Again, the pre-service teacher here refutes the false student conjecture using a specific case, that of 15/16 & 7/8. Maria here appears to be suggesting that, in order to generate a counterexample that refutes student C's conjecture, all one must do is create a pair of fractions each one piece short of whole. However, again, as such a procedure can only be inferred from what Maria said and wasn't stated explicitly by her, Maria's counterexample can not be coded semi-general or general.

The single semi-general counterexample encountered in this study was generated by pre-service teacher Ella in refuting student A's conjecture that, "7/8 is greater than 6/9 because 7 is greater than 6":

> So, if the denominator was, if it was 7 out of 20, and then here it was 6 out of 12, yeah, that's great that 7 is greater than 6, that all makes sense, but it depends on the denominator. So, for example, if the denominator under 7 is greater than the denominator under 6, then this 7 on 20, for example, is smaller than this 6 on 12.

Here, Ella hinted at a procedure to follow in generating additional counterexamples similar to the one she provided by stating that, if one were to create two fractions, one with a numerator of 7 and another with numerator of 6, and further, ensure that "the denominator under 7 is greater than the denominator under 6," the two fractions would constitute an adequate counterexample.

As Ella’s counterexample suggested a procedure for generating several, but not every counterexample (i.e., only those with numerators of 7 and 6), it was coded “semi-general,” not “general.” Following Ella’s procedure, one could generate a fraction pair like 7/10 & 6/8, in which “the denominator under 7 is greater than the denominator under 6,” and that would constitute an adequate counterexample. However, using Ella’s procedure, one could also generate another fraction pair that would fail to refute student A’s conjecture. For example, the fraction pair 7/10 & 6/9 could be generated following Ella’s procedure, as in this fraction pair, “the denominator under 7 is greater than the denominator under 6.” However, this fraction pair confirms rather than refutes student A’s conjecture, as the fraction with the larger numerator, 7/10, is actually the larger of the two fractions. Although Ella’s procedure is only sometimes successful, we found its attempt at generalization noteworthy.

While no general counterexamples were provided by pre-service teachers in refuting the false conjectures of students A and C, it is worth mentioning what sort of counterexample would have been classified as such. For example, suppose that Annabelle, the pre-service teacher who refuted student A’s conjecture with the counterexample of 4/8 & 1/2, had also described how any pair of equivalent fractions would successfully refute student A’s false conjecture. If this had been the case, Annabelle’s counterexample would have been coded “general,” as her response would have suggested a procedure to follow in generating an infinite number of adequate counterexamples for refuting student A’s false conjecture. As another example, suppose that Maria, the pre-service teacher who refuted student C’s false conjecture using the counterexample of 15/16 & 7/8, had gone on to say that any pair of fractions each one unit fraction short of making a whole could be used to refute student C’s conjecture. If Maria had accompanied her counterexample with a description of such a procedure, her counterexample would have been coded “general,” as this procedure would enable one to generate an infinite number of adequate counterexamples to refute student C’s false conjecture.

**Accessibility of Counterexamples**

As mentioned above, in coding pre-service teachers’ counterexamples using an operationalized framework for pedagogical power, we failed to capture additional and noteworthy distinctions in these counterexamples. While the vast majority of counterexamples provided by pre-service teachers were coded “specific,” there was great variety in these specific counterexamples in terms of their potential accessibility, or lack thereof, to a young child learning mathematics. In particular, specific counterexamples varied in terms of: a) mapping or the degree to which the numbers comprising the fractions in a given counterexample mirrored those in the student’s confirming example, 7/8 & 6/9, b) logic or how closely the reasoning employed in understanding a counterexample matched that employed by the student who authored the false claim being refuted, and c) complexity or the number of mathematical concepts conceivably involved in understanding a given counterexample.

While coding is still underway, some preliminary results from the coding of diverse subsamples of pre-service teachers can be reported here (Table 7). As an example, consider the counterexample of 6/8 & 5/6 provided by pre-service teacher Esther to refute student A’s false conjecture. In following the process for quantifying “mapping” described in the methods section of this paper, Esther’s counterexample was assigned a value of 2, as the difference between the numerators in the first fraction of the confirming example, 7/8, and Esther’s first fraction, 6/8, was 1, whereas the difference between the numerators in the second fraction of the confirming example, 6/9, and Esther's second fraction, 5/6, was 1. As the sum of these two differences (1 + 1 = 2) was 2, Esther's counterexample received a mapping value of 2. As the logic employed in comprehending Esther's counterexample was that the fraction with the larger numerator is the smaller fraction, Esther's counterexample received a logic score of 0, the strongest logic score possible. However, as a student would conceivably need to know several concepts in order to comprehend Esther's counterexample, the complexity score assigned to this particular counterexample was a 3. As it is not immediately apparent which of the two fractions in Esther's counterexample, 6/8 & 5/6, is the larger of the two, a student, or even an adult, for that matter, would likely need to convert the two fractions into equivalent fractions sharing a common denominator before the relative sizes of the fractions could be compared. For example, the fractions 6/8 & 5/6 could be converted into the fractions 36/48 & 40/48, respectively, which would enable one to see that 6/8 (i.e., 36/48), despite having the larger numerator, is actually the smaller of the two fractions. To get to the stage where these two fractions could be compared as such, however, would conceivably require a student to understand the following three concepts: a) converting fractions into equivalent fractions that share a common denominator facilitates their comparison, b) equivalent fractions represent equal quantities even though they are comprised of different numbers, and c) equivalent fractions can be made by multiplying the numerator and denominator by the same number (e.g., 6/8 x 6/6 = 36/48). While Esther's counterexample scored well in terms of mapping and logic, as a student would conceivably need to know 3 mathematical concepts before they could comprehend the counterexample, her counterexample scored less well in terms of complexity.

Table 7

*Mapping, logic, and complexity of specific counterexamples*

| Participant | Student Explanation (Student A or Student C) | Counter-example provided | Mapping (Disparity between example and counter example) | Logic (Agreement between student logic and logic employed by author of counter-example) | Complexity (Number of math concepts a student would need to understand to make sense of the counter-example) |
|---|---|---|---|---|---|
| Esther | A | 6/8 and 5/6 | 2 | 0 | 3 |
| Ella | A | 7/20 and 6/12 | 0 | 0 | 2 |
| Cynthia | C | 3/4 and 9/12 | 0 | 1 | 1 |
| Veronica | C | 5/8 and 1/2 | 4 | 2 | 2 |

*Note.* Only adequate counterexamples that were also coded specific ($R_2$) are included here.

By contrast, consider the counterexample of 7/20 and 6/12 provided by Ella to refute student A's false conjecture. As there is no difference between the numerators in the first fraction of Ella's counterexample and the first fraction in the confirming example, nor is there any difference between the numerators in the second fraction of Ella's counterexample and the second fraction in the confirming example, Ella's counterexample received a mapping score of 0. The logic underlying Ella's counterexample was that the fraction with the larger numerator is the smaller fraction. As such, Ella's counterexample received a logic score of 0, the best logic score possible. In order to comprehend Ella's counterexample, a student would conceivably need to comprehend that 7/20 is less than 10/20, which itself is equivalent to 1/2, and further, that 1/2 is equivalent to 6/12, therefore, 7/20 is smaller than 6/12. As such, a student would need to know two concepts in order to understand Ella's counterexample: 1) that of equivalent fractions and 2) that 1/2 is a useful benchmark for comparing fractions.

Turning to student C, consider the counterexample of 3/4 and 9/12 provided by pre-service teacher Cynthia to refute student C's false conjecture. In following the process for quantifying "mapping" specified in the methods section of this paper, this counterexample was assigned a mapping value of 0, as, like the confirming example of 7/8 & 6/9, the fractions in Cynthia's counterexample were also one and three pieces short of making a whole, respectively. As the logic employed in comprehending Cynthia's counterexample was that the fraction missing fewer pieces is the same size as the fraction missing more pieces, this counterexample received a "logic" score of 1. Finally, regarding complexity, both raters agreed that the sole concept a student would likely need to know in order to comprehend Cynthia's counterexample was that of equivalent fractions. As such, Cynthia's counterexample received a "complexity" score of 1.

By contrast, consider the counterexample of 5/8 & 1/2 provided by Veronica to refute student C's false conjecture. The fractions comprising Veronica's counterexample were three and one pieces short of making a whole, respectively. On the other hand, the confirming example, 7/8 & 6/9, consisted of two fractions that were one and three pieces short of making a whole, respectively. As such, Veronica's counterexample received a mapping score of 4 (i.e., $1 - 3 = |-2| = 2$ and $3 - 1 = 2 \Rightarrow 2 + 2 = 4$). The logic underlying Veronica's counterexample was that the fraction missing more pieces is the larger fraction, whereas student C's logic was that the fraction missing fewer pieces is larger. As such, Veronica's counterexample received a logic score of 2. In order to comprehend Veronica's counterexample, a student would likely need to know that 5/8 is greater than 4/8, and as 4/8 is equivalent to 1/2, 5/8 is greater than 1/2, even though 5/8 is missing more pieces. As a student like student C would need to comprehend two concepts in order to make sense of Veronica’s counterexample, those of equivalent fractions and the usefulness of the 1/2 benchmark, this counterexample received a complexity score of 2.

## Discussion

This study demonstrates that, unlike their counterparts at the secondary level (Peled & Zaslavsky, 1997), the pre-service elementary teachers in this sample didn’t tend to provide semi-general or general counterexamples when refuting students’ false claims. When asked to refute the false conjecture that two rectangles having congruent diagonals are congruent, 24/45 of the pre-service secondary teachers in Peled and Zaslavsky’s study provided adequate counterexamples. Of these adequate counterexamples, 7 were specific, 4 were semi-general, and 13 were general. By contrast, of the 21 adequate counterexamples provided by pre-service elementary teachers in refuting the conjectures of students A and C in the present study, all but one such counterexample was specific. This is concerning because specific counterexamples, while mathematically correct, are more likely to leave children with the impression that their false conjectures are correct except for some single pathological case. By consequence, elementary pre-service teacher educators may wish to dedicate time in teacher education courses to developing pre-service elementary teachers’ ability to generate more pedagogically powerful semi-general and general counterexamples.

An additional contribution of this study is the providing of a framework for distinguishing counterexamples according to their accessibility to a child. Existing research suggests that the design of counterexamples influences whether or not they lead students to abandon or cling to their faulty conjectures (Zazkis & Chernoff, 2008). According to our framework, counterexamples falling in the same category of Peled & Zaslavsky's framework can be further distinguished in terms of: a) the degree to which the numbers comprising a counterexample map onto the numbers comprising a confirming example, b) the manner in which the logic underlying a counterexample mirrors that employed by a student who has authored a faulty conjecture, and c) the mathematical complexity of a counterexample. While yet to be empirically validated, this new framework suggests that generating counterexamples composed of numbers that more closely resemble those constituting a particular confirming example are likely to be more effective in leading a student to abandon a faulty conjecture. Additionally, according to this framework, counterexamples that employ logic that more closely resembles that employed by a student in generating faulty conjecture would be more likely to support students in abandoning their false conjectures. Finally, of two counterexamples to some false student conjecture, the one that is least complex mathematically is, in theory, more likely to succeed in leading a student to abandon faulty reasoning.

Further work with school-aged children would need to be conducted in order to empirically verify such propositions. Accordingly, a next step in this line of inquiry could involve presenting elementary-school students with the task of generating their own conjectures for comparing fractions. Of these conjectures, those that are false could then be responded to using counterexamples falling at different points along the mapping, logic, and complexity continuums. The effectiveness of such counterexamples at leading students to abandon their false conjectures could then be assessed.

The present study examined counterexamples to students' false conjectures in one topic area, the comparison of fractions. However, the framework described here could be applied or at least used to guide one's work when generating counterexamples in a number of different mathematical domains. In today's reform mathematics classroom, teachers are to encourage students to generate their own strategies for solving problems of arithmetic, geometry, and a host of problems in other content areas, as well. For example, rather than demonstrate the standard procedure for solving, say, a multi-digit addition problem, or the formula for finding the area of a square, teachers today are to encourage students to devise their own methods and procedures for solving such problems. Undoubtedly, such attempts are likely to result in the generation of methods that may work in some cases, but fail to work in general (e.g., Erlwanger, 1973; Stafylidou & Vosniadou, 2004). As such, teachers today must not only be prepared to refute students' false conjectures for comparing fractions, but their conjectures in a range of other mathematical domains, as well.

Teachers today will undoubtedly encounter false student conjectures on a daily basis and, therefore, must be ready to respond with counterexamples that will support, rather than confuse or frustrate, their students. While ensuring that the counterexamples one generates in response to students' false conjecture are pedagogically powerful is likely to support students' thinking, as demonstrated by the present study, such counterexamples may not come so easily, especially to the novice teacher. By consequence, a second response to students' false conjectures could involve the crafting of counterexamples that not only do well to map those cases that confirm students' reasoning, but also mirror the logic utilized by students and are marked by the least complexity possible.

## References


Ball, D. L., Hill, H. C., & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator, 29*(1), 14-17, 20-22, 43-46.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, *59*(5), 389-407.

Common Core State Standards Initiative (2010). *Common Core State Standards for mathematics*. Washington, DC: Council of Chief State School Officers and National Governors Association Center for Best Practices. Retrieved from http://www.corestandards.org/Math/

Erlwanger, S. H. (1973). Benny's concept of rules and answers in IPI mathematics. *Journal of Children's Mathematical Behavior, 1*(2), 7-26.

Ginsburg, H. P., Jacobs, S. F., & Lopez, L. S. (1998). *The teacher's guide to flexible interviewing in the classroom: Learning what children know about math.* Needham Heights, MA: Allyn & Bacon.

Glaser, B. G., & Strauss, A. L. (1965). The constant comparative method of qualitative analysis. *Social Problems*, *12*(4), 436-445.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal*, *27*(1), 29-63.

Peled, I., & Zaslavsky, O. (1997). Counter-examples that (only) prove and counter-examples that (also) explain. *Focus on Learning Problems in Mathematics*, *19*(3), 49-61.

Phelps, G., & Gitomer, D. (2012). Content knowledge for teaching: Mathematics grades 4-5 assessment. *Educational Testing Service,* 1-20.

Potari, D., Zachariades, T., & Zaslavsky, O. (2009). Mathematics teachers' reasoning for refuting students' invalid claims. Paper presented at the Proceedings of CERME (France), Lyon (pp. 281-290). Lyon, France: CERME.

Stafylidou, S., & Vosniadou, S. (2004). The development of students' understanding of the numerical value of fractions. *Learning and Instruction*, *14*(5), 503-518.

Stylianides, A. J., & Ball, D. L. (2008). Understanding and describing mathematical knowledge for teaching: Knowledge about proof for engaging students in the activity of proving. *Journal of Mathematics Teacher Education*, *11*(4), 307-332.

Zaslavsky, O., & Peled, I. (1996). Inhibiting factors in generating examples by mathematics teachers and student teachers: The case of binary operation. *Journal for Research in Mathematics Education, 27*(1), 67-78.

Zazkis, R., & Chernoff, E. J. (2008). What makes a counterexample exemplary? *Educational Studies in Mathematics*, *68*(3), 195-208.