NCTM RESEARCH CONFERENCE 2016
April 11-13 • San Francisco

Paper Title: Constructing and Critiquing Arguments: Effect of an Instructional Sequence

Author(s): Sean Yee, yee@math.sc.edu, University of South Carolina

Yi-Yin (Winnie) Ko, Winnie.Ko@indstate.edu, Indiana State University

Sarah K. Bleiler, Sarah.Bleiler@mtsu.edu, Middle Tennessee State University

Justin D. Boyle, jboyle@bamaed.ua.edu, University of Alabama

Session Title: Constructing and Critiquing Arguments: Effect of an Instructional Sequence

Session Type: Brief Research Report

Presentation Date: Tuesday, April 12, 2016 9:15-9:45am

Presentation Location: San Francisco, California

Overview

Over the past decade, a significant emphasis has been placed on the teaching and learning of proof across content domains in mathematics. In particular, the *Common Core State Standards for Mathematical Practice* (2010) emphasize the need for students to regularly construct viable arguments and critique each other's reasoning in classrooms. This recommendation poses serious challenges because students and teachers at many grade levels have considerable difficulty actively engaging with the complexities of proof (e.g., Healy & Hoyle, 2000; Knuth, 2002; Knuth, Choppin, & Bieda, 2009; Ko & Knuth, 2013; Lo, Grant, & Flowers, 2007). In this paper, we share our before-during-after instructional sequence where students actively and communally create criteria for proof. As suggested in the *Principles to Actions* (NCTM, 2014), the instructors "facilitate meaningful mathematical discourse, support productive struggle in learning, and elicit and use evidence of student thinking" (p. 10) to study how students' arguments change with the instructional sequence.

Educational Significance

When learning proof, students typically observe their teacher presenting a polished and completed proof (Stylianou, Blanton, & Knuth, 2009). Within such learning environments, it is not surprising that students often come to believe that their teacher has sole authority to judge the validity of their proof (Harel & Sowder, 1998). It is also no surprise that students often experience difficulty both identifying and constructing proofs (Healy & Hoyles, 2000; Knuth et al., 2009). We wanted students to learn proof as a communally-negotiated and sense-making process. Therefore, we designed an instructional sequence that increased students' autonomy in actively engaging with proof.

Theoretical Framework

This study draws on Stylianides's (2007) definition of proof and Harel and Sowder's (1998, 2007) proof schemes. First, Stylianides's (2007) definition of proof incorporates a focus on the set of statements (i.e., definitions, axioms, theorems) and the appropriate forms of argumentations and representations accepted and understood within a particular mathematical community. As the goal of our instructional sequence emphasizes students' ownership of proof within the classroom, we chose Stylianides's (2007) definition of proof because of the importance of the study to see knowledge of proof as participatory within the classroom community. Second, Harel and Sowder (1998) define a proof scheme as "the individual's scheme of doubts, truths, and convictions" (p. 244). We used Harel and Sowder's (2007) external conviction, empirical, and deductive proof schemes to look for changes on students' arguments before and after participating in our instructional sequence. According to Harel and Sowder (1998, 2007), a student holding an external conviction proof scheme would be convinced by the form of an argument, a teacher, a textbook, or symbolic representations. A student holding an empirical proof scheme would be convinced by specific cases or individuals' mental images. A student holding a deductive proof scheme would be convinced by mathematical demonstration or logical progression of accepted statements. Together, Stylianides's (2007) and Harel and Sowder's (1998, 2007) frameworks allow us to investigate how the instructional sequence affected the construction of student arguments. To this end our research study asked, how did the students' written arguments differ before and after they participated in the instructional sequence?

Methods

Participants

All four authors implemented the same instructional sequence in their own classrooms at four different universities. Two of the courses were upper-division mathematics content courses, and the other two were secondary mathematics methods courses. A total of 57 students, who are undergraduate and graduate mathematics or secondary mathematics education majors, volunteered to participate in the study. All of the participants had completed calculus courses and introductory proof courses or were presently in introductory proof courses.

Instructional Sequence

Our instructional sequence was implemented in mathematics content and methods courses at the college level using the Sticky Gum Problem (SGP), depicted in Figure 1, and can be easily translated into K-16 mathematics classrooms with other proof-related tasks.

The Sticky Gum Problem

Ms. Hernandez came across a gumball machine one day when she was out with her twins. Of course, the twins each wanted a gumball. What's more, they insisted on being given gumballs of the same color. The gumballs were a penny each, and there would be no way to tell which color would come out next. Ms. Hernandez decides that she will keep putting in pennies until she gets two gumballs that are the same color. She can see that there are only red and white gumballs in the machine.

1) Why is three cents the most she will have to spend to satisfy her twins?
2) The next day, Ms. Hernandez passes a gumball machine with red, white, and blue gumballs. How could Ms. Hernandez satisfy her twins with their need for the same color this time? That is, what is the most Ms. Hernandez might have to spend that day?
3) Here comes Mr. Hodges with his triplets past the gumball machine in question 2. Of course, all three of his children want to have the same color gumball. What is the most he might have to spend?
4) Generalize this problem as much as you can. Vary the number of colors. What about different size families? Prove your generalization to show that it always works for any number of children and any number of gumball colors.

*Figure 1***.** The Sticky Gum Problem (Fendel, Resek, Alper, & Fraser,1996).

We chose the SGP because: it (1) involves multiple methods and variables; (2) is appropriate for middle-school through college students; (3) provides opportunities to create specific examples then move beyond empirical arguments to generalizations.

Our instructional design was organized around the SGP into the three components shown in Table 1: the before-class activity (offer and justify a solution to the SGP), the during-class activity (evaluate the five instructor-selected students' arguments individually, then small groups, then communally agree upon criteria for proof based on their evaluations), and the after-class activity (revise their before-class argument). Extended results of students' impressions of the activity can be found in previous articles by the authors of this paper (Bleiler, Ko, Yee, & Boyle, 2015; Boyle, Bleiler, Yee, & Ko, 2015).

Table 1

The Three-Components Instructional Sequence.

Before-Class Activity	• Each student solves the Sticky Gum Problem (SGP). • Each student submits the SGP electronically prior to the during-class activity. • The instructor chooses five distinct arguments submitted by their students.

During-Class Activity	• Each student individually evaluates the five distinct arguments
	• Students break into small groups of 2-4 students.
	    • Each group looks at the same five pre-selected classmate's arguments.
	    • Each group discusses and decides which of the five selected arguments are convincing and valid proofs.
	    • Each group determines how they decided whether each argument is a valid and convincing proof.
	    • Each group creates a list of five proof criteria through small and whole group discussions.
	• All small groups rejoin the entire class.
	    • The entire class has discussions, compares the small group criteria, and determines a class-wide communal set of five criteria for proof.

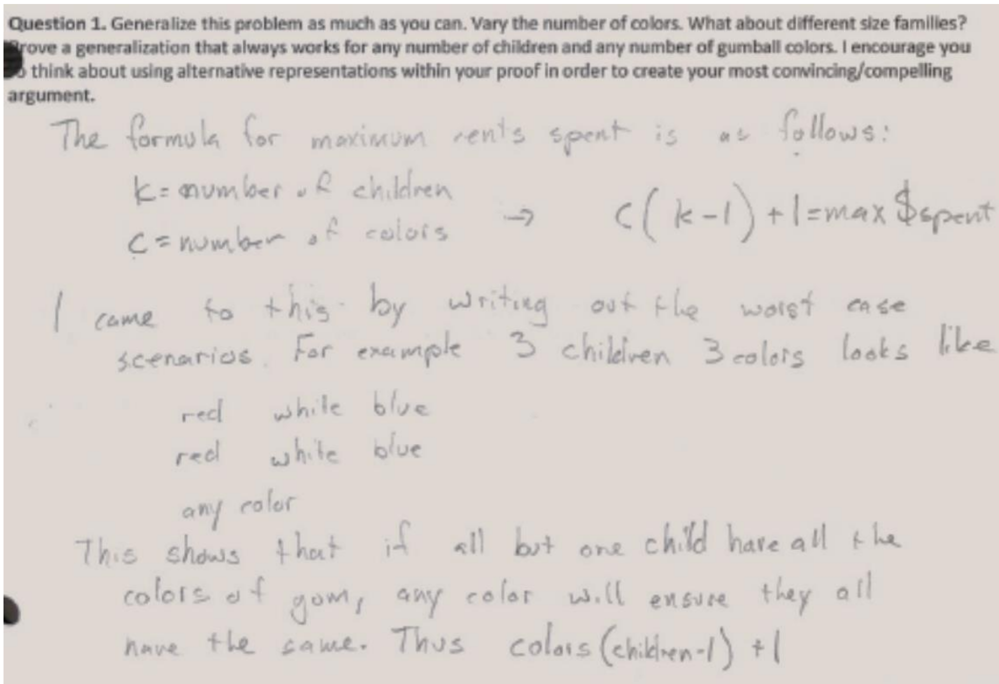
After-Class Activity	• Each student rewrites their argument for SGP to satisfy the communal class criteria and submits their new proof.

The five distinct students' arguments described in Table 1 were chosen for diversity to provoke conversation, not by their mathematical accuracy. Hence, this activity treated the instructor as a facilitator rather than an authority on what counts as proof. The small groups placed their criteria for proof in front of the class so that all students could identify the common themes across groups. The instructor then facilitated a classroom discussion with the explicit goal of creating five classroom-based criteria for proof.

All four authors collected all students' work throughout the instructional sequence as well as classroom artifacts. Each student's before-class and after-class arguments were analyzed to identify their appropriate proof scheme (Harel and Sowder, 1998, 2007). A total of 57 before-class and 57 after-class arguments were collected from the four classes, and each of the four authors analyzed the arguments from their own class. For internal reliability, all 114 arguments were independently evaluated by at least two authors. All variations in coding were discussed by the authors until consensus was found.

Coding Proof Schemes

As students were asked to revise their before-class argument to align with their communal criteria for proof, we coded each student's before-class and after-class arguments based on Harel and Sowder's (1998, 2007) proof schemes to identify their changes. Because evidence from the students' arguments frequently suggested overlap of the categories of external conviction, empirical, and deductive proof schemes, we created another three categories, empirical/external conviction, external conviction/deductive, and empirical/deductive. Figure 2 is Gabriel's before-class argument coded as Empirical/Deductive because we see an example used for support of the argument along with deductive explanation in nature. In fact, it could be easy to classify Gabriel's argument with an empirical proof scheme because the justification stems from one example with three children and three colors. However, there are elements of the deductive proof scheme, such as "worst case scenario" and the concluding remark, "if all but one child have all the colors of gum, any color will ensure they all have the same," shown in Gabriel's discussion as well. These are referencing the restrictive characteristics of the transformational proof scheme (Harel & Sowder, 1998)—a subset of the deductive proof scheme—because Gabriel presumed to understand the restrictions that transform the situation by assuming the worst case scenario and discussing what transformation would be necessary to satisfy the last child receiving the same color. Thus, Gabriel's work was coded this proof scheme as Empirical/Deductive.


Question 1. Generalize this problem as much as you can. Vary the number of colors. What about different size families? Prove a generalization that always works for any number of children and any number of gumball colors. I encourage you to think about using alternative representations within your proof in order to create your most convincing/compelling argument.

The formula for maximum cents spent is as follows:

k = number of children
c = number of colors → c(k-1)+1 = max $ spent

I came to this by writing out the worst case scenarios. For example 3 children 3 colors looks like

red white blue
red white blue
any color

This shows that if all but one child have all the colors of gum, any color will ensure they all have the same. Thus colors (children-1) + 1

Figure 2. Gabriel's before-class argument of the Sticky Gum Problem.

It is important to note that we are not looking for a mathematically correct justification, but rather interpreting the means in which students communicate their justification.

As shown in Figure 3, Chuck offers another example of the need to offer overlap in the proof schemes. Chuck follows a clear mode of argument representation (Stylianides, 2007) with axiomatic traits from the deductive proof scheme (Harel & Sowder, 1998) as seen with the stated accepted principles under the given.

PATTERN- # COLORS = Most gumballs that can be dispensed without a repeat.
CHILDREN = One less than the number of times all colors need to be dispensed
+ 1 extra to receive the final matching gumball

Given: C, X, K, P are in the set of Natural Numbers.
C is a specific # of different colors.
X is the number of pennies deposited.
K is a specific # of children who want the same color.
P is the most pennies required to have K number of repeat colors dispensed.
No more iterations of "deposit penny," "receive gumball" are required once P = X
Exactly 1 color is dispensed for <u>each</u> penny deposited.

Prove: $P = [C(K-1)] + 1$

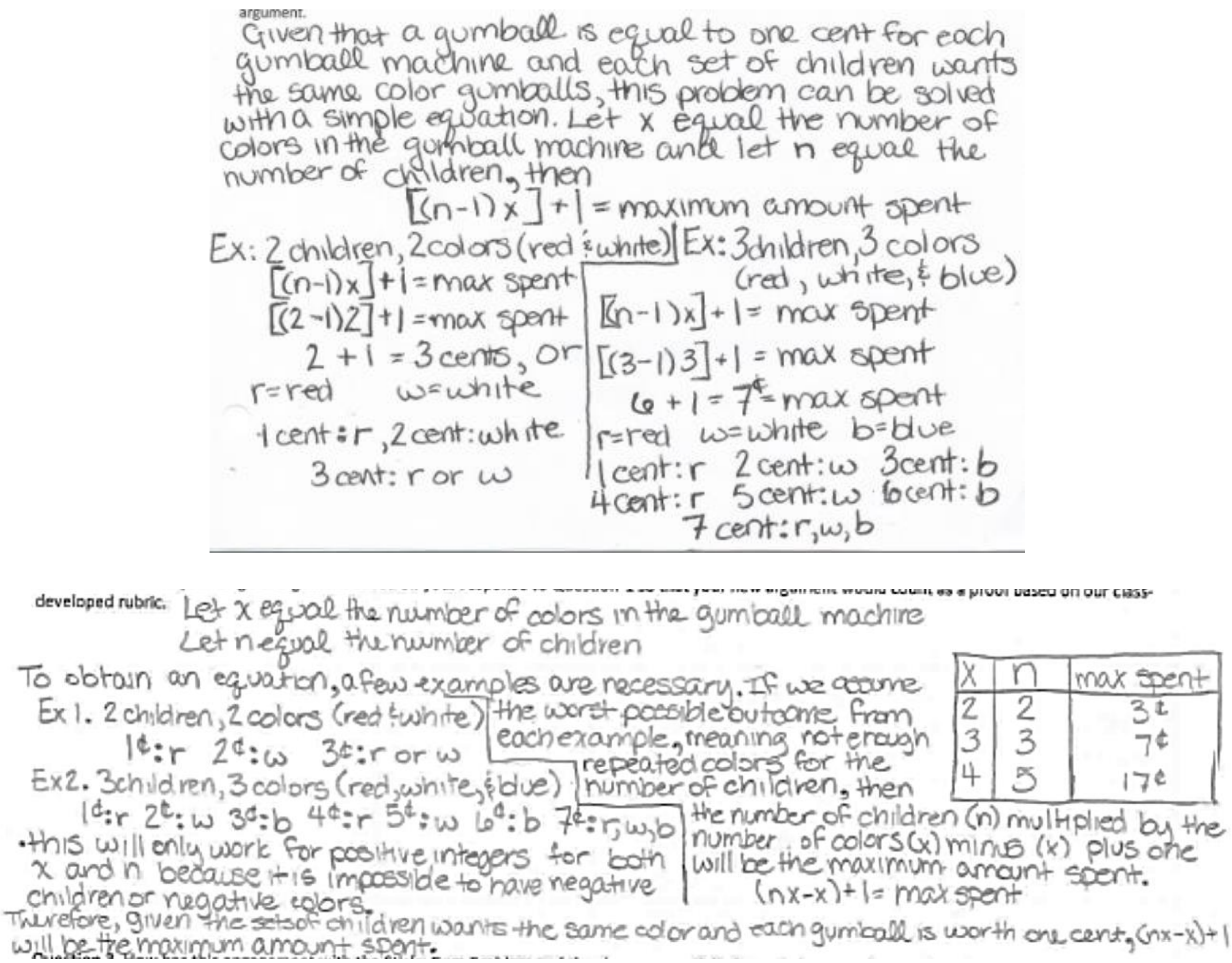
STATEMENT

1. Assume a color will not repeat before all other colors have been dispensed.
2. When X = C, all colors have been dispensed because of the assumption in Statement 1.
3. When X = C + 1, a repeat color will be dispensed because of Statement 2.
4. Assume a color will not repeat again until all other colors have been dispensed again.
5. When X = C * K, all colors will have been repeated K times because of the assumption in Statement 4.
6. When X = C * K – 1, only one more repeat is required to achieve P, because of the given information and Statement 5.
7. Therefore P = X when X = C(K-1) + 1, because of Statement 6.

Figure 3. Chuck's before-class argument justifying classified as using empirical and deductive proof schemes.

The concern is that the assumption in the statement one is not addressed in the argument. Specifically, why can we assume a color will not repeat before all other colors have been dispensed (this falls under the "worst case scenario in Phil's argument)? We see that there are ritual traits from the external conviction proof scheme embedded within this argument as well. Thus, Chuck's before-class argument was coded as External/Deductive.

Beth's work depicted in Figure 4 required some discussion among the researchers. Specifically, there was discussion with the after-class argument.


argument.
Given that a gumball is equal to one cent for each gumball machine and each set of children wants the same color gumballs, this problem can be solved with a simple equation. Let x equal the number of colors in the gumball machine and let n equal the number of children, then

$[(n-1)x]+1$ = maximum amount spent

Ex: 2 children, 2 colors (red & white)
$[(n-1)x]+1$ = max spent
$[(2-1)2]+1$ = max spent
$2+1$ = 3 cents, or
r=red w=white
1 cent: r, 2 cent: white
3 cent: r or w

Ex: 3 children, 3 colors (red, white, & blue)
$[(n-1)x]+1$ = max spent
$[(3-1)3]+1$ = max spent
$6+1$ = 7¢ = max spent
r=red w=white b=blue
1 cent: r 2 cent: w 3 cent: b
4 cent: r 5 cent: w 6 cent: b
7 cent: r, w, b

developed rubric. Let x equal the number of colors in the gumball machine
Let n equal the number of children
To obtain an equation, a few examples are necessary. If we assume the worst possible outcome from each example, meaning not enough repeated colors for the number of children, then the number of children (n) multiplied by the number of colors (x) minus (x) plus one will be the maximum amount spent.
$(nx-x)+1$ = max spent
Ex 1. 2 children, 2 colors (red & white)
1¢: r 2¢: w 3¢: r or w
Ex 2. 3 children, 3 colors (red, white, & blue)
1¢: r 2¢: w 3¢: b 4¢: r 5¢: w 6¢: b 7¢: r, w, b
• this will only work for positive integers for both x and n because it is impossible to have negative children or negative colors.

x	n	max spent
2	2	3¢
3	3	7¢
4	5	17¢

Therefore, given the sets of children wants the same color and each gumball is worth one cent, $(nx-x)+1$ will be the maximum amount spent.

Figure 4. Beth’s before-class (top) and after-class (bottom) arguments eventually coded as empirical and external/deductive respectively.

Author1 and Author4 agreed that the Beth’s before-class argument used an empirical proof scheme. The after-class argument was coded as Empirical/Deductive by Author1, while Author4 coded the after-class argument as Empirical/External. Both authors discussed their analysis of the revision and Author1 agreed that the narrative of assuming the worst possible outcome did not mandate a deductive argument on its own. Specifically, as the argument made this assumption without justifying its purpose, Author1 concluded that the argument lacked deductive structure but relied on ritualistic and external proof scheme for the same reasons stated by Author4. Thus agreement was resolved amongst the researchers that the before-class argument used an empirical proof scheme while the after-class argument used an empirical/external conviction proof scheme.

Results

Table 2 illustrates the coding of each student's before-class argument (rows) and after-class argument (columns). Colors have been added to illustrate blends. For example, empirical is red and external conviction is yellow, so empirical/external conviction is orange. Each cell in Table 2 illustrates the number of students that transitioned from one proof scheme to another. For instance, row 2 column 6 shows that six (out of the 57) students transitioned from empirical to deductive proof schemes.

Table 2

Changes in Students' Before-Class and After-Class Arguments.

Brefore-Class Proof Schemes (rows) After-Class Proof Schemes (columns) TOTAL	After-Class None	After-Class Empirical	After-Class Empirical/External Conviction	After-Class External Conviction	After-Class External Conviction/Deductive	After-Class Deductive	After-Class Empirical/Deductive	**Original Arguemnt Total**
Before-Class None	0	2	0	0	0	1	0	**3**
Before-Class Empirical	5	5	3	6	0	6	2	**27**
Before-Class Empirical/External Conviction	1	0	0	0	0	3	0	**4**
Before-Class External Conviction	2	1	1	3	0	1	0	**8**
Before-Class External Conviction/Deductive	0	0	0	0	0	2	0	**2**
Before-Class Deductive	1	0	0	1	0	4	0	**6**
Before-Class Empirical/Deductive	2	0	0	0	1	3	1	**7**
Revised Argument Total	**11**	**8**	**4**	**10**	**1**	**20**	**3**	**57**

The Before-Class None row and After-Class None column indicate that students did not submit their work for the before-class and the after-class activity, respectively. It is important to notice that most students' before-class arguments were empirical (27), while most students' after-class arguments were deductive (20). Thus, we see a significant shift away from empirical proof schemes towards deductive proof schemes after our class activity.

Kimmie's before-class and after-class arguments (Figure 5) together offer is prototypical of students transitioning from an empirical proof scheme to deductive proof scheme.

let n = amount of children and c = number of colors in machine
suppose n=2, p = number of pennies needed
then if c=2 then 3 pennies are needed.
c=3 then 4 pennies are needed.
c=4 then 5 pennies are needed.
c=5 then 6 pennies are needed
so if n=2 then add 1 to the number of colors to get how many pennies you need, so c+1=p
suppose n=3. focusing on the fact that worst case is 2 of each color comes out, the next ball will make 3.
if c=3 then 7 is needed because 3 are needed. two of each plus 1 is 7
c=4 then 9 is needed
c=5 11 is needed so c(2)+1
if there are more than three children expanding the equation will work
if 6 children and 3 colors were used 16 pennies is needed.
c(n-1)+1 = p

Conjecture: The maximum amount of pennies required to allow the number of children the same color gumball can be represented in the form p=c(n-1)+1 where p is the number of pennies, c is the amount of colors in the machine, and n is the number of children receiving gumballs. In order to satisfy the children the least amount of pennies is 1 too few of one of the colors, which is represented as (n-1). To show this for every color in the machine, multiply the minimum for one color by the total number of colors in the machine, c(n-1). When one more gumball is drawn, no matter what the next color is, it will satisfy the needs of the children, so add one penny to the amount. Therefore, c(n-1) +1, represents the maximum amount of pennies needed for each children to have the same color of gumball no matter how many children or different colors are in the machine.

Figure 5. Kimmie's before-class (hand written) and after-class (typed) arguments coded as empirical and deductive respectively.

Kimmie's before-class argument bases its justification around the given examples and justifies the generalized equation using examples. Thus we see Kimmie's before-class argument is clearly empirical while Kimmie's after-class argument uses deductive reasoning to convince the reader of the validity of their equation. Kimmie illustrates how a student's argument shifted after implementing the instructional sequence.

Of the 114 before-class (57) and after-class (57) arguments, the diagonal of Table 2 shows few students proof schemes remained the same. The main diagonal of Table 2 shows that five empirical proof schemes remained empirical (row2, column2), three external proof schemes remained external convictions (row4, column4), four deductive proof schemes remained deductive (row6, column6), and one empirical/deductive remained empirical/deductive. Altogether, 13 of the 57 proof schemes remained the same after the instructional sequence, demonstrating that the majority of students changed proof schemes after the instructional sequence.

Discussion

To determine how the instructional sequence affected students' arguments, we identified students' proof schemes as their "scheme of doubts, truths, and convictions" (Harel & Sowder, 1998, p. 244) by looking at their before-class and after-class arguments to the SGP. Our results show that the majority of students changed their argument after participating in the instructional sequence. Of those changed, the number of empirical proof schemes decreased while the number of deductive proof schemes increased. Also, the number of external conviction proof schemes remained about the same.

Our research results build upon prior research by demonstrating how offering students the opportunity to critique each other's reasoning and developing communal criteria for proof changes their written arguments. As shown in Table 2, there was a shift from empirical proof schemes to deductive proof schemes with our instructional sequence. Because students were asked to revise their arguments based on their classroom criteria for proof, developing a communal understanding of proof supported them in better communicating mathematical ideas through arguments. While previous research has shown that students are able to delineate between empirical and deductive arguments (Segal, 1999; Weber, 2010), our instructional sequence provides an opportunity for students to create their classroom community criteria for proof and to generate their arguments towards a more deductive proof scheme.

References

Bleiler, S. K., Ko, Y. Y., Yee, S. P., & Boyle, J. D. (2015). Community development and evolution of a rubric for proof writing. In C. Suurtamm & A. R. McDuffie (Eds.), *Annual Perspectives in Math Ed 2015: Assessment to Enhance Learning and Teaching* (pp. 97-108). Reston, VA: NCTM.

Boyle, J. D., Bleiler, S. K., Yee, S. P., & Ko, Y.-Y. (2015). Transforming perceptions of proof: A four-part instructional sequence. *Mathematics Teacher Educator*, 4(1), 32-70.

CCSSM (2010). Common Core State Standards for Mathematics. Washington, DC: Council of Chief State School Officers.

Fendel, D., Resek, D., Alper, L., & Fraser, S. (1996). *Interactive mathematics program year 1 – unit 2: The game of pig* (p.99). Emeryville, CA: Key Curriculum Press.

Harel, G. & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in collegiate mathematics education, III* (pp. 234-283). Washington DC: Mathematical Association of America.

Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 805-842). Charlotte, NC: Information Age Pub.

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education, 31*(4), 396-428.

Knuth, E. (2002). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education*, 33(5), 379-405.

Knuth, E., Choppin, J., & Bieda, K. (2009). Middle school students' production of mathematical justifications. In D. Stylianou, M. Blanton, & E. Knuth (Eds.), *Teaching and learning* *proof across the grades: A K-16 perspective* (pp. 153-170). New York, NY: Routledge.

Ko, Y.-Y., & Knuth, E. J. (2013). Validating proofs and counter examples across content domains: Practices of importance for mathematics majors. *Journal of Mathematical Behavior*, 32(2013), 20-35.

Lo, J. J., Grant, T., & Flowers, J. (2007). Challenges in deepening prospective teachers' understanding of multiplication through justification. *Journal of Mathematics Teacher Education, 11*(1), 5-22.

National Council of Teachers of Mathematics (NCTM). (2014). *Principles to actions*. Reston, VA: National Council of Teacher of Mathematics.

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. Washington, DC: Authors.

Stylianides, A. J. (2007). Proof and Proving in School Mathematics. *Journal for Research in Mathematics Education, 38*(3), 289-321.

Stylianou, D. A., Blanton, M. L., Knuth, E. J., & National Council of Teachers of Mathematics. (2009). *Teaching and learning proof across the grades : a K-16 perspective*. New York: Routledge.

Weber, K. (2010). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning*, 12(4), 306–336.