PRODUCTIVE USE OF EXAMPLES FOR PROVING: WHAT MIGHT THIS LOOK LIKE?

Orit Zaslavsky
New York University & Technion – Israel Institute of Technology
oritrath@gmail.com

Inbar Aricha-Metzer, New York University, arinbar@gmail.com
Pooneh Sabouri, New York University, psabouri@nyu.edu
Michael Thoms, New York University, mt1908@nyu.edu
Oscar Bernal, New York University, osbernal@nyu.edu

The study reported in this paper is part of a larger study on the roles of examples in learning to prove. We focus here on manifestations of students' productive use of examples for proving in the course of exploring conjectures and proving or disproving them. In this context, we define productive use of examples for proving as students' utterances that indicate that working with examples led them to realize and gain insights into some aspects of the key ideas for proving (or disproving) the conjecture. There were a total of 39 participants (12 middle school, 17 high school, and 10 undergraduate students). Each took part in an individual one-hour task-based interview. We identified 77 cases of productive use of examples, 41 based on an interviewer's provision of example(s) and 36 based on students' spontaneous generation of examples. These cases serve to characterize students' strengths that are not directly fostered in school.

Keywords: Reasoning and Proof

Proof and Proving in Mathematics Education

It is commonly agreed among mathematicians and mathematics educators that mathematical proof and proving are at the heart of mathematics, and that the activity of mathematically proving is dauntingly difficult even for most good undergraduate students. A continuing concern in mathematics education is that students do not sufficiently understand the nature of evidence and proof in mathematics and that they struggle with providing logically sound justifications and arguments to support the validity of mathematical conjectures or claims (e.g., Healy & Hoyles, 2000; Kloosterman & Lester, 2004; Knuth, Choppin, & Bieda, 2009). This concern has been guiding numerous studies, as it reflects a deficiency in one of the key elements of mathematics and mathematical practice (e.g., Harel & Sowder, 2007; Knuth, 2002; Sowder & Harel, 1998). Consequently, there have been calls for proof to play a more central role in mathematics education, by researchers (e.g., Ball, Hoyles, Jahnke, & Movshovitz-Hadar, 2002), as well as reform initiatives (the *Common Core State Standards for Mathematics* and the NCTM *Principles and Standards for School Mathematics*). However, despite these calls, research continues to indicate that students' understanding of proof is far from being satisfactory (Harel & Sowder, 2007; Healy & Hoyles, 2000).

A major source underlying students' difficulties in understanding proof and proving is related to their treatment of examples (e.g., Healy & Hoyles, 2000; Zaslavsky, Nickerson, Stylianides, Kidron, & Winicki-Landman, 2012). There is evidence that an inhibiting factor in students' proving (at all levels) is an over reliance on examples. They often infer that a general claim is true for all cases on the basis of checking just a number of examples that satisfy this claim. This tendency has been recognized as a stumbling block in the transition from inductive to deductive arguments, and the progression from empirical justifications to proof (e.g., Fischbein, 1987). This tension between empirical and formal aspects of proving suggests that understanding the logical relations between

Bartell, T. G., Bieda, K. N., Putnam, R. T., Bradfield, K., & Dominguez, H. (Eds.). (2015). *Proceedings of the 37th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. East Lansing, MI: Michigan State University.

examples and statements is a non-trivial task that is critical for proving. However, this kind of understanding is not usually explicitly addressed in the course of learning mathematics, in general, and learning to prove, in particular.

There have been attempts to help students learn the limitations of examples for proving in order to reduce their tendency to infer from examples more than is logically valid (e.g., Sowder & Harel, 1998; Stylianides & Stylianides, 2009; Zaslavsky et al, 2012). While these attempts address the limitations of examples for proving, they overlook the potential of example-based reasoning strategies in enhancing conjecturing and proving. In fact, less attention has been given to facilitating students' ability and inclination to build on the potential strengths of using examples for proving. More specifically, there is scarce research on using examples generically, i.e., in a way that allows to see the general through the particular, make sense of a mathematical statement, and gain insight into all or some of the main ideas of its proof (e.g., Knuth, Kalish, Ellis, Williams, & Felton, 2011; Leron & Zaslavsky, 2013; Mason & Pimm, 1984; Rowland, 2001). Mason and Pimm's (1984) terms of *generic example* and *generic proof* capture the essence of what we mean by using examples generically. Accordingly, "A generic example is an actual example, but one presented in such a way as to bring out its intended role as the carrier of the general." (ibid p. 287); and a "generic proof, although given in terms of a particular number, nowhere relies on any specific properties of that number." (ibid p. 284). Example-based reasoning strategies encompass this way of thinking with and through examples.

We believe that students' failure to engage productively in example-based reasoning strategies, to think about examples generically, and to analyze examples when engaging in activities related to proving, accounts for many of the difficulties they encounter in learning to prove. Our study stems from the stand that students should learn to use and analyze examples analytically and generically, not only in order to gain a better understanding of the conjectures (or statements) that they explore but also in order to learn to develop proofs (or dis-proofs) of these conjectures.

Very little research has focused on the nature of middle school, high school, or undergraduate mathematics students' thinking about and use of examples in generating, making sense of, and proving mathematical conjectures. Alcock and Inglis (2008) argue that such studies are needed in order to effectively develop instructional practices that foster the development of students' learning to prove. We aim at better understanding the nature of example use across grade levels, and in particular, how example use may support students' reasoning and proof development.

Zaslavsky (2014) distinguished between three settings of example use: spontaneous example use, evoked example production, and provisioning of examples. The spontaneous setting highlights what may come naturally to learners and experts, and how productive their choices and what they make of them are. The evoked example production allows us to study what choices learners make when pushed to use examples and also how productive they are. This setting has a strong diagnostic power, as it may evoke students' strengths as well as their weaknesses with respect to exemplification and proving. The provisioning of examples by a researcher allows us to examine what learners see in these examples, and in what ways they are able to build on the given examples to gain insights about how to justify or prove a claim. This setting also may shed light on possible mis-matches between intentions (of a teacher/researcher and a learner). In our study, we distinguish between example uses that involve student generated examples and those that involve researcher provided examples.

The Study

The study reported in this paper is part of a larger study of the roles of examples in learning to prove. Its purpose is to better understand the roles examples play in the development, exploration, and justification of mathematical conjectures, with the overarching goal being to help students appreciate the need to prove and to learn to prove. In this portion we focus on ways in which students use their own examples or examples provided by the researchers to support their claims about the validity of mathematical statements and conjectures. More specifically, we examine cases in which the use of examples can be considered productive for proving. For productive use of examples for proving we consider indications of gaining insights from an example or set of examples about the main idea(s) of a proof. It can be manifested, for instance, by a shift from no clue why a conjecture works (or doesn't work) to an articulation of an idea why. Some of the manifestations of this type of example-use may be seen as generic proving (Leron & Zaslavsky, 2013).

Data Collection

This study was based on individual task-based interviews with 12 middle school (MS) students, 17 high school students (HS), and 10 undergraduate (UG) mathematics majors. The interviews lasted approximately 1 hour and were comprised of a series of tasks in which participants were given the opportunity to conjecture and prove.

Task 2: The Sum of Consecutive Integers

Part 1:
This question involves consecutive numbers. For example, 2, 3, and 4 are consecutive numbers, but 2, 3, and 8 are not consecutive numbers.
Tyson came up with a conjecture about consecutive whole numbers that states: If you add any number of consecutive whole numbers together, the sum will be a multiple of however many numbers you added up. *At this point the interviewer suggests that the participant give an example of how the conjecture works for 5 consecutive whole numbers.*

Tyson thinks that this conjecture will always be true no matter how many consecutive numbers you use or which consecutive numbers you choose. So he thinks that if you add any 3 consecutive numbers, the answer will be a multiple of 3, or if you add any 6 consecutive numbers, the answer will be a multiple of 6, and so on.
Do you think the conjecture is true for any set of consecutive numbers, not just when you pick five consecutive numbers?

Part 2:
Let's come back to the Question 2 [i.e., Part 1 above] conjecture that the sum of five consecutive numbers is a multiple of 5.
At this point the interviewer says while writing the example: Another student had an idea of how to explain it. For the five consecutive numbers 5, 6, 7, 8, and 9, she decided to write the sum as (7-2)+(7-1)+7+(7+1)+(7+2), and writing it that way helped her to explain why the sum must be a multiple of 5. How do you think that helped her see why the conjecture is true for any five consecutive numbers?

The interview protocol for middle and high school participants included 8 tasks total each, and the one for undergraduate participants included 7 tasks total. Three tasks were shared across all participant populations. In this paper we focus mainly on one of these three shared tasks (Task 2 above): The Sum of Consecutive Integers Task. Similar versions of this task served researchers in other studies (e.g., Tabach et al., 2011). For several reasons (mainly due to time constraints, and protocols' modifications done after a number of interviews had been conducted), not all students got to engage in all the tasks that were included in the final interview protocols.

Data Analysis

While we started out looking for example-uses and focusing on whether each example-use was productive for proving, other categories emerged as we were analyzing the data, thus, in part, we used a grounded theory approach to analyze participant responses. The units of analysis were the tasks (except for Task 2, for which we coded each part separately). For each participant, we coded

Bartell, T. G., Bieda, K. N., Putnam, R. T., Bradfield, K., & Dominguez, H. (Eds.). (2015). *Proceedings of the 37th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. East Lansing, MI: Michigan State University.

his or her performance on each task according to several categories (productivity, example-source, proof-exhibition). We began by identifying cases in which we observed productive and non-productive example-use for proving. For example, in task 2 part 1, if a participant was able to use many numerical examples to determine that Tyson's conjecture was true only for odd numbers of integers but was not able to produce a legitimate argument as for why this was the case, we considered this "non-productive" for proving (although this activity was clearly productive for conjecturing). In order to be categorized as "productive" for proving, a participant had to use examples to make an argument that showed not only *that* Tyson's conjecture was true only for odd numbers of integers, but *why* this conjecture holds or does not hold, based on the parity of the number of integers involved. Instances of productive example use included participants' use of examples to generate an argument, and were related to a shift in their ability to provide a valid justification that would hold for any other such case.

Additionally, we looked at the source of the example, and distinguished between cases in which an example was provided by the student (spontaneously) or by the interviewer (non-spontaneously). This distinction is important as cases in which productive example-use was based on provided examples, may have pedagogical implications in the classroom.

For each participant we coded his or her performance on each task as productive (P), non-productive (NP), or indecisive. There were two main reasons for considering a case indecisive: (i) if a student came up with a proof but it was unclear whether the examples were helpful in reaching the proof; or (ii) if it was not clear whether an argument qualified as proof. Altogether, 222 cases were analyzed, of which 24 were indecisive.

We also coded cases according to whether or not a proof was exhibited (even a partial or informal one), and whether or not examples were used. When examples were used, we distinguished between cases that included just examples generated by the students (Exp. by St.) and cases where examples were provided also or solely by the interviewer (Exp. by Int.). Note that for those who completed Task 2 Part 1 with a full proof that the conjecture holds for all odd numbers and does not hold for even numbers, and used the same reasoning as in the prompt for Part 2, did not receive the second part (to eliminate redundancy).

Findings

Scope of Productive Example-Use for Proving

The findings in Table 1 include all cases that were coded either as productive or as non-productive (excluding the 24 indecisive cases). There were a total of 198 cases, 62 MS, 89 HS, and 47 UG. Of the 198 cases, 77 (39%) included productive use of examples for proving. Of these, more than half (41) the cases were based on examples provided by the interviewer.

Table 1: Distribution of Cases by Example-Use and Productivity for Proving in all Tasks

<table>
<tr><th rowspan="3">Grade Level</th><th colspan="3">Came up with a Proof (or partial proof)</th><th colspan="3">No Proof</th><th rowspan="3">Total</th></tr>
<tr><th colspan="2">Productive (P) Use of Examples for Proving</th><th rowspan="2">No Exp. (NE)</th><th colspan="2">Non-Productive (NP) Use of Examples for Proving</th><th rowspan="2">No Exp. (NE)</th></tr>
<tr><th>Exp. by St.</th><th>Exp. by Int.</th><th>Exp. by St.</th><th>Exp. by Int.</th></tr>
<tr><td>MS</td><td>10</td><td>13</td><td>2</td><td>27</td><td>10</td><td>0</td><td>62</td></tr>
</table>

Bartell, T. G., Bieda, K. N., Putnam, R. T., Bradfield, K., & Dominguez, H. (Eds.). (2015). *Proceedings of the 37th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. East Lansing, MI: Michigan State University.

HS	11	18	2	37	17	4	**89**
UG	15	10	3	12	4	3	**47**
Total	**36**	**41**	**7**	**76**	**31**	**7**	**198**

In terms of productivity, middle school and high school students performed similarly, as 37% of MS cases and 33% of the HS cases exhibited productive use of examples for proving, while the undergraduate students exhibited considerably more productive use of examples (53%).

Table 2: Productive and Non-Productive Example-Use in Task 2 (the Sum of Consecutive Integers)

Grade Level	Task	Productive (P) Use of Examples for Proving		Non-Productive (NP) Use of Examples for Proving		Total
		Exp. by St.	Exp. by Int.	Exp. by St.	Exp. by Int.	
MS	**Part 1**	4	0	6	0	**10**
	Part 2	0	4	0	2	**6**
HS	**Part 1**	4	0	13	0	**17**
	Part 2	0	7	0	5	**12**
UG	**Part 1**	5	0	3	0	**8**
	Part 2	0	5	0	1	**6**
Total		**13**	**16**	**22**	**8**	**59**

In table 2 we present the findings related to Part 1 and Part 2 of Task 2. Looking at the task as a whole, for this task there are larger differences between the extent of productive use of examples for proving between the three groups: MS – 50% (8 of the 16 cases), HS – 38% (11 of the 29 cases) and UG – 71% (10 of the 14 cases).

Figure 1 examines the trajectory of students who completed both parts of Task 2, and for which none of their performances was indecisive (this reduced the total number of cases in Table 2 by 15).

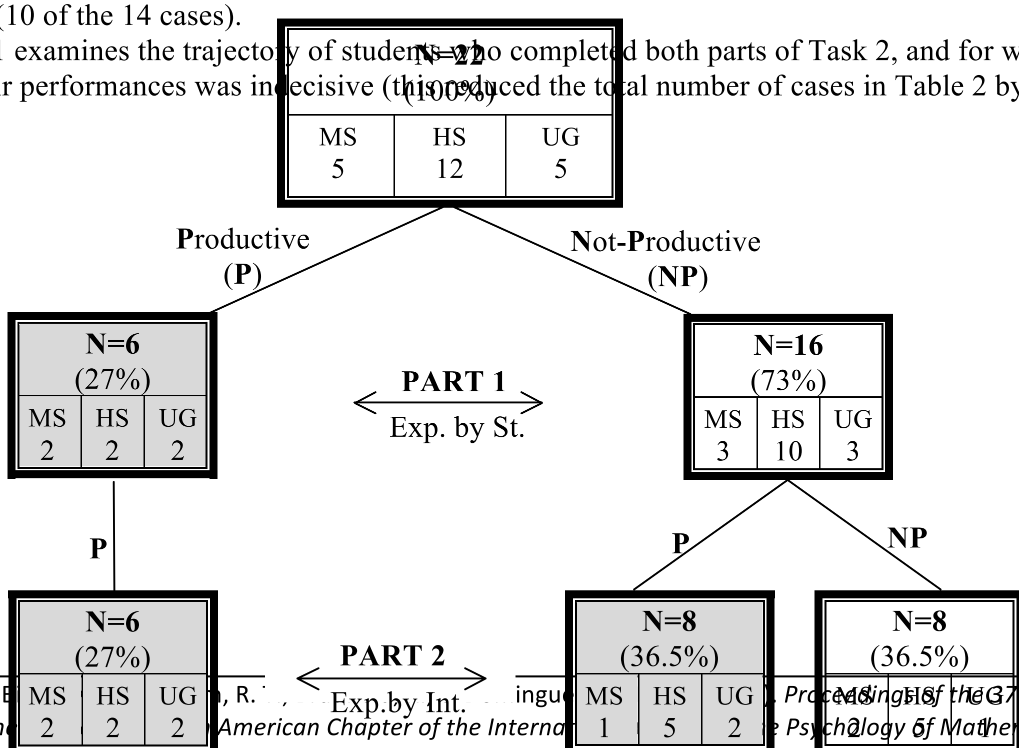

Figure 1: Paths of productive and non-productive uses of examples in Task 2
(... of students)

Bartell, T. G., B..., R. ..., ... lingu... Proceedings of the 37th annual me... American Chapter of the Interna... Psychology of Mathematics Education. East Lansing, MI: Michigan State University.

All of the cases that dealt with Task 2 were cases in which examples were explicitly used. In Part 1, the examples were generated by the students, while in Part 2 – a generic example was provided by the interviewer. Not surprisingly, students who used examples productively in Part 1, on their own, were also able to use the generic example productively. However, interestingly, half of the students who were not able to use examples productively in Part 1, were able to reason productively with the generic example provided by the interviewer in Part 2.

Characteristics of Productive Use of Examples for Proving

We turn to two cases that convey what productive use of examples for proving may look like.

Case #1: A HS student's use of his own examples productively. In Part 1 of Task 2, Sam tries three sets of examples: 1 + 2 + 3 + 4 +5 ; 2 + 3 + 4 + 5 +6 ; 3 + 4 + 5 + 6 + 7. He calculates the last two sums, and immediately is able to make an argument in support of the truth of the conjecture for any five numbers with the following observation:

"I tried a few examples, and then I realized that, well, if you add 1 to every number, then you're ultimately adding 5, because there's 5 numbers. And if the first- and the first- and if the first example 1, 2, 3, 4, 5 is- equals a multiple of 5, then by adding 5 to- to every case, it'll stay a multiple of 5." [10:43 - Time Stamp]

In other words, Sam is able to present a pseudo-inductive argument that emerges from the observation that the conjecture is true for a base case (1+2 +3+4+5), and the mechanism by which this sum changes from this case to the "next" case (where "next" is defined as increasing each term in the sum by 1) does not change the divisibility property of the sum with respect to 5. By looking at the sequence of his three examples Sam is able to both see and utilize modular reasoning when considering the divisibility properties of this sum, as increasing a number by multiples of 5 does not change its remainder (in this case, zero) upon division by 5. As Sam puts it: "if you subtract 5 from a multiple of 5, it'll still stay a multiple of 5."[13:00]

Sam is able to take advantage of the generality of this observation by answering a question that he himself had posed earlier in the interview, namely, whether negative integers were allowed in Tyson's conjecture. He is able to leverage his reasoning about the modular distribution of integers that are divisible by 5 into a correct claim that Tyson's conjecture works just as well for negative integers as it does for positive. In other words, Sam was able to extend the domain of the conjecture by creating an argument that relied solely on three numerical examples. In this case we do not consider each one of his three examples in isolation as "generic," as that clearly does not reflect his thinking. However, we consider all three seen in conjunction with each other as one generic example, as the insight that Sam gained from these examples was located in the relationship between them.

Sam is able to use this insight to create a legitimate argument for why Tyson's overall conjecture is incorrect for four consecutive integers. He reasons: "I thought of a number like 4, which is 1, 2, 3, 4, which adds up to 10. And then it's not a multiple of 4, so, and… even if you add or subtract from that, it'll always be uh, it won't be a multiple of 4."[16:40] In effect, he argues that his base case is a counter-example, and this counter-example does not just hold for the particular example of 1 + 2 + 3+4, but in fact holds for *any* four consecutive integers. Although he does not explicitly discuss remainders, we can interpret his argument as noticing that the remainder upon division by 4 is invariant under increasing a number by multiples of 4.

Case #2: A HS student's use of an interviewer's example productively. In Part 1 Isaac is clearly operating empirically, as he chooses a wide range of examples and uses them for verifying

Bartell, T. G., Bieda, K. N., Putnam, R. T., Bradfield, K., & Dominguez, H. (Eds.). (2015). *Proceedings of the 37th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. East Lansing, MI: Michigan State University.

that the conjecture is true (e.g., for 5 numbers), without being able to offer a logically valid explanation of *why* it is. He explains that it is because "I did a bunch of trials that go really far into the depths of numbers, including negatives which kind of sealed the deal for me, because negatives are really different from positives."[24:26]

During the second part of this task, the interviewer presented Isaac with a generic example, rewriting 5 + 6 + 7 + 8 + 9 as (7-2) + (7-1) + 7 + (7+1) + (7+2). Immediately, Isaac is able to see the generality within this particular example and apply it to a generic argument. Isaac uses this argument to explain why it must work for any odd number of consecutive integers, and also why it must not work for an even number of consecutive integers. He also produces a parallel algebraic representation, which is the first time that he has done so within this task. Perhaps this is due to the visual salience of the invariance of the 7's in the representation provided by the interviewer. Isaac's immediate response to the interviewer's prompt is reproduced below:

"Okay, so I can see now- this is pretty good proof for why it's, uh, for why it has to be a multiple of 5 or just a multiple of an odd number in general. Or, uh, so um, um the thing with this is, what happens-- ... Um, these numbers cancel each other out. The 2 cancels the 2, the 1 cancels the 1. And you just end up getting 7+7+7+7+7 and, um, and the reason this wouldn't work with an odd- with an odd pair, like if you added- if you added a, uh, 10 [The interviewee wrote +10 at the end of 5+6+7+8+9] to this and then you- and then you added plus 7 plus 3 [The interviewee wrote +(7+3) at the end of the (7-2)+(7-1)+7+(7+1)+(7+2)], then these would cancel out. Then you would be left with (7×5) + (7+3) plus... Yeah, plus (7+3), which would just give you, um, 3- 3 numbers off of what you want. So, yeah. And I guess you could do it with an equation, using x... (x-2) + (x-1) + (x) + (x+1) + (x+2) + (x+3). These cancel each other out, these do it as well, and this is just left there as a kind of like, almost like it ruins the party or something. So, yeah."

He later uses the word "symmetry" to explain this argument. In other words, Isaac was initially "stuck" and the generic example provided by the interviewer helped Isaac create a deductive argument as for why Tyson's conjecture was only true for odd numbers of integers.

Concluding Remarks

While the vast majority of studies on students learning to prove focus on their difficulties and suggest ways to address these difficulties, our study identifies numerous cases of students treating examples generically on their own. Moreover, these cases capture shifts from not being able to explain why a mathematical statement is true (or false) to being able to see clearly why it must work in general. Thus, the findings add to our understanding of processes by which students may learn to prove, and at the same time suggest that for some students under certain conditions this may come (almost) naturally with minimal interference. In other words, these cases could be inspiring for teachers who want to build on students' strengths.

Acknowledgments

The research was supported in part by the National Science Foundation under grant DRL-1220623 (Eric Knuth, Amy Ellis, & Orit Zaslavsky, principal investigators). The opinions expressed herein are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Alcock, L., & Inglis, M. (2008). Doctoral students' use of examples in evaluating and proving conjectures. Educational Studies in Mathematics, 69, 111–129

Ball, D., Hoyles, C., Jahnke, H., & Movshovitz-Hadar, N. (August, 2002). *The teaching of proof*. Paper presented at the International Congress of Mathematicians, Beijing, China.

Fischbein, E. (1987). *Intuition in Science and Mathematics*. D. Reidel Publishing Company, Dordrech, the Netherlands

Harel, G. and Sowder, L. (2007). Toward a comprehensive perspective on proof. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*, (pp. 805-842). Greenwich, CT: Information Age Publishing.

Kloosterman, P., & Lester, F. (2004). *Results and interpretations of the 1990 through 2000 mathematics assessments of the National Assessment of Educational Progress*. Reston, VA: National Council of Teachers of Mathematics.

Knuth, E. (2002). Secondzary 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.

Knuth, E., Kalish, C., Ellis, A., Williams, C., & Felton, M. (2011). Adolescent reasoning in mathematical and non-mathematical domains: Exploring the paradox. In V. Reyna, S. Chapman, M. Dougherty, & J. Confrey (Eds.), *The adolescent brain: Learning, reasoning, and decision making* (pp. 183-209). Washington, DC: American Psychological Association.

Leron, U., & Zaslavsky, O. (2013). Generic Proving: Reflections on scope and method. *For the Learning of Mathematics, 33*(3), 24-30.

Mason, J., & Pimm, D. (1984). Generic examples: Seeing the general in the particular. *Educational Studies in Mathematics, 15*(3), 277-290.

Rowland, T. (2001). Generic proofs in number theory. In S. Campbell and R. Zazkis (Eds.), *Learning and teaching number theory: Research in cognition and instruction,* pp. 157-184. Westport, CT: Ablex Publishing.

Sowder, L., & Harel, G. (1998). Types of students' justifications. *Mathematics Teacher, 91,* 670–675.

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

Tabach, M., Levenson, E., Barkai, R., Tsamir, P., Tirosh, D., Dreyfus, T. (2011). Secondary teachers' knowledge of elementary number theory proofs: the case of general-cover proofs. *Journal of Mathematics Teacher Education, 14,* 465-481.

Zaslavsky, O. (2014). Thinking with and through examples. In Liljedahl, P., Nicol, C., Oesterle, S., & Allan, D. (Eds.), *Proceedings of the Joint Meeting of PME 38 and PME-NA 36*, *v.1,* pp. 21-34. Vancouver, Canada: PME.

Zaslavsky, O., Nickerson, S. D., Stylianides, A., Kidron, I., & Winicki-Landman, G. (2012). The need for proof and proving: Mathematical and pedagogical perspectives. In G. Hanna and M. de Villiers (Eds.) *Proof and proving in mathematics education: The 19th ICMI Study, 15*, 215-229.