Education and Science

Vol 50 (2025) No 222 45-66

Original Article

Investigating Student Reasoning When Faced with a Mathematical Statement that Contains both Confirming and Contradicting Examples

Zülfiye Zeybek Şimşek [1], Elif Kılıçoğlu [2]

Abstract

The purpose of this study is to address the ways in which mathematically gifted students reason when faced with both confirming and contradicting examples for a mathematical statement. By addressing this issue, this study aims to investigate the types of examples, generalizations and justifications that students construct after confronting confirming and contradicting examples for the statements. Eight students who enrolled in a Science and Art Center volunteered to participate in a semi-structured individual interview. The results indicated that the types and the purposes of suggested examples varied among the students. Research investigating student reasoning suggests that students' justification schemes reflect their current view of the collection of examples that are considered as sufficient for the validation of a mathematical generalization. This study revealed that the types of examples were informative regarding the types of generalizations and arguments that were constructed by the students.

Keywords

Example types
Generalizations
Gifted students
Justifications
Mathematical reasoning

Article Info

Received: 06.21.2023
Accepted: 01.10.2025
Published Online: 04.25.2025

DOI: 10.15390/EB.2025.12962

Introduction

The importance of constructing valid arguments and critiquing the reasoning of others has been emphasized as an essential component for mathematics education from kindergarten through high school (Common Core State Standards for Mathematics [CCSSM], 2010; National Council of Teachers of Mathematics [NCTM], 2000). Mathematics educators have supported the incorporation of reasoning and proving as essential components of mathematics education at all grade levels by demonstrating that early elementary students can engage in such activities successfully (Komatsu, 2010; Stylianides & Ball, 2008). The many calls to make reasoning and proving central to students' daily mathematical practices at all levels (CCSSM, 2010; NCTM, 2000) could explain the increasing studies focusing on these skills. Although many studies focus on students' ability to construct viable arguments and/or evaluate others' arguments, there are comparatively fewer studies that focus on students' processes for refuting invalid statements and critiquing the reasoning applied in invalid statements (Yopp, 2015; Yopp, Ely, Adams, Nielsen, & Corwine, 2020; Zeybek Şimşek, 2021). Critiquing the reasoning applied in an invalid mathematical claim is an intellectually demanding task (Giannakoulias, Mastorides, Potari, &

[1] Tokat Gaziosmanpaşa University, Faculty of Education, Department of Maths and Science Education, Türkiye, zulfiye.zeybek@gop.edu.tr
[2] Hatay Mustafa Kemal University, Faculty of Education, Mathematics and Science Education, Türkiye, elifacil@mku.edu.tr

Zachariades, 2010; Zeybek Şimşek, 2021) and is influenced by knowledge of argumentation, the mathematics register, and methods for handling conceptual insights (Yopp, 2015; Yopp et al., 2020).

Studies demonstrate that students as well as teachers provide various responses to invalid generalizations (e.g., Balacheff, 1991; Giannakoulias et al., 2010; Yopp, 2015; Zeybek Şimşek, 2021). For instance, Balacheff (1991) has reported that once a student encounters a counterexample, his/ her actions could vary from modifying the original conjecture's condition (or the definition) to simply ignoring the counterexample as not sufficient. Similarly, Yopp (2015) has documented imprudent claims developed by pre-service teachers after a counterexample was identified. All these findings may indeed be related to the idea of cognitive conflict and how students handle cognitive conflict once they encounter a counterexample. As many researchers, we believe that learning occurs through cognitive conflict when students encounter new information that contradicts with previously formed mental structure which results in either changing existing mental structure or developing a new one (Gal, 2019; Piaget, 1975; Zaskis & Chernoff, 2008). Thus, cognitive conflict created by encountering a counterexample might serve as an essential mechanism to feed intellectual curiosity as well as reasoning skills of learners. For instance, Stylianides and Stylianides (2022) nominate purposefully selecting instructional tasks to promote cognitive conflict to foster students' intellectual curiosity which constitutes an essential step for developing reasoning and proving skills. However, Gal (2019) argues that not all students are ready to deal with cognitive conflict and difficulties with logical thinking ability might hinder students from dealing with such conflicts.

Some researchers see further exploration of false mathematical statements as an opportunity to investigate learners' logical thinking ability while others hold different perspectives. For instance, Komatsu (2010) has documented that fifth graders could deal with cognitive conflict once they encounter a counterexample and could attain genuine mathematical processes with refutations. Yopp (2015), on the other hand, mentions the limitations of such activities by stating "attempts to go beyond the existence of a counterexample, including making claims about classes of counterexamples and claims about cases that confirm to the original claim, can lead to problematic responses when a counterexample would have sufficed" (p. 79). Although the attempts that go beyond suggesting a counterexample could lead to problematic responses, we believe that such attempts are akin to genuine mathematical processes and therefore, could be essential to better conceptualize students' mathematical reasoning processes.

This paper suggests that mathematical statements that allow for the construction of both confirming and contradicting examples can cause cognitive conflict and initiate attempts to go beyond the identification of counterexamples. Given the substantial body of research indicating that mathematically gifted students differ from their peers in their ability to generalize and their desire to identify patterns and relationships (Leikin, 2021; Sriraman, 2004), it proposes that gifted students can vary in their approaches once both confirmations and contradictions are identified. That is, mathematically gifted students might demonstrate various strategies, including constructing generalizations and arguments that go beyond suggesting a counterexample. Thus, this paper explores the following research questions:

1. When faced with mathematical statements for which both confirming and contradicting examples could be found, what types of reasoning do mathematically gifted students use?
    a. What types of examples do they construct?
    b. What types of generalizations do they recognize?
    c. What types of arguments do they construct to justify their generalizations?
2. What are the relationships between the examples that the gifted students suggested and the generalizations and the arguments that they constructed?

Theoretical Background

Examples

Recent studies have highlighted the importance of studying examples (e.g., Alcock, 2004; Alcock & Weber, 2010; Ellis et al., 2019). More specifically, these studies focus on the different roles and usages of examples for the processes of justifying, generalizing, and proving (e.g., Ellis et al., 2019). Examples are an essential approach for students to make sense of conjectures (Alcock, 2004), support generalizing acts (Goldenberg & Mason, 2008), and be encouraged to analyze structural relationships, which is crucial for proof construction (Goldenberg & Mason, 2008; Pedemonte & Buchbinder, 2011).

Buchbinder and Zaslavsky (2009) investigate different types of examples with respect to determining the truth value of mathematical statements. They classify examples as confirming, non-confirming, contradicting, or irrelevant and argue that these classifications differed based on whether the mathematical statement was universal or existential. Thus, being cognizant about the various types of available examples in the process of proving should play a crucial role in mathematical reasoning. However, this paper argues that being able to identify the different roles of examples depends on students' referent knowledge and is closely related to their justification schemes. Balacheff (1991) echoed a similar theme, stating that "the existence of a referent knowledge (the scientific knowledge or the knowledge to be taught) gives the right to decide whether a fact is contradictory or not with respect to this knowledge" (p. 2).

Watson and Mason (2005) argue that examples are not isolated rather they should be perceived as instances of a class of potential examples, which they referred as example spaces (p. 51). Stylianides and Stylianides (2009) conceptualize a link between learners' example spaces and the notion of justification schemes, which they referred as example spaces for validation. They have presented major justification schemes proposed by Harel and Sowder (1998) and corresponding example spaces for validation in increasing levels of mathematical sophistication (see Stylianides & Stylianides, 2009 for further details). Stylianides and Stylianides (2009) state: "A student's justification scheme reflects his or her current view of the collection of examples that are considered as sufficient for the validation of a mathematical generalization, that is, it reflects the student's personal example space for validation" (p. 320). If students' justification schemes reflect their example space, then the examples constructed by students should also be informative regarding the students' justification schemes and their referent knowledge. As Watson and Mason (2005) view examples as "illustrations of concepts and principles" (p. 3), learner generated examples are seen as illustrations of their justification schemes in this study. The example spaces corresponding to justification schemes proposed by Stylianides and Stylianides (2009) are adopted to analyze students' proposed examples in this study (see Table 1).

Table 1. Example Types and Example Spaces Corresponding to Justification Schemes

Confirming Examples			**Contradicting Examples**		
Naive Empirical Example	**Crucial Experiment Example**	**Conventional Example**	**Naive Empirical Example**	**Crucial Experiment Example**	**Conventional Example**
Students consider a few confirming examples that are convenient to check or randomly chosen.	Students consider confirming examples that are selected based on some kind of strategy.	Students consider all confirming examples in the domain of a mathematical statement.	Students consider a few contradicting examples that are convenient to check or randomly chosen.	Students consider contradicting examples that are selected based on some kind of strategy.	Students consider all contradicting examples in the domain of a mathematical statement.

Generalizations

Mason, Burton and Stacey (2010) define the process of generalizing as "moving from a few instances to making guesses about a wide class of cases" (p. 8). Dörfler (1991) uses generalization as both "an object and a means of thinking and communicating" (p. 63). There is no doubt that the process of generalizing is a crucial component of mathematical learning and should be at the center of mathematics classrooms at all levels (Blanton, Levi, Crites, & Dougherty, 2011; Mason, Burton, & Stacey, 2010). Yet, studies investigating the processes of students' generalizing often report their difficulties in recognizing, employing, and constructing general statements (English & Warren, 1995). Although what type of knowledge counts as general may differ among students or mathematicians, to capture a mathematical relation from a given set or a common element across cases and to transport it to a new set or adjust an idea to incorporate a larger range of phenomena requires complex way of thinking and reasoning (English & Warren, 1995).

Dörfler (1991) classifies generalizations into two categories: empirical and theoretical. According to Dörfler, empirical generalization is based on detecting the common features or qualities of objects, which is why it is considered problematic. Theoretical generalization is, in contrast, constructed through abstracting the essential qualities such as relations among objects and starts with a "system of action" (Dörfler, 1991, p. 71). Harel describes students' pattern generalization in a similar fashion as "result pattern" and "process pattern" generalizations: "in process pattern generalization, students focus on regularity in the process, whereas in result pattern generalization, they focus on regularity in the result" (2001, p. 11). This paper uses a similar distinction to describe students' behavior when they encounter both confirming and contradicting examples. When the students see the regularity in the process and extract this to transform it into more general reasoning, it is categorized a "process pattern generalization" in this study. Conversely, when students only observe the regularity in the results with no further investigation of the underlying structure, it is called "result pattern generalization."

Argumentation

Mathematical reasoning and proof are viewed as an essential component for deep mathematical learning (e.g., Harel & Sowder, 1998; NCTM, 2000; Stylianides & Stylianides, 2009). Yet, research indicates that students of all levels of education have serious issues with constructing arguments and tend to rely on specific examples to determine the validity of mathematical statements (aka empirical proof scheme) (e.g., Healy & Hoyles, 2000). Stylianides (2007) states that the main difference between invalid mathematical reasoning (empirical reasoning) and valid way of reasoning (deductive reasoning) lies in the modes of argumentation. That is, while empirical arguments provide a nonsecure method to verify the truth of a mathematical generalization by treating only a proper subset of all the cases, proofs provide conclusive evidence by considering all cases covered in the domain of a generalization.

Harel (1998) refers to the desire for a more secure validation method as an intellectual need. Thus, the main question of this study is whether students feel the need to provide conclusive evidence for treating all elements in the domain of a generalization. Stylianides and Stylianides (2009) adopt justification schemes based on the distinction of whether participating students feel an intellectual need for more secure mathematical reasoning by drawing upon the existing related literature (e.g., Balacheff, 1988; Harel & Sowder, 1998). A similar approach was followed in this study when evaluating the arguments constructed by participating students (see Table 2).

Table 2. Argument Types and Their Characteristics

Empirical Argument	Non-Empirical Argument	Conventional Argument
Students construct arguments based on the confirming evidence offered by a few cases (Balacheff, 1988).	Students recognize empirical arguments as insecure methods for validating a mathematical generalization, but construct arguments that deviate from mathematical proofs (Stylianides & Stylianides, 2009).	Students recognize the necessity of and can construct proofs as a secure method for validating a mathematical generalization (Stylianides & Stylianides, 2009).

Mathematically Gifted Students

Researchers have demonstrated that mathematically gifted students show greater patience and persistence during problem solving processes compared to their peers (Budak, 2012) and they are more inclined to demonstrate intellectual curiosity and creativity during problem solving processes (Hong & Aqui, 2004). Given that mathematically gifted students often go beyond just finding an answer to grasp for structural relations and search for multiple strategies (Gorodetsky & Klavirb, 2003), the tasks that require constructing multiple strategies, generalizations and/or justifications might then better serve to their creativity (Berg & McDonald, 2018). Similarly, Leikin (2021) argues that mathematically gifted students are more flexible and creative when required to solve problems in multiple ways. Researchers also indicate that mathematically gifted students' ways of thinking and reasoning are similar and parallel to those of mathematicians (Leikin, 2021; Sriraman, 2004). That is, mathematicians attempt to form an intuition about truth of a mathematical statement by consciously trying to construct examples and counterexamples prior to adopting more formal methods of establishing truth (Alcock & Inglis, 2008; Sriraman, 2004). Thus, it could then be hypothesized that mathematically gifted students also show tendency of constructing various types of examples, generalizations as well as justifications to form their intuition while verifying mathematical statements. All these characteristics of mathematically gifted students constitute the reasons for why they were selected as the participants of this study.

Methods

Phenomenography research is a method used in educational studies to depict what students perceive from the same concept (Trigwell, 2006). This reveals the diversity in students' perceptions and allows researchers to map them qualitatively (Marton, 1986). This method is a technique widely used in educational research to analyze what different individuals understand or how they perceive specific concepts (Wihlborg, 2004). Over time, this method has evolved and developed as a tool used in educational research to understand why some students learn better than others. This was a phenomographic study in that it focused on revealing how students reason or experience cognitive conflict, categorizing their reasoning processes by investigating their similarities and differences, and revealing the connections between these categories.

Participants

Given that the purpose of this study was to delve into the process of reasoning about mathematical statements that allow both conforming and contradicting examples, it was hypothesized that mathematically gifted students might have more desire and drive to construct, refine, and explore mathematical generalizations and arguments. Thus, the participants consisted of eight students enrolled at a Science and Art Center in Turkey, where they were selected based on their Wechsler Intelligence Test scores.

Science and Art Centers are the institutions that provide education at primary and secondary levels in Turkey, where the education practices of the gifted are carried out. The Centers were established in 1997 with the Decree-Law on Special Education to support gifted students according to their abilities in addition to the education they receive in formal primary and secondary schools (Ministry of National Education [MoNE], 2022). Students in these centers should be nominated mostly by their elementary school teachers to be accepted by these schools. The Centers administer a two-step test to candidates before accepting the students. One of the tests is the Total Ability Test and the other one is the Wechsler intelligence test (WISC-R). The WISC-R test is an individual intelligence test developed to determine the mental performance of individuals (Wechsler, 1975). Students with a score of 130 and above according to the WISC-R test are considered as gifted and have the right to register with the Centers (Akkanat, 2004).

The participants of the study were determined using a purposive sampling method. Purposive sampling is a technique widely used in qualitative research to obtain information-rich data for the effective use of limited resources (Patton, 2002). Since the purpose of the study was mainly to investigate students' reasoning skills, it was necessary to recruit students who could express themselves and were unafraid of explaining their thoughts. Therefore, the authors contacted the principals and the mathematics teachers of a Science and Art Center in Turkey and informed them about the aims of the study. The Center's mathematics teacher was asked to nominate students whose communication skills were suitable for the purpose of the study. Students with both good communication skills and good math scores were nominated by their math teachers and volunteered to participate in the study. Murat, Kemal, Melike, and Nilay attended eighth grade while Binnur, Ceylin, Demet, and Reyyan (all pseudonyms) attended seventh grade during the time this study was conducted. Although the participating students were at different grade levels, it should be noted that comparing the reasoning skills of the students at different grade levels was beyond the scope of this study. All required permissions were obtained from the participating students' parents.

Task-Based Interviews

A semi-structured interview protocol was designed to investigate student reasoning when faced with mathematical statements that allowed for the construction of both contradicting and confirming examples.

Do you think it is true or false? Explain your reasoning.

Mathematical Statement 1: "The sum of three consecutive numbers is always an odd number."
Mathematical Statement 2: "If the difference between nominator and denominator is smaller, then the fraction is bigger."
Mathematical Statement 3: "At least one diagonal of a quadrilateral cuts the quadrilateral in two triangles with the same area."
Mathematical Statement 4: "If the perimeter of a rectangle then increases its area also increases."

Figure 1. The interview tasks

The interview tasks (Figure 1) included four mathematical statements, which were designed by reviewing the existing literature (e.g., Ball, Hoyles, Jahnke, & Movshovitz-Hadar, 2002; Ma, 1999; Zaskis & Chernoff, 2008). According to the curriculum implemented in Turkey, students learn all the underlying concepts in these interview tasks (i.e., fractions, quadrilaterals, perimeter, and area formulas) in the fifth and sixth grades (MoNE, 2018). Thus, the tasks were designed to fall within the conceptual reach of the students given that the underlying concepts had already been covered in their education. However, the tasks could still cause cognitive conflict because both confirming and contradicting examples could be suggested for each statement.

The interviews were conducted with each participant by one of the authors, and all interviews were video recorded. The researchers have not worked with or observed this group of students before. The role of the researchers was to probe students to reflect on their thoughts during the interviews.

During the individual interviews, the participants were presented with each mathematical task one by one, and they were provided enough time to work on each task. All interviews took approximately 45-60 minutes. The participants were provided with a tablet computer, paper and pencils during the interviews and they were encouraged to draw figures and write down the examples, generalizations or justifications that they explained on the paper, or the tablet provided to them at the beginning of the interviews. All papers and screenshots were collected for data analysis. Various probing questions were employed to better conceptualize the verification processes of the participants. The examples that the participants constructed while verifying the presented mathematical statement consisted of some of the probing questions (e.g., Why do you suggest those examples?, If I ask you to suggest another example, what would it be?). When the participants constructed generalizations regarding the examples that they suggested during the interviews such as "these examples show that it [the statement] would not always hold true" or "if we try out the numbers up to 10, then it will continue in the same way", the interviewer probed the participants by making a reference to their generalizations to achieve a deeper understanding of their use of examples. In addition, the arguments that the participants constructed were also employed as the probing questions (e.g., Would you convince your teacher with this argument?, Do you think that you justified your claim?).

Data Analysis

The data analysis began with transcribing the individual interviews and reviewing the participants' responses to each mathematical statement. A constant comparative method (Glaser & Strauss, 1967) was utilized to illustrate the participants' reasoning when constructing examples, generalizations, and justifications as follows: (1) the authors independently reviewed all the responses and identified the examples, generalizations, and justifications that the participants constructed during the interviews; (2) the authors independently coded the students' examples, generalizations, and justifications using coding schemes constructed after reviewing the existing literature (Table 2); and (3) the authors compared their coding and identified any mismatches in their schemes, after which they worked together to generate new codes or adjust existing codes.

The coding of the participants' responses occurred in three phases. In the first phase, the examples constructed by the participant for each task were coded in two categories as *Confirming*—when the participants suggested examples that illustrated the cases in which the task would hold true—and *Contradicting*—when the participants recognized that the tasks were false and then suggested examples that illustrated the cases in which the task would not hold true. Later, the types of contradicting and conforming examples were coded according to the notion of example spaces for validation as described by Stylianides and Stylianides (2009) (see Table 1). For instance, Nilay stated: "I wanted to see if it [the statement] would be true for a quadrilateral with all sides equal, opposite sides equal and none of the sides equal" to explain her strategy for considering a square, a rectangle and a trapezoid as examples. Given that she suggested the examples that were selected based on a strategy and she referred to that strategy explicitly, the examples were coded as crucial experiment examples.

In the second phase of the coding process, the participants' generalizations identified were coded in two main categories as result pattern generalization and process pattern generalization and two subcategories for each main category as true and false. When the students recognized regularity of the results of the examples that they constructed with no purpose of further investigating the underlying structure within these examples, the generalizations were coded as result pattern generalizations. When the students did not only see the generality through the examples, but they also recognized and captured a relationship across these examples, the generalizations were coded as process pattern generalizations. For instance, Reyyan claimed: "If there are more even numbers than odd numbers, then the result would come out as an odd number or vice versa" for the mathematical statement 1. As evident in her statement, she not only recognized that the statement was false, but she also captured a true relationship across the examples she suggested previously. Thus, her generalization constituted an example of true process pattern generalization.

In the last phase of the coding process, the arguments that the participants constructed were coded as empirical, non-empirical and conventional based on whether a secure or non-secure method was employed to verify the truth of the generalization from a mathematical standpoint (see Table 2). For instance, Murat justified he statement [Mathematical Statement 1] as follows: "There will be three consecutive numbers. In this case, it is either even-odd-even or odd-even-odd. If we do odd-even-odd, we add two odd numbers, and it will be an even number. Then, even number and even number will make an even number". Murat knew that adding odd-even-odd numbers would come out as an even number which would contradict the statement. His justification was not built upon using examples. Yet, his justification failed to provide an insight into why adding two odd numbers would come out as an even number. All these caused his justification to be classified as non-empirical argument.

Analyzes were carried out simultaneously by two researchers, and coding reliability was calculated, which shows coding similarities and differences. In order to obtain inter-coder reliability, the reliability formula specified by Miles and Huberman (1994) was employed ($Reliability = Consensus\ /\ (Consensus + Disagreement)\ x\ 100$). Obtaining a reliability percentage of at least 70% between two coders is necessary for the reliability of data analysis process (Yıldırım & Şimşek, 2006). In this study, the coding reliability percentage was found to be 0.94. The coders were in complete agreement about whether students proposed confirming and contradicting examples for the statements, or whether the generalizations they proposed were true or false. They differed only in a few cases on the types of contradicting examples that the students suggested. For example, Murat stated: "…There will be three consecutive numbers. In this case, it is either even-odd-even or odd-even-odd. If we do odd-even-odd, we add two odd numbers, which will be an even number. Then, we add an even number to an even number. Let's say these numbers are 1,2,3. Adding 1 to 3 is 4, and then adding 2 to the sum will be 6, which is an even number. That's why I say wrong!" One of the researchers coded the student's proposed examples as crucial experiment examples, on the grounds that the student chose the examples strategically. However, the other researcher stated that the student used examples as a tool to clarify the general idea and moreover suggested a way to construct example clusters. Thus, the examples suggested by Murat in this case should be considered as conventional examples. The discussions continued until a full consensus was reached, and Murat was included in the conventional example column as a result (see Table 3).

Findings

In this section, the types of examples and generalizations that students constructed for each mathematical statement will be shared first and then, the types of arguments constructed by the students for their generalizations will be documented. Later, the relationships between the examples suggested by the students and the generalizations and arguments that they constructed will be focused.

The Types of Examples and Generalizations Constructed by the Students

Table 3 below displays the example and generalization types constructed by the students for the first mathematical statement. As can be seen in the table, all students were able to come up with both confirming and contradicting examples for the statement and then to conclude that the statement was false. Although all students were able to recognize that the statement would not always hold true, not all students were able to generalize regarding the domain of the statement correctly. Given that the students were not limited to construct only one generalization, one student, Reyyan, constructed a true and false process pattern generalization based on the examples constructed.

Table 3. The Types of Examples and Generalizations Constructed for Mathematical Statement 1

Confirming Example			Contradicting Example			True Result Pattern Generalization	True Process Pattern Generalization	False Process Pattern Generalization
N.E.	C.E.	C.	N.E.	C.E.	C.			
Binnur	Demet	Kemal	Binnur	Demet	Kemal	Demet	Demet	Reyyan
Ceylin	Reyyan	Melike	Ceylin	Reyyan	Melike	Murat	Murat	
		Nilay			Murat	Melike	Melike	
		Murat			Nilay	Nilay	Nilay	
						Reyyan	Reyyan	
						Kemal	Kemal	
						Binnur		
						Ceylin		

Note. N.E stands for Naive Empirical Example, C.E. stands for Crucial Experiment Example and C stands for Conventional Example

Two students, Binnur and Ceylin, constructed both confirming and contradicting examples for the first statement. However, both students did not see the generality into the specific examples and could not capture common elements across cases. The examples proposed by these students rather seemed randomly chosen. These two students were able to conclude that the statement was false after constructing a contradicting example. Yet, they had no purpose of further investigating in which cases the statement would hold true.

Binnur: I tried out some examples. When I add 1,2,3, the sum is 6. So, it is a false statement!

Interviewer: If I ask you to suggest another example, what would it be?

Binnur: For instance, we could add 4,5,6. The sum would be 15. When we add 1,2,3, it becomes 6, but when we add 4,5,6, it becomes 15. But it's still wrong.

In the excerpt above, Binnur was able to suggest valid contradicting and confirming examples. Yet, she had no purpose of further investigating common elements in these examples. Rather, she provided the examples of 1, 2, 3, which seemed randomly chosen, and refuted the statement right after encountering the contradicting examples by stating: "…So, it is a false statement!". Not attempting to investigate the common elements among the examples nor selecting examples strategically constituted the reasons of why the examples were coded as naive empirical examples. Ceylin acted similarly.

Murat suggested the same examples—1, 2, 3. However, he used the examples to illustrate the process pattern he recognized. He stated: "…There will be three consecutive numbers. In this case, it is either even-odd-even or odd-even-odd. If we do odd-even-odd, we add two odd numbers, which will be an even number. Then, we add an even number to an even number. Let's say these numbers are 1,2,3. Adding 1 to 3 is 4, and then adding 2 to the sum will be 6, which is an even number. That's why I say wrong!" As evident in his statement, Murat construed a more general idea and used the examples only as a tool to describe this idea. His statement: "… Let's say these numbers are 1, 2, 3. Adding 1 to 3 is 4, and then adding 2 to the sum will be 6" demonstrates that he construed the general idea (If we do odd-even-odd, we add two odd numbers, which will make an even number) first and then tried to illuminate his idea by using these specific examples and adding these numbers in this specific order. Murat not only suggested confirming and contradicting examples, but he also demonstrated a way of constructing clusters of examples. Thus, his examples represented conventional examples in this case.

Demet: I think it is false, because 0,1, 2 are consecutive numbers but no, it did not happen! The sum is 3 and it is an odd number. 1, 2, 3 are consecutive numbers but the sum of these three numbers is 6. It is not an odd number. 6 is an even number. It says the sum is an odd number. That's why I said it is false!

Interviewer: Why did you choose these examples?

Demet: I wanted to start from the beginning. That's why I started with 0.

As can be seen above, Demet thought that the statement would be false from the start and attempted to demonstrate this idea with an example which she thought would be a contradicting example. Yet, she confronted a cognitive conflict when the example she suggested served as a confirming example: "I think it is false, because 0, 1, 2 are consecutive numbers but no, it did not happen! The sum is 3 and it is an odd number." After confronting a cognitive conflict, she tended to suggest another example —1, 2, 3— and she concluded that the statement was false. Demet stated that she chose 0, 1, 2 since she wanted to start with the number 0 purposefully, which is why the examples she suggested were classified as crucial experiment examples. Reyyan was the other student who chose the examples she provided based on a strategy.

Reyyan: I tried out some of the consecutive numbers in my mind and the result came out mostly an odd number. But I tried 5, 6, 7 and the sum was 18. 18 is an even number, so it is false!

Interviewer: Why did you pick those numbers out?

Reyyan: I always try to do the numbers up to 10. Because the numbers will continue based on the number 10. Later, it will continue as 11, 12, 13. So if we try out the numbers up to 10, then it will continue in the same way.

Reyyan claimed that since the last digit of a number would always be from 0 to 9, it would suffice to check out the sum of three consecutive numbers up to ten. She argued that the same pattern would continue for bigger numbers. Reyyan's effort of checking out the numbers up to ten constituted her strategy while suggesting examples for the statement (coded as crucial experiment examples). Later when asked why adding two odd numbers makes an even number, she stated: "For instance, 1 and 3 make 4. 3 and 5 add up to 8 or 5 and 7 add up to 12. It will continue like this. We look at the last digit in bigger numbers. Last digit will always be 1, 2, 3, 4, 5, 6, 7, 8, 9 or 0. Two odd numbers should always make an even number." The last digit generalization constructed by Reyyan would indeed be valid while adding two odd numbers and/or identifying odd and even numbers; however, Reyyan was not able to construct a structural relationship among her strategy with the statement [Mathematical Statement 1]. Rather she simply generalized that the last digits would always be the same in each number without making it apparent why and how this generalization could be applicable. Furthermore, the last digit strategy would not be useful to identify whether adding three consecutive numbers would result in odd or even number. Reyyan should have based her strategy off of the number of even or odd numbers instead. Thus, her last digit generalization was coded as false process pattern generalization. Later, Reyyan was able to reach a generalization which was coded as true process pattern generalization by further analyzing the structural relationships among the examples she suggested as evident in her following statement: "If there are more odd numbers then the result will come out as an even number or vice versa."

The students further struggle with deciding whether the second statement would always hold true. Table 4 demonstrates the results regarding the responses to the mathematical statement 2 cumulatively below.

Table 4. The Types of Examples and Generalizations Constructed for Mathematical Statement 2

Confirming Example			Contradicting Example			True Result Pattern Generalization	True Process Pattern Generalization	False Process Pattern Generalization
N.E.	C.E.	C.	N.E.	C.E.	C.			
Demet	Melike	Kemal	Binnur	Demet	-	Kemal	Kemal	Kemal
Binnur	Murat		Nilay	Kemal		Melike		Melike
Nilay			Reyyan	Melike		Nilay		Nilay
Ceylin				Ceylin		Ceylin		Ceylin
Reyyan				Murat		Reyyan		Reyyan
						Binnur		
						Demet		
						Murat		

Most students struggled with determining when the second statement would hold true and failed to capture common elements across cases correctly, which resulted in a false process pattern generalization as evident in Table 4 above.

> *Ceylin: I chose a proper and an improper fraction to show that this [the statement] is false. I chose one proper and one improper fraction of the same denominator. For instance, let these fractions be 2/5 and 7/5. The difference between the nominator and denominator in 2/5 is 3 and the difference is 2 in 7/5. The smaller difference in this case is in 7/5 and it is bigger. This supports the statement! But I wanted to find a case that did not support it [the statement]. Let it be 2/1, or 2/3 and 9/3, let's try these out. The difference between the nominator and denominator in 2/3 is 1 and 6 in 9/3. So, the smaller difference is in 2/3 but 9/3 is bigger than 2/3. This now makes it false. So, I think the statement is not always true.*

Ceylin argued that the statement was false in the cases of comparing one proper and one improper fraction of the same denominator. Although she encountered a contradicting example (the fractions of 2/5 and 7/5) that caused a cognitive conflict, she dismissed the example and tried to find out a confirming example instead. Other students showed a similar tendency of generalizing incorrectly by arguing that the statement would be false for improper fractions. For instance, Melike stated: "The statement would be true for comparing proper fractions, but it would be opposite if the fractions were improper fractions."

Kemal, on the other hand, was the only student who correctly investigated in which cases the statement held true. He stated: "It [the statement] could be true or false for proper fractions. But if two fractions that we compare have the same denominator and they both are proper fractions then the statement would be true. Let's say it is 2/10 and the other is 9/10. It would always be true for the fractions of the same denominator. The reason is that the difference between the nominator and denominator is smaller. Then, it means that the fraction is close to a whole and it would be bigger."

It was observed that all students proposed examples that either contradicted or confirmed the statement. However, one of the findings of the study is that not all students were able to successfully identify the structural relationships between the examples they suggested. Although the participants concluded that the statement was not always true after encountering a contradicted, they failed to determine under which conditions the statement would be true or false.

Table 5. The Types of Examples and Generalizations Constructed for Mathematical Statement 3

Confirming Example			Contradicting Example			True Result Pattern Generalization	True Process Pattern Generalization	False Process Pattern Generalization
N.E.	**C.E.**	**C.**	**N.E.**	**C.E.**	**C.**			
Demet	Nilay	Kemal	Demet	Nilay	Kemal	Kemal	Kemal	Demet
Binnur		Melike	Binnur		Melike	Melike	Melike	Nilay
Ceylin		Murat	Ceylin		Murat	Nilay	Murat	
Reyyan			Reyyan			Reyyan		
						Ceylin		
						Binnur		
						Demet		
						Murat		

All students were able to suggest examples that contradicted or confirmed mathematical statement 3. However, not all students determined the structural relationships among these examples successfully. As a result, while all students successfully concluded that statement 3 was false after confronting a contradicting example, most of the students failed to determine the structural pattern among the cases in which the statements would hold true. Two students—Demet and Nilay— for instance stated a relationship which was erroneous. Nilay stated: "It is not correct! It is true for a square and a rectangle, but not true for a trapezoid…I wanted to see if it [the statement] would be true for a quadrilateral with all sides equal, opposite sides equal and none of the sides equal… whether this statement is true depends on the side lengths of the quadrilaterals. If the corresponding sides are proportional, then it is true!" Nilay chose the examples she considered strategically (crucial examples) and reached a false generalization by arguing that the statement would be true if the side lengths of the quadrilaterals were proportional. When asked why the areas of the two triangles formed by a diagonal in a trapezoid would be different, Nilay answered: "The angles are different [pointing to the corresponding interior angles of the triangles] and the sides are not proportional." She regarded the situations of similarity as a determining condition for the statement. Unlike Nilay, Kemal considered triangle congruency and area formula of triangles in acquiring triangles of the same area.

Kemal: At least one diagonal of a quadrilateral could cut it into two triangles with the same area in regular quadrilaterals. But it didn't say regular quadrilateral here. For example, the shape I just drew was not a regular quadrilateral and it [the diagonal] did not split the area in half. The area of a triangle is half of the base times height. The bases of the triangles [referring to the triangles constructed by the diagonal in Figure 2] are the diagonal here, but the heights are different so are the areas.

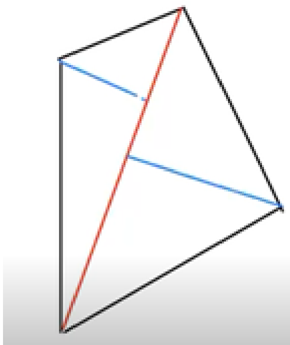

Figure 2. A contradicting example constructed by Kemal for the mathematical statement 3

Kemal further explained what he meant by regular quadrilaterals by saying: "…. In these quadrilaterals [regular quadrilaterals], the triangles are the same, the sides of them are the same so are the areas." Kemal named the quadrilaterals in which diagonals cut the quadrilateral into two congruent triangles as regular quadrilaterals such as a rectangle, a square or a parallelogram. Then, he generalized that the areas of the triangles formed by a diagonal of the quadrilateral to which he referred as regular quadrilateral should be the same given that the triangles would be congruent. Although his definition of a regular quadrilateral was erroneous, his generalization that he based off of triangle congruency held true.

Most students struggled to determine whether the presented mathematical statement would hold true as evident in Table 6. Kemal, Melike and Binnur were the only students who correctly concluded that there would not always be a constant relationship between the perimeter and area of a rectangle.

Table 6. The Types of Examples and Generalizations Constructed for Mathematical Statement 4

Confirming Example			Contradicting Example			True Result Pattern Generalization	False Result Pattern Generalization	True Process Pattern Generalization	False Process Pattern Generalization
N.E.	**C.E.**	**C.**	**N.E.**	**C.E.**	**C.**				
Demet	-	Kemal	-	Melike	Kemal	Kemal	Demet	Kemal	Demet
Nilay						Melike	Nilay	Binnur	Nilay
Ceylin						Binnur	Ceylin		Ceylin
Murat							Murat		Reyyan
Reyyan							Reyyan		
Melike									
Binnur									

Among these three students, Kemal and Binnur were also able to construct a true generalization regarding when the statement would be false. Binnur claimed: "If both sides of a rectangle get increased so does its area. But if one side of the rectangle is decreased, then the statement might be false, that is, the area might decrease as a result." Although Binnur was correct about her claim, her claim was dissociated from formal analysis and, as a result, solely remained based on her intuition. She attempted to construct a contradicting example by stating: "Let's say the length is 10 cm and the width is 5 cm. If we decrease the length by 2 cm or let's say it will be decreased by 7 cm, then the length will be 3 cm. The width will increase and be 7 cm, and its area will change". However, the example she suggested served as a confirming example instead. Binnur did not attempt to perform formal mathematical analysis, nor did she attempt to suggest a contradicting example. Rather, her intuition seemed self-evident to her, and she was very confident that her intuition was true.

Kemal constructed a true generalization as follows: "… it is wrong. Here's the thing. When one side length is the same in two rectangles, the other side length of the rectangle should be increased in measurement to get the perimeter increased. In this case, it is true. But if we decrease one of the sides, then it might not always hold true. So, for example, let's say the sides of a rectangle would be 3 and 2. If I increase one side and make it a 5 and decrease the other side, let's make it 1, the perimeter would be increased, but the area would decrease."

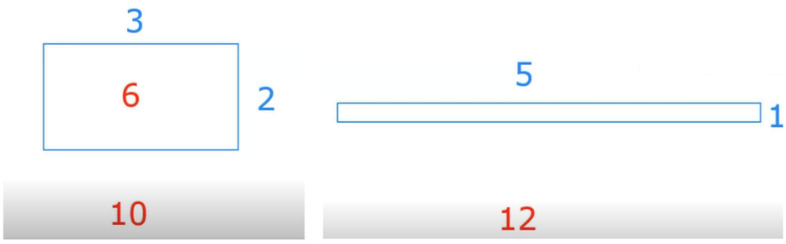

Figure 3. A contradicting example constructed by Kemal for the mathematical statement 4

Most students argued that there is a constant relationship between the two measures, which resulted in a false generalization for mathematical statement 4. For instance, Nilay stated: "Perimeter of a rectangle is the sum of all four sides. In this case, when perimeter increases, the side lengths should also increase. The multiplication of two side lengths of a rectangle gives us the area of the rectangle. Since the side lengths increase so does the area." Nilay, like the other students, proposed a false generalization and justified her generalization based on what Stavy and Tirosh (1996) refer to as "the intuitive rule: More A- More B". Although this generalization might hold true for the case of increasing both or only one pair of opposite sides of a rectangle, it does not hold true all the time (see Figure 3).

The Types of Arguments Constructed by the Students

Table 7 below displays all types of arguments constructed by the students cumulatively. As can be seen in the table, the students struggled with constructing arguments to justify their generalizations. They either constructed no argument or constructed arguments that deviated from being a general mathematical argument.

Table 7. Types of Arguments Constructed by the Students

	No Argument	False Argument	Empirical Argument	Non-Empirical Argument	Conventional Argument
M.S.1	Melike Binnur Nilay Ceylin		Demet Reyyan	Murat Kemal	
M.S.2	Melike Binnur Ceylin Nilay Demet Murat Reyyan				Kemal
M.S.3	Ceylin	Nilay	Demet Binnur Reyyan	Murat	Kemal Melike
M.S.4	Binnur	Demet Nilay Ceylin Reyyan Murat	Melike		Kemal

Note. M.S. stands for Mathematical Statement

For mathematical statement 3, Demet correctly concluded that the statement would not always hold true for all quadrilaterals and was able to construct a valid contradicting example as in Figure 4. Yet, she did not know how to justify that none of the diagonals of the quadrilateral she drew would cut the quadrilateral in two triangles of the same area.

Figure 4. A contradicting example constructed by Demet for the mathematical statement 3

Demet: For instance, we could check the side lengths of the triangles. We could only be certain if we measure the lengths of the sides or know the side lengths. The triangles in this shape [referring to the quadrilateral in Figure 4] will look different, one will look smaller, and the other one will look bigger. We cannot be sure right now without any measurement results.

Demet argued that without knowing the side lengths of the triangles, whether the areas were the same could not be justified. Demet's tendency to rely on the measurement results (i.e. side lengths) to justify her generalization was classified as empirical argument. Melike, on the other hand, constructed an argument based on the area formula of a triangle to justify her generalization, which was classified as conventional argument.

Melike: This statement would be true for regular quadrilaterals but would be false for irregular ones. Since in a regular quadrilateral one diagonal cuts the quadrilateral in two congruent triangles but it might not happen in irregular ones. For instance, I drew a general quadrilateral like this, and the sides and the angles are different. It would not hold true for it.

Figure 5. A contradicting example constructed by Melike for the mathematical statement 3

Interviewer: How would you justify that?

Melike: The heights of these triangles would be different, but the bases are the same, the diagonal. Thus, the areas would be different in this case.

Whether there is a relationship between the types of examples and generalizations constructed by the students with the types of arguments will be addressed next.

The Relationship Between the Types of Examples with the Types of Generalizations and Arguments Constructed by the Students

The results of this study demonstrated that the students were successful at suggesting contradicting examples for most of the statements and reaching the generalization of the sameness in the results of these examples (coded as result pattern generalization). However, the students were not as successful at reaching process generalizations and constructing mathematical arguments (see Table 3, 4, 5, 6 and 7). We hypothesized that the types of examples that the students construct to verify the statements could inform the types of generalizations and arguments constructed by the students. The findings of this study supported this hypothesis by demonstrating that the students who suggested the types of examples coded as crucial experiment and conventional examples were more successful at generating process pattern generalizations and constructing non-empirical and conventional arguments. On the contrary, the students who suggested examples that were randomly selected (coded as naive empirical examples) struggled to identify the structural relationships among the examples and to construct arguments. To better demonstrate the link between the types of examples and the generalizations and/or arguments constructed by the students, the cases of Binnur and Kemal would be described more in depth in the following section.

Table 8. Types of Examples, Generalization, Argument Constructed by Binnur

	Confirming Examples	Contradicting Examples	Result Pattern Generalization	Process Pattern Generalization	Argument
M.S.1	N.E.	N.E.	True	-	No argument
M.S.2	N.E.	N.E.	True	-	No argument
M.S.3	N.E.	N.E.	True	-	No argument
M.S.4	N.E.	N.E.	True	True	No argument

Table 8 displayed that Binnur suggested examples that were coded as naive examples for the mathematical statements. Table 8 also displayed that while Binnur constructed true result pattern generalizations, she could not construe process pattern generalizations except the last mathematical statement. That is, Binnur was successful at the process of constructing a generalization by looking at several cases and identifying the sameness of the results among these cases. However, the generalizations in the eyes of her occurred by not dwelling in the particularities of the examples suggested. When it comes to constructing arguments, as can be seen in Table 8, Binnur failed to construct an argument at all.

Table 9. Types of Examples, Generalization, Argument Constructed by Kemal

	Confirming Examples	Contradicting Examples	Result Pattern Generalization	Process Pattern Generalization	Argument
M.S.1	C.	C.	True	True	Non-empirical
M.S.2	C.	C.E.	True	True + False	Conventional
M.S.3	C.	C.	True	True	Conventional
M.S.4	C.	C.	True	True	Conventional

Kemal, on the other hand, suggested mostly conventional examples and he was more successful to reach process pattern generalizations and to construct arguments coded as conventional arguments. All these results of the students will be discussed under the lights of the current literature in the following part.

Discussion

This paper investigated the types of reasoning that mathematically gifted students employed when faced with mathematical statements for which both conforming and contradicting examples could be suggested. It was hypothesized that the statements that contained both contradicting and confirming examples could constitute a productive venue to investigate the attempts that go beyond suggesting a counterexample. Thus, the purpose of this study was twofold: (1) investigating the examples, generalizations, and justifications constructed by the students while verifying the correctness of mathematical statements for which both confirming and contradicting examples could be suggested and (2) investigating whether there was a relationship between the examples suggested by the students and their generalizations and arguments.

Researchers suggest that employing patterns that do not always hold true could be essential for teaching students to recognize the limitations of empirical arguments as a valid way of proving (e.g., Ball et al., 2002). This study showed that employing such patterns could also be essential for teaching students to construct various types of examples, investigate the structural relationships between examples, and construct generalizations and arguments. The mathematical statements provided a productive venue for the participating students to construct various types of examples, generalizations, and arguments. Thus, such statements could be efficient for implementing recommendations for creating mathematics classrooms in which students explore, construct, or refine mathematical conjectures and use a variety of reasoning to justify or disprove them (CCSSM, 2010; NCTM, 2000). Many researchers have advocated for classroom environments in which students can engage in such activities as rich learning opportunities that mathematically gifted students require to meet their intellectual needs (Berg & McDonald, 2018; Leikin, 2021; Sriraman, 2004). Some might argue that constructing various types of examples, generalizations, or arguments might not be detected if these tasks were implemented by a different group of participating students. Although these concerns are important and the results of this study cannot be generalized to all students, such mathematical tasks could still serve as a rich learning opportunity to foster students' mathematical thinking and reasoning skills.

Employing counterexamples as an instructional tool to create cognitive conflict to achieve conceptual change has been widely adopted (Gal, 2019; Zaskis & Chernoff, 2008). Zaskis and Chernoff (2008) demonstrated that the convincing power of different types of counterexamples when creating and confronting cognitive conflict can vary. This study showed that encountering a contradicting example did not always result in the students confronting a cognitive conflict or creating a conceptual change. For instance, when Ceylin encountered a contradicting example for her generalization regarding mathematical statement 2, she said: “This [the examples she suggested] supports the statement! But I wanted to find a case that did not support it [the statement]. Let it be 2/1, or 2/3 and 9/3, let’s try these out!” The examples Ceylin suggested as contradictions confirmed the statement. However, Ceylin attempted to suggest other examples that supported her initial claim instead of abandoning or modifying it. Similarly, Binnur generalized that “if one side of the rectangle is decreased, then the statement might be false, that is, the area might decrease as a result.” Binnur attempted to suggest an example to confirm her generalization by stating: “Let’s say the length is 10 cm and the width is 5 cm. If we decrease the length by 2 cm or let’s say it will be decreased by 7 cm, then the length will be 3 cm. The width will increase and be 7 cm, and its area will change.” However, the example she suggested contradicted her claim. The conflicts that emerged from the unexpected results of Ceylin and Binnur’s suggested examples thus did not serve as a bridge to a mathematically sound claim (Zaskis & Chernoff, 2008). When looking at the examples they suggested for the mathematical statements, they were usually randomly selected most of the time (coded as naive examples). Gal has indicated various parameters, such as difficulties with formal reasoning or a poor understanding of conceptual data, as reasons students may not confront or resolve a cognitive conflict in the ways envisioned by educators (2019, p. 241). Further analyzing the participating students’ examples, the students’ behavior while suggesting examples or confronting/resolving a cognitive conflict was similar. Both Ceylin and Binnur disregarded the examples they had suggested rather than abandoning and/or modifying either their claims or the examples. In other words, they did not show any tendency toward further analyzing the underlying reasons for the unexpected results. Similarly, they demonstrated a lack of intellectual curiosity while selecting examples for verifying or supporting their claims. Hadamard (1945) has called the ability to discern or to choose “mathematical creativity.” The participating students who thought more rigorously while choosing examples could thus be referred to as more creative. Thus, mathematical creativity could be counted as an important parameter for confronting or resolving a cognitive conflict.

The findings of this study demonstrated that intuition also plays an essential role in confronting or resolving cognitive conflict. Ben-Zeev and Star described intuition from a classical intuitionist perspective as “the answer becomes self-evident immediately” (2001, p. 5). For mathematical statement 4, 5 students—Demet, Nilay, Ceylin, Reyyan, and Murat—generalized that a constant relationship between the perimeter and the area of a rectangle, so that if the perimeter of a rectangle increases, so does its area. Stavy and Tirosh (1996) labeled this intuitive rule “More A–More B.” Nilay reasoned: “The multiplication of two side lengths of a rectangle gives us the area of the rectangle. Since the side lengths increase, so does the area.” Nilay, like the other four students, constructed her argument for her false process pattern generalizations based on this intuitive “More A–More B” rule. Thus, it would not be wrong to state that intuition also plays an essential role in constructing arguments and should be considered a parameter for students’ tendency to respond inconsistently to mathematical tasks. Although Yopp (2015) would regard such responses as problematic, false arguments constitute an essential aspect of students’ mathematical reasoning. Péter-Szarka (2012) has argued that false responses indicate learners’ mathematical reasoning skills and constitute an essential part of their reasoning.

The findings also demonstrated that the role of the different examples mattered in the students' generalization acts. The students mostly construed true result pattern generalizations for mathematical statements. That is, they were able to conclude that the statements were false when faced with contradicting examples. However, the students who constructed conventional examples were able to construe not only result pattern generalizations but also process pattern generalizations. That is, those who employed conventional examples while verifying the correctness of the mathematical statements were more successful at unpacking the statement condition. For instance, Kemal generalized that Mathematical Statement 2 "would always be true for the fractions of the same denominator. The reason is that the difference between the nominator and denominator is smaller. Then, it means that the fraction is close to a whole and it would be bigger." The examples chosen were informative regarding students' generalization acts and the types of generalization they constructed. During the mental processes of constructing examples, especially those coded as conventional examples, many associated processes were brought into play. Thus, the examples suggested by the students not only illuminated their cognitive processes but also the generalization types that they constructed. The mathematical practice of "look for and make use of structure" (CCSSM, 2010, p. 8) was more evident when the students constructed conventional examples.

Stylianides and Stylianides (2009) posited that students' justification schemes reflect their example spaces for validation. The findings of this study demonstrated that the types of examples suggested by the students were also informative regarding students' justification schemes along with the types of generalization. That is, all participating students were able to suggest valid contradicting examples for the mathematical statements (except for the last mathematical statement) and to conclude that the statements were false. However, not all students were able to see the common elements or to identify the structural relationships among these examples and nor were they able to refine the original mathematical statements. As a result, not all students successfully constructed valid mathematical arguments for their generalizations (see Table 7). For instance, Ceylin and Binnur mostly suggested examples coded as naive examples while verifying the presented mathematical statements and they were less successful at constructing arguments for their generalizations. On the contrary, the students who employed mostly conventional examples while verifying mathematical statements were more successful at constructing non empirical and conventional arguments. Given that justifying is a very demanding task, and it is influenced by different parameters such as knowledge of argumentation (Harel & Sowder, 1998; Stylianides & Stylianides, 2009) and skill with the mathematics register (Epp, 2003; Mata-Pereira & da Ponte, 2017), we believe that it is important to know the informative factors regarding students' justification behaviors.

Suggestions

The purpose of this study was to address the ways in which mathematically gifted students reason when faced with both confirming and contradicting examples for a mathematical statement. It was hypothesized in the study that the statements in which both confirming and contradicting examples could be suggested might cause cognitive conflict and support intellectual curiosity of the students to propose various examples, generalizations and justifications, the types of attempts to go beyond solely proposing a counterexample. The findings of the study demonstrated that employing such mathematical statements was indeed essential to get the students to construct various types of examples, to investigate the structural relationships between the examples, and to construct generalizations and arguments. This study was conducted with eight mathematically gifted students which could also serve as a limitation. Therefore, it could be suggested to conduct the study with different groups (i.e. students with different academic achievements) and in different settings (i.e. small group interviews) to see whether mathematical statements that contain both confirming and contradicting examples could foster students' example, generalizations and justification construction processes.

The findings of this study also demonstrated that the participants dealt with cognitive conflict caused by encountering a counterexample differently. For instance, Ceylin and Binnur chose to ignore the counterexample while Kemal investigated further to understand why the examples contradicted the statement. Studies nominated various factors including difficulties with formal reasoning, lack of intuition or poor understanding of the data for not confronting a cognitive conflict or resolving the conflict as envisioned by educators. We believe that lacking a good understanding of the concepts involved in the statements might also hinder students' processes of confronting a cognitive conflict. However, investigating the effects of content knowledge on the ways of dealing with cognitive conflict was beyond the scope of this study since the data collected would fall short of making such claims. Thus, it could be suggested to conduct studies to further delve into this issue.

Furthermore, the findings of this study demonstrated that the types of examples suggested by the students were informative regarding students' justification schemes along with the types of generalization. Thus, it could be suggested that the types of examples constructed by learners should be considered as an instructional and/or research tool to promote as well as to analyze students' justification schemes.

References

Akkanat, H. (2004). Üstün veya özel yetenekliler. In M. R. Şirin, A. Kulaksızoğlu, & A. E. Bilgili (Eds.), *Seçilmiş makaleler kitabı* (pp. 167-193). İstanbul: Çocuk Vakfı.

Alcock, L. J. (2004). Uses of example objects in proving. In M. J. Hoines & A. B. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (pp. 17-24). Bergen, PME 28.

Alcock, L., & Inglis, M. (2008). Doctoral students' use of examples in evaluating and proving conjectures. *Educational Studies in Mathematics*, *69*(2), 111-129. doi:10.1007/s10649-008-9149-x

Alcock, L., & Weber, K. (2010). Undergraduates' example use in proof construction: Purposes and effectiveness. *Investigations in Mathematics Learning, 3*(1), 1-22. doi:10.1080/24727466.2010.11790

Balacheff, N. (1988). Aspects of proof in pupils' practice of school mathematics. In D. Pimm (Ed.), *Mathematics, teachers and children* (pp. 216-235). London: Hodder & Stoughton.

Balacheff, N. (1991). Treatment of refutations: Aspects of the complexity of a constructivist approach to mathematics learning. In E. von Glaserfeld (Ed.), *Radical constructivism in mathematics education* (pp. 89-110). Hingham, MA: Kluwer.

Ball, D. L., Hoyles, C., Jahnke, H. N., & Movshovitz-Hadar, N. (2002). The teaching of proof. In L. I. Tatsien (Ed.), *Proceedings of the International Congress of Mathematicians* (pp. 907-920). Beijing: Higher Education Press.

Ben-Zeev, T., & Star, J. (2001). Intuitive mathematics: Theoretical and educational implications. In B. Torff & R. J. Sternberg (Eds.), *Understanding and teaching the intuitive mind: Student and teacher learning* (pp. 29-56). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Berg, D. H., & McDonald, P. A. (2018). Differences in mathematical reasoning between typically achieving and gifted children. *Journal of Cognitive Psychology, 30*(3), 281-291. doi:10.1080/20445911.2018.1457034

Blanton, M. L., Levi, L., Crites, T., & Dougherty, B. J. (2011). *Developing essential understandings of algebraic thinking for teaching mathematics in grades 3-5*. Reston, VA: National Council of Teachers of Mathematics.

Buchbinder, O., & Zaslavsky, O. (2009). A framework for understanding the status of examples in establishing the validity of mathematical statements. In M. Tzekaki, M. Kaldrimidou, & H. Sakonidis (Eds.), *Proceedings of the 33rd Conference of the International Group for the Psychology of Mathematics Education* (pp. 225-232). Thessaloniki: Aristotle University of Thessaloniki.

Budak, I. (2012). Mathematical profiles and problem solving abilities of mathematically promising students. *Educational Research and Reviews, 7(*16), 344-350. doi:10.5897/ERR12.009

Common Core State Standards for Mathematics. (2010). *Common core state standards mathematics.* VA: Author.

Dörfler, W. (1991). Forms and means of generalization in mathematics. In A. J. Bishop, S. Mellin-Olsen, & J. Dormolen (Eds.), *Mathematical knowledge: Its growth through teaching* (pp. 61-85). Dordrecht: Springer Netherlands.

Ellis, A. B., Ozgur, Z., Vinsonhaler, R., Dogan, M. F., Carolan, T., Lockwood, E., … Zaslavsky, O. (2019). Student thinking with examples: the criteria-affordances-purposes strategies framework. *The Journal of Mathematical Behavior,* 53, 263-283. doi:10.1016/j.jmathb.2017.06.003

English, L., & Warren, E. (1995). General reasoning processes and elementary algebraic understanding: Implications for instruction. *Focus on Learning Problems in Mathematics, 77*(4), 1-19.

Epp, S. S. (2003). The role of logic in teaching proof. *The American Mathematical Monthly*, *110*(10), 886-899. doi:10.1080/00029890.2003.11920029

Gal, H. (2019). When the use of cognitive conflict is ineffective-problematic learning situations in geometry. *Educational Studies in Mathematics,* 102, 239-256. doi:10.1007/s10649-019-09904-8

Giannakoulias, E., Mastorides, E., Potari, D., & Zachariades, T. (2010). Studying teachers' mathematical argumentation in the context of refuting students' invalid claims. *The Journal of Mathematical Behavior*, 29, 160-168. doi:10.1016/j.jmathb.2010.07.001

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New Brunswick, NJ: Aldine Transaction.

Goldenberg, P., & Mason, J. (2008). Shedding light on and with example spaces. *Educational Studies in Mathematics, 69*(2), 183-194. doi:10.1007/s10649-008-9143-3

Gorodetsky, M., & Klavirb, R. (2003). What can we learn from how gifted/average pupils describe their processes of problem solving?. *Learning and Instruction, 13*(3), 305-325. doi:10.1016/S0959-4752(02)00005-1

Hadamard, J. W. (1945). *Essay on the psychology of invention in the mathematical field*. Princeton, NJ: Princeton University Press.

Harel, G. (1998). Two dual assertions: The first on learning and the second on teaching (or vice versa). *The American Mathematical Monthly, 105*(6), 497-507. doi:10.1080/00029890.1998.12004918

Harel, G. (2001). The development of mathematical induction as a proof scheme: A model for DNR-based instruction. In S. Campbell & R. Zaskis (Eds.), *Learning and teaching number theory* (pp. 185-212). Norwood, NJ: Ablex Publishing Corporation.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in collegiate mathematics education III* (pp. 234-283). Providence, RI: American Mathematical Society.

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education, 31*(4), 396-428. doi:10.2307/749651

Hong, E., & Aqui, Y. (2004). Cognitive and motivational characteristics of adolescents gifted in mathematics: Comparisons among students with different types of giftedness. *Gifted Child Quarterly, 48*(3), 191-201. doi:10.1177/001698620404800304

Komatsu, K. (2010). Counter-examples for refinement of conjectures and proofs in primary school mathematics. *Journal of Mathematical Behavior, 29*(1), 1-10. doi:10.1016/j.jmathb.2010.01.003

Leikin, R. (2021). When practice needs more research: The nature and nurture of mathematical giftedness. *ZDM-Mathematics Education*, 53, 1579-1589.

Ma, L. (1999). *Knowing and teaching elementary mathematics. Teachers' understanding of fundamental mathematics in China and the United States*. Erlbaum: Mahwah.

Marton, F. (1986). Phenomenography - a research approach to investigating different understandings of reality. *Journal of Thought, 21*(3), 28-49. Retrieved from https://www.jstor.org/stable/42589189

Mason, J., Burton, L., & Stacey, K. (2010). *Thinking mathematically* (2nd ed.). Harlow: Pearson.

Mata-Pereira, J., & da Ponte, J. P. (2017). Enhancing students' mathematical reasoning in the classroom: teacher actions facilitating generalization and justification. *Educational Studies in Mathematics, 96*(2), 169-186. doi:10.1007/s10649-017-9773-4

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: SAGE.

Ministry of National Education. (2018). *Ortaokul matematik dersi (1, 2, 3, 4, 5, 6, 7 ve 8. sınıflar) öğretim programı*. Ankara: MEB.

Ministry of National Education. (2022). *Science and art centers directive*. Retrieved from https://orgm.meb.gov.tr/meb_iys_dosyalar/2022_12/06214921_BIYLIYM_VE_SANAT_MERKEZLERIY_YOYNERGESIY.pdf

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

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: SAGE.

Pedemonte, B., & Buchbinder, O. (2011). Examining the role of examples in proving processes through a cognitive lens: The case of triangular numbers. *ZDM Mathematics Education, 43*(2), 257-267. doi:10.1007/s11858-011-0311-z

Péter-Szarka, S. (2012). Creative climate as a means to promote creativity in the classroom. *Electronic Journal of Research in Education Psychology*, *10*(28), 1011-1034. doi:10.25115/ejrep.v10i28.1547

Piaget, J. (1975). *The child's conception of the world*. Totowa, NJ: Littlefield, Adams.

Sriraman, B. (2004). Gifted ninth graders' notions of proof: Investigating parallels in approaches of mathematically gifted students and professional mathematicians. *Journal for the Education of the Gifted*, *27*(4), 267-292.

Stavy, R., & Tirosh, D. (1996). Intuitive rules in science and mathematics: The case of 'more of A - more of B'. *International Journal of Science Education, 18*(6), 653-667. doi:10.1080/0950069960180602

Stylianides, A. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education*, *38*(3), 289-321. doi:10.2307/30034869

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, 307-332. doi:10.1007/s10857-008-9077-9

Stylianides, A. J., & Stylianides, G. J. (2022). Introducing students and prospective teachers to the notion of proof in mathematics. *Journal of Mathematical Behavior,* 66, 100957. doi:10.1016/j.jmathb.2022.100957

Stylianides, G. J., & Stylianides, A. J. (2009). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education,* *40*(3), 314-352. doi:10.5951/jresematheduc.40.3.0314

Trigwell, K. (2006). Phenomenography: An approach to research into geography education. *Journal of Geography in Higher Education, 30*(2), 367-372. doi:10.1080/03098260600717489

Yopp, D. A. (2015). Prospective elementary teachers' claiming in responses to false generalization. *The Journal of Mathematical Behavior,* 39, 79-99. doi:10.1016/j.jmathb.2015.06.003

Yopp, D. A., Ely, R., Adams, A. E., Nielsen, A. W., & Corwine, E. C. (2020). Eliminating counterexamples: A case study intervention for improving adolescents' ability to critique direct arguments. *The Journal of Mathematical Behavior,* 57, 100751. doi:10.1016/j.jmathb.2019.100751

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

Wechsler, D. (1975). Intelligence defined and undefined: A relativistic appraisal. *American Psychologist*, *30*(2), 135-139. doi:10.1037/h0076868

Wihlborg, M. (2004). Student nurses' conceptions of internationalisation in general and as an essential part of Swedish nurses' education. *Higher Education Research & Development*, *23*(4), 433-453. doi:10.1080/0729436042000276459

Yıldırım, A., & Şimşek, H. (2006). *Sosyal bilimlerde nitel araştirma yöntemleri* (6^{th} ed). Ankara: Seçkin Yayıncılık.

Zaskis, R., & Chernoff, E. J. (2008). What makes a counterexample exemplary?. *Educational Studies in Mathematics, 68*(3), 195-208. doi:10.1007/s10649-007-9110-4

Zeybek Şimşek, Z. (2021). "Is it valid or not?": Pre-service teachers judge the validity of mathematical statements and student arguments. *European Journal of Science and Mathematics Education, 9*(2), 26-42. doi:10.30935/scimath/10772