Seventh-grade students' perceptions of qualities in a mathematical argument

Sigrid Iversen

Seventh-grade students' perceptions of qualities in a mathematical argument

Sigrid Iversen

Norwegian University of Science and Technology, Norway; sigrid.iversen@ntnu.no

This paper reports on a task where seventh-grade students evaluated five pre-written arguments designed to display various proofs and non-proof arguments. The analysis focuses on what students described as qualities of the arguments. The results indicate that students appreciate arguments they perceive to understand, arguments that are short, and arguments containing text. Thus, the task approach holds the potential to unravel what students perceive as the qualities of an argument. However, there is no clear relationship between the features of the arguments and what students perceive to be qualities. Further investigation should occur to see how the task can be improved to better display aspects of valid mathematical arguments to help students appreciate these and, in turn, be able to produce mathematical proofs.

Keywords: Primary education, mathematics, argumentation, proof.

Introduction

Proof holds a prominent role in mathematics, and researchers and policymakers worldwide increasingly appreciate its importance for mathematics learning (Stylianides et al., 2017). In Norway, 'Reasoning and argumentation' is one of six core elements in the new curricula implemented since autumn 2020. It states that the students should prove that their solutions to mathematical tasks are valid (Kunnskapsdepartementet, 2020). This formulation introduces proving to the primary school curricula in Norway (age 6 to 13), but it is not clear how it could be implemented into practice (Valenta & Enge, 2020). Reasoning and proving in primary education (ProPrimEd) is an intervention-based project that aims to answer this call by developing research-based materials to help teachers implement proving into their teaching. A mathematical proof is here defined using Stylianides' (2007) definition of proof: a kind of argument that uses forms of reasoning and expression that are mathematically valid and suitable for a specific classroom community and uses true statements accepted by the same community. By this definition, it is assumed that students at all grade levels can engage meaningfully in the practice of proving.

This paper reports on a lesson in the intervention conducted in grade 7 (age 12-13). The aim was to prompt students to become aware of the qualities of a good mathematical argument as an entry into work with proving. Lannin (2005) recommends that "research should examine the types of tasks that encourage students to examine the variety of justifications and generalisation strategies that other students use" (p. 254). Thus, this study examines the potential of this task approach. In addition, possible connections between students' evaluation and the designed arguments are explored to see whether the task can help students become aware of the features of valid arguments. The research question is: *What are seventh-grade students' perceptions of qualities of a mathematical argument?*

Theoretical framework and related literature

G. Stylianides' (2008) framework defines categories to analyse students' reasoning and proving activities, following A. Stylianides' (2007) definition. The framework distinguishes between proofs and non-proof arguments, where proofs are demonstrations or generic examples. Without a specific example, a demonstration draws on the properties of and relations between mathematical objects to show why a conjecture is true. This can be done using variables or other means of representing mathematical objects. For example, a random even number could be represented as 2*n* or "pairs of shoes", depending on the community. Counterexamples, contradictions, proofs by induction and proofs by exhaustion are also considered demonstrations. A generic example draws on a particular example and explains the underlying mechanisms to show why a conjecture must be valid for all cases. The affordances of using generic examples to help students move from showing that something is true towards showing why it is true is widely recognised (see e. g. Aricha-Metzer & Zaslavsky, 2019). This suggests that generic examples are a promising entry into work with proving at the primary level.

In G. Stylianides' (2008) framework, a non-proof argument is either an empirical argument or a rationale. An empirical argument consists of showing that a conjecture holds in some cases without showing why, hence providing "inconclusive evidence for the truth" (G. Stylianides, 2008, p. 12). A rationale is introduced as a fourth category to capture arguments not covered by the three former types. It is neither an empirical argument nor a proof but an attempt to prove that either lacks reference to accepted statements or uses statements that are not accepted by the community. In this sense, a rationale can be seen as a proof that misses some of the steps or content needed to convince a given community. The categories described in this section provide the backdrop for the five pre-written arguments presented in the Methods section.

How students perceive mathematical arguments have been investigated earlier, for example, by Bieda and Lepak (2014) and Healy and Hoyles (2000). Both studies show that students are likely to accept empirical arguments as proof. In Bieda and Lepak's (2014) study, the students were the same age as those in this study and had no documented proving experience. They were given two examples of arguments to consider, one empirical non-proof argument and one proof, and were instructed to decide which argument they preferred. In addition, they were asked to describe how the one they did not prefer could be amended to be more convincing. The results indicated that students were inclined to prefer examples accompanied by explanatory text. They both had a numeric example to show that a conjecture holds and text explaining why. The students in Healy and Hoyles' (2000) study had undergone teaching of proving and were given several arguments to consider, such as empirical arguments and proofs, using various modes of representation (e. g., everyday language and algebraic symbols). Their results indicated a discrepancy between what kinds of proofs students themselves would produce and what proof they believed would get the best mark by an evaluator. The students in the study had more success evaluating proofs written in words instead of algebraic notation and found them more convincing. The authors inferred that students' informal and narrative argumentation should be exploited to develop their proof competence. The present study draws on these results by 1) using a variety of informal representations such as contexts, drawings, and narrative explanations, and 2) prompting students to reflect on their proof conceptions by asking them to evaluate and choose among a set of arguments. Both studies described above applied interviews and surveys as data collection methods, while this study will take a different approach by observing group work without the presence of a teacher. This difference allows for insight into the potential of the task.

Method

This study is a single instrumental case study, where the researcher focuses on an issue and uses a bounded case to illustrate it (Creswell & Poth, 2018). Here, the case consists of students who work on a proof-related task. The issue explored is the task's potential to increase students' awareness of the qualities of a good mathematical argument. The study was conducted in spring 2021. A class consisting of 19 seventh-grade students participated, and their regular teacher taught the lesson. According to the teacher, who had taught the group for three consecutive years, the students had not met the term argumentation explicitly in their mathematics instruction. Therefore, the study provides insight into their first meeting with this theme, and this case is thus instrumental in exploring this task as an entry into argumentation. However, previous observation and descriptions given by the teacher suggested that the class was in the habit of showing their work, that is, in detailed writing, when they worked on tasks. The data was collected through video recordings of the students working in groups of three to four, giving five groups. Because of limited access to cameras, three out of five groups were chosen to be videotaped based on the level of verbal interaction observed in earlier lessons. The data material consisted of verbatim transcripts of the video recordings of the group discussions and the groups' written responses, including the groups that were not filmed. Therefore, one group was neither recorded nor given any written reasons for their opinions and is not included in the data material. Hence, the number of participating students was 16, whereby 12 were video recorded, and four submitted a shared written response.

The task, shown in Figure 1 below, was presented to the class by the teacher in plenary along with five pre-written arguments, with no additional information given.

> Mira, one of the students in a class, said:
>
> *The five-times table is easy. When you want to figure out what a number multiplied by five is, you can just multiply it with ten and then divide the result by two.*
>
> Several of Mira's classmates were uncertain that this could be true for all possible numbers. Some said that it could be true, but they were uncertain of why. They explored Mira's conjecture and made their own arguments.
>
> Your task is to
>
> 1. Choose the argument that you think is the best one, and write down two reasons why
> 2. Choose the argument that you think is the worst, and write down two reasons why
>
> Be aware that an argument should be such that it makes us more certain that a conjecture is true for all numbers and that we understand why it is true.

Figure 1: The given task

The arguments, shown in Figure 2 below, were crafted to demonstrate different arguments based on G. Stylianides' (2008) framework and were designed to be perceived as the work of a student their age. Abi's argument is a proof in the form of a generic example, while the rest are non-proof arguments. Hannah's and Inga's arguments are empirical arguments using a few examples, with the distinction that Hannah gives numeric examples while Inga uses a drawing to show an example. Leo's argument uses larger numbers and refers to using a calculator. All three are empirical arguments, while Belma's argument is a rationale, as it contains a part of an argument but lacks logical connections and details to be convincing. None of the arguments is of the form demonstration, which emphasised the difference between generic examples and empirical arguments.

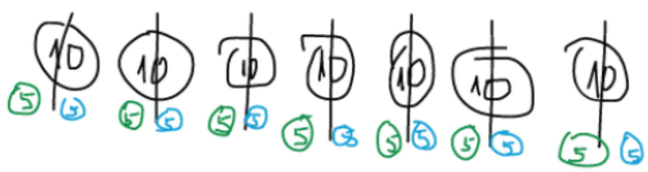
Name	Argument
Hannah (empirical argument)	Yes, it is correct. $2 \cdot 5 = 10 \quad 2 \cdot 10 = 20 \qquad 7 \cdot 5 = 35 \quad 7 \cdot 10 = 70 \qquad 8 \cdot 5 = 40 \quad 8 \cdot 10 = 80$ half half half
Belma (rationale)	Yes. $6 \cdot 5 = 30$ $6 \cdot 10 = 60$ It happens because 5 is the half of 10. $5 + 5 = 10 \qquad 10 : 2 = 5$
Inga (empirical argument, with drawing)	$4 \cdot 10$  $4 \cdot 5$ $4 \cdot 5$ half
Leo (empirical argument, larger numbers)	$3 \cdot 10 = 30 \qquad 26 \cdot 20 = 260 \qquad 268 \cdot 10 = 2680$ $3 \cdot 5 = 15 \qquad 26 \cdot 5 = 130 \qquad 268 \cdot 5 = 1340$ It is always this way. I have used my calculator to check many cases.
Abi (generic example)	It is correct, because: For example, if we want to calculate $7 \cdot 5$. That means to find out how much 7 fives are if added together. We can start by working out what 7 tens are:  Next, we can split each ten into two fives:  7 tens = 7 fives + 7 fives $7 \cdot 10 = 7 \cdot 5 + 7 \cdot 5$ Thus, $7 \cdot 5$ is the half of $7 \cdot 10$! If we use another number, not 7, then it will be just the same. It will only be a different number of tens and fives.

Figure 2: The arguments that were given to the students (translated from Norwegian)

The students were given copies of the arguments and the task and worked on it for 20 minutes before being called back to a whole-group discussion. While the groups worked with the task, the teacher and two researchers observed the groups.

The unit of analysis was a student contribution, either written or verbal. The contributions that regarded one of the arguments and described a positive or negative feature were collected. This gave 36 utterances. The further analysis was performed as an inductive qualitative content analysis (Mayring, 2015). The aim was to understand the different perceptions and their magnitude in the data material. The utterances were coded inductively to capture the feature it addressed. Codes describing related features were then collected into overarching categories. For example, the code "short, positive" and the code "long, negative" both belong to the category "short", as they both suggest the perception that an argument should be short.

Findings

The analysis of the 36 utterances resulted in seven categories, as shown in Table 1 below. In the following, each category is elaborated on in order of appearance in the table.

Table 1: Overview of utterances by category

Categories	Explanation	Short	Text	Order	Drawing	Examples	Warrants
Frequency	15 (42 %)	7 (19 %)	5 (14 %)	3 (8 %)	3 (8 %)	2 (6 %)	1 (3 %)

The most frequent category, explanation, considered utterances related to understanding or explanation. It applied whenever a student stated that an argument was explained well or was easy to understand. Both Hannah's, Leo's, and Belma's arguments were said to be easy to understand, and some students claimed that Hannah "…explained it really well". One student spoke about Inga: "Really bad explanation. I did not understand what she meant. She just drew." Abi's argument is criticised: "…is hard to understand because it is long and messy". Another negative remark about it is that "It is so much strange going on here at once", indicating that it was considered complex by the student and might represent something the student was not used to seeing. Other students appreciated Abi's argument: "Because he explains, for example, that the tens are divided into two". This utterance indicates that a student noticed an essential feature of the generic argument. Other students said, "It has both writing and drawing. Very good explanation. Everyone can understand this. No difficult words were used". Thus, Abi's argument was either valued for its thorough explanation or was not appreciated because of its length and complexity. These examples show that both the short and the more elaborate arguments could be explained well. Hence, what students mean when they say that something explains well or is easy to understand is unclear.

The category short considers utterances about the length of the argument. Hannah's argument was appreciated because "It is simple and short", and about Belma's argument, some said that it was positive "…that she used only one example". "Leo's and Hannah's arguments are good because they did not have too much text to read and understand", while "Abi's argument is hard to understand because it is long and messy". The students appeared to value short arguments because they took up little space and took little effort to read.

Several students mentioned the presence of explanatory text, especially regarding Inga's and Abi's arguments. Abi's argument was valued because it had both writing and drawing, as indicated by the quote in the previous section. Inga's argument was the only one that none of the students preferred, and one student said: "The others have written text. That is why hers is the worst because it is easier to explain by writing than by drawing". Other utterances to support this view are: "She does not explain what it is that she has drawn", and "but she does not explain what she does". Hence, a short argument was not necessarily appreciated if the students did not find the content satisfying.

The four least frequent categories are the order of the argument, the use of drawings, the number of examples, and the warrants. Concerning the order of the argument, one student gave all the utterances, for example: "Leo just starts. Now I don't know, if I start with the first, I don't know (if it is true)". The student seemed to believe that the conclusion, whether the conjecture is valid or not, should be stated at the beginning of the argument. Two utterances were about the number of examples: "There are more examples", "checked on many numbers". This indicates an appreciation of examples and is related to the features that the arguments were meant to display. However, the task intended that the students recognise these arguments as mere examples and not convincing arguments. These utterances suggest the opposite outcome of what was intended. The use of drawings is also mentioned by a few students, either saying that it is good to use a drawing or that the quality of the drawing affected the quality of the argument. The last category, warrants, captured an utterance where the student, in a critical tone, said, while reading from Leo's argument: "I have used my calculator to check many cases. Ok?" indicating that this did not strengthen the argument.

There were few direct references to the features that the arguments were designed to display. For instance, no student commented that Hannah's and Leo's arguments only showed that the conjecture was valid for some examples or that Belma's explanation was incomplete. Instead, as shown above, some remarks suggested that it is good to have many examples. The two utterances appreciating Abi's explanation for being thorough are other examples that indicate a possible awareness of how this argument differs from the rest. Except for these few exceptions, the data shows little awareness of the features of the pre-written arguments.

Discussion

This study offers insight into how students perceive mathematical arguments for general conjectures and suggests that features like the explanation, length, and the presence of text are the qualities that the students in this group value most. However, there is no apparent relationship between students' perceptions and the features that the arguments were designed to display. These findings, along with a discussion on the methodological approach and the task's design, are addressed below.

The appreciation of empirical arguments is evident in this study, as in previous studies (Bieda & Lepak, 2014; Healy & Hoyles, 2000). The inclination to prefer explanatory text is also evident, as Bieda and Lepak (2014) also found. However, the data show that students' reactions are more nuanced. The notion of 'explain' seems to hold divergent meanings, where explanation appears to be a feature connected to whether the mere mathematical content of the argument makes sense or is possible for the reader to understand. This discrepancy might be related to the class habits, where there is an emphasis on showing one's work. To clarify how one has found the correct answer to a mathematical task. Thus, explanation, and in its extension conviction, might concern the presentation of a solution. This perspective is not compatible with assessing arguments for general conjectures. There seems to be a gap to fill to bring the students' attention to the difference between evaluating a written task solution and evaluating whether an argument shows that a conjecture must be valid for all cases. It can be understood as a necessary shift in the socio-mathematical norms in the group, concerning what can count as an acceptable mathematical explanation (Yackel, 2002). At the more practical level, the results emphasise the importance of a well-orchestrated classroom discussion where issues like the difference between showing that and explaining why are addressed. Teachers can benefit from exploring teacher moves to support students' argumentation by pressing to justify why something works. Such actions are suggested by Martino and Maher (1999), who describe questioning that can prompt students' justification when they work on mathematical problems.

Methodologically, this study provides a new lens into students' evaluation of arguments by unravelling how students in groups act without the influence of a teacher or a researcher. The results suggest that students can both explore and verbalise what they perceive to be qualities of arguments but that the nature of these qualities is often distant from what would be accepted by the mathematical community. A limitation of this approach is that it makes it impossible to get further insight into the students' meaning of the words 'explain' and 'understand', which frequently occurs in the data material. A follow-up interview where students are asked to elaborate on their conceptions of these notions could therefore be done to enrich the understanding of the case. As discussed in the previous section, this could provide further insight into how the gap between evaluating a written solution and evaluating an argument could be filled.

This study explores the potentials and challenges of a task where students evaluate others' work, which is an approach recommended by Lannin (2005). The results indicate no clear relationship between argument design and the students' evaluation, but that the task offers an entry point into discussing the qualities of a good mathematical argument. Further study should be made to explore improvements in the task. First, one possible approach is to use fewer arguments. In this task, one could reduce the number of arguments to three: one empirical example, one rationale, and one generic example, the distinction between showing that and explaining why could be highlighted in this way. Second, asking the students to argue for the conjecture themselves before being presented with the pre-written arguments should be explored to see if it might influence how they perceive the arguments. Third, one could consider the suitability of the conjecture. Durand-Guerrier et al. (2012) warn that too simple conjectures can obscure the need for proving. Therefore, it should be explored if conjectures of different complexity have different affordances in this task. Last, an extension of this task could be to find ways to highlight the deductive nature of proofs. A possible approach here is to use valid arguments where the order of the steps is altered and ask students to reorganise the steps to make the argument logical and convincing. Exploring these possibilities could be further steps toward finding fruitful ways to engage primary school students in proving.

References

Aricha-Metzer, I., & Zaslavsky, O. (2019). The nature of students' productive and non-productive example-use for proving. *Journal of Mathematical Behaviour, 53,* 304–322. http://dx.doi.org/10.1016/j.jmathb.2017.09.002

Bieda, K. N., & Lepak, J. (2014). Are you convinced? Middle-grade students' evaluations of mathematical arguments. *School Science and Mathematics*, *114*(4), 166–177.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications, Inc.

Durand-Guerrier, V., Boero, P., Douek, N., Epp, S. S., & Tanguay, D. (2012). Argumentation and proof in the mathematics classroom. In G. Hanna & M. de Villiers (Eds.), *Proof and proving in mathematics education: New ICMI study series 15* (pp. 349–367). Springer.

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

Kunnskapsdepartementet. (2020). *Core elements*. https://www.udir.no/lk20/mat01-05/om-faget/kjerneelementer?lang=eng

Lannin, J. K. (2005). Generalization and justification: The challenge of introducing algebraic reasoning through patterning activities. *Mathematical Thinking and Learning*, *7*(3), 231–258. http://dx.doi.org/10.1207/s15327833mtl0703_3

Martino, A. M., & Maher, C. A. (1999). Teacher questioning to promote justification and generalization in mathematics: What research practice has taught us. *Journal of Mathematical Behavior*, *18*(1), 53–78. https://doi.org/10.1016/S0732-3123(99)00017-6

Mayring, P. (2015). Qualitative content analysis: Theoretical background and procedures. In A. Bikner-Ahsbahs, C. Knipping, & N. Presmeg (Eds.), *Approaches to qualitative research in mathematics education: Examples of methodology and methods* (pp. 365–380). Springer.

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

Stylianides, G. J. (2008). An analytic framework of reasoning-and-proving. *For the Learning of Mathematics*, *28*(1), 9–16.

Stylianides, G. J., Stylianides, A. J., & Weber, K. (2017). Research on the teaching and learning of proof: Taking stock and moving forward. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 237–266). National Council of Teachers of Mathematics.

Valenta, A., & Enge, O. (2020). Bevisrelaterte kompetanser i læreplanen LK20 for matematikk i grunnskolen. *Acta Didactica Norden*, *14*(3). http://dx.doi.org/10.5617/adno.8195

Yackel, E. (2002). What we can learn from analysing the teacher's role in collective argumentation. *The Journal of Mathematical Behaviour*, *21*(4), 423–440. https://doi.org/10.1016/S0732-3123(02)00143-8