



HAL
open science

Structural characteristics of generic arguments using (counter)examples produced by primary students: A Toulmin analysis

Kayo Miura, Yusuke Shinno

► **To cite this version:**

Kayo Miura, Yusuke Shinno. Structural characteristics of generic arguments using (counter)examples produced by primary students: A Toulmin analysis. Thirteenth Congress of the European Society for Research in Mathematics Education (CERME13), Alfréd Rényi Institute of Mathematics; Eötvös Loránd University of Budapest, Jul 2023, Budapest, Hungary. hal-04410970

HAL Id: hal-04410970

<https://hal.science/hal-04410970>

Submitted on 22 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Structural characteristics of generic arguments using (counter)examples produced by primary students: A Toulmin analysis

Kayo Miura¹ and Yusuke Shinno²

¹Hiroshima University, Graduate School of Humanities and Social Sciences, Japan

²Hiroshima University, School of Education, Japan; shinno@hiroshima-u.ac.jp

In this paper, we characterise generic arguments based on examples and counterexamples by primary students in terms of argumentative structures. To accomplish this, we (1) employed a theoretical framework based on Balacheff's levels of validation of proof and (2) conducted a Toulmin analysis adapted from the argumentation analysis described by Reid and Knipping. In sum, the results showed similarities and differences between example-based and counterexample-based arguments in terms of the complexity of their argumentative structures. Based on this analysis, we discuss what distinguishes generic arguments from other types of arguments (e.g., empirical and general).

Keywords: Argumentation, counterexamples, generic examples, primary education.

Introduction

A *generic example* refers to a particular case, but is used as a representation of generality (Mason & Pimm, 1984). This notion was further conceptualised by Balacheff (1988), who framed it as an argument that conveys the characteristic properties and structures of all possible cases without relying on characteristics that are solely applicable to the particular case(s) under consideration. In the literature on proof and proving, researchers have paid special attention to this notion, as it is often seen as an essential argument in the transition between empirical and mathematical proofs (Reid & Knipping, 2010). However, there is no consensus on what constitutes a generic example (also known as a generic argument or generic proof/proving) due to its complexity (e.g., Leron & Zaslavsky, 2013; Reid & Valleo Vargas, 2018). Using Balacheff's theoretical framework and Toulmin's analytical model, this study focused on the structural characteristics of generic arguments, with attention to areas that distinguish them from other types of arguments. Based on the idea of *generic disproof* (Zaslavsky, 2018), this study also developed a framework for analysing example-based arguments to prove a statement and counterexample-based arguments to disprove a statement.

Theoretical framework

Balacheff (1988) proposed a framework consisting of *knowledge*, *formulation*, and *validation*. The levels of validation of proof are particularly well-known. This framework allows us to analyse how (counter)example-based arguments are associated with the transition between pragmatic and mathematical proofs. Referring to previous works (Miyakawa, 2002; Reid & Knipping, 2010), each level can be summarised as follows:

- *Naïve empiricism*: in which the truth of a result is asserted after verifying several cases
- *Crucial experiment*: in which a statement is verified on the basis of testing a special case that is chosen to be not too special (e.g., choosing an extreme case to explore the extent of its validity)
- *Generic example*: in which a special case is used to stand for all possible cases on the basis of the characteristic properties and structures of its class

- *Thought experiment*: in which the representation is detached from special cases and actions; instead, it is based on operations or transformations that involve the characteristic properties and structures often explained by natural language¹.

The current study operated within this framework to characterise and analyse different types of arguments to prove and disprove a statement (Table 1). In principle, we considered a statement to be a general statement (or universal proposition).

Table 1: Framework for characterising arguments based on Balacheff's levels

Balacheff's levels of validation	Proof of a statement	Disproof of a statement
Naïve empiricism	Empirical argument by examples	Empirical argument by counterexamples
Crucial experiment		
Generic example	Generic argument by examples	Generic argument by counterexamples
Thought experiment	General argument by ordinary language	General argument by ordinary language

In terms of structural characteristics, no essential differences exist between the first two levels (as demonstrated later). Therefore, both correspond to empirical arguments in Table 1, although they are comparatively distinct. We categorised arguments by counterexamples based on Peled and Zaslavsky (1997), who proposed three levels in terms of explanatory features, including specific, semi-general (generic), and general counterexamples (Miura & Shinno, 2022). Unlike an empirical argument by examples, one by counterexamples is an acceptable mathematical proof, as a specific counterexample disproves a universal proposition. An argument made via a specific counterexample is therefore not a naïve empiricism but lies “in-between” empirical and generic arguments (Reid & Knipping, 2010, p. 133), as a prover carefully chooses a (counter)example to refuse a statement, which is seemingly associated with the nature of the crucial experiment. Although a specific counterexample shows that a statement is false without reason, semi-general (generic) and general counterexamples both offer explanations (Peled & Zaslavsky, 1997). From the research perspective, it is valuable to link an “argument that proves” with an “argument that disproves”, but few studies have attempted this (e.g., Lee, 2016). Thus, a basic question is as follows: *What characterises the different arguments (empirical, generic, and general) produced by primary students for proving and disproving a statement?* We addressed this by conducting an argumentation analysis using Toulmin's model, which allowed us to understand and compare different arguments based on argumentative structures.

¹ According to Balacheff (1988), this remains marked by anecdotal temporality, but the operations and relationships that form the basis of the proof are designated otherwise than by the results of their use, which is the case for the generic example (p. 165). Balacheff took an example from Cauchy's proof of the intermediate value theorem in 1821, which was mainly written in natural language with fewer symbols. The thought experiment can therefore be understood in the same way as the “discursive generic example” expressed in ordinary natural language.

Methodology

Argumentation analysis

Toulmin's (1958) work has widely been used to study mathematical arguments made by students and/or teachers in different ways (e.g., Knipping, 2008; Pedemonte & Balacheff, 2016). Focusing on argumentative structures in arguments written by primary students, this study adapted the Toulmin model from a method of *argumentation analysis* (AA) described by Knipping and Reid (2019). AA was first proposed to analyse complex arguments in classrooms from local and global argumentative structures. Although AA employs a reduced Toulmin scheme in terms of Claims/Conclusions, Data, Warrants, and Backings, we also used other auxiliary elements (Qualifiers and Rebuttals). An argument aims to establish a claim (C) based on specific data (D) and general warrants (W). Here, a warrant is produced to support the relationship between the data and the claim. If further justification is needed, then a backing (B) is produced to support the warrant. A qualifier (Q) expresses the degree of conviction or confidence to hold the conclusion, while a rebuttal (Reb) states the conditions under what extent the conclusion would (not) hold. We also incorporated the element of refutation (Ref) into the AA from Knipping and Reid (2019). Our method employed three stages to reconstruct written arguments as argumentation structures: 1) assign roles (C, D, W, Q, Reb, B, Ref) to statements; 2) group these into "steps" in which the C of one step is used as D for the next step (a chain "C/D" is called an *intermediate conclusion*, per Knipping and Reid (2019)); 3) interpret the levels of arguments (Table 1) in terms of the entire structure. Although AA permits the analysis of both local and global argumentation structures (the latter consists of different argumentation "streams" composed of different steps), we did not need to do this because we collected written data from students rather than using classroom data.


Context and task

This study investigated how primary students produced arguments by examples and counterexamples. In the Japanese mathematics curriculum, "proof" is an official secondary school content, but there is no explicit opportunity to informally learn this at the primary level. It is therefore important for researchers and teachers to develop a curriculum and task design for proof and proving across grades. To explore potentials and challenges, we designed a set of tasks for 90 Japanese primary students (grades 5–6, ages 10–12). After receiving ethics approval from a committee at our affiliated institution, we implemented these tasks in written form during June 2021 via survey; this consisted of three tasks, including (#1) an introductory task to understand the problem situation, (#2) a proving task with a true statement, and (#3) a disproving task with a false statement (This task (#3) was chosen from a national assessment test conducted among 7th graders in Japan in 2005). Figure 1 shows tasks #2 and #3. This material is well-known as a "number pyramid", wherein students must calculate numbers in each row as follows: add adjacent numbers in the first (bottom) row, find and add two numbers in the second row, and find a number in the third (top) row. Task #1 consists of three sub-tasks, including (#1-1) finding the number in the third row when the three numbers in the first row

are given (e.g., $[3,3,3 \rightarrow 12]^2$), (#1-2) expressing the number in the third row as an addition form (e.g., ‘ $3+3+3+3=12$ ’), and (#1-3) writing a general property to answer the question, ‘If all the three numbers in the first row are the same, then what can you state about the number in the third row?’ The participants were familiar with the number pyramid itself, but not necessarily as a (dis)proving task.

To complete the tasks, all the participating students wrote in their native language (Japanese). We then translated their answers into English. In this paper, we could not consistently consider how ‘the same’ students behaved in tasks #2 and #3 (only one student was considered in this regard). Although this is a limitation of this study, the results of the present analysis show the potential for further analysis tracing individual students’ works.

2 Yama looked at three examples below and considered the following “statement”.




“If the numbers at both ends of the first row are the same, then the number in the third row is always even.”

Is Yama’s idea true? Choose “true” or “not true”. Also, explain the reason why you think so in the box below.

A. True B. Not true

3 Kimu looked at three examples below and considered the following “statement”.



“If the middle number in the first row is even, then the number in the third row is always even.”

Is Kimu’s idea true? Choose “true” or “not true”. Also, explain the reason why you think so in the box below.

A. True B. Not true

Figure 1: Proving and disproving tasks (translated by the authors)

Results and analysis

Results

Using the Table 1 framework, we briefly summarised responses from participants (N=90). For task#2, 73 students ($\approx 81\%$) chose “A. True” correctly with writing reasons, while for task#3 only 37 students ($\approx 41\%$) chose “B. Not true” with writing reasons. It seems that a considerable number of students failed to respond to the disproving task. Although it is worth analysing students’ incorrect choices and their difficulties (e.g., the order of tasks (#2 and #3) might influence students’ behaviours), in this paper, we focus on the participants’ responses which have been classified into the three types of arguments for each task in Table 2.

Table 2: A summary of students’ responses (N=90)

	Example-based (task#2)	Counterexample-based (task#3)
Incorrect choice	10	47
Empirical arguments	31	17
Generic arguments	20	13

² In students’ writings, they often drew a diagram of the number pyramid to represent some instances. For concision, we write $[3,3,3 \rightarrow 12]$ to show that the three numbers in the first row are (3,3,3) and the number in the third row is 12.

General arguments	22	7
Others	7	6

Example-based arguments

In task #2 (Figure 1), Yama's statement is a claim (C), with three examples as data (D) supporting C. Participant S24 (G5) wrote why the statement was true as follows:

S24 (G5): I think it is right. I thought this was correct because when I tried it with other diagrams, like $[8,5,8 \rightarrow 26]$ I thought this was correct because it also appears [as even] when I try other numbers like $[2,5,2 \rightarrow 14, 7,6,7 \rightarrow 26, 4,8,4 \rightarrow 24]$

The additional examples used by S24 are also data (D2) that support C. Moreover, the underlined text works as a naïve empiricism to C. We therefore interpreted this as an empirical argument. Participant S43 (G5) wrote the following:

S43 (G5): There are two sets of the same number. For example, if the numbers in the first step are 3, 2, 3, it is expressed as $3+2+2+3$. Since there are two 3s and two 2s, the two different numbers are added and doubled so that $(3+2) \times 2 = 10$. An even number is a number divisible by 2, so the number in the third row is always even if the numbers at both ends of the first row are the same.

Figure 2 shows a layout of S43's argument, which creates an intermediate conclusion (C/D) referring to characteristic structures included in the task. W1 and W2 convey the general properties. The example used in this argument was taken from the original problem, and works as a generic example to explain. We interpreted such an argumentative structure as a generic argument.

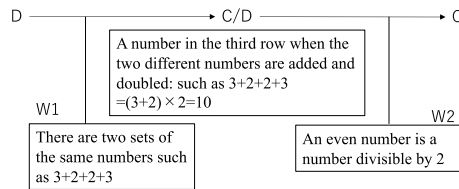


Figure 2: Layout of S43's generic argument by example

Participant S28 (G6) wrote the following:

S28 (G6): Right + Middle [in the first row] = Right in the second row, Left + Middle = Left in the second, and all [two] numbers in the second rows are the same. Multiplying them by 2 means that [the third row] will be an even number

Figure 3 shows a layout of S28's argument, which creates C/D in the layout of a generic argument (Figure 2). W1, W2, and B refer to the structure and properties supporting C, but are expressed as ordinary words rather than numerical examples. Thus, we interpreted this as a general argument.

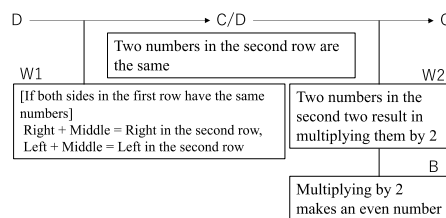


Figure 3: Layout of S28's general argument by ordinary language

Counterexample-based arguments

In task #3 (Figure 1), Kimu's statement is C1, with three examples as D supporting C. To refuse this statement, another claim (C2) may need to negate C1. Participant S43 (G5) wrote the following:

S43 (G5): It just happens by chance³ that the third column is even. Such like the case [3,4,4→15], the even number in the middle number [in the first row] does not necessarily make that the number in the third row is even.

S43's argument a new data-claim relation (D2, C2) with a counterexample to refuse the original statement (C1). We considered the words 'just happens by chance' as a qualifier (Q) to C1. S43's argument is mathematically acceptable based on a specific counterexample (D2), but does not convey the reason why the statement is false. We therefore interpreted this as an empirical argument.

Participant S12 (G6) wrote the following:

S12 (G6): In the above case, it just happens by chance [or sometimes] that the ends of the first row were odd-numbered together and even-numbered together. If one breaks apart the odd and even numbers, the third row will be odd. For example, [see Figure 4]

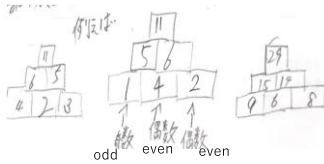


Figure 4: S12's drawing

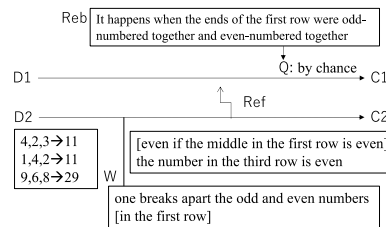


Figure 5: Layout of S12's generic argument by counterexamples

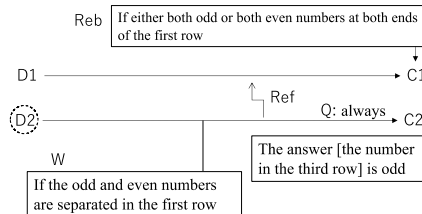


Figure 6: Layout of S42's general argument by ordinary language

Figure 5 shows a layout of S12's argument, which includes the D2-C2 relation as well as W, Reb, and Q. The warrant supporting C2 and the rebuttal (with Q) conditioning C1 refer to the characteristic task structures and explain why C1 is refused. In other words, the participant's understanding of the structures reveals how to generate a counterexample. Thus, we interpreted this as a generic argument.

Participant S42 (G6) wrote the following:

S42 (G6): All three examples [above] are either both odd or both even numbers at both ends [of the first row], so if the odd and even numbers are separated, the answer is always odd. So, I think it is incorrect.

³ The original Japanese word 'たまたま(tamatama)', translated 'by chance', was interpreted as an unlikely or less plausible qualifier.

Figure 6 shows a layout of S42's argument, which explains why C1 is refused using ordinary language rather than numerical examples. Hence, D2 is implicit data supporting C2. The warrant and the rebuttal refer to the characteristic structures, which are similar to those in S12's generic argument.

Discussion and conclusion

In terms of argumentative structures, most arguments consisted of a single argumentation step, but some generic arguments included intermediate conclusions (C/D) for the next step (e.g., Figure 2). Although not shown in this paper, some generic arguments by counterexamples also referred to properties such as "the two numbers in the second row can be either even/odd or odd/even". This feature of generic arguments is consistent with Balacheff's definition of generic examples. By contrast, empirical arguments use a simple structure that often consists of D and C without other elements to support C. The structures of general arguments are very similar to those of the generic argument, albeit with different representations. Comparing example-based and counterexample-based arguments, the latter involves a more complex structure because it requires the inclusion of two parallel D-C relations as well as a refutation (and correction) of the original statement. Then, generic arguments by counterexamples explain the strategic method to generate a counterexample based on the properties and structures of the problem. Different generic arguments may be reconstructed as various layouts, but this paper only provides one layout due to space constraints. These arguments sometimes involve C/D, B, Q, and Reb conditioning C. Although general arguments also convey the reason why the statement is false, their structures are less complex than those of generic arguments. This is likely because general arguments do not involve numerical counterexamples.

In conclusion, the framework based on Balacheff's levels allowed us to identify different arguments (empirical, generic, and general) for both proving and disproving a statement. However, this paper offers a relatively limited empirical analysis of the collected data because we could only present illustrative cases. There are some analyses one might undertake. First, the analysis focusing on how 'the same' students behaved in different tasks consistently (or inconsistently) is promising, although only in the case for S43 are the related analyses available in this paper. Second, the order of the tasks (#1, #2, #3) might also affect the students' behaviours (if a disproving task is presented before a proving task, might students behave differently?). Third, this study targeted the structural aspects of argumentation in terms of a Toulmin analysis, but future research should further enhance our understanding of the nature of proof by considering other aspects (such as linguistic, functional, epistemological). Regarding the linguistic aspects, for example, we considered the difference between generic and general arguments, but we paid scant attention to different informal representations (such as drawings, and ordinary Japanese words) included in students' writings (as well as our translation from Japanese to English). How can such representations (and the translation issues) affect the structural aspects of argumentation? Further research is needed to address these limitations.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Numbers JP19H01668 and JP20K02913.

References

- Balacheff, N. (1988). Aspects of proof in pupils' practice of school mathematics. In D. Pimm (Ed.), *Mathematics, Teachers and Children* (pp. 216–235). Hodder and Stoughton.
- Knipping, C. (2008). A method for revealing structures of argumentation in classroom proving processes. *ZDM - Mathematics Education*, 40, 427–441. <https://doi.org/10.1007/s11858-008-0095-y>
- Knipping, C., & Reid, D. (2019). Argumentation analyses for early career researchers. In G. Kaiser, & N. Presmeg (Eds.), *Compendium for early career researchers in mathematics education* (pp.3–31). Springer. https://doi.org/10.1007/978-3-030-15636-7_1
- Lee, K. (2016). Students' proof schemes for mathematical proving and disproving of propositions. *The Journal of Mathematical Behavior*, 41, 26–44. <https://doi.org/10.1016/j.jmathb.2015.11.005>
- Leron, U., & Zaslavsky, O. (2013). Generic proving: Reflections on scope and method. *For the Learning of Mathematics*, 33(3), 24–30.
- Mason, J., & Pimm, D. (1984). Generic examples: Seeing the general in the particular. *Educational Studies in Mathematics*, 15, 277–289. <https://doi.org/10.1007/BF00312078>
- Miura, K., & Shinno, Y. (2022). A study on students' explanations by counter-examples in elementary school mathematics: Characterising students' responses based on a survey using “number pyramid” as a task. *Journal of Science Education in Japan*, 46(3), 243–257. (In Japanese) <https://doi.org/10.14935/jssej.46.243>
- Miyakawa, T. (2002). Relation between proof and conception: The case of proof for the sum of two even numbers. In A. Cockburn & E. Nardi (Eds.), *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education* (pp. 21–26). PME.
- Pedemonte, B., & Balacheff, N. (2016). Establishing links between conceptions, argumentation and proof through the enriched Toulmin model. *The Journal of Mathematical Behavior*, 41, 104–122. <https://doi.org/10.1016/j.jmathb.2015.10.008>
- Peled, I., & Zaslavsky, O. (1997). Counter-example that (only) prove and counter-example that (also) explain. *FOCUS on Learning Problems in Mathematics*, 19(3), 49–61.
- Reid, D. A., & Knipping, C. (2010). *Proof in mathematics education: Research, learning and Teaching*. Sense Publishers. <https://doi.org/10.1163/9789460912467>
- Reid, D., & Valero Vargas, E. (2018). When is a generic argument a proof? In A. Stylianides & G. Harel (Eds.), *Advances in mathematics education research on proof and proving* (pp. 239–251). Springer. https://doi.org/10.1007/978-3-319-70996-3_17
- Toulmin, S. (1958/2003). *The uses of argument*. Cambridge University Press.
- Zaslavsky, O. (2018). Genericity, conviction, and conventions: Examples that prove and examples that don't prove. In A. Stylianides & G. Harel (Eds.), *Advances in mathematics education research on proof and proving* (pp. 283–298). Springer. https://doi.org/10.1007/978-3-319-70996-3_20