Journal on Mathematics Education
Volume 13, No. 4, 2022, pp. 605-630

Examples and generalizations in mathematical reasoning – A study with potentially mathematically gifted children

Simone Jablonski*, Matthias Ludwig

Department of Mathematics and Computer Science Education, Goethe University Frankfurt, Germany
*Correspondence: jablonski@math.uni-frankfurt.de

Received: 11 August 2022 | Revised: 14 December 2022 | Accepted: 16 December 2022 | Published Online: 17 December 2022

Abstract

Mathematical arguments are central components of mathematics and play a role in certain types of modelling of potential mathematical giftedness. However, particular characteristics of arguments are interpreted differently in the context of mathematical giftedness. Some models of giftedness see no connection, whereas other models consider the formulation of complete and plausible arguments as a partial aspect of giftedness. Furthermore, longitudinal changes in argumentation characteristics remain open. This leads to the research focus of this article, which is to identify and describe the changes of argumentation products in potentially mathematically gifted children over a longer period. For this purpose, the argumentation products of children from third to sixth grade are collected throughout a longitudinal study and examined with respect to the use of examples and generalizations. The analysis of all products results in six different types of changes in the characteristics of the argumentation products identified over the survey period and case studies are used to illustrate student use of examples and generalizations of these types. This not only reveals the general importance of the use of examples in arguments. For one type, an increase in generalized arguments can be observed over the survey period. The article will conclude with a discussion of the role of argument characteristics in describing potential mathematical giftedness.

Keywords: Examples, Longitudinal Study, Mathematical Giftedness, Mathematical Reasoning, Typology

How to Cite: Jablonski, S., & Ludwig, M. (2022). Examples and generalizations in mathematical reasoning – A study with potentially mathematically gifted children. *Journal on Mathematics Education, 13*(4), 605-630. http://doi.org/10.22342/jme.v13i4.pp605-630

Only when you can explain something, have you understood it. A sweeping statement that, when expressed in different terms, finds scientific acceptance. The ability to formulate arguments is not only a central learning goal of mathematics teaching, but is also an important basis for understanding mathematical contexts (Hanna, 2000). Mathematical arguments are not only the building blocks of content but are also methodological tools for deeper understanding. Justifying connections or dealing with assumptions are only two examples of mathematical argumentation activities.

Mathematical reasoning is sometimes considered in the diagnosis, characterization and promotion of mathematical giftedness in students from an outcome-focused perspective through the characterization of the formulated arguments (Heinze, 2006; Maddocks, 2018; Sowell et al., 1990). The relevance of arguments for mathematics education can also be extended to the context of potential mathematical giftedness. This reveals two gaps in research, which are elaborated as follow:

1. the question of how the arguments of potentially mathematically gifted children can be descriptively characterized.
2. a longitudinal description of change in these characteristics.

The question of how to describe the arguments of potentially mathematically gifted children arises from various models of potential mathematical giftedness and how they integrate mathematical reasoning. For example, Käpnick (1998) does not explicitly characterize potentially mathematically gifted third and fourth grade children in terms of the arguments they formulate. Nevertheless, he mentions giftedness-specific abilities that can theoretically have an influence on mathematical reasoning and the resulting arguments. For example, particular skills in recognizing and indicating mathematical structures may have an influence on the discovery of particularities and connections (Fritzlar & Nolte, 2019), whereby, such a discovery would form the basis for the formulation of an argument (Amielia et al., 2018). Furthermore, it is unclear to what extent a particular intuition in potentially mathematically gifted children (Fuchs, 2006; Sriraman, 2004) influences how arguments are linguistically formulated. It is, therefore, unclear whether the arguments of potentially mathematically gifted children can be described in a standardized way, and which possible characteristics can be observed, and which is the research focus of this article.

Furthermore, the need for a longitudinal perspective arises from the predominantly cross-sectional view of potential mathematical giftedness. Despite the consensus that the development of potential giftedness is a multi-year process (Käpnick, 1998), characterizations of mathematical giftedness are usually selective, i.e., for the grade in which the diagnosis is made. Such a selective analysis with regard to mathematical reasoning can be found here in isolated cases (Heinze, 2006; Maddocks, 2018; Sowell et al., 1990). Brunner (2019) concludes, after a corresponding review of the literature that previous studies on reasoning, though not specific for giftedness, have generally been conducted as cross-sectional analyses. Nevertheless, indicators from developmental psychology (Piaget, 1928) on the one hand, and research into giftedness (Käpnick, 1998; Sjuts, 2017) on the other, suggest possible changes in the way that mathematical arguments are formulated between the ages of nine and twelve. For example, the ability for stating general structures emerges during the fifth and sixth grades.

The emergence of this ability to state general structures will be the focus of this article and will be applied to the field of mathematical reasoning. The mathematical arguments of potentially mathematically gifted children are examined longitudinally from a product-oriented perspective regarding, firstly, the use of examples, and secondly, generalized formulations. The aim is to develop a descriptive building block for describing the arguments of potentially mathematically gifted children and to take a crucial step towards a longitudinal analysis. The aim is to make a scientific contribution to the description of potentially mathematically gifted children about mathematical reasoning – a contribution which can be used as a basis for further empirical studies as well as for practical considerations on diagnosis and support. This research goal is pursued as part of a qualitative longitudinal study with potentially mathematically gifted children, which, with its theoretical basis, methodology and results, forms the focus of the article.

Mathematical Arguments: Definition and Positioning from a Product-oriented Perspective

Formulating arguments is a fundamental communicative activity of human beings. Habermas (1984-87) defines reasoning as a type of speech in which participants address contentious claims to validity and attempt to redeem or criticize them with arguments. An argument contains reasons that are linked in a systematic way to the claim to validity of a problematic statement (van Eemeren et al., 1996). This definition allows us to identify four aspects that characterize reasoning in general: Formulating an argument is an activity that takes place in a *social context*, more specifically in a joint oral or written conversation, communication, or interaction with corresponding participants. The starting point for an argument is a *disagreement* about a point of view. Its aim is to *speak for or against a point of view*, to justify something, and/or to increase or decrease the acceptance of a point of view (van Eemeren et al., 1996). For this purpose, reasoning is built up in a *systematic way*. Reasons are presented based on the claim to validity of a disputed viewpoint, so that a gradual and seamless reduction to already recognized statements is created.

These aspects are also relevant for mathematical reasoning, but with one reservation: Mathematical reasoning in a mathematics lesson is usually distinguished from everyday arguments by the absence of a real point of disagreement, which, in this list of aspects, goes hand in hand with the starting point and the goal of reasoning. This divergence highlights the need for a contextual definition of reasoning. *Mathematical reasoning* can be located more in terms of problem solving, explaining and gaining deeper understanding (Baker, 2003). This special characteristic is also accompanied by the establishment of a starting point for argumentation: starting points that require reasoning activities, e.g., questioning statements, justifying discoveries, and formulating an argument. An obvious basis for mathematical argumentation is reasoning tasks that present assertions or call for extreme cases to be considered. If, for example, the consideration of extreme cases is chosen as the starting point for reasoning, an argument might be prompted by *when does a certain case occur?* or *does it always apply? Why or why not?*

When emphasizing a product-oriented perspective, the result of an argumentative activity is referred to in the following article as the *argumentation product*. This includes all statements produced during a reasoning activity, either in writing, orally or non-verbally. Considering their systemic structure, argumentation products are structured by two elements in the context of the article: *discovery* and *justification*. Both elements are necessarily linked in the argumentation product because a discovery lacks certainty without justification. From the students' perspective, what is to be discovered is something new and is not predetermined by the teacher or the task (Amielia et al., 2018). The article defines a discovery in the context of an argumentation product as a statement that is elaborated on and formulated by the students, e.g., based on a mathematical problem. A justification is defined as a statement that is intended to substantiate the formulated discovery by means of generally accepted statements (Toulmin, 2003).

Mathematical Arguments in the Context of Mathematical Giftedness

As stated in the introduction, the focus of this article is on the description of the mathematical arguments of potentially mathematically gifted children. At this point, the relationship to potential mathematical giftedness is examined in greater detail, with the help of the definitions of mathematical reasoning that consider a product-oriented perspective. For this purpose, *giftedness* is understood as an individual performance potential that can develop into a visible above-average performance (Käpnick, 1998). This emphasizes the distinction between giftedness as a potential achievement, and achievement as a visible outcome, which gives rise to the term *potential giftedness*. It is also defined as a domain-specific characteristic. Giftedness in the sense of this article refers to the ability and performance domain of mathematics.

Definitions of giftedness using an IQ score to diagnose, and label seem inappropriate in this understanding of giftedness; firstly, because of the distinction between giftedness and achievement, and secondly, because of the emphasis on domain-specific giftedness. At the same time, this positioning raises the question of appropriate methods for diagnosing potential mathematical giftedness. In the work of Käpnick (1998), this labelling is effected by means of special characteristics of potentially mathematically gifted children (Figure 1, left side). Based on an empirical comparative study, the catalogue of characteristics lists the characteristics of giftedness for children in the third and fourth grades in which significant differences between potentially mathematically gifted children and a comparison group were found. Sjuts (2017) developed an extended catalogue of characteristics for the fifth and sixth grades under comparable conditions. Figure 1 (right side) lists the additional characteristics for children in these grades.

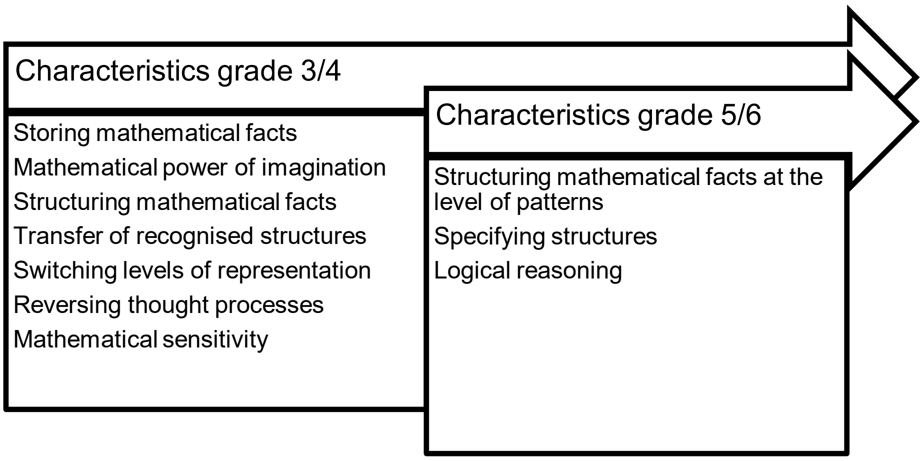

Figure 1. Mathematics-specific characteristics of potentially gifted children in grades 3/4 (Käpnick, 1998; left) and 5/6 (Sjuts, 2017; right)

There are various characterizations used in giftedness research to describe potential mathematical giftedness in the style of Käpnick (1998) and Sjuts (2017). Here, focus is placed on the role of mathematical reasoning in the various characterizations. Durak and Tutak (2019) assume – irrespective of possible peculiarities in the formulation of arguments – that there are no differences between potentially mathematically gifted and normally gifted pupils of primary school age in terms of their basic need for argumentative justifications. In Käpnick's (1998) catalogue of characteristics for potentially mathematically gifted children in the third and fourth grades, peculiarities in formulating arguments are not explicitly included in the characteristics of giftedness, although it remains unclear whether it was relevant in the task selection of the study. In contrast, mathematical intuition is taken into account in the description of typical problem-solving processes (Fuchs, 2006; Sriraman, 2004). Intuition is described as a spontaneous, largely unconscious mental process. Potential forms of this may include sudden insights or ideas for solutions, as well as fragmentary or diffuse justifications or explanations of the problem's solution (Fuchs, 2006). While intuition may yield the correct result when solving problems, the path to the solution often remains implicit for observers. To paraphrase using the structure of an argument: While intuition can be useful in formulating a discovery, it appears to skip the step of explicit reasoning. Gutierrez et al. (2018) propose the hypothesis that "mathematically gifted students […] tend to show unusual paths of reasoning" (p. 170). The authors justify their observation with the special abilities of potentially mathematically gifted children, e.g. "the abilities to identify patterns and relationships among different elements, generalize and transfer mathematical ideas or knowledge from one context to another, or invert mental processes of mathematical reasoning" (Gutierrez et al., 2018, p. 170). This list of facilitating skills also changes the role of mathematical arguments in catalogues of characteristics such as Käpnick's. Special abilities in reasoning might thus be justified theoretically by empirically proven and favorable abilities or preconditions. The list of favorable factors put forward by Ufer et al. (2008) includes mathematical knowledge, metacognitive skills, and the organization of hypotheses and statements. This perspective on mathematical reasoning legitimizes a justification of possible connections with the help of the characteristics of potentially mathematically gifted children, for example:

1. A creative approach to problem solving, whereby the children find different ways of solving problems and play with mathematical relationships and numbers (Assmus & Fritzlar, 2022; Fritzlar & Nolte, 2019; Käpnick, 1998). Creativity in the sense of Creative Mathematical Reasoning seems particularly relevant for discovering relationships together with the use of mathematical knowledge (Joklitschke et al., 2022).
2. Special skills in structuring mathematical facts, through which children can recognize mathematical patterns and form classes (Käpnick, 1998; Sjuts, 2017). This seems to be particularly relevant for discovering relationships, but also for organizing hypotheses and statements, as well as for using strategies of metacognitive control and organization.[1]

In addition to this indirect connection, there are also models of potential mathematical giftedness that explicitly list selected aspects of mathematical reasoning as a special skill. For example, Heinze's (2005) model of giftedness lists the "need for plausible, mathematical explanations and striving for knowledge" and "the ability to formulate exact and complete justifications of mathematical facts" (p. 295) as characteristics of potentially mathematically gifted children of primary school age. The latter in particular would also influence the formulated products of argumentation and their characteristics. Logical reasoning also plays a role in what Sjuts (2017) refers to as an explicit characteristic of giftedness.

It remains unclear from the various perspectives how the mathematical arguments of potentially mathematically gifted children can be described descriptively and whether the children adopt different approaches. The remainder of this article will explore such a description, using one characteristic, namely the role of examples and generalizations in formulating products of reasoning. The next section justifies this choice following an introduction to the theory.

Examples and Generalizations in Mathematical Arguments

Argumentation and proof are understood as two specific forms of justification that refer to different settings and thus partly follow different rules and use different means (Pedemonte, 2002). The concepts are thus related, in particular through the assumption that formal-deductive reasoning is based on a deductive approach with formally correct arguments (Pedemonte, 2002). Argumentation is thus a decisive, propaedeutic aspect of proof. In order to establish the connection and distinction between examples and generalizations in the formulation of arguments, argumentation is first considered in this propaedeutic relationship to mathematical proof. Harel and Sowder (1998) use proof schemes to describe typical procedures (hereafter referred to as types) with which a person tries to convince themselves and others of a mathematical statement. The authors distinguish between three different types: external persuasion, empirical evidence, and analytical evidence. What is interesting in the distinction between arguments with examples and arguments with generalizations is, on the one hand, the empirical type, in which the inductive procedure is described as based on one or more examples – in the sense of

[1] The selection of these two characteristics is intended to illustrate a possible connection by way of example. This is not to rule out further connections between reasoning and the characteristics of potentially mathematically gifted children.

calculations with fixed numerical values, or figures with fixed side lengths and angles. The starting point for proof is an individual case or concrete example against which a statement is tested. The analytical type, on the other hand, proceeds deductively by using generally valid ideas and facts, hereafter generalizations, to provide proof (Harel & Sowder, 1998).

The two types do not have to stay separate from each other. Reid and Knipping (2010) use the following distinction to classify the different roles of examples in proving:

1. Empirically: concrete examples are used for proof without any further classification of them.
2. Generically: concrete examples are used for proof, and these are deemed to represent a class.
3. Symbolically: no concrete examples are used, but words with semantic meaning are used for proof.
4. Formally: No concrete examples or words with semantic meaning are used for proof.
5. Looking back to Harel and Sowder (1998), a generic approach can hence turn an empirical proof into an analytical one.

The remaining article applies the distinction between different types of proof to the level of argumentation and, in particular, to the level of argumentation products. An empirical argument is based on one or more examples, whereas an analytical argument is formulated using mathematical rules and facts. A generic argument describes an analytical argument that has been developed with the help of examples. Along the lines of proof schemes, the concept of argumentation type describes a recurring, regular procedure that a person uses when reasoning and which is reflected by characteristics in the argumentation products that they formulate.

Apart from a descriptive description, the use of examples in arguments has both potential and risks. Examples, if chosen skillfully, bring vividness, comprehensibility and validity as an empirical examination (Komatsu, 2017). Under certain circumstances, they are the starting point for discovering a mathematical relationship in the first place and for deriving a justification from it. Last but not least, its relevance to defining the partial competence of "recognizing the general in an individual case and thus gaining insight into why something is always and by necessity valid or must be valid and, on this basis, being able to develop a valid generalization" (Brunner, 2019, p. 327; translated) is evident. Various authors cite empirical argumentation in the sense of argumentation where the general validity of a statement is erroneously derived from individual examples as a possible problem (Nussbaum, 2011). Unlike generic reasoning, a generally valid argument is not developed here through the example, but the example is used as an empirical derivation of generality. Studies on this show that children in primary school do not usually argue at a general level, and only begin making increasing use of mathematical rules as proof for their arguments in higher secondary school (Koleza et al., 2017). Still, the process of generalizing is seen as a stage of early algebraic thinking, being related to numeric pattern generalization from arithmetic (Kieran et al., 2016; Sari & Ng, 2022).

Once again, the question arises as to which considerations can be applied in connection with potential mathematical giftedness. On the one hand, the use of special intuition in potentially mathematically gifted children raises the question of whether there is a need for examples and generalizations – what is discovered seems intuitive and not worth further explanation. On the other hand, potential factors influencing mathematical reasoning are also relevant for reasoning using examples and rules. While Gutierrez et al. (2018) speak relatively superficially of students' ability to "identify patterns and relationships among different elements" and "generalize and transfer mathematical ideas", the transition from "structuring mathematical facts" (Käpnick, 1998) to "structuring at the pattern level" or "specifying a structure" (Sjuts, 2017) in particular describes a potential influencing factor on generalizations in mathematical arguments. The latter differs from the former by explicitly using representatives to structure (Sjuts, 2017), which is fundamental to a generic approach to reasoning. Here, this also prompts a progression over time or changes in the use of examples and generalizations – if it does not (yet) seem to be a special feature with potentially mathematically gifted children in the third and fourth grades, a generalized indication of a structure and a representative, and hence possibly also generalizations in arguments, play a role in the fifth and sixth grades.

Research Question

From the theoretical outline, the roles of mathematical reasoning in the various characterizations of potential mathematical giftedness can be summarized as follows: explicit mention as a characteristic of potentially gifted children (Heinze, 2005; Maddocks, 2018; Sowell et al., 1990), mention of sub-areas of reasoning as a mathematics-related giftedness characteristic (Gutierrez et al., 2018; Sjuts, 2017), listing of characteristics that could influence the formulation of argumentation products (Käpnick, 1998) and no further analysis as a giftedness-specific characteristic (Durak & Tutak, 2019). These different positions justify taking a closer look at the argumentation products of potentially mathematically gifted children, describing them descriptively and, where appropriate, identifying types of reasoning among the children. The discussion below will focus on a description of empirical and analytical arguments.

Looking back at the comparison of Käpnick's (1998) catalogue of characteristics for the third and fourth grades with Sjuts' (2017) for grades five and six reveals a possible change in the reasoning products of potentially mathematically gifted children over time. Although logical reasoning does not feature in the catalogue for third and fourth grade children, it is included as a giftedness-specific characteristic for fifth and sixth grade children. Structuring skills are also related to the level of patterns from the fifth grade onwards, i.e., they take place at a general, possibly also rule-guided level. This might be relevant when formulating generalizations.

As derived from the theoretical background, selective observations in cross-sectional studies cannot describe individual changes in the products of (empirical and analytical) argumentation. However, this seems to be relevant; firstly, for characterizing the products of argumentation in the context of a long-term giftedness profile, and secondly, for changing the products of argumentation in general (Brunner, 2019). Results on the development of children's thinking are available outside the field of mathematics education. Piaget's stage theory (Piaget, 1928) and the LOGIK study (Schneider & Bullock, 2009) describe cognitive changes in logical thinking and also give indicators for changes in how children formulate general arguments through developmental psychological presuppositions, but without reference to mathematical arguments. The fact that mathematical reasoning and possibly even recurring types – in the sense of long-term characterizations of the products of reasoning (Harel & Sowder, 1998) – change with increasing age is indeed indicated by developmental psychological descriptions and results from research on giftedness. However, direct conclusions are absent due to different study settings and selective analyses. This is where interest arises in focusing not only on the description of argumentation products, but also on long-term change – still subject to a restriction to examples and generalizations. This gives rise to the underlying research question of the study: *Which types can be characterized in potentially mathematically gifted children in relation to the use of examples and generalizations in the products of argumentation?* It is the aim of this article to contribute to the description on how potentially mathematically gifted children relate to mathematical reasoning.

METHODS

Sampling and Period of the Study

Indicators of changes in the products of argumentation – in the use of examples and generalizations – are investigated following an exploratory approach. A longitudinal study with a constant sample and survey method was chosen because of the changes over time. The sample consists of 37 children from the enrichment program Junge Mathe-Adler Frankfurt ("Young Maths Eagles Frankfurt"), (Jablonski & Ludwig, 2021). At the start of the study, all children who were in the third and fourth grade at the time, were invited to participate. Children from the new third grade in the subsequent school term were also added half a year later to allow for panel mortality. Table 1 gives an overview of the characteristics of the sample.

The children are selected for the Mathe-Adler program through nomination by their mathematics teachers on the basis of Käpnick's (1998) catalogue of characteristics. An indicator task test is completed by the children during the first weeks of participation in order to identify the special mathematics-specific characteristics for the third and fourth grade (cf. Figure 1, left) (Fuchs & Käpnick, 2009). This is less about performance-oriented use of the tasks, instead, it is more about the procedures used for solving them. The score achieved on the test does not directly lead to exclusion from the program but is taken as an occasion for further assessment steps (e.g., interviews and discussions with parents). At this point, the score serves to characterize the sample as *potentially mathematically gifted* as defined in the theoretical background. Fuchs and Käpnick (2009) give an average score of 14 points for a group of potentially mathematically gifted children in the third and fourth grades, with 27 maximum achievable points. Table 1 allows this to be compared with the children in the sample.

Table 1. Sample characterization

Group	Number of children fully interviewed	Observation window Grade (half-year) Start; End	Observation window average ages Start; End	Score achieved in indicator test Minimum; Maximum; Average; Median
1	11	4 (2); 6 (1)	10 Y. 3 M.; 11 Y. 9 M.	8; 24; 14,9; 15,5
2	13	3 (2); 5 (1)	9 Y. 6 M.; 11 Y.	13; 26; 19,4; 20
3	13	3 (1); 4 (2)	8 Y. 11 M.; 10 Y. 5 M.	10; 24; 16; 16

The longitudinal study was conducted between April 2018 and May 2020 (see Figure 2). An initial pilot study with 13 children initially took place in April 2018. The main study began for groups 1 and 2 in May 2018 and group 3 in November 2018. In the main study, four surveys were conducted with all children – each with a time interval of six months between the two survey points and a total observation period of 18 months.

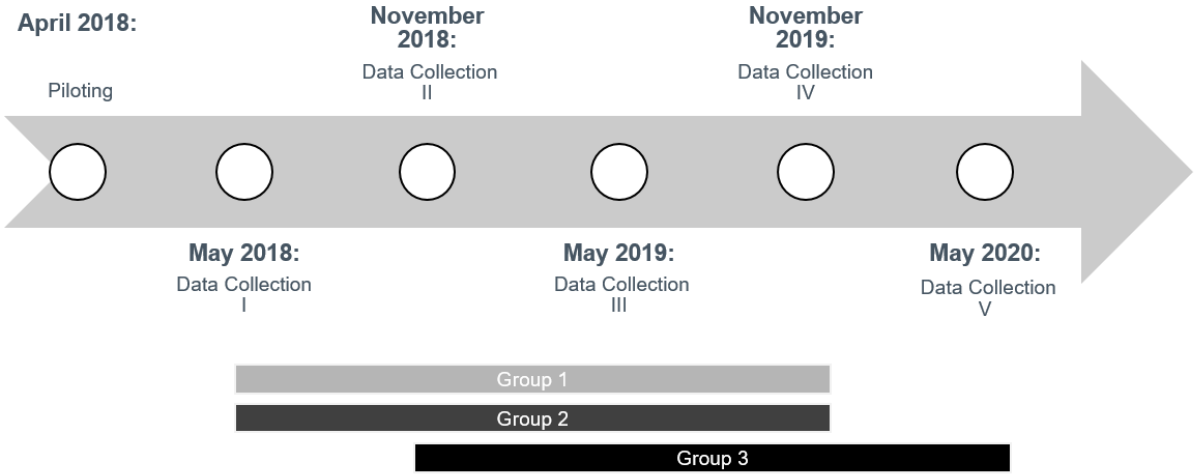

Figure 2. Timeline of the data collection

Classification of the Survey Tool

The children's argumentation products, as defined in the theoretical framework, and possible changes were recorded in individual one-to-one interviews. Oral reasoning was considered advantageous to written reasoning, in that clarity and precision are improved with immediate prompting or by asking follow-up questions. In addition, the fact that "children of primary school age sometimes still find it very difficult to overcome linguistic barriers, to organize their thoughts and to write them down independently" (Bezold, 2009, p. 89; translated) suggests that "oral ability in the area of mathematical reasoning [...] is still significantly higher at primary school age than ability in written reasoning" (Brunner, 2019, p. 328; translated).

The interviews are based on justification tasks from the field of arithmetic with a reference to its potential for reasoning and generalization in the context of early algebraic thinking (Kieran et al., 2016; Sari & Ng, 2022). The basic formats for the interview tasks are the number wall with three basic stones and the non-square number grid with two arrow numbers (see Figure 3, cf. also Bezold (2009); Moor (1980); Wittmann and Müller (1990)). The task formats and tasks are based on a teaching unit by Bezold (2009) with a focus on discovery learning. This study showed the potential to elicit both examples and generalizations. Thus, the tasks are suitable in the sense of the study. The piloting in 2018 including 13 students confirmed this statement.

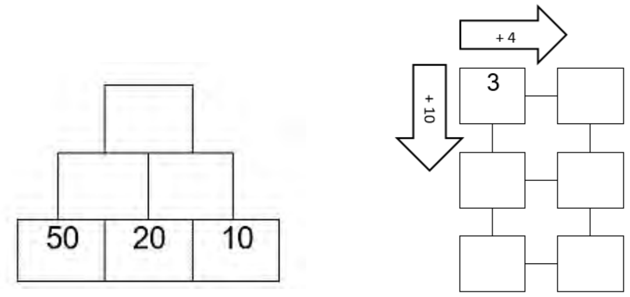

Figure 3. Examples of task formats “number wall” and “number grid”

An interview refers exclusively to one of the two formats. These task formats were chosen because; firstly, the number correlations they contain offer opportunities for reasoning, and secondly, they should not present any computational hurdles to focus on reasoning.

The process of each interview is controlled by standardized guidelines providing the tasks, follow-up questions and space for notes. An interview contains five tasks with prompts for argumentation, although only four of the tasks are relevant to the research question discussed in this article. In task 1, the reasoning trigger is generated by a (false) statement after an example task like Figure 3 has been calculated.

1. [Number wall] " Three foundation stones of a number wall always produce the same capstone, no matter how I arrange the foundation stones"
2. [Number grid] "Swapping the two numbers in the arrows does not change the target number."

These false statements were chosen to allow the children an opportunity to argue using a counterexample. In task 2, the children calculate differently arranged number walls/number grids and should eventually recognize the statement as false and comment upon it. Arguments are then made based on the interrelationships between the basic elements and the outcome. Task 3 involves the question of how the basic elements of a number wall or number grid need to be arranged to maximize the result (Figure 4 above). Task 4 asks how the result of a number wall or number grid changes when a basic element is changed (see Figure 4 below).

The trigger for reasoning in tasks 1 and 2 is the verification of an assertion. Task 3 focuses on the search for the maximum of the result depending on the arrangement of the basic elements. In task 4, a corresponding reasoning trigger arises from the change of the result by rearranging the basic elements. All tasks are based on examples so that the children are given the opportunity to use examples in the reasoning products. The guideline allows for the children to first formulate discoveries and, if necessary, – if not formulated independently – the interview prompts for a justification (e.g., *Why is it like that?*). After a justification has been provided, the guidelines allow for further follow-up questions in the form of generalizations (e.g., *Is it always like this?*). This step is intended to initiate analytical arguments – if they have not already been formulated independently. This selection of tasks, together with the follow-up questions on generalization, is intended to provide the opportunity for empirical and analytical reasoning in equal measure.

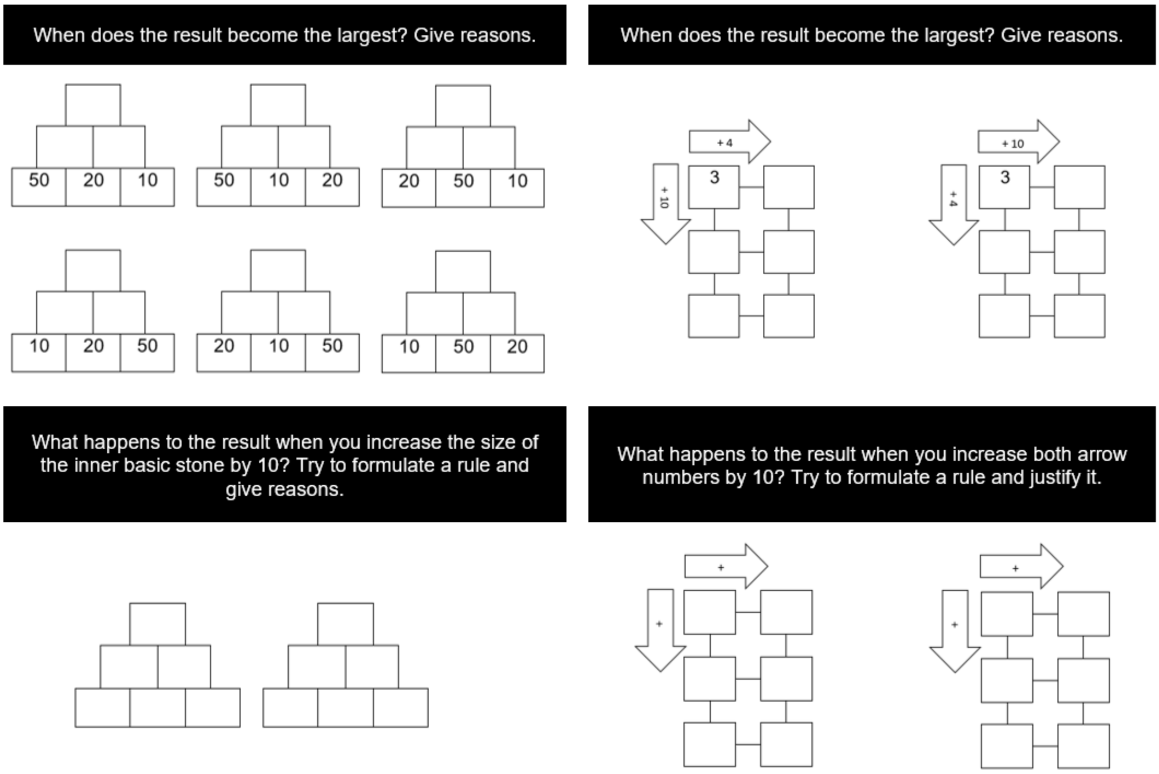

Figure 4. Overview of the reasoning tasks from the "number wall" and "number grid" interviews

Performing Data Collection

The interviews took place in parallel with the Mathe-Adler sessions; the children were taken out of the session individually for a period of 15 minutes. Trained student assistants supervised the process. The children were first informed about the aim of the interview and then asked for their consent. The interviews were recorded using an audio device so that the conversation between the child and the interviewer could subsequently be transcribed. The children were given the appropriate task material at the start of the interview. The tasks were visible on the material and were also introduced by the interviewer following pre-determined guidelines.

Four different variants of the interview were used to exclude the effect of recollection as far as possible: (a) number wall addition, (b) number grid multiplication, (c) number grid addition and (d) number grid multiplication. This meant that each child completed four interviews with different task formats. However, the structure and the reasoning triggers were identical. The order of interview formats (a) to (d) were systematically varied over the survey period to rule out variations due to possible differences in the level of difficulty. Likewise, no feedback on the content of the solution to the task or methodological advice on the argumentation procedure was given to minimize potential learning and memory effects due to the structurally identical tasks. The Mathe-Adler sessions took place at their usual frequency between the data collection exercises. There was no special focus on promoting reasoning over other skills.

Approach to Data Analysis

The transcripts prepared using the audio recordings and written materials form the basis for analysing the products of the argumentation. In total, four interviews, each consisting of four tasks, were conducted with 37 children. An argumentation product here describes the combination of all statements on a task.

This meant that 148 products of argumentation were analyzed for each of the four data collection exercises. An analytical framework for individual oral interviews was developed for the analysis using both the theoretical considerations, as well as the interviews from the pilot study. Based on the product-oriented perspective, the analysis of the argumentation products was carried out by coding the children's explicit statements. Categories were selected in line with the qualitative content analysis according to Mayring (2014) – following a deductive procedure – that contribute to answering the research question. First, the scheme calls for the products of argumentation to be coded structurally. Following the theoretical introduction in chapter 2.1, the term discovery includes any verbalized observation of the children relating to the task. Justification is considered to be all statements made by the children who (were supposed to) justify their conclusion concerning the discovery made, based on the task and the task material.

The generality category was chosen to assess the children's argumentation products in terms of the use of examples and generalizations. Its theoretical origin can be found in the distinction between empirical and analytical in Harel and Sowder (1998), and in Reid and Knipping’s (2010) listing of empirical, generic, symbolic and formal. The study's analytical framework first distinguishes between the empirical and analytical characteristics according to Harel and Sowder (1998) for discoveries and justifications. The most frequently occurring category while a conversation in a task is coded regarding a possible shift from the example to the general rule, as occurs in generic argumentation. Here, the generic approach is not initially distinguished due to the focus on the outcome, which does not emphasize the actual process of development, but is considered within the framework of the discussion. The formal approach without semantic meaning does not seem to have any particular relevance for the oral survey setting.

This category is supplemented by completeness to be able to describe a content-related component of the empirical or analytical reasoning. Following the distinction made in Heinze's argumentation analysis (2005, 118ff.), student responses were coded by an inability to answer, statements without explanatory value and statements with explanatory value. Additionally, the two subcategories: (1) incomplete (in the sense of a statement without all necessary mathematical information or explanatory value) and (2) complete (in the sense of a statement with all necessary mathematical information or explanatory value) are distinguished for the discoveries or justifications in the study's analytical framework.

Table 2. Categories of the analysis framework for the formulation of a discovery

Classification of the discovery	**Empirical: The discovery contains a reference to concrete figures or an example calculation.**	**Analytical: The discovery refers to general facts without giving specific examples.**
Incomplete: The discovery does not provide all necessary mathematical information.	"I notice that the 50 is in the middle."	"The number walls have different results."
Complete: The discovery provides all relevant mathematical information.	""The biggest result occurs when 50 is in the middle."	"I notice that the biggest number is in the middle and that's why a bigger result occurs."

Coding was carried out by the author and several trained student assistants. The analytical framework was empirically confirmed in the study by good values of intercoder reliability (generality κ = 0.71 and completeness κ = 0.72). Tables 2 and 3 describe the categories as well as their assignment rules with examples from the interviews of the pilot study – for the element's discovery and justification respectively. The question, *When does the capstone of the number wall become largest?* is fundamental for a number wall example with the capstones 10, 20 and 50. The prompts provided by the guidelines for justifying the discovery (*Why is this so?*), and for general reasoning (*Is this always so?*), must also be considered here for the context in which the argumentation products are formed.

Table 3. Categories of the analytical framework for formulating a justification

Classification of the justification	Empirical: Equivalent to the discovery	Analytical: Equivalent to the discovery
Incomplete: The justification does not contain any explanatory value.	"Because 50 is in the middle."	"Because the highest number is in the middle."
Complete: The justification contains explanatory value.	"Because 50 is in the middle and is then added to the brick on one side and the one on the other side."	"[...] because the largest number is in the middle and the largest number is added to both."

The four sample statements in Table 3 are all coded as discoveries because the children formulate an independent observation concerning the task in each of these statements. The distinction between (a) empirical and (b) analytical results from the integration of the concrete material – in the case of (a), it is the naming of 50, in the case of (b), the reference to general representatives such as the result and the largest number. The distinction between incomplete and complete refers to the necessary mathematical information contained in the statements. Their assessment varies with the task and the information previously determined as necessary. This coding is performed in a similar way in Table 3 for the justifications.

The analytical framework serves as the basis for coding the children's argumentation products. The research question requires that types be identified using examples. The methodological approach is based on Kluge's (2000) typology. Such type of formation is based on the requirements for internal homogeneity and external heterogeneity, so that cases of one type differ to a minimum extent, and cases of different types differ to a maximum extent. The attribute range for type formation results from the combination of the categories presented – generality of the argument and completeness of the argument – including the respective binary expression: incomplete and empirical, incomplete, and analytical, complete and empirical, complete and analytical.

The assignment to one of the four dimensions is first made for each child for each of the four data sets using empirical boundaries. To this end, the four argumentation products from an interview are analyzed in terms of their completeness and validity. The proportion of complete and generalized elements is then described in percentage terms. Here, the same empirical boundary is chosen for each of the four data sets. This is intended to enable an individual description of change and allocation, without possible changes in the group as a whole having any influence on it. In a final step, a timeline is created for each child with the assignment in all data sets, and the children are grouped according to their longitudinal classifications. The children who can be assigned three times to a combination of characteristics (1, 2, 3 or 4) are classified first. This approach is limited to one of the two categories and one of its expressions, e.g., analytical, for the remaining children. In a final step of the grouping, tendencies are considered, i.e., the change in a comparative dimension, e.g., from empirical in the first data collection to analytical in at least the third and fourth data collection.

RESULTS AND DISCUSSION

Quantitative Results of Coding

Table 4 summarizes the results of coding all the children's *discoveries*. N describes the number of task prompts for a discovery that were asked in connection with a data set. No corresponding statement was formulated for the no discovery classification.

Table 4. Overview of discoveries for all data sets

N = 148	t_1		t_2		t_3		t_4	
	Emp.	Ana.	Emp.	Ana.	Emp.	Ana.	Emp.	Ana.
No discovery	4.8%		0%		0%		0%	
Incomplete discovery	35.9%	16.8%	27.2%	21%	9.2%	19.9%	7.2%	19.5%
Complete discovery	2%	40.5%	4.5%	47.3%	5.7%	65.2%	7%	66.3%
Total	**37.9%**	**57.3%**	**31.7%**	**68.3%**	**14.9%**	**85.1%**	**14.2%**	**85.8%**

Table 4 first shows that the children use examples in formulating their discoveries, with the number of analytical discoveries predominating and increasing across all data sets. At the beginning of the survey, the proportion was 57.3 %, rising to 85.8 % within the 18-month observation period. A potential interaction between completeness and generalization can also be identified: if coding tends to be incomplete in the case of empirical discoveries, the opposite can be observed in the case of an analytical discovery.

Table 5 shows a similar overview of the *justifications* formulated. Since follow-up questions were asked in the interview, and justifications were specifically initiated using the guidelines, only the argumentation products in which such prompting occurred are considered here. This number of prompts is described as N_I for each data collection. The proportions shown therefore refer to the children's prompted justifications.[2]

Table 5. Overview of justifications for all data sets

	t_1 (N_I = 114)		t_2 (N_I = 131)		t_3 (N_I = 130)		t_4 (N_I = 123)	
	Emp.	Ana.	Emp.	Ana.	Emp.	Ana.	Emp.	Ana.
No justification	5.3%		4.6%		8.5%		0.8%	
Incomplete justification	46.2%	15.8%	55.3%	9.2%	29.4%	10%	29.3%	6.4%
Complete justification	12.6%	20.1%	11.7%	19.2%	11.3%	40.8%	16.2%	47.3%
Total	**58.8%**	**35.9%**	**67.0%**	**28.4%**	**40.7%**	**50.8%**	**45.5%**	**53.7%**

The increasing tendency of analytical discoveries seems to be reflected in the children's justifications. When considering the entire period, the share of analytical justifications increases from 35.9% to 53.7%. Nevertheless, this increase is not across the board, as the use of examples initially increases in t_2. What is also like the coding results of the discoveries is that of the feature combinations incomplete and empirical, as well as complete and analytic, seem to be particularly pronounced. Comparing the two tables, the first thing that stands out is that justifications are more often formulated empirically, and discoveries are more often formulated analytically.

The results first give an overview of the children's argumentation products across groups. Below, individual changes in argumentation products are described through the characterization of types. Due

[2] It should be noted here that the proportion of independently formulated justifications is low for all groups and data sets. The focus at this point is, therefore, on the prompted justifications. Reference is made to Jablonski and Ludwig (2021) for an analysis of the independent justifications.

to their interplay in the context of an argument, the elements of discovery and justification are initially considered together in the context of type formation and summarized under the generic term argument.

Characterization of Types

When all the data from the interviews is considered, the following empirical limits to the assignment of children emerge. In the completeness category, 54 % of all elements examined are incomplete, and 46 % of all elements examined are complete. In the general validity category, 43% are empirical and 57% analytical (Jablonski & Ludwig, 2021). These empirical limitations provide the basis for ranking children for each of the four data sets. Figure 5 shows the classification for t_1 in the form of a scatter plot. The x-axis shows the proportion of analytical elements, the y-axis the proportion of complete elements of a child and the lines represent the empirical boundaries.

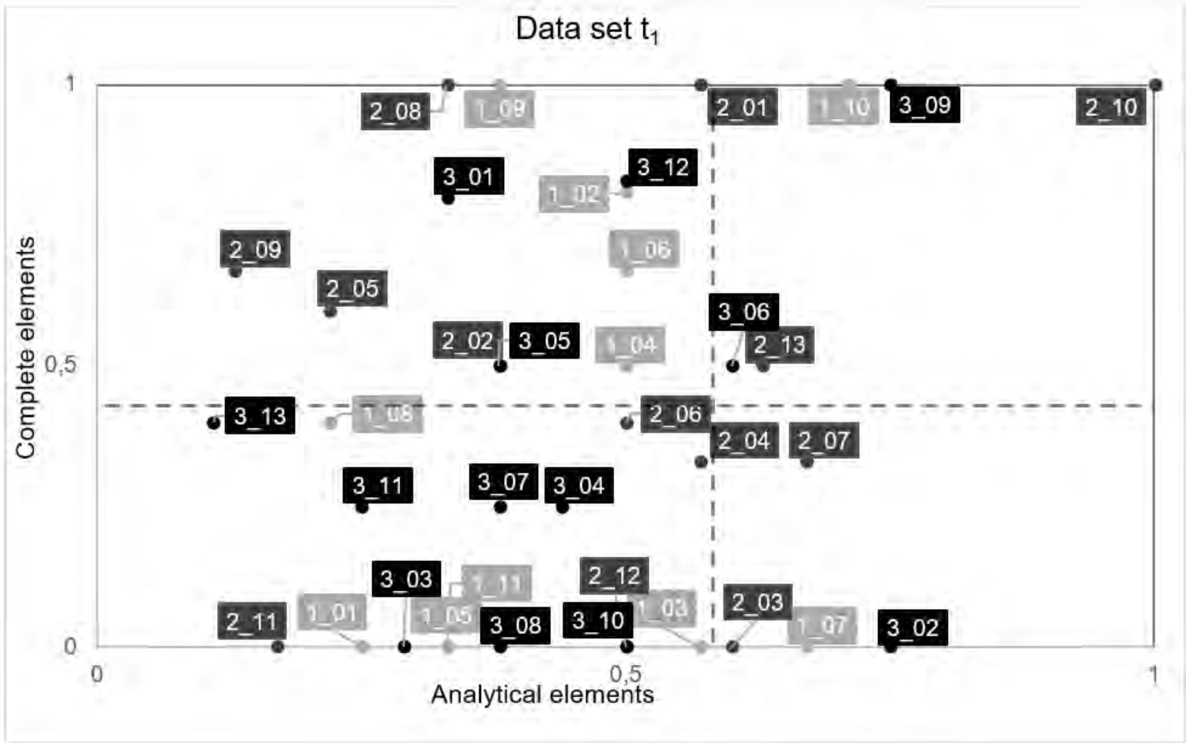

Figure 5. Scatter plot for the first data collection

Figure 6 refers to t_4. Each child has a fixed number starting with the associated group in all data sets. The scatter plots thus render the classification of each child visually and also provide an overview of any changes across groups. Comparing the scatter plots for t_1 and t_4, we see an increase in generalized elements across the groups, which was already observed in Tables 4 and 5.

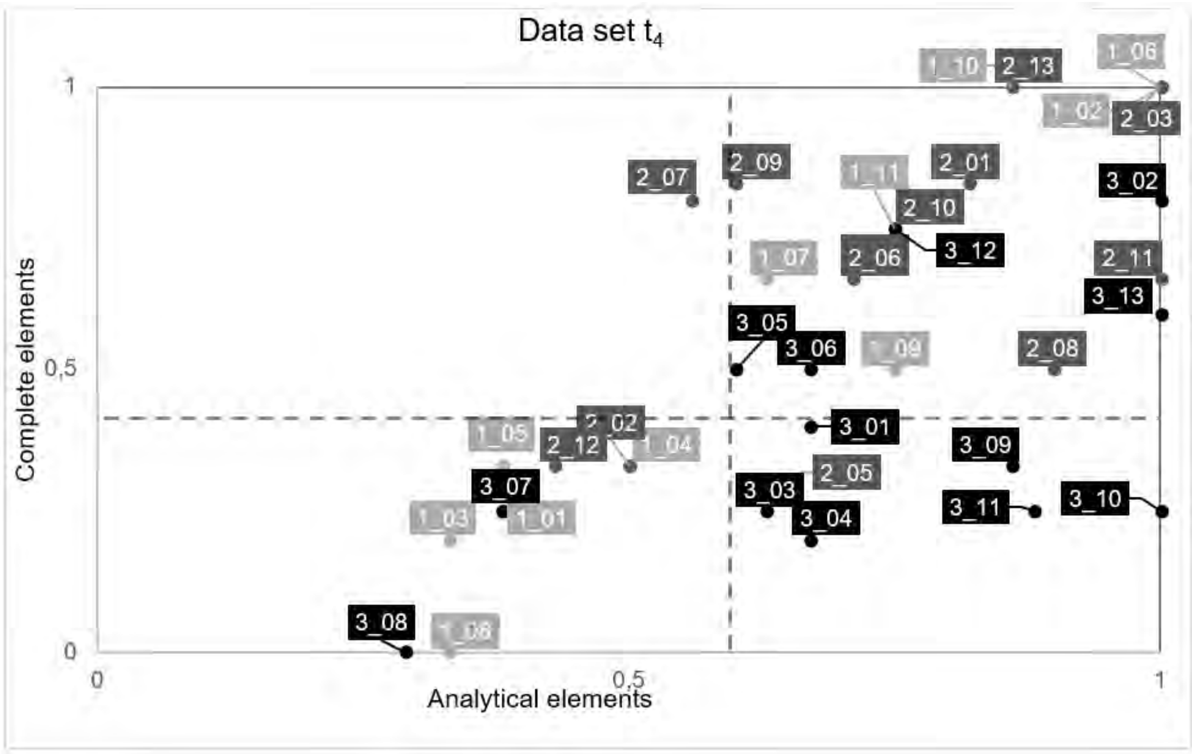

Figure 6. Scatter plot for the fourth data set

Grouping the individual children longitudinally yields the following six types, to which 34 of the 37 interviewed children can be assigned (see Figure 7). At this point, the criterion of internal homogeneity comes into play, in that the children of a case are considered like each other.

Type 1: Analytical arguments

- Type 1.1: Stable analytical
- Type 1.2: Stable analytical and complete

Type 2: Incomplete arguments

- Type 2.1: Stable incomplete
- Type 2.2: Stable incomplete and empirical

Type 3: Changing over time from empirical to analytical arguments

Type 4: Changing over time from incomplete to complete arguments

Figure 7. Summary of the types formed

Children whose argumentation products go beyond concrete numerical examples or the material at hand are classified in type 1. These children consistently refer to general mathematical facts and/or the structure of the task format. Type 1.1 describes children who put forward analytic arguments in a stable manner over the survey period, but the content of which cannot be classified as either stable incomplete or stable complete. This applies to six children. Type 1.2[3] includes children who, in addition to a stable analytical style of reasoning, are also able to formulate stable complete discoveries and justifications, i.e., statements that contain all necessary information. This applies to five children.

Type 2 includes children who argue in a stable, incomplete way. Two subtypes are also formed in this case. Type 2.1 includes those children who provide discoveries and justifications over the entire observation period, where information or the explanatory value required for completeness is missing. No uniform statement on the use of examples and generalizations can be made for them. In contrast, children in type 2.2 not only argue incompletely, but also in a stable empirical way. A total of ten children are assigned to type 2, four of them to type 2.1 and six to type 2.2.

The ten children assigned to type 3 begin with an empirical style of argumentation in t_1 and possibly also in t_2, using concrete numerical examples and referring directly to the material. The children increasingly formulate their discoveries and reasons beyond concrete examples and use general facts from t_3. No clear statement can be made about the completeness of the discoveries and justifications.

Type 4 is also variable in nature. This type is empirically weakly represented and is assigned to three children. These children are similar in their style of reasoning, in that they initially argue incompletely in t_1 and possibly also in t_2. From t_3 at the latest, they formulate discoveries and justifications that increasingly contain all relevant information or explanatory value. No statement can be made about the use of examples and generalizations for these children.

The differences between the types in terms of their external heterogeneity become clear through their characterizations. This is illustrated by the changes over time in the cases of the different types in Figures 8 and 9. The four survey classifications of all cases of a type are sorted into a diagram using color gradients. Apart from isolated survey results, two fixed ranges emerge for types 1 and 2, where the cases are classified across all data sets. With type 1, it is the area analytical, with type 2, incomplete. The two types can be considered disjunctive throughout the survey process, aside from individual data sets.

Types 3 and 4 also differ when different categories are considered. In contrast to types 1 and 2, they show a tendency to change over the survey period, which is illustrated by the color gradients and corresponding direction of the arrows. The sample does not provide other combinations of comparative dimensions, e.g., analytical, and incomplete, and are, therefore, not listed.

Below, case studies are used to present the three types 1, 2.2 and 3, which show particular relevance for the research question through characterization by means of empirical or general argumentation products. The prototypical procedure of the children in various reasoning activities is described for this purpose.

[3] It should be noted here that type 1.2 is not a subset of type 1.1 in this overview. Children from type 1.2 also fulfil all the characteristics of type 1.1 in their approach to reasoning. Nevertheless, the focus here is on the distinction between the two types, which is why the more comprehensive type is always assigned in terms of its characteristic features. The same applies to types 2.1 and 2.2.

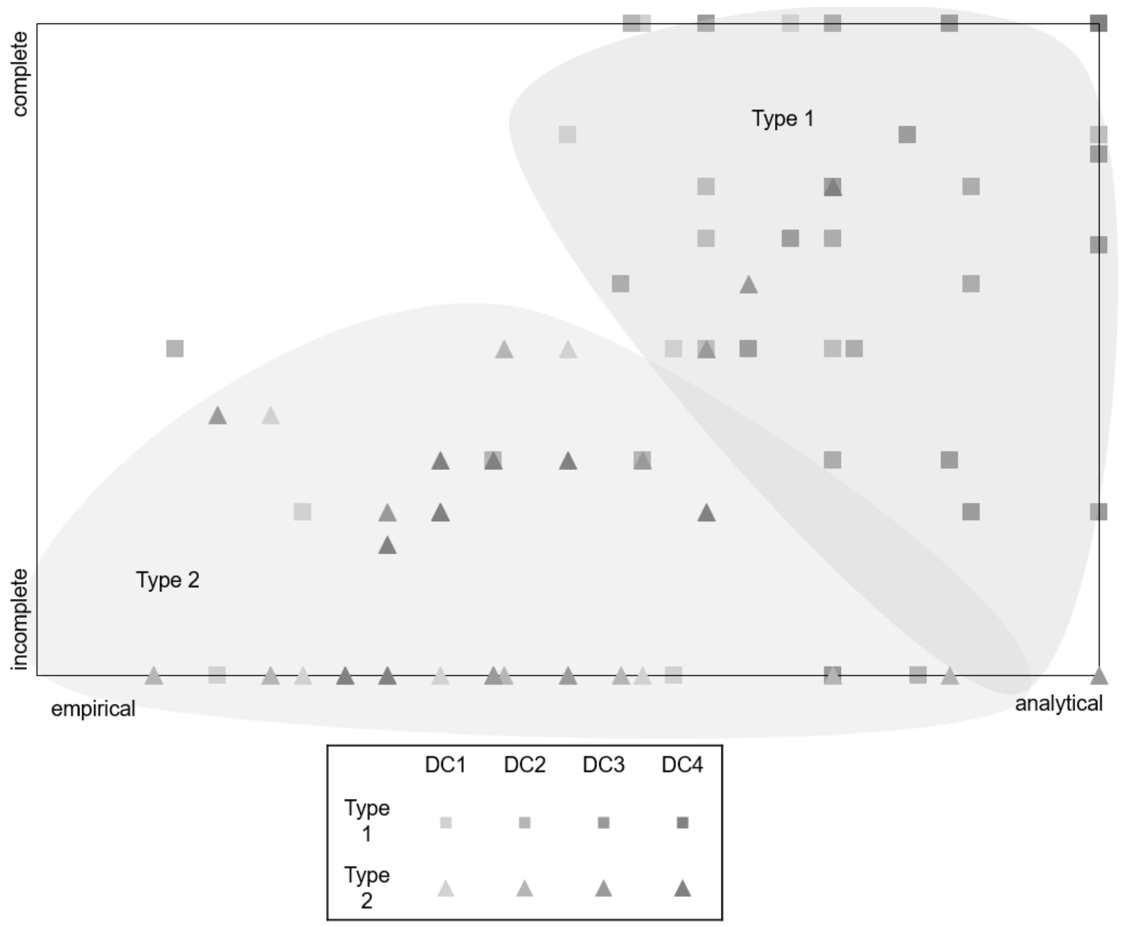

Figure 8. Changes over time within the formed types 1 and 2

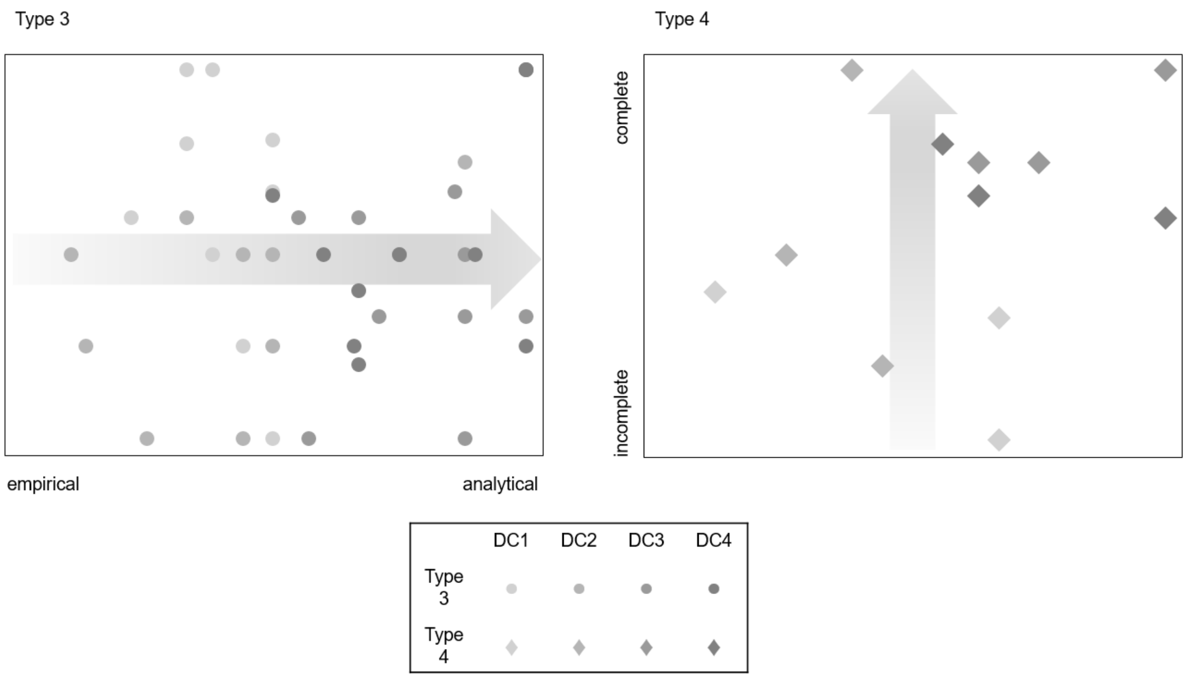

Figure 9. Changes over time within the formed types 3 and 4

Case Studies for the Types

Type 1 – Stable analytical arguments: The following comparison of two argumentation products at the time of the first and fourth data surveys characterizes the stable analytical formulation of a child's arguments. In both cases, it is the argument product that is formulated in task 3. In t_1, the task initially focuses on the question, *When does the target number of the number grid become largest?* In this and all following transcripts, "I" describes the person interviewing, and "B" the child being interviewed.

B: I think where the larger number points downwards.
I: Very good and why is that?
B: Because (...) basically you actually follow the path downwards. And if the smaller number is there, then you also move down in smaller steps. So, you take two steps down and one step to the side, no matter how you calculate.
(2_10_t_1, Number grid addition)

Child 2_10 first formulates a general discovery in response to the task. When asked why this is true, the child justifies why the result is the largest by using the calculation path in the number grid. This not only refers to a single way of calculating, but to " any way you calculate". Such argumentation without reference to the concrete example is also evident in data set t_4 in response to the corresponding question, *When does the capstone of the number wall become the largest?*

B: When the largest number is in the middle.
I: Why?
[…]
B: Because whenever the largest number is in the middle, it refers back to the two next to it, and that makes the results larger. However, for this to happen, the largest of these first three numbers must really be in the middle.
(2_10_t_4, Number wall multiplication)

In t_4, the child again formulates a general discovery about the position of the largest number and a general justification, which, as in the first data set, concerns a step-by-step approach to the calculation or the construction of the arithmetic format.

When checking the false claim of task 1, the children of this type either use a counter-example, which in this case can correspond to a complete argument by refutation, or they also argue completely and analytically, as child 2_10 does in t_2: "I think if I put the biggest one in the middle, the capstone will be higher, because then I multiply the numbers that are on the outside with the ones that are on the inside and the one that is on the inside is the biggest, then I think the biggest result should come out." (2_10_t_2, Number wall multiplication)

Again, general terms are used to describe the elements and the calculation path. Type 1 children use virtually no examples, either in their discoveries or in their use of assertions. It is striking that one in two children's stable analytical argumentation is accompanied by a stable complete style of reasoning which is hence, assigned to type 1.2.

Type 2.2 – Stable incomplete and empirical arguments: The argumentation product of task 4 in t_1 is presented to show the approach to reasoning of this type. The starting point is the question of what happens to the capstone of a number wall if you increase the size of the middle capstone by 10. The child calculated the two number walls shown in Figure 10 and, together with the interviewer, discovered a change of 20 in the capstone.

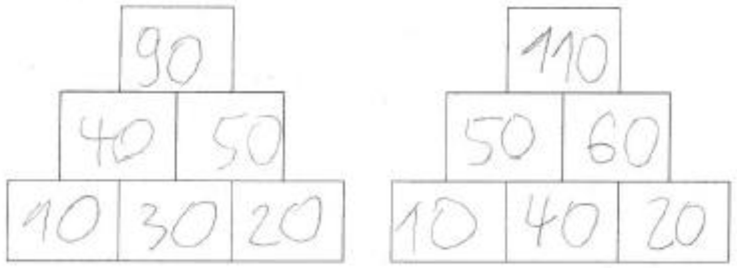

Figure 10. Number walls for the transcript of child 1_11

B: (Points to the middle foundation stones of the number walls) I think because (...) if you add the two together, it equals 20.
I: Both? [...] So you mean the middle foundation stones? What do I get when I add 30 and 40?
B: [...] No, I don't mean the middle ones, but the 40 and the 50 together are equal to 10, the 50 and the 60 together are 10 and then 10 plus 10 is 20.
(1_10_t_1, Number wall addition)

The empirical procedure becomes clear in two ways in this justification of child 1_11, which is based on a comparison of the numbers in the middle row: first, the child uses the task material by pointing to the numbers, which suggests an implicit reference to the particular example. Second, the concrete example is also reflected in the explicit formulation, where numbers are not generalized, but are described by their concrete values. In so doing, the child does not formulate a general rule - even when asked.

This approach is characteristic of type 2.2 children. Both discoveries and justifications are usually formulated with reference to the material and with the help of concrete numerical examples. The children involved in this case used examples to illustrate and explain their statements. With child 1_11, the procedure is characterized by showing and naming distinctive arithmetic steps through concrete numbers. Nevertheless, the justification remains incomplete – even when asked – and comprehensibility is therefore hampered.

When reproducing and verifying assertions during task 1, the children also find it difficult to formulate a suitable counter-example – despite the mainly empirical approach to reasoning. Although there is a greater use of examples in type 2.2, they are used less for actual justification or a generic approach than for illustrating implicit and non-linguistic justifications. Overall, the children seem to have problems using appropriate examples as justification.

Type 3 – Changing over time from empirical to analytical arguments: The increasing tendency of analytical discoveries and justifications over the survey period will be illustrated by the example of task 3 in a comparison of t_2 to t_3. First, we consider the argumentation product that emerged during the second survey in response to the question, *When does the target number of a number grid become largest?* Figure 11 shows the corresponding task material.

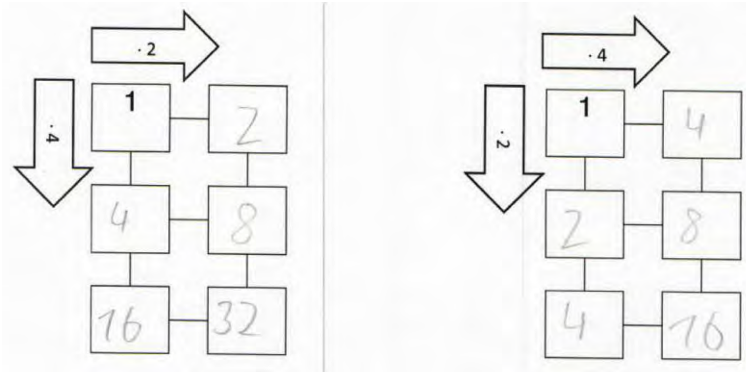

Figure 11. Number grid for the transcript of child 2_11

I: OK, so we found out that the number became smaller. How will the target number be largest at the bottom?
B: (Points to the arrow number downwards of the number grid of task 3) When the larger number is down here.
[…]
I: Why? Why is it so important that it is in the down arrow and not to the right?
B: If you have a 4 here instead of a 16, you must take double here, in this case four times. If you take double here, it is clearly more.
(2_10_t_2, Number grid multiplication)

Child 2_08 primarily uses empirical justifications in t_1 and t_2, which is shown in this section. While the discovery is formulated in general terms, an implicit reference to the concrete example is nevertheless recognizable by pointing to the arrow number. Similarly, the justification (the child's last statement) is related to a concrete numerical example used by the child uses to compare the change step by step. Here, the idea of empirical argumentation also emerges, as a generally formulated discovery is justified using a single example. The change in argumentation behavior is evident in the interview at time t_3, which, in turn, is portrayed in a representative manner for this case based on task 3.

I: OK, we've now found that the arrangement of the foundation stones changes the result of the capstone. When does the capstone become largest?
B: When the largest number is in the middle.
I: Very good, why?
B: Because you then take the largest number twice. [...] The largest number is then used the most often and so it is always calculated several times.
(2_10_t_3, Number wall multiplication)

The present calculation no longer seems to play a role in t_3. The child formulates both the discovery, "If the largest number is in the middle", and the reasons, e.g. "Because you then take the largest number twice", without direct reference to the calculation, i.e., analytically. Instead of numerical examples, child 2_08 uses general terms at this point – like type 1 – such as largest number, which can be applied to any version of the task format. Furthermore, a description of the format's structure replaces the step-by-step comparison of the calculation. This procedure is also repeated in data set t_4 for the number grid, which can be seen, for example, through the justification: "Because we calculate the number twice, because we have a 2-by-3 field. That's why we calculate the one arrow pointing to the right only once, but the arrow pointing down twice".

This tendency towards the use of more analytical elements in justifying and generalizing connections and number relationships is not borne out in the comprehension and verification of assertions, or if it is, it only occurred in individual cases. Most type 3 cases used complete counterexamples consistently or recurrently.

CONCLUSION

The aim of this study was to answer the research question of how the argumentation products of potentially mathematically gifted children can be described in terms of empirical and analytical elements, and what longitudinal changes occur. The cross-group coding of the argumentation products into categories of generality as well as of completeness permit three key observations as follows.

1. The role of examples in argumentation: Empirical arguments across all data sets play a role in both the formulation of discoveries as well as the corresponding justifications. This proportion is higher for justifications than for discoveries. Firstly, this seems to demonstrate the potential of examples as illustrations of facts. Secondly, this number illustrates the danger of empirical argumentation with more generalized discoveries than justifications (Nussbaum, 2011). This hypothesis is further supported by the combinations occurring, primarily analytical and complete and empirical and incomplete. Hence, it appears that (just) using examples does not generally contribute to the completeness of discoveries and justifications.
2. The increase in generalizations: Across all 37 children interviewed, the formulation of generalizations – both in the formulation of discoveries and in the formulation of justifications – increases over the course of data collection. This result is initially in line with previous theoretical evidence for such a change (Piaget, 1928; Sjuts, 2017).
3. Examples and generalizations for describing types of argumentations: Type formation also emphasizes the impressions of the first two observations. Of the six types identified, three are characterized through their use of examples and generalizations.

The children of the first type (*stable analytical*) do not generally use examples to formulate their discoveries and justifications. Because of their general formulations of discoveries and justifications, this type resembles a deductive (Harel & Sowder, 1998) or symbolic approach (Reid & Knipping, 2010). An exception is the formulation of counterexamples when reacting to a false claim. Case study child 2_10 took a step-by-step approach to justification in each instance, which, representative of the totality of children of this type, suggests a potential connection to special structuring skills (at the pattern level) (Käpnick, 1998; Sjuts, 2017).

Type 2.2 children (*stable incomplete and empirical*) are characterized by their repeated and consistent use of examples in their argumentation products. They mainly use examples to describe and illustrate their reasoning. Their approach corresponds to the characterizations of an empirical (Reid & Knipping, 2010) or inductive (Harel & Sowder, 1998)) formulation. Nevertheless, their arguments regularly remain incomplete and implicit. One possible connection to the intuition of potentially mathematically gifted children (Fuchs, 2006) is conceivable in this case, although a starting point for argumentation was explicitly revealed through follow-up questions in the interviews. Here, then, intuition would be understood less in terms of a lack of need, in that an argument is expected, than in terms of a lack of clarity about what is expected in the context of a complete argument or generalization. Their difficulties in choosing a suitable counterexample are particularly interesting, which would, at first, seem plausible with regard to the characterization of the type.

Type 3 children (*changing over time from empirical to analytical*) are strongly represented empirically and transition over time from empirical to analytical reasoning. It was, therefore, possible to observe that they increasingly formulated more general facts in their arguments and in doing so, broke free from concrete numerical examples. These children confirm the observation across groups regarding an increase in generalizations, in addition to the level of individual children. The change in type, therefore, describes a transition from the empirical to the analytical level of proof schemes (Harel & Sowder, 1998). A generic approach (Reid & Knipping, 2010) might be particularly relevant in this transition over time. It is conceivable that the increase in analytical arguments does not emerge independently of examples but based on examples. This remains open for the time being due to the study's focus on products. Nevertheless, opportunities exist for further longitudinal studies in the field of argumentation processes.

Regarding considerations of argumentation characteristics associated with mathematical giftedness, a high degree of heterogeneity and inter-individual differences in the formulation of arguments are evident, despite the selection of potentially mathematically gifted children. The type formation and argumentation characteristics emphasize that potential giftedness does not automatically lead to a homogeneous group, and that the problematic use of examples, e.g. in the context of empirical reasoning (Nussbaum, 2011), cannot be ruled out. Connections to mathematics-specific giftedness characteristics can be assumed, even if the formulation of complete and analytical arguments is not clearly related to potential mathematical giftedness in all cases of the study. The cross-group increase in analytical arguments speaks for the theoretical considerations on the influence of further giftedness-specific characteristics, e.g., structuring skills.

The findings emerged from an exploratory approach and should be interpreted accordingly. The small sample size means that the findings cannot be generalized, but instead require confirmation in quantitative studies. Observation of a control group is a useful extension to the hypothetical considerations of links to mathematics-specific giftedness characteristics and, finally, to modelling potential mathematical giftedness. Furthermore, the findings need to be interpreted in the context of specific task selection. Especially for the distinction between empirical and analytical, this seems relevant, as the tasks are initially based on concrete examples. The follow-up questions and direct prompts for general reasoning in the guide were intended to ensure that general argumentation seems desirable to the children here. Nevertheless, a possible discrepancy arises at this point between the expectations for the guidelines from a research perspective, and what the children assume to be expected, e.g., generalization based on the question, *Is it always like this?* It is important to note that the examples of tasks do not initially show this intention clearly. It must also be stressed that the study only analyzed the children's spoken words and that any non-verbal reflections – including those related to intuitive approaches – were disregarded.

What is more, the findings were obtained in the context of the "Junge Mathe-Adler Frankfurt" enrichment program and should be interpreted with this restriction in mind. Even though there was no deliberate focus of argumentation during the support sessions, and there was no feedback on the content and methodology of the interviews, the enrichment program itself can be considered an influencing factor. The children may have changed their way of reasoning because of the support work and their increased experience with mathematics contents and procedures. The same applies to the children's ongoing mathematics lessons. However, this limitation can be regarded as less relevant since there were usually only the same for a maximum of two children in the sample. In addition, although precautions were taken to minimize the effects of recollection, it is possible that the children became accustomed to the structure and question types used during the interviews and remembered after which answer no further questions were asked. A possible memory effect, however, does not seem to have any particular significance for the evaluation of this research question, since coding always included the highest-occurring argumentation product formulated after initiation.

The study presented here provides a basis for the following ideas for further research: The descriptions and changes in the use of examples and generalizations in mathematical arguments cannot be represented one-dimensionally, despite being restricted to potential mathematical giftedness. Nonetheless, typing in particular has shown that it is possible to structure, group and simplify the change processes. At this point, apart from a quantitative examination, a further interest arises for more complex case studies that focus on a procedural perspective as well as further sub-areas of mathematical reasoning, thereby rendering the relationship between giftedness and reasoning as tangible.

Declarations

Author Contribution	:	SJ: Conceptualization, Writing - Original Draft, Formal analysis, and Methodology, Editing and Visualization; ML: Writing - Review & Editing, Validation and Supervision
Funding Statement	:	This research was funded by Dr. Hans Messer Stiftung and Stiftung Polytechnische Gesellschaft Frankfurt, Germany.
Conflict of Interest	:	The authors declare no conflict of interest.

REFERENCES

Amielia, S. D., Suciati, S., & Maridi, M. (2018). Enhancing students' argumentation skill using an argument driven inquiry-based module. *Journal of Education and Learning (EduLearn)*, *12*(3), 464–471. https://doi.org/10.11591/edulearn.v12i3.8042

Assmus, D., & Fritzlar, T [Torsten] (2022). Mathematical creativity and mathematical giftedness in the primary school age range: an interview study on creating figural patterns. *ZDM – Mathematics Education*, *54*(1), 113–131. https://doi.org/10.1007/s11858-022-01328-8

Baker, M. (2003). Computer-Mediated Argumentative Interactions for the Co-Elaboration of Scientific Notions. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to Learn: Confronting Cognitions in Computer-Supported Collaborative Learning environments* (pp. 47–78). Springer, Dordrecht. https://doi.org/10.1007/978-94-017-0781-7_3

Bezold, A. (2009). *Förderung von Argumentationskompetenzen durch selbstdifferenzierende Lernangebote [Promotion of argumentation skills through self-differentiating learning opportunities]*. Kovac.

Brunner, E. (2019). Förderung mathematischen Argumentierens im Kindergarten: Erste Erkenntnisse aus einer Pilotstudie [Promoting Mathematical Reasoning in Kindergarten: Initial Findings from a Pilot Study]. *Journal für Mathematik-Didaktik*, *40*(2), 323–356. https://doi.org/10.1007/s13138-019-00146-y

Durak, T., & Tutak, F. A. (2019). Comparison of Gifted and Mainstream 9th Grade Students' Statistical Reasoning Types. In M. Nolte (Ed.), *Including the Highly Gifted and Creative Students – Current Ideas and Future Directions* (pp. 136–143). WTM. https://doi.org/10.37626/GA9783959871327.0

Fritzlar, T [T.], & Nolte, M. (2019). Research in mathematical giftedness in Germany-Looking back and ahead. In M. Nolte (Ed.), *Including the Highly Gifted and Creative Students – Current Ideas and Future Directions* (pp. 8–20). WTM. https://doi.org/10.37626/GA9783959871327.0

Fuchs, M. (2006). Procedures of mathematically potentially talented third and fourth classes in problem-free empirical studies on typing specific problem-handling styles. *Journal für Mathematik-Didaktik*, *27*(2), 165–166. https://doi.org/10.1007/BF03339035

Fuchs, M., & Käpnick, F. (2009). *Mathe für kleine Asse. Klasse 3/4. Band 2 [Math for little aces. Grade 3/4. volume 2]*. Cornelsen.

Gutierrez, A., Benedicto, C., Jaime, A., & Arbona, E. (2018). The Cognitive Demand of a Gifted Student's Answers to Geometric Pattern Problems. In *Mathematical Creativity and Mathematical Giftedness* (pp. 169–198). Springer, Cham. https://doi.org/10.1007/978-3-319-73156-8_7

Habermas, J. (1984-87). *The theory of communicative action: Reason and the rationalization of society*. Polity; Beacon Press.

Hanna, G. (2000). Proof, explanation and exploration: An overview. *Educational Studies in Mathematics*, *44*(1/2), 5–23. https://doi.org/10.1023/A:1012737223465

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In E. Dubinsky, A. Schoenfeld, & J. Kaput (Eds.), *Research in Collegiate Mathematics Education* (pp. 234–283). American Mathematical Society.

Heinze, A. (2005). *Lösungsverhalten mathematisch begabter Grundschulkinder - aufgezeigt an ausgewählten Problemstellungen [Solving behavior of mathematically gifted elementary school children - demonstrated by selected problems]*. LIT.

Heinze, A. (2006). Lösungsverhalten mathematisch begabter Grundschulkinder — aufgezeigt an ausgewählten Problemstellungen [Solving behavior of mathematically gifted elementary school children - demonstrated by selected problems]. *Journal Für Mathematik-Didaktik*, *27*(1), 79–80. https://doi.org/10.1007/BF03340105

Jablonski, S., & Ludwig, M. (2021). Changes in mathematical reasoning - A longitudinal study with mathematically gifted children. In M. Inparsitha, N. Changsri, & N. Boonsena (Eds.), *Proceedings of the 44th Conference of the International Group for Psychology of Mathematics Education* (Vol. 3, pp. 91–100). PME.

Joklitschke, J., Rott, B., & Schindler, M. (2022). Notions of creativity in mathematics education research: A systematic literature review. *International Journal of Science and Mathematics Education*, *20*(6), 1161–1181. https://doi.org/10.1007/s10763-021-10192-z

Käpnick, F. (1998). *Mathematisch begabte Kinder. Modelle, empirische Studien und Förderungsprojekte für das Grundschulalter [Mathematically gifted children. Models, empirical studies and support projects for the primary school age]*. Peter Lang GmbH.

Kieran, C., Pang, J., Schifter, D., & Ng, S. F. (2016). *Early Algebra*. Springer International Publishing. https://doi.org/10.1007/978-3-319-32258-2

Kluge, S. (2000). Empirically grounded construction of types and typologies in qualitative social research. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *1*(1). https://doi.org/10.17169/fqs-1.1.1124

Koleza, E., Metaxas, N., & Poli, K. (2017). Primary and secondary students' argumentation competence: a case study. In T. Dooley & G. Gueudet (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education* (pp. 179–186). DCU Institute of Education and ERME.

Komatsu, K. (2017). Fostering empirical examination after proof construction in secondary school geometry. *Educational Studies in Mathematics*, *96*(2), 129–144. https://doi.org/10.1007/s10649-016-9731-6

Maddocks, D. L. S. (2018). The identification of students who are gifted and have a learning disability: A comparison of different diagnostic criteria. *Gifted Child Quarterly*, *62*(2), 175–192. https://doi.org/10.1177/0016986217752096

Mayring, P. (2014). *Qualitative content analysis: Theoretical foundation, basic procedures and software solution*. SSOAR. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-395173

Moor, E. de. (1980). *Wiskobas bulletin. Leerplanpublikatie 11*. IOWO.

Nussbaum, E. M. (2011). Argumentation, dialogue theory, and probability modeling: Alternative frameworks for argumentation research in education. *Educational Psychologist*, *46*(2), 84–106. https://doi.org/10.1080/00461520.2011.558816

Pedemonte, B. (2002). Relation between argumentation and proof in mathematics: cognitive unity or break? In J. Novotná (Ed.), *Proceedings of the 2nd Conference of the European Society for Research in Mathematics Education* (Vol. 2, pp. 70–80). Charles University, Faculty of Education.

Piaget, J. (1928). *Judgment and reasoning in the child*. Routledge. https://doi.org/10.4324/9780203207260

Reid, D., & Knipping, C. (2010). *Proof in mathematics education. Research, learning and teaching*. Sense Publisher.

Sari, P., & Ng, S. F. (2022). Exploring quantitative relationship through area conservation activity. *Journal on Mathematics Education*, *13*(1), 31–50. https://doi.org/10.22342/jme.v13i1.pp31-50

Schneider, W., & Bullock, M. (Eds.). (2009). *Human development from early childhood to early adulthood: Findings from a 20 year longitudinal study*. Psychology Press.

Sjuts, B. (2017). *Mathematisch begabte Fünft- und Sechstklässler [Mathematically gifted fifth and sixth graders]*. WTM.

Sowell, E. J., Zeigler, A. J., Bergwall, L., & Cartwright, R. M. (1990). Identification and description of mathematically gifted students: A review of empirical research. *Gifted Child Quarterly*, *34*(4), 147–154. https://doi.org/10.1177/001698629003400404

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. https://doi.org/10.4219/jeg-2004-317

Toulmin, S. E. (2003). *The uses of argument*. Cambridge University Press. https://doi.org/10.1017/CBO9780511840005

Ufer, S., Heinze, A., & Reiss, K. (2008). Individual predictors of geometrical proof competence. In O. Figueras, J. L. Cortina, S. Alatorre, T. Rojano, & A. Sepúlveda (Eds.), *Proceedings of the Joint Meeting of the 32nd Conference of the International Group for the Psychology of Mathematics Education and the XX North American Chapter* (pp. 361–368). PME.

van Eemeren, F. H., Grootendorst, R., Snoeck Henkemans, A. F., & Blair, J. A. (1996). *Fundamentals of argumentation theory: A handbook of historical backgrounds and contemporary developments*. L. Erlbaum. https://doi.org/10.4324/9780203811306

Wittmann, E., & Müller, G. (1990). *Handbuch produktiver Rechenübungen. Vom Einspluseins zum Einmaleins [Handbook of productive arithmetic exercises. From the multiplication table to the multiplication table]*. Klett.