CONTENT AND FORM – ALL THE SAME OR DIFFERENT QUALITIES OF MATHEMATICAL ARGUMENTS?

Sarah Ottinger[1], Ingo Kollar[2], Stefan Ufer[1]

[1]University of Munich (LMU), [2]University of Augsburg

Students entering academic mathematics programmes struggle with various challenges in their transition from secondary school to tertiary education. One challenge is the strong focus on formal-deductive argumentation and proof in university mathematics. Producing acceptable mathematical arguments requires both, the ability to find deductive lines of arguments as well as skills to communicate these arguments with precision. We present a study with N=159 students at the transition from secondary to tertiary education that examines how the quality of mathematical arguments and of different formal aspects of their presentation are interrelated. We discuss implications for research as well as for support of students at the beginning of their mathematics study.

INTRODUCTION

A substantial amount of students give up studying mathematics during their first year at university (Heublein, 2014). Possible reasons for the high drop-out rate might be that the character of mathematics as a scientific discipline changes dramatically in the transition from school to university. This is not primarily a change of topics, but there is a shift toward an increased depth in the subject, with respect to the understanding and use of formal mathematics (Clark & Lovric, 2008). In tertiary mathematics courses, abstract concepts, formally presented arguments and proofs play a central role. Students are exposed to the emphasis on multiple representations of mathematical objects and on the precision of mathematical language (Clark & Lovric, 2008). Our study is situated in the transition phase from secondary to tertiary education with a specific focus on mathematical argumentation and proving, and the use of formal representations to communicate mathematical arguments.

Mathematical argumentation, i.e. to generate arguments for or against a mathematical conjecture and to convince oneself as well as the mathematical community about their validity, comprises empirical exploration (e.g., Koedinger, 1998), logical deductions and the ability to deal consciously with formal-symbolic representations and mathematical language (Epp, 2003). Several studies indicated that students at all levels have great difficulty with the task of proof construction (e.g., Healy & Hoyles, 1998; Ufer, Reiss, & Heinze, 2008). Even students who want to pursue undergraduate courses in mathematics at university often show poor proof-writing attempts, which may consist of little more than a few disconnected calculations or are characterised by an imprecise or incorrect use of mathematical words or phrases (Epp, 2003). There has been much research pointing to reasons for these deficiencies (e.g., Selden & Selden, 2011). Models of the proving process suggest to differentiate two idealized subprocesses of proving when searching for explanations (Selden & Selden, 2009): Firstly, students have to find adequate arguments and organize them into a deductive chain mentally. Secondly, they have to communicate their arguments and proofs in a formally correct way according to mathematical standards.

The content of mathematical arguments

Identifying a conclusive chain of mathematical arguments is a complex problem solving process that relies on several individual prerequisites, like knowledge of heuristic strategies (Schoenfeld, 1985) and conceptual mathematical knowledge (Ufer et al., 2008). Moreover, methodological knowledge on the nature of proofs (e.g., Healy & Hoyles, 1998) is necessary to direct this search process. For example, evaluating the truth or falsity of mathematical statements requires knowledge about the role of examples and counterexamples (Koedinger, 1998). During the proof construction process, students have to identify relations between mathematical concepts, and select those for which they see a chance to support them by acceptable mathematical arguments and organize them in a conclusive deductive chain.

When analysing students' proof skills, research has often focused on the *content of students' arguments* that become visible in students' work, deliberately disregarding the *formal quality* of the presentation of these arguments (e.g., Healy & Hoyles, 1998; Reichersdorfer, Vogel, Fischer, Kollar, Reiss, & Ufer, 2012). Even though this is a reasonable choice when viewing proof from a problem solving perspective, the adequate presentation of arguments is also a relevant goal of most university mathematics programmes (Epp, 2003).

The form of mathematical arguments

Engelbrecht (2010) points out that students have to be able to communicate their arguments in a "subject-specific, scientific language". When thinking about the quality of a specific mathematical argument, however, the use of a specific formal notation or corresponding mathematical language constructs (like "Let x be…", "For all y…") is certainly not a necessary feature for the validity and acceptability of a proof, even if this feature occurs in many mathematical texts. On the other hand and more generally, the precise communication of mathematical ideas is a decisive criterion. This means that, if a specific formal notation or specific mathematical language is used, it must be used in a precise and correct way.

However, there is a wide basis of research documenting that students have problems to use formal notations and specific mathematical language in a correct way: (1) Students' difficulties in *using logical symbols* correctly are well documented (Epp, 2003). One reason for this might be that logical statements can be interpreted differently in formal and informal settings. For instance, in informal settings, the statement "Some A are B." is taken to imply that "Some A are not B.", but in mathematics, this implication is not valid (Epp, 2003). (2) Clement (1982) reported that a large proportion of university engineering students have problems translating relationships expressed in spoken language into corresponding *mathematical* *expressions,* and vice versa. A famous example is the statement "There are 8 times as many people in China as there are in England". Some students seem to treat variables as symbols for objects or persons, writing 8C=E in this case. Comparable problems might be identified in symbolizing relationships like the divisibility of two integers. (3) Connected to this, students often have trouble with using variable symbols correctly. For example, they fail to understand that the value of a variable can be arbitrary, but fixed and does not change its value within one algebraic expression. Some also fail to introduce the meaning of the variable symbols they use. Epp (2011) noted that, alongside the emphasis on mechanical procedures at school, the meaning of variables as unknown quantities with specific properties, such as in functions or as expression for universal statements may be obscured. (4) Students' problems with quantifiers are also well-documented (e.g., Dubinsky & Yiparaki, 2000; Epp, 2003; Selden & Selden 2011). It seems to be a challenge for students to understand that the meaning of a statement is influenced by the order of the quantifiers, or to know the scope of a quantifier. Selden and Selden (1995) see students' difficulties in interpreting implicit quantifiers (i.e. expressed in words, not symbols) as a significant barrier for proof construction.

Even though newer studies take into account the content of students' arguments as well as their formal quality (e.g., Selden & Selden, 2009, 2011), the relation between the two has rarely been studied. In some works, the two quality aspects seem to be treated as fairly separated, as if skills in the formal presentation of arguments are something that is necessary primarily *after* a conclusive chain of arguments is found (e.g., Engelbrecht, 2010). While the skill to use some formal aspects might – in this sense – be fairly independent of students' skills to find conclusive chains of arguments, this needs not to be held for all formal aspects. Some works emphasize a stronger connection between, for example, understanding the language of logic (as different from everyday language) and logical notation, and the understanding of logical structures themselves (e.g., Epp, 2003). This is in line with theories that emphasize an epistemic function of language use (Sfard, 2008), which assumes that (mathematical) thinking is at least partly structured by the mental use of language. Following this line of argument, not being able to use formal language, notations or representations correctly might reflect and also cause a deficient understanding of the arguments that are constructed and presented in a proving or argumentation process. Thus, it remains an open question, which aspects of formal quality of students' arguments are connected to the content quality of these arguments, and which are less related to it.

GOALS OF THE STUDY AND RESEARCH QUESTIONS

Although undergraduate students' problems in constructing mathematical proofs and generating rigorous mathematical argumentations have been reported in many studies (e.g., Selden & Selden, 2011), there have been little attempts to study how the content quality of mathematical arguments and their formal quality are interrelated. To fill this gap, the present study addresses the following questions: (1) Which difficulties of mathematical argumentation regarding content and formal quality can be identified?

(2) Do content quality and different dimensions of formal quality of students' arguments form a one-dimensional construct, or is it necessary to differentiate multiple quality dimensions of students' mathematical arguments?

DESIGN AND METHODS

N=159 incoming students (72 female) from a regular mathematics programme, financial mathematics programme and a mathematics teacher education programme with an average age of 19.67 years (SD = 3.18) from two German universities took part in our study, which was embedded in a voluntary two-week preparatory course for university mathematics. Daily lectures and tutorials about elementary number theory as well as about other basic topics such as sets, functions and relations were included in this course. On day four, students worked for 45 minutes on mathematical argumentation problems from elementary number theory on their own adapted from Reichersdorfer et al. (2012). These comprised technical proof skills (e.g., "Show that for all natural numbers, a and b the following statement is true: If 15 divides (10a-5b) then 3 divides (2a-b).", 5 items), flexible proof skills (e.g. "Prove the following statement: The product of three consecutive even numbers is divisible by three.", 4 items) and conjecturing skills (e.g. "Prove or refute the following statement: If the sum of two natural numbers is even, then the product of these two numbers is always even.", 4 items with correct and false statements, each).

To score the *content quality* of students' argumentations a four-level coding was applied. For this, we analysed the mathematical ideas visible in the students' solution, disregarding their formal presentation as much as possible. We scored no or irrelevant trials with score zero, partially correct solutions including less than half of all central arguments required with score one, solutions including more than half of all central arguments but with small methodological errors (like an incorrect proof structure) with score two and completely correct solutions with score three.

Coding schemes for different aspects of formal quality were developed based on data from prior studies: *Symbolizing divisibility* (e.g., use of the symbol |) was coded on two levels (0: incorrect, 1: correct). A three-level coding was applied to score the *use of logical symbols* (e.g., ⇔ or ⇒; 0: using logical notations, although no logical statement is made, 1: use of incorrect logical symbols for logical statements, 2: correct), *symbolizing definitions* (*"Let x be 3…", =:, :=* , *:*⇔) (0: not symbolizing of definitions, although necessary, 1: incorrect, 2: correct), and the *use of variables* (0: inconsistent or incorrect, 1: correct and consistent, but without systematic introduction, 2: completely correct). The *use of quantifiers* (universal quantifiers and existential quantifiers) was coded on four levels (0: no use of quantifier, although necessary, 1: incorrect use of a single quantifier, 2: correct use of single quantifiers, but problems with the use of consecutive quantifiers, 3: correct). If a certain formal notation or corresponding language constructs were not used in a student solution, the respective value was coded as missing value. The only exception was if the corresponding aspect would have been required to communicate the argument according to the mathematical standards of the course. If this was the case and the corresponding aspect did not occur, this was coded with the lowest score (0). All arguments were coded by two independent raters and interrater reliability for each part of the test was found to be good (Mean of ICC=.86, SD =.08).

RESULTS

Descriptive results for the content quality of arguments can be found in Table 1.

	Technical proof skills	Flexible proof skills	Conjecturing skills (true)	Conjecturing skills (false)
Mean quality score	1.37 (.75)	1.24 (.80)	1.30 (.73)	1.68 (.85)

Table 1: Means (and standard deviations) of the content quality of arguments

On average, less than half of all arguments required to completely solve the items were present. The findings further support prior results (Reichersdorfer et al., 2012), that students have less trouble with refuting false statements than to solve technical proof tasks, tasks that require flexible proof skills, or conjecturing tasks for true statements. As regards our research here, we see substantial variation in students' proof skills. For space restrictions, we will not differentiate the different task types in the further analysis, even though this might be an interesting direction to pursue.

Table 2 presents how often formal quality aspects were coded in students' solutions, as well as presents means and standard deviations of the standardized quality scores for the different aspects of formal quality. As might be expected from the type of tasks, symbols for defining mathematical objects occurred comparably rarely (24.8%), while variables were used in 84.4% of the solutions. It was, nevertheless, possible to write arguments of high content quality without using variables. We would like to repeat that not using a certain formal notation or corresponding language construct did only result in coding as "incorrect (0)", if the corresponding formal aspect would have been necessary to communicate the students' solution according to the norms of the course.

	Symbolizing divisibility	Use of logical symbols	Symbolizing definitions	Use of variables	Use of quantifiers
Cases	63.5%	59.3%	24.8%	84.4%	40.9%
Mean score	.85 (.36)	.77 (.41)	.53 (.25)	.71 (.32)	.53 (.46)

Table 2: Number of cases coded, means (and standard deviations) of the standardized quality scores of the use of symbolic notations and formal representations

Results indicate that symbolizing definitions and the use of quantifiers caused the most problems, followed by the use of variables and the use of logical symbols. We identified the following difficulties in the use of symbolic notations and formal representations: In 9.4% of all solutions, an incorrect symbolizing of divisibility could be observed. Students showed an incorrect order of symbols or wrote "$a|b$" even though a did not divide b. In 12.1%, students applied logical symbols invalidly. For example, they used the implication symbol to delineate different statements, even though no valid implication could be established between the two statements. They marked valid logical relations by the use of an incorrect symbol in 3.4% of all solutions. In 2.6 % of all solutions, definitions were not made explicit at all, although the meaning of a symbol had been changed. In 18.4%, definitions were made explicit, but using a wrong symbol. For instance, some students marked a definition only by using the usual equal sign. In 6.4% of all solutions, variables were used inconsistently, for instance, representing the sum of consecutive even numbers by (2k) + (2m). In 36.1%, variables were used without a systematic introduction that explained what they stood for. We found that in 14.5% of all solutions, students did not use quantifiers or verbal quantifications, even it would have been necessary. In 6.7%, single quantifiers or verbal quantifications were used incorrectly, for example introducing a variable x, with a statement like "∃ x..." instead of "∀x…". In less than 1% of all solutions, students used single quantifiers correctly, but still showed problems with the order of consecutive quantifiers.

	Factor 1	**Factor 2**
Quality of arguments	.445*	.142
Symbolizing divisibility	.706*	-.037
Use of logical symbols	.676*	.010
Symbolizing definitions	.091	.178*
Use of variables	.004	.582*
Use of quantifiers	-.037	.352*

Table 3: Geomin rotated factor loadings

To analyse how the quality of arguments and the quality of different formal aspects of their representations are interrelated, we used exploratory factor analysis. Missing values in codings of formal quality were accounted for using the Full Information Maximum Likelihood (FIML) method. Each single task solution represented one case. The resulting hierarchical structure of the data (solutions nested in students) was also accounted for statistically analysis. Principal components analysis was used because the primary purpose of this study was to identify and later compute composite scores for the factors. Initial eigenvalues indicated that the first two factors explained 32.67% and 18.5% of the variance in all quality codings. The two factor solution was preferred because of our previous theoretical considerations and because it showed a significantly better model fit than the one factor solution (χ^2 (9) = 40.946, $p<.001$). Table 3 contains the Geomin rotated factor loadings for all quality criteria. The two factors were correlated significantly ($r=.40$, $p<.01$).

DISCUSSION

The goal of this study was to identify students' difficulties of mathematical argumentation and proving, and to analyse how the quality of the content of students' arguments and the formal quality of their presentation are interrelated. Firstly, our study replicates results that finding adequate arguments and communicating arguments with formal precision is a great challenge for students at the secondary-tertiary transition in mathematics (e.g., Clark & Lovric, 2008; Selden & Selden, 2009). In particular, when longer arguments have to be produced, students at the transition show similar problems as were reported for secondary students (Ufer et al., 2008) to find and describe conclusive chains of multiple deductive arguments. Regarding the formal quality of students' arguments, our sample shows evidence of all those problems that are documented in the literature, e.g., use of mathematical symbols, use of variables and quantifiers, and explicating definitions (e.g., Epp, 2003; Selden & Selden, 2011).

Apart from this, our study is to our knowledge the first that systematically studies relations between the content of students' arguments and their formal presentation. There are good theoretical arguments to assume that some of the formal aspects are quite unrelated to the content quality of an argument (Engelbrecht, 2010). Nevertheless, there are also theoretical reasons to assume that some formal aspects might be connected to the content quality of an argument (Epp, 2003; Sfard, 2008). We took an explorative approach to study these relations, and our analyses indicate that two dimensions of argument quality can be distinguished in our sample. One of these dimensions is substantially related to the content quality of students' arguments, but also to higher scores on *symbolizing divisibility* and *using logical symbols* for the respective arguments. Both of these formal aspects address relations between mathematical ideas (numbers and statements). The other dimension, largely unrelated to content quality, described the *use of variables* and *quantifiers* and – less pronounced – *symbolizing definitions*. These formal aspects seem to be more relevant to clarify the meaning of the mathematical objects used in an argument.

Of course our study was restricted to a specific educational setting and mathematical content. Nevertheless, our results indicate that not all, but some aspects of formal argument quality go along with the quality of the argument to be presented itself. If these results can be sustained, they might offer fruitful information to conceptualize student support in the learning of mathematical argumentation and proof. In particular, it might be possible to address some aspects (e.g., variables, quantifiers) separately in form of general behavioural schemata (Selden & Selden, 2009), while for others (e.g., logical symbols) a deeper connection to the underlying argument content will be necessary.

References

Clark, M., & Lovric, M. (2008). Suggestion for a theoretical model for secondary-tertiary transition in mathematics. *Mathematics Education Research Journal, 20*(2), 25–37.

Clement, J. (1982). Algebra Word Problem Solutions: Thought Processes Underlying a Common Misconception. *Journal for Research in Mathematics Education, 13*(1), 16–30.

Dubinsky, E., & Yiparaki, O. (2000). On student understanding of AE and EA quantification. In E. Dubinsky, A.H. Schoenfeld, & J. Kaput (Eds.), *Research in collegiate mathematics education IV*. (pp. 239-286). Providence, RI: AMS.

Engelbrecht, J. (2010). Adding structure to the transition process to advanced mathematical activity. *International Journal of Mathematical Education in Science and Technology, 41*(2), 143-154.

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

Epp, S. (2011). Variables in mathematics education. In P. Blackburn, H. van Ditmasch, M. Manzano & F. Soler-Toscano (Eds.), Tools for Teaching Logic (pp. 54-61). Berlin/ Heidelberg: Springer.

Heublein, U. (2014). Student Drop-out from German Higher Education Institutions. *European Journal of Education 4*, 497-513.

Healy, L. & Hoyles, C. (1998). *Justifying and Proving in School Mathematics: Technical Report on the Nationwide Survey*. London: Institute of Education.

Koedinger, K. R. (1998). Conjecturing and argumentation in high school geometry students. In Lehrer, R. and Chazan, D. (Eds.), *New Directions in the Teaching and Learning of Geometry.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Reichersdorfer, E., Vogel, F., Fischer, F., Kollar, I., Reiss, K., & Ufer, S. (2012). Different collaborative learning settings to foster mathematical argumentation skills. In T. Tso (Ed.): *Proceedings of the 36th Conference of the International Group for the Psychology of Mathematics Education*, Vol. 3, 345-352.

Schoenfeld, A. H. (1985). *Mathematical problem solving*. New York: Academic press.

Selden, J. & Selden A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics, 29,* 123-151.

Selden, J., & Selden, A. (2009). Teaching proving by coordinating aspects of proofs with students' abilities. In M. Blanton, D. Stylianou, & E. Knuth (Eds.), *The learning and teaching of proof across the grades* (pp. 339-354). London: Routledge/ Taylor & Francis.

Selden, A., & Selden, J. (2011). Mathematical and non-mathematical university students' proving difficulties. In L. R. Wiest & T. D. Lamberg (Eds.), *Proceedings of the 33rd annual conference of the North American chapter of the International Group for the Psychology of Mathematics Education* (pp. 675–683). Reno, NV.

Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing.* Cambridge, UK: Cambridge University Press.

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 PME 32 and PME-NA XXX, Vol. 4* (pp. 361-368). Morelia, Mexico: Cinvestav-UMSNH.