Proof validation aspects and cognitive student prerequisites in undergraduate mathematics

Daniel Sommerhoff, Stefan Ufer, Ingo Kollar

Angaben zur Veröffentlichung / Publication details:

Sommerhoff, Daniel, Stefan Ufer, and Ingo Kollar. 2016. "Proof validation aspects and cognitive student prerequisites in undergraduate mathematics." In Proceedings of the 40th Conference of the International Group for the Psychology of Mathematics Education, PME40, Szeged, Hungary, 3-7 August 2016, edited by C. Csikós, A. Rausch, and J. Szitányi, 219–26. Szeged: PME. http://www.lulu.com/shop/csaba-cs%C3%ADkos-and-attila-rausch-and-judit-szit%C3%A1nyi/proceedings-of-the-40th-conference-of-the-international-group-for-the-psychology-of-mathematics-education-volume-4/paperback/product-22902041.html.

PROOF VALIDATION ASPECTS AND COGNITIVE STUDENT PREREQUISITES IN UNDERGRADUATE MATHEMATICS

Daniel Sommerhoff[1], Stefan Ufer[1], Ingo Kollar[2]

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

Proof validation is an important skill to acquire for university students and is also essential as a monitoring activity during proof construction. Our study analyzes students' difficulties with three aspects of proof validation as well as the influence of domain-specific and domain-general cognitive student prerequisites (CSP) on proof validation skills. Results indicate that students' proof validation skills depend on the type of error in a purported proof and are influenced by conceptual mathematical knowledge and metacognitive awareness. Overall domain-general and -specific CSPs affect performance to roughly the same degree, whereas generative CSPs like problem solving skills have no contribution. These results question the current way of teaching the concept of proof primarily by proof construction exercises.

INTRODUCTION

Proof construction has been a focus of university mathematics for a long time and constitutes a research focus within mathematics education. Yet, students often get in touch with proofs in other ways, e.g. they engage in proof comprehension when reading textbook proofs or in proof validation when reading and judging the correctness of their own or other students' proofs or potentially erroneous lecture notes. Mastering these activities per se is essential for students, but is even more so since proof validation, i.e. the ability to evaluate individual arguments and entire proofs, is a crucial monitoring activity while constructing proofs (Selden & Selden, 2003). Apart from proof construction, research thus increasingly focuses on proof comprehension and proof validation (Healy & Hoyles, 2000; Inglis & Alcock, 2012; Selden & Selden, 2003; Weber & Mejía-Ramos, 2011).

Prior research revealed that even university students have severe problems in validating proofs (Selden & Selden, 2003). Gaining effective means of fostering students' proof validation skills therefore is of utmost importance. A prerequisite to design instruction that is effective for the acquisition of proof validation skills at the university level is a better understanding of different aspects of proof validation skills and of their relation to cognitive student prerequisites (CSP). The present study therefore explores students' proof validation skills in two ways: We analyze which types of errors in purported proofs are easy to detect for students and which pose difficulties. In addition, we explore the dependency of proof validation skills on various domain-specific and domain-general CSPs.

2016. In Csíkos, C., Rausch, A., & Szitányi, J. (Eds.). *Proceedings of the 40th Conference of the International Group for the Psychology of Mathematics Education,* Vol. 4, pp. 219–226. Szeged, Hungary: PME.

PROOF VALIDATION

"Reading" proofs comprises three main activities, each having different goals (Selden & Selden, 2015). *Proof comprehension* is the activity of reading a proof (e.g. when studying a textbook proof) that is known to be true with the aim of understanding it. *Proof validation* refers to reading a proof and trying to judge its correctness. The third related activity is *proof evaluation*, which is not only aimed at assessing the correctness of a proof, but also at evaluating the proof regarding multiple other properties, e.g. its clarity or convincingness.

Amongst these three skills, proof validation is closest related to mathematical proof construction skills, because validation is essential during proof construction for checking individual inferences as well as the overall structure and conclusiveness of a constructed proof. Due to this status as a monitoring activity, similar activities can be found in many domain-general frameworks for argumentation or problem solving, e.g. as *evidence evaluation* and *drawing conclusions* (Fischer et al., 2014) or *looking back* (Polya, 1945), or in modern self-regulation frameworks (De Corte et al., 2011).

Recent studies underline the importance of proof validation and unveiled clear differences in proof validation behavior, e.g. between experts and novices (Inglis & Alcock, 2012; Weber & Mejía-Ramos, 2011). While novices tend to focus on surface features of proofs and individual inferences (*zooming in*), experts rather focus on the high-level structure (*zooming out*) and skim proofs to grasp the overall structure before zooming in and looking at details.

Individual prerequisites of proof validation

Proof validation requires different knowledge facets and skills; e.g., judging the correctness of a proof is hardly possible without a sufficient mathematical knowledge base. Proof validation can therefore be seen as a complex cognitive skill that depends on several CSPs that can roughly be divided into two parts: domain-specific and domain-general prerequisites (c.f. Figure 1). Prior research on proof construction (Chinnappan, Ekanayake, & Brown, 2011; Schoenfeld, 1985) indicates that both could influence students' proof validation skills, but their relative impact is still unclear. Yet, in contrast to proof construction, which requires the generation of own, multi-step arguments, we view proof validation as a non-generative, more evaluative activity so that the impact of rather generative CSPs is questionable.

On the domain-specific side, it is assumed that students' conceptual and procedural mathematical knowledge base as well as mathematical strategic knowledge (Weber, 2001) impact proof construction skills. Transferring this to the non-generative activity of proof validation, at least the influence of procedural knowledge is questionable, but also the approach strategies for mathematical proofs encoded in mathematical strategic knowledge might not be relevant. On the domain-general side, various constructs, including problem solving skills (Chinnappan et al., 2011) and general inferential reasoning skills, likely influence proof construction. Again, the influence of students' problem solving skills on the non-generative proof validation is questionable. Finally, prior research (e.g. Yang, 2012) suggests an influence of metacognitive awareness on students' proof validation skills since proof validation can be seen as cognitively demanding, requiring students to reflect the given proof and their proof validation process on various levels.

Apart from the individual relations to proof validation skills, knowledge of the contribution of domain-specific *vs*. domain-general CSPs can be utilized to effectively foster proof validation skills: E.g., a high impact of domain-general prerequisites would support the inclusion of instructional support for more general skills, while a low impact would support trainings mostly focusing on conceptual knowledge of the corresponding proof content.

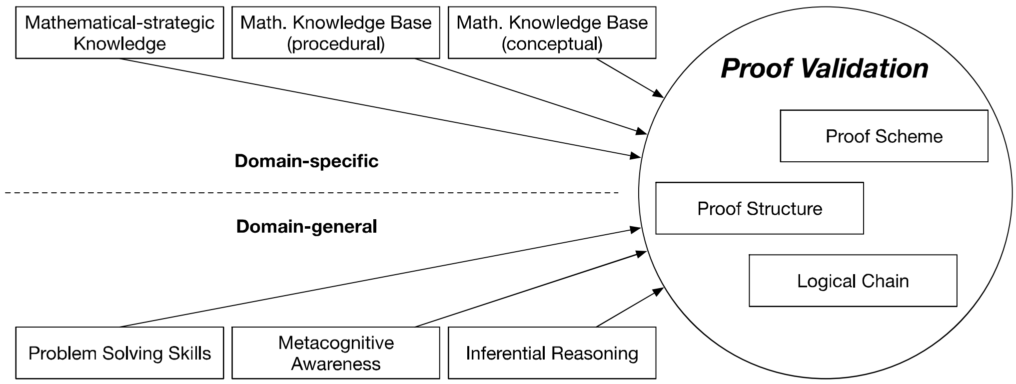

Figure 1: Conceptual framework of CSPs and aspects of *proof validation skills*

Aspects of proof validation

Proof validation is concerned with finding errors in purported proofs. Yet mathematical proofs can contain different kinds of errors (e.g. unsupported inferences, wrong use of definitions or cyclic argumentation) that refer to different aspects of a proof and that are not equally easy to detect (Healy & Hoyles, 2000). Accordingly, these different aspects of proof can be used to examine proof validation skills more closely, i.e. to differentiate between students' proof validation skills to detect specific types of errors. Heinze and Reiss (2003) put forward three aspects of *methodological knowledge* that are inherent to every proof and can be used to structure the different kinds of errors: *Proof scheme* refers to the kinds of reasoning used within each argument of a proof, e.g. inductive and deductive inferences or reference to an authority. *Proof structure* refers to the overall argumentative, logical structure of a proof. For a linear, direct proof, this structure should begin with the given premises and end with the statement that has to be shown. Finally, *logical chain* focuses on individual inferences within a proof. In order to obtain a correct logical chain, the premises for each step have to be proven beforehand or be part of the theoretical basis and no unknown or unproven statements may be used.

AIM AND RESEARCH QUESTIONS

The goal of this study was to identify university students' skills and problems in proof validation regarding different aspects of proofs and to explore the impact of domain-specific and domain-general CSPs on the proof validation skills. We focused on the following questions:

1. Are there differences in students' proof validation skills regarding the detection of the three different errors types? Are students able to relate the reasoning for their judgment to the (in-)correct aspects of the purported proofs?
2. What is the influence of students' domain-general and domain-specific cognitive prerequisites on their proof validation skills?
3. Does this influence depend on the error type contained in a purported proof?

SAMPLE AND METHOD

66 mathematics university students (24 male, 41 female, 1 NA; $M_{\text{age}} = 21.19$) who had finished their first semester participated in the study, which is part of a larger investigation aimed at fostering students' mathematical proof skills.

To examine students' proof validation skills, they were asked to judge the correctness of four purported proofs of one proposition (closed format) and to explain their decision (open format). To assure an unbiased validation, they were told that fellow students had created the proofs. Students were given liberal but limited time to think about the purported proofs. The proposition was taken from elementary number theory to ensure that a potential lack of advanced mathematical knowledge did not hinder the proof validation. Two days later the students were asked to judge four purported proofs of another proposition in order to validate the initial findings.

Proposition

The product of three arbitrary consecutive integers is divisible by 3.

Martin's proof

Let $a \in \mathbb{Z}$ be an arbitrary integer. Accordingly the two consecutive integers can be written as $a+1$ and $a+2$.
We are interested in the product

$$a \cdot (a+1) \cdot (a+2)$$

Expanding the term yields:

$$a \cdot (a+1) \cdot (a+2) = a \cdot (a^2+3a+2) = a^3+3a^2+2a$$

Since the proposition should hold for an arbitrary integer a, the statement has to hold independently of a. We therefore only look at the coefficients of the term a^3+3a^2+2a. For the sum of these coefficients we get:

$$1+3+2=6$$

Accordingly since $3|6$ holds, it also holds that $3|a^3+3a^2+2a$ resp. $3|a \cdot (a+1) \cdot (a+2)$. Thus the product of three arbitrary consecutive integers is divisible by 3 and we have proven the proposition.

Figure 2: Proposition 1 and the purported proof with an error in the *logical chain*

For both propositions, one proof was correct and each of the three other proofs contained an error corresponding to one of the three error types. A translated version of proposition 1 and the purported proof containing an error in the logical chain are shown in Figure 2. Obviously the conclusion that a^3+3a^2+2a is divisible by 3 independently of a when the sum of the coefficients 1+3+2 is divisible by 3 is both wrong and unwarranted. Nevertheless, the step is deductive in nature since Martin seems to refer to some general rule for this (without stating it explicitly).

To measure their CSPs, students were given paper and pencil tests measuring their mathematical knowledge base (conceptual and procedural), mathematical strategic knowledge, inferential reasoning skills (Inglis & Simpson, 2008), metacognitive awareness (Schraw & Dennison, 1994) and problem solving skills (four non-mathematical problem solving tasks). The tests used closed and open items. Two raters coded the open items following theory-based coding schemes. The interrater reliability was $\kappa > .76$ ($\kappa_{Mean} = .93$; $SD = .09$). All scales had an acceptable internal consistency with $\alpha_{Mean} = .70$ ($SD = .10$), only the internal consistency for mathematical strategic knowledge was a bit low with $\alpha = .58$ (4 items).

RESULTS

With a total of 59.1 % correct answers, students' overall performance in judging the correctness of proofs was moderate, yet significantly greater than chance ($t(259) = 3.29$, $p < 0.001$). Comparing the different purported proofs (correct proof and proof with errors in the proof scheme, proof structure or logical chain) a Conchran's Q test determined significant ($\chi^2(3) = 70.97$, $p < .001$) differences between the solution rates (cf. Figure 3, left; dark-grey). Pairwise comparisons between the four purported proofs with a Bonferroni correction yielded significant ($p < .05$) differences for all comparisons except for correct proof *vs*. proof scheme.

Students were quite accurate in judging the correct proof as correct (81.8 %) and the purported proof with an inductive proof scheme as wrong (86.4 %). On the other hand, students performed about chance on the proof containing an error in the logical chain (45.5 %) and significantly below chance ($t(64) = -5.54$, $p < 0.001$) in the proof containing an error in the proof structure (22.7%). The delayed test with proposition 2 showed similar patterns regarding students' judgments (cf. Figure 3, left; light-grey) as well as for the Conchran's Q test ($\chi^2(3) = 28.88$, $p < .001$).

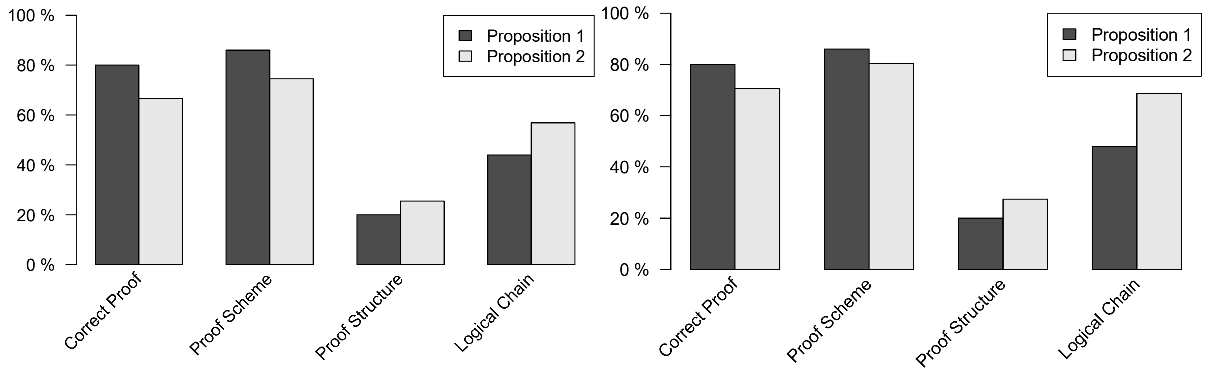

Figure 3: Solution rates without (left) and with consideration of explanations (right)

Students' explanations for their decisions revealed that some students marked faulty proofs as correct, although they had spotted the error or weakness of the proof. A typical statement was "Except for the part with 'only looking at the coefficients' the proof seems to be correct". Counting all answers as correct that showed that the error had been detected yielded the same result patterns for both propositions although numbers change slightly (cf. Figure 3, right).

The number of students stating an explanation for their decisions varies widely between the purported proofs (43.9 % correct proof, 74.2 % proof scheme, 25.8 % proof structure, 66.7 % logical chain). For the proofs with an error in the proof scheme students were best in finding the correct reason for judging the proof as wrong (75.5 % of the given explanations), for proof structure the worst (29.4 % of the given explanations). The same pattern was observed for proposition 2.

We employed a generalized linear mixed-effects model (GLMM) analysis using the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) to analyze the influence of the CSPs on proof validation. The model includes all six CSPs as well as the aspects of proof validation as fixed effects and the participant as a random effect. The model explained 36.8 % of the variance in students' proof validation performance by the CSPs and the aspects of proof validation. The model shows that, compared to identifying a correct proof, it is much harder to identify errors in proof structure ($b = -2.91$, $p < .001$) and logical chain ($b = -1.52$, $p < .001$) but easier to identify errors in the proof scheme ($b = 0.35$, $p > .05$). Of the CSPs, only the conceptual mathematical knowledge base and metacognitive awareness showed significant relations (c.f. Figure 4; stand. regression weights $\beta = 0.39$ and $\beta = 0.33$ respectively) to students' proof validation skills. Employing the GLMM on data from both propositions yields similar results.

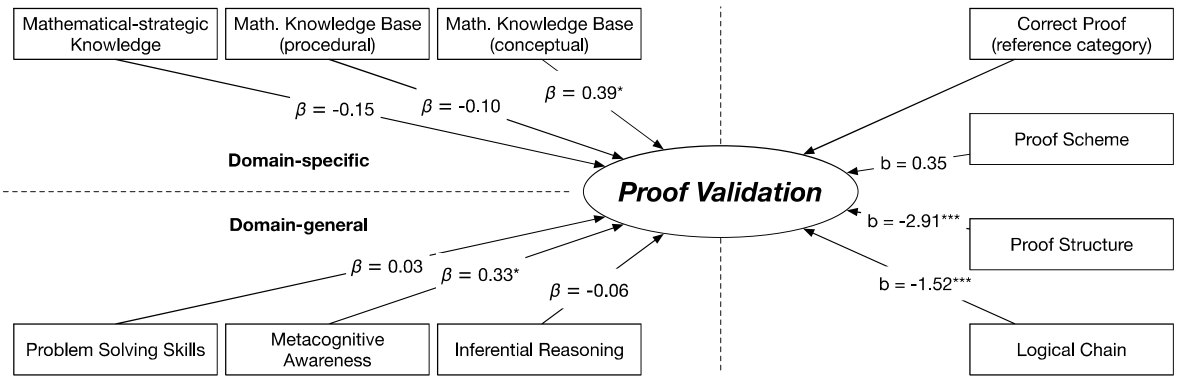

Figure 4: GLMM of CSPs and aspects of proof validation skills

For proposition 1, an analysis of interaction effects revealed that students' conceptual knowledge supports the detection of a wrong proof structure ($p < .05$) more than the identification of the other proofs. On the other hand, conceptual knowledge shows a weaker connection to identifying wrong proof schemes as compared to identifying the other proofs ($p < .05$). Finally, the impact of metacognitive awareness is stronger for detecting errors in the proof scheme as compared to evaluating the other proofs ($p < .05$). Again, the data from the second proposition showed a similar pattern.

DISCUSSION

The results of our study focusing on the aspects of proof validation skills reveal clear differences in students' performance. Students have little problems identifying correct proofs and refusing inductive proof schemes, but perform poor when confronted with other error types. The low success rate in finding errors in the overall logical structure of the proof can be seen as additional evidence for the results that mathematics students often focus too narrowly on individual inferences (*zooming in*) (Mejía-Ramos & Weber, 2014; Weber & Mejía-Ramos, 2011). Yet, students were also not excellent at finding errors in the *logical chain* that refers to the individual inferences and the *zooming in*. The analysis of students' explanations adds to this: How come students mark proofs as correct although they were able to identify errors? And why do only few students give reasons for their judgments although they were explicitly prompted? One answer might be problems in understanding the proofs and giving suitable reasons. Alternatively, the wrong judgments despite finding the errors could also be a side effect of good proof constructions skills and students' insight that the proof could be tackled with a similar argument. Further evidence on students' thoughts and views is needed here, e.g. from interview or think-aloud studies.

So far, the overall results on the aspects of proof validation resemble those from prior research, e.g. on secondary students in the area of geometry, indicating some generalizability of these results over content area and age groups. Yet a replication with propositions from multiple content areas would be beneficial to further assure the validity and generalizability of the results.

The results regarding the influence of CSPs on proof validation skills indicate complex relations. Both the domain-specific as well as domain-general CSPs showed significant relations to students' proof validation skills of similar magnitude. Therefore, domain-general interventions, e.g. for metacognitive awareness, could have positive effects on students' proof validation skills and, vice versa, interventions on proof validation skills might be expected to transfer to skills in other domains to a certain extent. Yet, the results from the GLMMs show, that according interventions have to be created carefully: Amongst the CSPs, neither the mathematical-strategic knowledge nor the prerequisites referring to generative activities (procedural mathematical knowledge and problem solving) show a significant relation to proof validation skills. This missing relation of generative activities to proof validation is plausible. Beyond that, it indicates that proof validation might offer a better entry into the learning of proof than proof construction activities, since proof validation seems to be dependent on fewer prerequisites, in particular generative skills. This would be an alternative to the current university teaching style of mathematical proof, which is often mostly based on proof construction. Although a potential impact of proof validation on proof construction was not studied here, there are first results showing a significant connection of proof validation and construction skills. Thinking of proof validation as one prerequisite for proof construction also warrants such an approach. Still, more research, in particular from intervention studies would be required to support this strategy.

Overall, our approach using CSPs and aspects of proof validation was applied successfully and yielded several interesting implications for research on as well as for the teaching of mathematical proof. The fact, that proof validation seems to depend less on some of the CSPs than proof construction (e.g. problem solving) and that proof validation skills are needed for proof construction underlines the idea of "validation before construction".

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1).

Chinnappan, M., Ekanayake, M., & Brown, C. (2011). Specific and General Knowledge in Geometric Proof Development. *SAARC Journal of Educational Research*, *8*, 1–28.

De Corte, E., Mason, L., Depaepe, F., Verschaffel, L., Zimmerman, B., & Schunk, D. (2011). Self-regulation of mathematical knowledge and skills. In B. J. Zimmerman & D. H. Shunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 155–172). NY: T & F.

Fischer, F., Kollar, I., Ufer, S., Sodian, B., Hussmann, H., Pekrun, R., … Eberle, J. (2014). Scientific reasoning and argumentation: Advancing an interdisciplinary research agenda. *Frontline Learning Research*, *4*, 28–45.

Healy, L., & Hoyles, C. (2000). A Study of Proof Conceptions in Algebra. *Journal for Research in Mathematics Education*, *21*(4), 396–428.

Heinze, A., & Reiss, K. (2003). Reasoning and Proof: Methodological Knowledge as a Component of Proof Competence. In M. A. Mariotti (Ed.), *International Newsletter of Proof Competence* (Vol. 4).

Inglis, M., & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education*, *43*(4), 358–390.

Inglis, M., & Simpson, A. (2008). Conditional inference and advanced mathematical study. *Educational Studies in Mathematics*, *67*(3), 187–204.

Mejía-Ramos, J., & Weber, K. (2014). Why and how mathematicians read proofs: further evidence from a survey study. *Educational Studies in Mathematics*, *85*(2), 161–173.

Polya, G. (1945). *How to Solve it: A New Aspect of Mathematica Method*. Princeton, NJ: Princeton University Press.

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

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, *19*(4), 460–475.

Selden, A., & Selden, J. (2003). Validations of Proofs Considered as Texts: Can Undergraduates Tell Whether an Argument Proves a Theorem? *Journal for Research in Mathematics Education*, *34*(1), pp. 4–36.

Selden, A., & Selden, J. (2015). A comparison of proof comprehension, proof construction, proof validation and proof evaluation. In *KHDM Proceedings Hannover Germany December 2015*.

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics*, *48*(1), 101–119.

Weber, K., & Mejía-Ramos, J. (2011). Why and how mathematicians read proofs: an exploratory study. *Educational Studies in Mathematics*, *76*(3), 329–344.

Yang, K. L. (2012). Structures of cognitive and metacognitive reading strategy use for reading comprehension of geometry proof. *Educational Studies in Mathematics, 80*(3), 307-326.