RESEARCH ON MATHEMATICAL ARGUMENTATION: A DESCRIPTIVE REVIEW OF PME PROCEEDINGS

Daniel Sommerhoff, Stefan Ufer, Ingo Kollar

University of Munich (LMU)

Mathematical argumentation and proof (MA&P) traditionally are major topics of mathematics education in secondary and tertiary education. Although many studies focus on MA&P it remains unclear how they contribute to a coherent understanding of MA&P processes. We have analysed PME research reports focusing on MA&P published 2010 to 2014 to determine the different prerequisites as well as goals of argumentation and proving processes investigated within these reports. Results indicate that research on MA&P covers a broad range of processes, sub-skills and knowledge facets, but that individual reports predominantly address only singular aspects. A holistic approach to MA&P, taking into account the whole process or multiple sub-skills, is rare. We discuss implications for future research of MA&P.

INTRODUCTION

Mathematics is a proving science, and mathematical argumentation and proof (MA&P) therefore are central to mathematical activity (Ubuz, Dincer, & Bulbul, 2012). Many standard documents worldwide put forward MA&P as one central goal of mathematics learning (CCSSI, 2010), especially in secondary and tertiary education. Accordingly, mathematics education research has traditionally approached this field from various perspectives. It is widely agreed that MA&P comprise complex skills that integrate diverse individual cognitive prerequisites and different kinds of mathematical activities. From our understanding, an aim of MA&P research must be an increasingly coherent understanding of these diverse facets, since otherwise effective support of MA&P processes is not possible.

The purpose of this review is to analyse current research on mathematical argumentation and proof in secondary and tertiary education. To cover the diversity of MA&P extensively, we have based our analysis on existing theoretical frameworks of scientific reasoning which highlight prerequisites, processes and goals of MA&P: Predictors for mathematical argumentation skills (Ufer, Heinze, & Reiss, 2008), epistemic activities in scientific argumentation (Fischer et al., 2014), and argumentative and proving activities (Giaquinto, 2005).

THEORETICAL FRAMEWORK

Argumentation and proof

According to Balacheff (1999), there are two meanings of argumentation within the field of mathematics. Thus, mathematical argumentation can be considered a discursive activity aimed at convincing a listener. On the other hand, based rather on

2015. In Beswick, K.., Muir, T., & Wells, J. (Eds.). *Proceedings of 39th Psychology of Mathematics Education conference*, Vol. 4, pp. 193-200. Hobart, Australia: PME.

Toulmin's view, argumentation is an activity which is aimed at the generation, exploration and validation of conjectures and hypotheses in terms of their objective and individual rationality (Pedemonte, 2007). For our review we adopt this second view. Accordingly, mathematical proof is seen as a more formal form of mathematical argumentation, which is subject to (mostly implicit, social, and possibly changing) norms of the mathematical community. This difference between argumentation and proof is nicely put by Pedemonte (2008, p. 385):

> "There is a 'structural gap' between argumentation and proof because in argumentation inferences are based on content while in proof they follow a deductive scheme (data, claim, and inference rules)."

Sub-skills and knowledge facets

The success of mathematical argumentation and proving depends on individual prerequisites like domain-general and domain-specific knowledge facets, beliefs and more overarching skills. Over the last decades researchers have proven a variety of such sub-skills and knowledge facets to be predictive for mathematical argumentation skill (e.g., Ufer, Heinze, & Reiss, 2008), which therefore often are called predictors. For this review we adopt a framework of predictors worked out in Ufer, Heinze and Reiss (2008) that a) is well based on research from the last decades, b) is not limited to a specific mathematical area, and c) allows separation of domain-specific and more domain-general predictors. The framework contains six main predictors. *Methodological knowledge* is knowledge of the nature and the functions of proof as well as the acceptance criteria for a valid proof (Healy & Hoyles, 2000). *Mathematical knowledge base* consists of basic conceptual and procedural knowledge in the field of mathematics (Ufer et al., 2008). Mathematical strategic knowledge is knowledge about cues within mathematical tasks and problems that indicate which concepts and representation systems can be used productively (Weber, 2001). *Problem-solving skills* consist of domain-general and domain-specific problem solving skills and strategies (Schoenfeld, 1985). Finally there are *beliefs* about the mathematical content and nature of mathematics (Leder, Pehkonen, & Törner, 2002; Schoenfeld, 2010) as well as *affective aspects* like emotions and motivation towards mathematics (Hannula, 2006).

Similar approaches to consider complex skills together with various predictors can also be found for self-regulated learning (De Corte, Verschaffel, & Eynde, 2000), mathematical problem solving in general (Schoenfeld, 1985) or mathematical proof in geometry (Chinnappan, Ekanayake, & Brown, 2011), with very similar predictors.

Epistemic Activities

Besides their predictors, we describe MA&P processes by analysing their sub-activities with a framework that has been proposed by Fischer et al. (2014). It describes eight such "epistemic activities" (Table 1) from an interdisciplinary viewpoint that allow comparisons among different domains and topics. The idea is that cognitive aspects of individual MA&P processes can be described in terms of these basic activities. Albeit the linear presentation of these activities, they do not need to occur in this specific order, can be iterated in cycles and not necessarily are all present in an argumentative process.

Epistemic Activity	Description
Problem identification	Perceiving a mismatch concerning the explanation of a problem and building a problem representation
Questioning	One or more initial questions are identified
Hypothesis generation	Possible answers to the questions are derived from models, theoretic frameworks, …
Construction and redesign of artifacts	Development of a prototypical object, axiomatic system or another object used in order to work on the problem
Evidence generation	Evidence for the hypothesis is generated
Evidence evaluation	Evaluating evidence according to some norms
Drawing conclusions	Integrating different pieces of evidence, reevaluating the initial claim considering the new evidence
Communicating and scrutinizing	Sharing and discussing individual reasoning and argumentation within a community

Table 1: Overview of epistemic activities (Fischer et al., 2014).

Argumentative and proving activities

Not only the individual cognitive sub-activities within an argumentative process can be distinguished, but also the overall goal of the reasoning process with reference to task contexts. Mejia-Ramos and Inglis (2009) introduced a framework of argumentative and proving activities based on work by Giaquinto (2005). They divide argumentative activities associated with mathematical proof into the three categories *construction of novel arguments*, *reading arguments* and *presenting arguments*, each with a few sub-categories. Even though this distinction sounds very similar to some of the epistemic activities, it refers to the *overall goal* of MA&P processes, not the sequence of activities within this process.

AIM AND RESEARCH QUESTIONS

The goal of this review is to analyse which aspects of MA&P have been investigated in the last 5 years within the PME community, and to identify patterns that might yield directions for future research in understanding and supporting MA&P as a complex individual skill. The review was therefore guided by the following questions:

- To which extent does research on MA&P consider the different predictors, sub-activities, and goals of MA&P processes?
- Which combinations of predictors and epistemic activities are being

considered in MA&P research? Can research gaps be identified with regard to a comprehensive understanding of MA&P processes?

THE CURRENT STUDY

Literature selection, coding and analysis

We decided to restrict our review to PME research reports (RRs) published from 2010 to 2014, because we considered them to be a fair representation of latest good-quality, international mathematics education research. All 782 RRs of the PME proceedings from 2010 to 2014 were selected as data basis for the review. This selection bears the danger of overlooking research that is not published within the PME proceedings, but includes research that is of good quality and is not limited to journal publications, thus giving a more extensive picture of the activities in the community. A similar approach was taken by Matos (2013) for his literature review.

Based on an initial coding of the research topic and grade level, we selected those 129 RRs for detailed analysis which studied MA&P and which were situated in secondary or tertiary education. The focus on reasoning and argumentation in secondary and tertiary education is due to the fact that it differs considerably from that in pre-primary and primary education, particularly proof is rather non-existent.

A coding scheme was created to categorize these RRs according to the predictors investigated, the epistemic activities studied, and the type of reasoning activity (according to its goal) in the study. Reading the RRs completely, we coded the main research foci of each report with respect to the three theoretical frameworks. For each predictor we coded if it was a variable central to the RRs (e.g., it was the sole focus of the report), if it was considered substantially (e.g., it was analysed together with other predictors), if it was only mentioned (e.g., as a variable to be controlled), or if it did not occur at all. The goals of MA&P processes were coded in the categories *argument construction*, *argument reading*, and *argument presentation* where possible, but also codes *not explicit* and *multiple goals* were introduced. Moreover, we coded for each epistemic activity if it was focused in the report. The notion of "focused" is very important to understand the whole coding process. For example, if participants of a study were talking or discussing a problem only for purposes of the study (e.g., to foster collaboration or as a "thinking aloud" technique) this would neither be coded as the activities *proof presentation* or *communication and scrutinizing*, nor as the goal *proof presentation*. After several steps of refinement, the coding reliability reached an acceptable level with a mean inter-rater reliability of $\kappa_{Mean} = 0.77$ (SD = 0.15). Except for the interrater reliabilities of the epistemic activities *drawing conclusions* ($\kappa = 0.56$) and *communicating and scrutinizing* ($\kappa = 0.46$) all IRRs were acceptable (above $\kappa = 0.64$).

In an additional step the results of the descriptive analysis were backed up by considering examples of reports from the respective categories in order to ensure coding validity and to gain qualitative insight.

RESULTS

A total of 532 (68%) articles were situated in secondary (44%) or tertiary (24%) education and 160 (20%) RRs focused on MA&P. The intersection of both groups contained 129 (16%) RRs, which met the inclusion criteria, and were coded in detail. Comparing the research methods of these RRs (57% qualitative, 26% quantitative, 11% mixed methods, and 6% theoretical) to the ones found by Matos (2013), the selected reports are quite representative for PME RRs in terms of research methods. The same holds for the RRs' distribution of participants (Figure 1, left side).

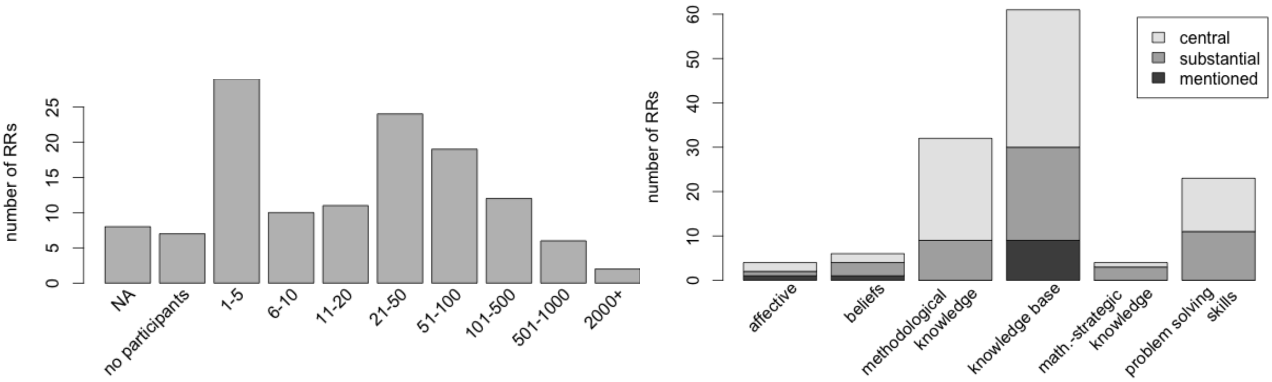

Figure 1: Distribution of number of participants (left) and use of predictors (right).

Starting our analysis with the predictors, *mathematical knowledge base* was studied by far most often (47% of the RRs; Figure 1, right side). Only 17% considered *methodological knowledge* and 18% *problem solving*. *Strategic knowledge*, *beliefs* and *affective aspects* were studied even less frequently (3%, 5%, and 3%, resp.). All in all only 22% of the RRs considered at least two of these predictors simultaneously, over two thirds of these cases focused on the predictor *mathematical knowledge base* in combination with any one predictor.

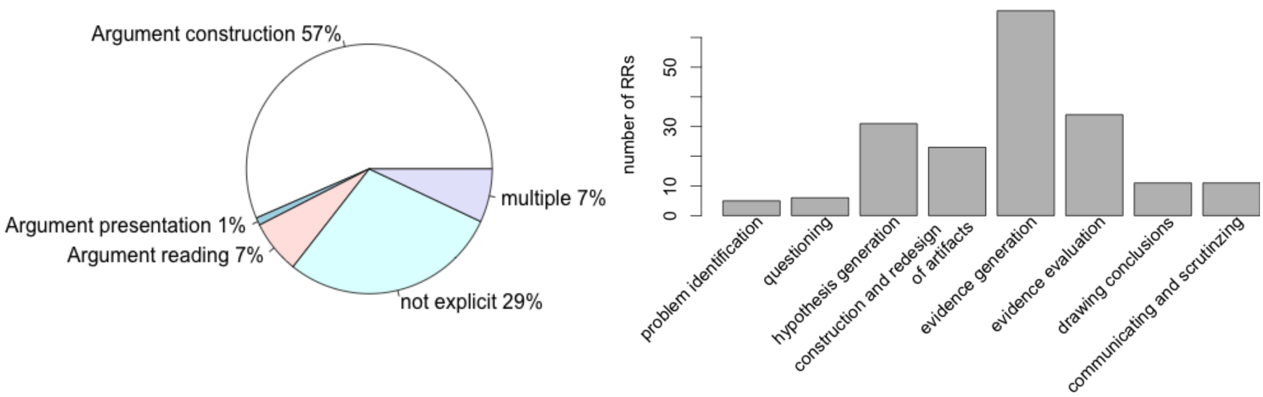

Figure 2: Frequencies of argumentative (left) resp. epistemic activities (right)

Regarding the goals of the argumentative processes (Figure 2, left side), almost 60% of the RRs focused on *argument construction*, 1% on *argument presentation* and 7% on *argument reading*. The number of reports including two or more of these goals is also low with 7%. Almost a third of the RRs (29%) could not be associated with one of these three activities, for example because they were theoretical. These results resemble those of Mejia-Ramos and Inglis (2009), who found *argument construction* in 64% of their sample, but no contribution on *argument presentation*.

In line with this focus on the dominating goal of *argument construction*, *evidence generation* was the most frequently studied epistemic activity (Figure 2, right side), followed by *hypothesis generation* and *evidence generation*. Nevertheless, all epistemic activities were studied at least in some form in some RR. A qualitative analysis of the RRs focusing at least one epistemic activity (96 of 129 RRs) revealed four main clusters (named A, B, C, D) of RRs, a finding also supported by a cluster analysis. Two of these clusters (A, D) focus on one epistemic activity only, the others (B, C) on several. Cluster A focuses solely on *evidence generation* and constitutes the largest cluster with 32% of the 96 RRs. A representative of this cluster is a RR on unjustified assumptions in geometry proofs, where students' written geometry proofs were analysed for these assumptions. The second largest cluster with 30% of the RRs is cluster B, the conjecturing cluster that focuses on the activities of *hypothesis generation*, *construction and redesign of artefacts* and *evidence generation*. A representative of this cluster is a videotaped interview study of the ways successful provers use examples when exploring and proving conjectures given to them. The third biggest cluster with 24% of the RRs is cluster C, the "complete" process cluster, which incorporates RRs looking at multiple epistemic activities at once. A representative of this cluster is a RR on the role of dynamic geometry on the process of exploration, conjecturing and proving geometrical problems. Finally, the smallest cluster with only 14% of the RRs is cluster D, the evaluation cluster, which focuses on the epistemic activity of *evidence evaluation*. A representative of this cluster is an eye-tracking study of the role of pictures while reading proofs.

These clusters also differ in the sample sizes and the applied research methods. The mean sample size in cluster A is 85, whereas the other clusters have mean sample sizes of 50 and below, with cluster D having the smallest mean sample size of 32. Although all clusters predominantly contain qualitative RRs, the percentages are especially high in the complete (C) and conjecturing (B) clusters with 77% resp. 79%. Apparently a qualitative approach is used more often when having a wider perspective on MA&P and/or focusing on several epistemic activities.

Data also reveal a strong connection between the processes and goals of MA&P investigated. Thus, RRs with a focus on *argument construction* predominantly studied the activities of *hypothesis* and *evidence generation*, whereas the RRs on *argument reading* or *presentation* focused exclusively on *evidence evaluation*. Especially in the case of *argument presentation* this seems surprising as a focus on *communicating and scrutinizing* would be obvious.

DISCUSSION

The aim of our review was to analyse the inclusion and combination of different predictors, sub-activities, and goals of MA&P in research on MA&P in PME and how it contributes to a comprehensive understanding of MA&P. The results reveal that perspectives on MA&P are often restricted to very specific aspects such as single epistemic activities or one or few predictors. Initially, such more focused analyses are necessary as a first approach to better understand complex skills. Nevertheless, MA&P require the coordination of multiple processes and knowledge facets. Even though taking a broader perspective of MA&P poses major methodological problems, e.g., in terms of sample size or time for testing or analysis, it is important to find ways to study the complex interactions of the often disconnected aspects described in existing research. This may include studies comparing the influence of different predictors or research on the coordination of different epistemic activities during MA&P processes.

We also find that MA&P are mostly researched in situations where *argument construction* is the main goal of the activity. This may be one reason why certain epistemic activities resp. their combinations are studied in more detail than others. However, Meija-Ramos and Inglis (2009) suggested that *argument presentation* and *argument comprehension* may be more important in learning settings than *argument construction*. Certainly, the relative importance of different goals of MA&P and different epistemic activities has to be seen in conjunction with the overall aims of mathematics instruction that may be more focused on *argument construction*. Still, we cannot expect to gather a comprehensive understanding of MA&P while having blind spots in our research.

Despite these imbalances and potential research gaps it must be underlined that, with over 20% of the RRs focusing on argumentation and proof, we have a sound basis of research on the separate aspects of MA&P. Thus, it may be time to build on that basis and to start studying the relations and interactions between the different facets of MA&P in order to obtain a coherent picture as well as more detailed knowledge how to foster MA&P effectively.

References

Balacheff, N. (1999). Is argumentation an obstacle? Invitation to a debate... *International Newsletter on the Teaching and Learning of Mathematical Proof*. Retrieved from http://www.lettredelapreuve.it/OldPreuve/Newsletter/990506Theme/990506ThemeUK.html

CCSSI. (2010). Common Core State Standards for Mathematics. Retrieved from http://www.corestandards.org/assets/CCSSI_Math Standards.pdf

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., Verschaffel, L., & Eynde, P. O. (2000). Chapter 21 - Self-Regulation: A Characteristic and a Goal of Mathematics Education. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 687–726). San Diego: Academic Press.

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.

Giaquinto, M. (2005). Mathematical Activity. In P. Mancosu, K. Jørgensen, & S. Pedersen (Eds.), *Visualization, Explanation and Reasoning Styles in Mathematics* (Vol. 327, pp. 75–87). Springer Netherlands.

Hannula, M. S. (2006). Affect in mathematical thinking and learning. *New Mathematics Education Research and Practice*, 209–232.

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

Leder, G. C., Pehkonen, E., & Törner, G. (2002). *Beliefs: A hidden variable in mathematics education?* (Vol. 31). Dordrecht: Kluwer Academic Publishers.

Matos, J. F. (2013). Trends and opportunities for research methodologies within the PME community. In A. Lindmeier & A. Heinze (Eds.), *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 51–66). Kiel, Germany: PME.

Mejia-Ramos, J. P., & Inglis, M. (2009). Argumentative and proving activities in mathematics education research. In F.-L. Lin, F.-J. Hsieh, G. Hanna, & M. de Villiers (Eds.), *Proceedings of the ICMI Study 19 conference: Proof and Proving in Mathematics Education* (Vol. 2, pp. 88–93). Taipei, Taiwan

Pedemonte, B. (2007). How can the relationship between argumentation and proof be analysed? *Educational Studies in Mathematics*, *66*(1), 23–41.

Pedemonte, B. (2008). Argumentation and algebraic proof. *ZDM*. *The International Journal on Mathematics Education*, *40*(3), 385–400.

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

Schoenfeld, A. H. (2010). *How we think : a theory of goal-oriented decision making and its educational applications*. New York: Routledge.

Ubuz, B., Dincer, S., & Bulbul, A. (2012). Argumentation in undergraduate math courses : A study on proof generation. In T.-Y. Tso (Ed.), *Proceedings of the 36th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 163–170). Taipeh, Taiwan: PME.

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* (Vol. 1, pp. 361–368). Morelia, Mexico: PME.

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