



HAL
open science

Introduction to the papers of TWG21: Assessment in mathematics education

Paola Iannone, Michal Ayalon, Jeremy Hodgen, Michiel Veldhuis

► **To cite this version:**

Paola Iannone, Michal Ayalon, Jeremy Hodgen, Michiel Veldhuis. Introduction to the papers of TWG21: Assessment in mathematics education. CERME 10, Feb 2017, Dublin, Ireland. hal-01949292

HAL Id: hal-01949292

<https://hal.science/hal-01949292>

Submitted on 9 Dec 2018

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

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

TWG21: Assessment in mathematics education

Introduction to the papers of TWG21: Assessment in mathematics education

Paola Iannone¹, Michal Ayalon², Jeremy Hodgen³ and Michiel Veldhuis⁴

¹Mathematics Education Center, Loughborough University, UK; p.iannone@lboro.ac.uk

²Department of Mathematics Education, University of Haifa, Israel, mayalon@edu.haifa.ac.il

³School of Education, Nottingham University, UK, ttzjh@exmail.nottingham.ac.uk

⁴Freudenthal Group - FSW, Utrecht University, Netherlands, m.veldhuis@uu.nl

Keywords: Summative assessment, formative assessment, validity, feedback, assessment methods.

Introduction

Given the prominence that the general education research ascribes to the impact of assessment on teaching and learning it is surprising that there had been so far no TWG dedicated to assessment of mathematics at CERME. TWG21, which met for the first time at CERME10 in Dublin, aimed to fill this gap. Given that this was a new group we decided to focus on assessment of mathematics considered broadly in order to gauge where the interest of the mathematics education community lies in this field, which encompasses very many different aspects. Although traditionally assessment has been discussed across many TWGs at CERME, TWG21 aimed to bring researchers together who have an interest in this topic and can, for the lack of a common forum, at times feel isolated. To reflect the landscape in the general literature we called for papers investigating the nature of assessment and its effects on student learning making use of a wide range of methodologies, from large quantitative and mixed methods study to small investigative qualitative studies. We were delighted to have 24 papers and one poster discussed at the conference. In what follows, we have grouped the papers in thematic clusters to reflect the variety of submissions regarding both focus and methodology. We conclude with some reflections on the working of the group and some suggestions for the directions this group can take in future CERME conferences.

Thematic clusters

We identified six overarching themes that could serve as an organizing tool for the papers submitted to TWG21. Below, we describe each of these themes in turn.

Different approaches to assessment: Papers in this theme considered the affordances, drawbacks and validity of innovative assessment, both for students and for teachers. Davies proposes comparative judgment at university level as a new way of assessing students. In his paper, he investigates issues connected to the validity of this method for assessing conceptual understanding in mathematics. Lemmo and Mariotti investigate the issues connected with transitions of tasks from a paper and pencil form to an electronic form. They challenge the view that students employ similar solving strategies in both environments and find that indeed students solve the task differently in the two modalities. Teledhal investigates the validity of narrative accounts as an assessment tool for problem solving and concludes that those accounts do not offer enough details of the problem-solving process to be a valid tool for assessment. Dahl describes the perceptions of a group of science students (engineers, mathematicians, and other sciences) for group oral assessment. She finds that students

across disciplines agreed that a group exam gives less differentiation of grades compared to an individual exam. Finally, Reit discusses whether the validity of teachers' intuitive assessment practices is supported by empirical findings and shows that a sequential consideration of thought structures in a solution approach leads to reasonable results and may justify its application in school due to its straightforward implementation, especially when assessing modelling tasks.

In service and pre-service teachers' views: A second important theme that emerged from the submissions to TWG21 was related to teachers' views, beliefs, and use of assessment methods, both during their training and in their professional practice. Hofmann and Roth report on a study aimed at fostering preservice teachers' diagnostic skills with a focus on students' abilities, problems and misconceptions with graphs of functions. They explore the affordances of two tools for promoting diagnostic skills: video analysis and task analysis. Pratt and Alderton analyse English mathematics teachers' assessment approaches in the context of the current changes in assessment policy in the UK. To this end they use a Foucauldian analysis of teachers' discourse to sketch the power structures involved. They find that the official removal of the levels only superficially affected teachers' practices and teachers still relate these to the 'old' language of attainment levels. Kaplan and Haser investigate 27 preservice middle school mathematics teachers' purposes in planning the assessment and their views and suggestions about the assessment part of a lesson plan. Findings of the study indicate that purposes underlined by preservice teachers in preparing the assessment part of the lesson are similar across the sample and they all related to the teacher actions.

Professional development: Papers in this cluster addressed the role of professional development in fostering teachers' (both in service and pre-service) competences in assessing student understanding. Grapin and Sayac investigate the use of external (e.g. researcher-created) assessment tasks by primary school mathematics teachers and teachers' practice by using an activity theory perspective. They find that teachers design tests with low levels of complexity and did not invest much in assessment as a professional activity. Pilet and Horoks present analytical tools to characterize assessment activities as part of teachers' practice in algebra. The authors exemplify why high school teachers came to consider assessment as a potential lever to enhance both the students' learning in mathematics and the teachers' development. Initial results indicate that the teachers developed better indicators to select the students' productions that they will use for the discussion after a task, but that they use they make of these products hasn't improved. In her theoretical paper, Andersson argues that the addition of the dimension Teacher Instruction (ATI) as a key strategy to the five key strategies proposed in Wiliam and Thompson's (2007) framework of formative assessment could facilitate the analysis of teachers' use of formative assessment activities and improve the guidance and support of teachers' implementation of high quality formative assessment practice. Finally, in this group Santos and Domingos investigate portfolio assessment in geometry for pre-service teachers through the lenses of activity theory and procepts. They find students engage in qualitative different pathways when solving these problems.

Formative assessment/feedback: We received many papers discussing formative feedback and the submissions in this group spanned from primary to upper secondary school with focuses both on teachers' use of formative assessment and students' engagement with such assessment. Chanudet investigates the assessment of problem solving by using a grid of criteria. The paper focuses on the use that teachers make of such tool to facilitate formative assessment and offers the example of the

practice of one teacher where she analyses instances of formative feedback occurring in this classroom. Zhao, Van den Heuvel-Panhuizen, and Veldhuis investigated the effects on student achievement of supporting Chinese primary mathematics teachers' use of classroom assessment techniques. In this experimental study, the intervention consisted of teachers participating in workshops on the use of these techniques and using them in their classrooms. Results indicate that the students of teachers that gained more insight about their students from using the techniques, improved their mathematics achievement scores more than other students. Gurhy focuses on Irish students' perspectives on the use of assessment for learning in primary school. Findings indicated that students were positive about the feedback in, and practices of, assessment for learning, became more confident and expressed a feeling of enjoyment related to this. Two related papers reported findings from FaSMEd, a European project on the use of technology for formative assessment. In the first paper, Cusi, Morselli and Sabena analyse a teacher's strategies to provide feedback during class discussion. They identify five strategies: revoicing, rephrasing, rephrasing with scaffolding, relaunching, and contrasting. In the second paper, the authors describe how materials were designed to facilitate technology-enhanced formative assessment practices. They then show how the design framework can be used to analyse the implementation of technology-enhanced materials. They argue that materials designed in this way, combined with the functionality of technology, enhance a teacher's capacity to activate Wiliam & Thompson (2007) formative assessment strategies.

Task design: Three papers were dedicated to this theme. O'Brien and Ní Ríordáin describe the development, design, and theoretical underpinning of a diagnostic test for algebra. The test is aimed at lower secondary students in Ireland and is intended to help teachers identify the causes of students' errors. The authors discuss their reasons for adopting this approach. Beck investigates students' written solutions from CAS-allowed exams. Based on the analysis of students' solutions a descriptive model for assessing these solutions is set up. The paper also discusses how formative assessment could help students develop their competencies in communicating mathematics. Moomaw investigates the validation of a constructivist game- and story-based measure (Teddy Bear Picnic) for pre-school mathematics. In this measure, pre-school pupils are assessed while playing several interactive games. Psychometric tests show that the test appears to be a valid and reliable measure of pupils' level of mathematical development.

Large-scale/standardized tests: Finally, we received several papers addressing issues related to the use and design of large nationwide standardized tests. Garuti, Lasorsa, and Pozio describe the development of items for national assessment in Italy. They show how both quantitative and qualitative analysis can be used to improve the psychometric properties of items, whilst also improving their validity in terms of appropriate and relevant mathematical content. Ferretti and Gambini investigate the persistence of certain misconceptions in the transition between school and university. They focus on properties of powers and analyse two Italian nationwide databases to find that indeed certain misconceptions persist across this transition. Drüke-Noe and Kühn analyse characteristics of statewide exams in eight countries through task analysis and find that the cognitive demands of most competences needed to solve these tasks are rather low with the only the competence 'working technically' being often assessed. Cunningham, Shiel, and Close investigate the relation between the current Junior Certificate mathematics examination in Ireland for Grade 9 to the PISA and TIMSS frameworks. Their findings show that the Junior Certificate examination is moving closer in the direction of the PISA approach, but this is also motivated by the comprehensive

reform in mathematics in this country. Finally, Olande investigates how Grade 9 students solve an item involving the interpretation of graphs. Using student responses to an item from the national test in Sweden, his analysis shows that only a very small proportion of students use graphical reasoning in their solutions.

Conclusions

In the process of preparing for this new group at CERME10 we were impressed not only by the variety of work we received but also by the methodological variety of the papers that spread from small qualitative case studies to large statistical surveys. The theoretical frameworks employed were also varied, from Activity Theory to Foucauldian analysis. We believe this variety to be sign of a growing interest in mathematics education for assessment; not only in the sense of validation of large scale tests, but also in terms of the effect that assessment has on teachers' actions in the classroom and as such on student learning. This variety, however, can also be sign of a field which has yet to find its unifying themes: the presence of a forum for discussion like TWG21 can therefore help define these emerging unifying themes. Validity of assessment for example – although ubiquitous in many papers – was hardly explicitly addressed. Indeed, in the final session of our group which was dedicated to reflecting on the group experience with an eye to future meetings, we observed some issues which at times have hindered communication. One of those was the lack of uniformity in definitions of recurring terms or sometimes the lack of clear definitions at all. It was felt that agreement on definitions of basic terms is important for communication and collaboration, and the lack of this clarity of definitions can be again a manifestation of a developing and growing field. We also noticed the absence of papers discussing the impact of assessment methods on student learning, a theme which is very much present in the assessment literature. The final reflection of the group concerned the presence of *mathematics* in the research presented. The group felt that in a topic such as assessment it may be easy to lose the focus on the mathematics assessed and instead discuss generic assessment research. While assessment research in general education is obviously very important to the work of this group, all participants felt that the focus should be on the mathematics assessed, and that indeed it may be a difficult balancing act not to replicate research and constructs that are already used in the general assessment literature and keep the focus on the fact that we aim to use these findings and constructs to investigate the assessment of mathematics. Although this balancing act might make for a difficult enterprise, we are confident that in the coming CERMEs we will be able to continue discussing general assessment issues such as validity, but always with a clear focus on the mathematics to be assessed and its didactics.

References

William, D., & Thompson, M. (2007). Integrating assessment with learning: what will it take to make it work? In C. A. Dwyer (Ed.), *The Future of Assessment: Shaping Teaching and Learning* (pp. 53–82). US. Mahwah, NJ: Lawrence Erlbaum Associates.