Intelligent Analysis and Data Visualisation for Teacher Assistance tools: the case of exploratory learning

Manolis Mavrikis[1], Eirini Geraniou[2], Sergio Gutierrez Santos[3], Alexandra Poulovassilis[3]

[1] UCL Knowledge Lab, Institute of Education, University College London, UK

[2] Department of Curriculum Pedagogy and Assessment, Institute of Education, University College London,

[3] Birkbeck Knowledge Lab, Birkbeck University of London, UK

m.mavrikis@ucl.ac.uk, e.geraniou@ucl.ac.uk, sergut@dcs.bbk.ac.uk, ap@dcs.bbk.ac.uk

Brief biography of authors:

Dr Manolis Mavrikis is a Reader in Learning Technologies at the UCL Knowledge Lab. He holds a B.Sc. in Mathematics from University of Athens, Greece with an emphasis in education, M.Sc. with distinction in Informatics and Ph.D. in Artificial Intelligence in Education from the University of Edinburgh. His research interests developed over 20 years of experience, lie at the intersection of learning sciences, human–computer interaction and artificial intelligence. Manolis is currently an editor for the British Journal of Education Technology and director of the MA in Education and Technology at UCL.

Dr Eirini Geraniou is an Associate Professor of Mathematics Education at the UCL Institute of Education. Before joining the UCL Institute of Education, she taught maths in secondary school and university level. She has worked on a number of research projects (MiGen, Metafora, italk2learn, MC2), whose aim was to design, develop and evaluate digital tools for teaching and learning mathematics. She is currently Mathematics Subject Lead for the Teach First Programme and the lead for the master's module on digital technologies for mathematical learning.

Dr Sergio Gutierrez Santos was a Lecturer in Computer Science at Birkbeck at the time this research was done. He has a PhD in computer science (2007) from University Carlos III of Madrid. He has left academia for a life in industry but his research interests continue to centre on artificial intelligence (especially emergent technologies) and its application to problems in teaching and learning, in particular in relation to exploratory learning and creative thinking.

Prof Alexandra Poulovassilis is Professor of Computer Science at Birkbeck and Director of the Birkbeck Knowledge Lab. Her interests are in data integration, querying, analysis, visualisation, and personalisation, and she has published widely in these areas. She has held numerous research grants, many of them pursuing interdisciplinary research in collaboration with domain experts from education, the arts, and the sciences. The goal of such research is to effectively support learning communities in capturing, organising, discovering and sharing knowledge.

Intelligent Analysis and Data Visualisation for Teacher Assistance tools: the case of exploratory learning

Abstract

While it is commonly accepted that Learning Analytics tools can support teachers' awareness and classroom orchestration, not all forms of pedagogy are congruent to the types of data generated by digital technologies or the algorithms used to analyse them. One such pedagogy that has been so far underserved by Learning Analytics is exploratory learning, exemplified by tools such as simulators, virtual labs, microworlds and some interactive educational games. This paper argues that the combination of intelligent analysis of interaction data from such an exploratory learning environment (ELE) and the targeted design of visualisations has the benefit of supporting classroom orchestration and consequently enabling the adoption of this pedagogy to the classroom. We present a case study of learning analytics in the context of an ELE supporting the learning of algebra. We focus on the formative qualitative evaluation of a suite of Teacher Assistance tools. We draw conclusions relating to the value of the tools to teachers and reflect with transferable lessons for future related work.

Keywords: teacher support, artificial intelligence, learning analytics, exploratory learning

1. INTRODUCTION

Exploratory Learning Environments (*ELEs*) are a particular type of computer-based learning environment where the focus is on students' exploration of the knowledge domain. Examples of ELEs include simulators, virtual labs for science topics, some interactive educational games, and more generally microworlds that especially in mathematics and computing, adhere to a constructionist pedagogy (c.f. Papert & Harel, 1991; Healy & Kynigos, 2010). Compared to *Intelligent Tutoring Systems* (*ITSs*), ELEs pose ill-defined, open-ended tasks that are primarily promoting the development of conceptual rather than procedural knowledge. Recognising the need for considerable guidance to ensure learning in such open-ended activities (Kirscher et al., 2006; Kynigos, 1992; Mayer, 2004) and the potential for engagement and learning (Noss and Hoyles, 1996; Grawemeyer et al., 2017; Rummel et al., 2016), efforts have focused on intelligent components that provide feedback and support to students (e.g. Gutierrez Santos et al., 2012b; Amershi and Conati, 2009; Roll et al., 2010).

However, be it due to the lack of maturity of the technology, or potentially the scepticism around it, *artificial intelligence* (*AI*) cannot (and should not) replace key pedagogical strategies in this context as this may violate some of the constructionist principles upon which exploratory learning is founded (Mavrikis et al. 2013a). The silver lining is that this constraint also offers a response to criticisms of the field of educational technology, particularly the application of AI as wedded to an instructional pedagogy (c.f. du Boulay this issue, Wilson & Scott, 2017). The response is strengthened when we shift our attention from adaptive feedback or other recommendations for students to supporting the difficult role of the teacher as 'facilitator' or 'orchestrator' (Trouche, 2004; Hoyles et al., 2004). This role would be relatively easy in one-to-one student-tutor interaction, but scaling it up to the number of students in a typical class poses several orchestration challenges, further compounded by the use of technology (Dillenbourg, 2013, Prieto et al. 2015) and the open-ended nature of the exploratory tasks (Mavrikis et al., 2013).

While *Learning Analytics* (*LA*) tools have been generally framed as tools to support teachers' awareness, reflection and sense-making to support decision-making (Verbert, 2013; Holstein et al., 2017), as we review in Section 2, there has been little attention to exploratory environments. Martinez-Maldonado (2016) draws a clear connection with orchestration, conceptualising learning analytics as tools that can be used to address practical orchestration challenges in classrooms such as managing, adapting, scaffolding and assessing learning activities (c.f. Prieto et al., 2015, Rodriguez-Triana et al., 2018). Although this body of research helps us shift away from a narrow data-centric view of analytics towards a consideration of the pedagogical and human factors in play, some authors still point out how LA have not generally been designed with a focus on teachers' needs (Holstein et al., 2017; Dillenbourg, 2013). We agree and, in addition, argue that, by their nature, these analytics, particularly when simply quantifying interaction from digital tools, tend to favour instructionist pedagogies as we review in Section 2.

Accordingly, the overarching aim of the work presented in this paper is to identify ways to support teachers in orchestrating exploratory learning while adhering to a constructionist pedagogy. Our case study is the MiGen system, which includes an intelligent microworld designed to support 11-14 year old students' development of algebraic ways of thinking. As part of the MiGen system, we have co-designed with teachers a suite of visualisation and notification *Teacher Assistance* (*TA*) tools.

Our earlier work described the architectural design and implementation of an early version of the TA tools, focusing specifically on the Student Tracking tool (Pearce-Lazard et al., 2010; Gutierrez Santos et al., 2012). In Gutierrez Santos et al. (2017) we described also a Grouping Tool that makes recommendations to teachers about how to group students for productive discussion. In Mavrikis et al. (2016) we presented a high-level summative evaluation focusing particularly on teachers' perceptions of the tools in relation to supporting real-time decision making. In contrast, the present paper takes a reflective stance and presents the whole suite of the TA tools, their pedagogical rationale, and the design process in a summative manner. We present the usage scenarios of the TA tools and how we incorporated teachers' requirements. In addition, we reflect on a formative evaluation with teachers through a rich description of the qualitative findings. We then discuss these findings and draw transferable conclusions relating to the value of such tools and lessons for future related work.

2. BACKGROUND AND RELATED WORK

2.1 The MiGen system, eXpresser, and algebraic generalisation

The MiGen system is an intelligent, exploratory environment to support 11 to 14-year-old students learning of algebraic generalisation (Noss et al., 2012). Using a mathematical microworld called the *eXpresser*, students are asked to construct two-dimensional tiled models and associated algebraic rules. In order to build their model, students need to create 'building blocks' out of unit-square coloured tiles depending on their perception of the model's structure, and to repeat each building block in order to form a 'pattern'. The algebraic rules they are asked to construct relate to the number of tiles of each colour required to paint each pattern and their model overall. Figure 1 shows a sample model students can construct.

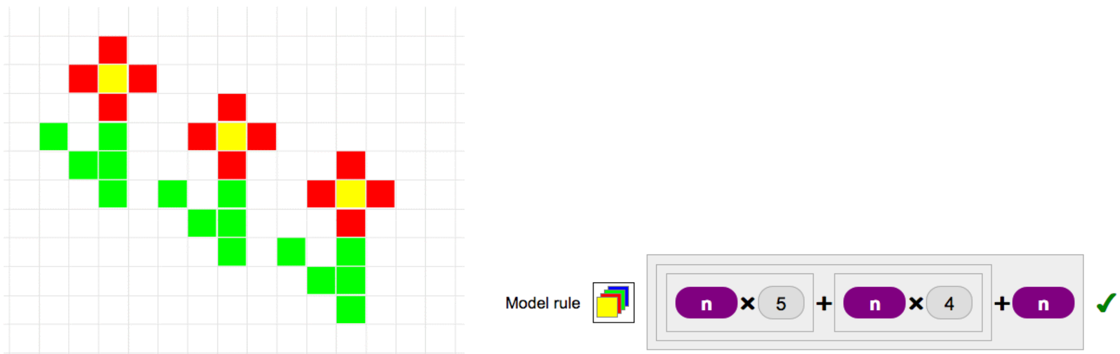

Fig. 1: An example model and rules that students may construct in eXpresser. They can do so by creating 'building blocks' to generate the centres of the flowers, the petals, and the stalks, which they will then repeat to make the yellow, red and green patterns. They will be nudged towards deriving general rules for *n* number of flowers. This is challenging as an 'animation' feature applies different random values to the variables used by the student in order to check the generality of their rules. In the case here, the rule is 5n (green tiles) plus 4n (red tiles) plus n (yellow tiles).

ELEs such as MiGen's eXpresser have the potential to support students' exploration while at the same time fostering progressive building of knowledge. For example, eXpresser has been specifically designed to support students with some well-known and researched challenges on algebra (see Mavrikis et al., 2012 for details).

The exploratory nature of the tasks undertaken with eXpresser requires that personalised feedback is provided to students as they construct their solutions, and not just at the end of the task. This feedback includes prompts to help students engage with a task, guide them towards successful completion of the task, and generalise their solutions (Gutierrez Santos et al., 2012b). The feedback is generated by a component of the MiGen system, the *eGeneraliser*, based on automatic analysis of students' actions in the eXpresser. The aim of the feedback is to balance students' freedom to explore while at the same time providing sufficient support to ensure that learning is being achieved.

2.2 Related Work

To our knowledge, MiGen's TA tools represent one of the earliest works aiming to support teachers' orchestration of exploratory learning during the use of digital tools in the classroom (c.f. an early publication on this topic, Pearce-Lazard et al., 2010).

The trend towards learning analytics has of course grown exponentially the last few years (Garcia et al., 2012; Zaldivar et al., 2012; Holstein et al., 2017; Martinez-Maldonado et al., 2017) particularly in higher education and around the development of dashboards (Schwendimann et al., 2017). However, there is limited work in school settings (c.f. Schwendimann et al., 2017) and, with a few isolated exceptions, they do not focus on exploratory learning explicitly. For example, a recent systematic literature review shows a small percentage of papers applying to pedagogical approaches of problem- or inquiry- and even games-based learning, that have similarities with learning through exploratory learning environments (Mangaroska & Giannakos, 2018).

In the last few years, several initiatives have started to focus on ELEs, including approaches building on the work described here and aiming to help teachers' reflection on the use of ELEs for science and mathematics learning integrated in larger platforms such as Metafora (Dragon et al., 2013) and MCSquared (Mavrikis & Karkalas, 2016). Related work in the field has aimed to inform teachers of students' progress and need for help in the context of computer programming labs (Gutierrez Rojas et al., 2012). Amir et al. (2013) targeted a virtual lab environment for chemistry and designed visualisations that provide insights regarding students' problem-solving processes. In contrast, Gueraud et al. (2009) and Ben-Naim et al. (2008) focused mostly on the statistics of students' interactions, e.g. how often did a student produce a certain kind of indicator, as a way to understand students' behaviour. Our own research emphasis has been informed by detailed requirements analysis with our teacher collaborators, who attributed more value to real-time information about students' progress in order to support their classroom orchestration.

This focus on orchestration creates a high synergetic potential between the work reported here and the use of learning analytics tools in Computer Supported Collaborative Learning (CSCL) and Inquiry-based learning (IBL) approaches. Notable examples in this area that relate to our work include Sergis et al. (2019) that investigate how to support teacher guidance and reflection during inquiry tasks, Martinez-Maldonado et al. (2015) that aimed explicitly to enhance teacher awareness of small group work with interactive surfaces, and van Leeuwen (2015) where teachers regulated collaborating groups in simulation situations. For a review in the broader area of Natural User Interfaces see Martinez-Maldonado et al. (2017). From a technical point of view, the use of rules in Voyiatzaki et al. (2008) to find specific landmarks in CSCL interactions bears some similarity to our detection of the interaction indicators that underlie our TA tools, although we have employed a wider range of AI techniques, including rule-based reasoning, case-based reasoning and pattern-matching (see further discussion in Section 4 below).

3. DESIGN AND EVALUATION METHODOLOGY

We have previously outlined our iterative design methodology both in the first generative phase of ideation (Pearce-Lazard et al., 2010; Gurierrez Santos et al., 2012) and in later stages of prototyping, testing and evaluation (Mavrikis et. al., 2013), recognising from the outset the need for collaboration with teachers from early on and throughout the process. Revisiting our work, our design and evaluation approach for the TA tools resembles the LATUX workflow put forward for designing and deploying awareness tools in the classroom (Martinez-Maldonado et al., 2015). That model identifies an initial problem identification phase, followed by a series of iterative formative evaluation stages. In relation to our exploratory learning context, the key challenge we faced in designing the TA tools was to overcome teachers' lack of experience with exploratory learning tools, let alone with using also learning analytics tools. We recognised that in-depth understanding of teachers' requirements for the TA tools necessitated observing and analyzing situations of actual usage context. We, therefore, adopted a 'contextual design' approach (Holtzblatt & Beyer, 1997) to engineer situations that gave teacher participants the opportunity to experience first-hand what it would mean to have access to such tools and, therefore, to offer deeper insights in subsequent one-to-one interviews with the research team (Mavrikis et. al, 2013). This approach helped address the twofold challenge that expert teachers do not necessarily have the data literacy skills or understanding of the potential of learning analytics to contribute to design decisions or to provide feedback without first using prototypes of the TA tools themselves.

Our Teacher Advisory Group on the MiGen project comprised around 20 mathematics teachers and educators from a broad spectrum of secondary schools in the London area, who attended regular project team meetings and gave their input throughout the project. However, the time that the teachers had available to use prototypes of the TA tools in their classrooms was limited, and collaboration with a core group of 4 teachers and their schools played a prominent role in this respect. In Mavrikis et al. (2016) we presented a method that allowed us to conduct evaluation studies of the TA tools in our premises with several teachers at a time, in a way that provided them with a realistic experience of using the tools in the classroom. For completeness, we describe this method again in Section 6.1, and then focus specifically on the formative evaluation phase, which was not discussed in Mavrikis et al. (2016). We describe in Section 6.2 the feedback received from teacher participants during the formative evaluation, which led to the summative evaluation and the final versions of the TA tools, described in Sections 4 and 5.

4. THE TEACHER ASSISTANCE TOOLS

MiGen's TA tools include the *Student Tracking* (ST), *Classroom Dynamics* (CD), *Goal Achievements* (GA) and also a Grouping Tool (see Figure 2). The ST, CD and GA tools were developed to address a set of Usage Scenarios that were iteratively identified during the first phase of ideation with our teacher advisory group, including classroom piloting of an early version of the ST tool with two teachers (Gurierrez Santos et al. 2012):

1. Finding out which students need the teacher's immediate help.
2. Finding out which students are progressing satisfactorily towards completing the task and which students may be in difficulty.
3. Finding out which students are currently disengaged from the task.
4. Identifying common conceptual and procedural difficulties that students are facing in order to provide more explanation to the class as a whole.
5. Finding out which students have finished the task.
6. Finding out which students have achieved which task goals.
7. Providing appropriate support and guidance to individual students: (i) during the lesson, and (ii) after the lesson.
8. Reflecting on the achievements of the class and planning the next lesson.

A description of these Usage Scenarios (USs), and how they are supported by the ST, CD and GA tools, is presented in Appendix 2. We refer to Gutierrez Santos et al. (2012) for a detailed description of the ST tool and to Mavrikis et al. (2016) for a detailed description of the CD and GA tools. All three of these tools rely on the detection or inference of a set of key *indicators* as students are interacting with the eXpresser. Lower-level, task-independent indicators are detected by the eXpresser itself, while higher-level, task-dependent indicators are inferred by the eGeneraliser component. Identification of the full set of indicators that are meaningful for teachers was achieved through an iterative process undertaken with our teacher collaborators during the first phase of ideation, resulting in the development of a variety of computational techniques to track approximately 50 different indicators (Gutierrez Santos et al., 2012). For the task-dependent indicators inferred by the eGeneraliser, these techniques include AI techniques such as: case-based reasoning, to compare the student's evolving solution with the set of possible solutions; rule-based reasoning to determine if the student has coloured their model in a general way; similarity matching to determine sequences of repetitive actions ('rhythm') in the student's construction; and a combination of all three to determine if a task goal is being achieved – see (Gutierrez Santos et al., 2012) for details. Working with our teacher collaborators, we were able to classify each of the 50 indicators as being positive, negative or neutral with respect to a student's constructive interaction with the system and achievement of the task learning goals.

All indicator occurrences are submitted to the MiGen Server for storage in the MiGen database as they are detected by the eXpresser or inferred by the eGeneraliser. The TA tools receive real-time information from the MiGen server relating to such events and each TA tool presents visually a selection of this information to the teacher (see Noss et al., 2012 for details of the MiGen system architecture).

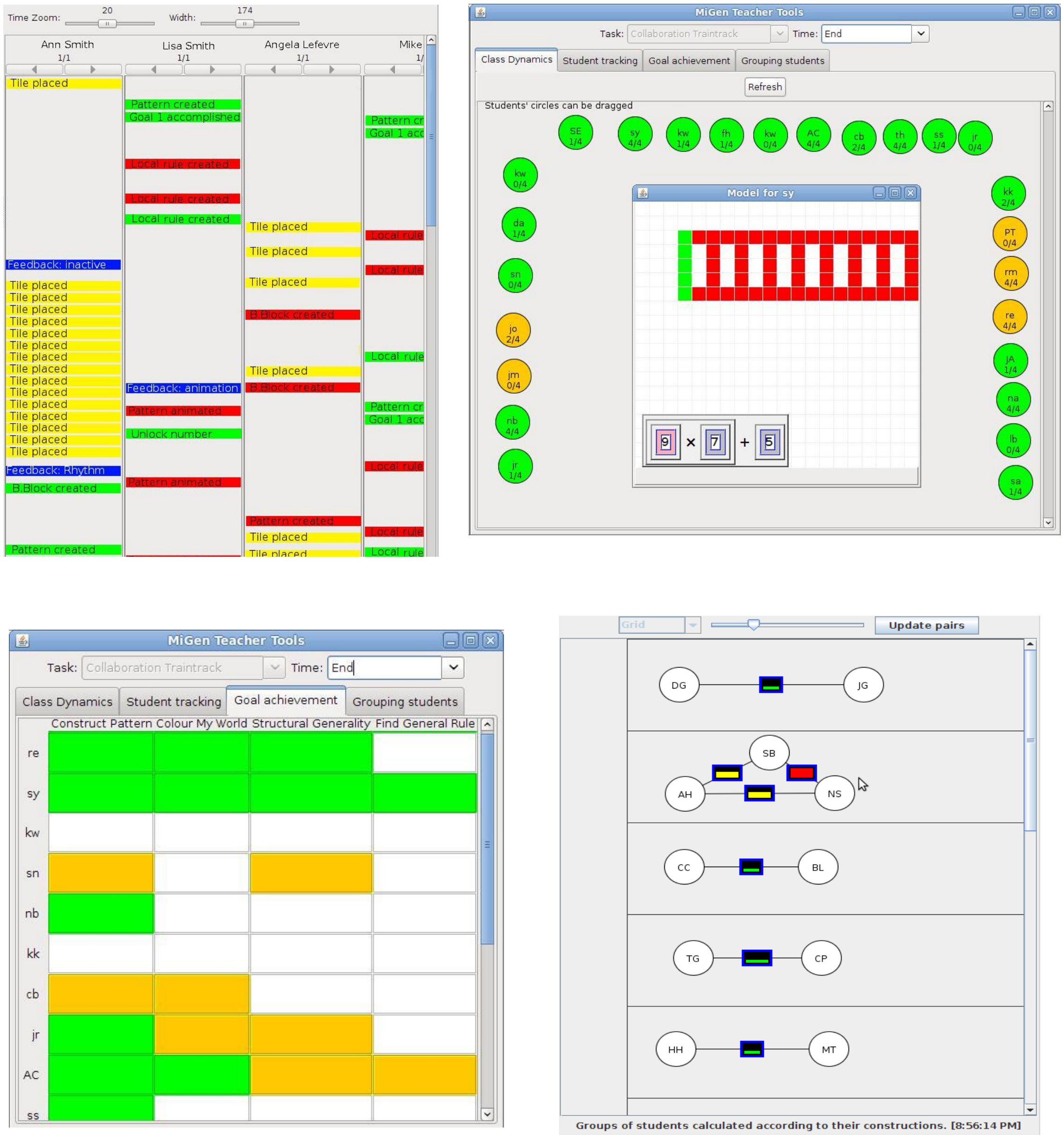

Figure 2: Student Tracking tool (top left). Class Dynamics tool (top right), with the students sitting at desks by the walls; the teacher has clicked on one of the students to see their current construction and rule. Goal Achievements tool (bottom left). Grouping tool (bottom right).

The ST tool monitors and displays the occurrence of all indicators, colour-coded (blue for system feedback and red/yellow/green for negative/neutral/positive indicators), in chronological order in a top-down timeline for each student. The CD and GA tool displays are driven by the subset of indicators that relate to students' current activity status, waiting for help from the teacher, and goal achievement status. The CD tool gives the teacher an at-a-glance view of which students are currently engaged with the task and who may be in difficulty and in need of help, representing each student in the classroom by a colour-coded circle containing the student's initials (green for a student who is working actively on the task, amber for possible inactivity, red for students who need the teacher's help). Within each circle, the proportion of task goals currently achieved by that student is also shown. The circles can be moved by the teacher so as to match the students' spatial positioning in the classroom. The GA tool shows a tabular display of students and task goals, each row showing the progress of one student in completing the task goals. Cells are colour-coded white/green/yellow, indicating respectively that a task goal has not been achieved/is achieved by the student's current construction/was achieved in the past but is not achieved by the student's current construction.

A 'time-stop' functionality allows the user to select a specific point in time with respect to which the TA tools' visualisations are generated, the default being the current time. This functionality allows teachers to see this information relating to a particular point in the past in order to better understand the context of a particular situation. Teachers can see what task goals were being accomplished and what solution approaches were being adopted by students, so as to assign additional homework or plan the next lesson. Due to the historical nature of the data stored in the MiGen database, this "looking back" could extend to earlier lessons, so for example the teacher could examine how a whole course has progressed to inform the design of future deliveries of the course. The time-stop functionality also allows the TA tools to be used by for research purposes to visualise students' interaction data arising from each classroom trial.

A fourth tool, the Grouping Tool (Figure 2, bottom right), addresses a ninth US that emerged after the formative evaluation described here: pairing students for productive discussion of their solution approaches at the end of a task. The design and evaluation of the Grouping Tool occurred independently of the other three tools and is discussed in (Gutierrez-Santos et al., 2017). In brief, the Grouping Tool undertakes pairwise comparisons of different students' solution approaches and proposes to the teacher suggested pairings of students (plus one triplet in the case of an odd number of total students) who can be asked to discuss their solutions at the end of the task. The aim is to put together students who have taken *different* solution approaches, so that they might compare these and consider if they are mathematically equivalent.

5. EXAMPLE USE OF THE TA TOOLS

In order to facilitate readers' understanding of the ST, CD and GA tools, we now briefly describe how they are used in a typical classroom session. At the start of the session, the teacher introduces the lesson and instructs students to open the eXpresser on their computers and to read about the current task within the system (e.g. the task may be to construct a model like the one in Fig. 1). While they are doing this, she opens up the TA tools on her computer, typically a tablet. For the first few minutes of the lesson, the teacher walks around the classroom to make sure that students are focusing on the task at hand and that they understand the task goals. Once students have begun working on the task, she can take a step back and use the TA tools to monitor students' progress (US 1-3).

Most of the time, the teacher will have the CD tool selected for display. If any students show as amber, she approaches them and encourages them to resume working on the task. Some students may call out to the teacher for help, or may raise their hands. The teacher encourages them to first seek help from the system: “If the system cannot help you, then I will come to you” she tells them, knowing that students in such a situation will automatically appear red in the CD tool.

If a student does appear red in the CD tool, the teacher goes to the student to help, since she knows that this is a situation where the system’s intelligent support cannot help the student any further (US 7i). If more than one student appears red, she clicks on those students’ circles to see their current models and rules, so that she can prioritise helping the students who seem to be having the most difficulty and also see if there are common misconceptions occurring, in which case she can pause the lesson and provide appropriate additional guidance to the whole class (US 4).

From time to time, the teacher looks also at the GA tool. Knowing which students have accomplished all the task goals allows the teacher to offer them additional activities (US 5 and 6). Other students may be advancing more slowly; the teacher can use this information to set them additional homework so that they can catch up with their peers if they need more time than is available in the lesson (US 7ii). If the GA tool shows that many students are not achieving a particular task goal, the teacher can interrupt the lesson to help all the students at the same time by clarifying a goal that may be unclear or by providing additional guidance to help students’ understanding (US 4).

At the end of the lesson, the teacher can use the ST tool to examine in detail how specific students have interacted with the eXpresser. For example, if she explained to a student during the lesson how to relate two patterns by using the same variable, she can check whether the student did this right away or required a period of ‘trial and error’ to understand the concept. Likewise, she can use the GA tool ‘frozen’ at specific times to assess how well the class has progressed with the task goals and whether some concepts will require reinforcement as extra homework or in a subsequent lesson (US 7ii and 8).

6. FORMATIVE EVALUATION

6.1 Method

The main aim of this phase was to test the full suite of prototype TA tools against the USs so as to inform final changes. This process resembles the ‘higher fidelity prototyping’ stage of the LATUX workflow, but we aimed at obtaining input from several teachers at a time. Apart from general feedback on use of the tools, the overarching research question was: “Can teachers use the tools effectively and how do they perceive their correspondence to the usage scenarios?”. We also wanted to find out what challenges teachers may face in using the tools to help us design any subsequent training.

We held a 3-hour evaluation session with 26 trainee Maths teachers from the Postgraduate Certificate of Education (PGCE) programme at the Institute of Education, split into two parallel groups of 13 for logistical reasons. Our decision to work with these teachers was both pragmatic (a sample we had access to) but also strategic in that in our experience they are more "tech savvy" and more likely to adopt and advocate the use of educational technology in their school when they return from their training[1].

Each participant had an installation of the MiGen system running on their computer. In the first half of the session, participants were introduced to the MiGen system as a whole and asked to work through several construction examples using eXpresser so as to gain familiarity with how students might use it in a lesson and the kinds of feedback the system would give to students. This was followed by a 15 minute break and a 30 minutes session where each of the TA tools was introduced using real student interaction data drawn from one of the classroom trials undertaken previously[2].

Using the time-stop functionality, the research team 'froze' the display of the data at 10 minutes into the lesson and gave a brief explanation to the participants of the information being shown in each tool. For the final hour of the session, participants were asked to move the display of the TA tools on their computer forwards, firstly to 30 minutes into the lesson, and then to 5 minutes prior to the end of the lesson. For each of these time points, they were asked to answer a list of questions relating to US 1–6 (listed in Appendix 3). At the end of the session, they were asked to complete a similar questionnaire (also listed in Appendix 3), this time relating to the full set of USs, where for each question we asked participants how they would tackle the corresponding US with and without the TA tools. Two participants did not complete either of the questionnaires. We summarise below participants' answers with respect to each of the USs. Table 1 provides further information and indicative comments of the 24 responders with respect to each of the USs.

6.2 Results

US 1. Finding out which students need immediate help

For both time points (30 minutes into the lesson and 5 minutes before the end), 23 participants demonstrated an appreciation of the CD tool. Most responses for the 30 minutes time point related to 'keeping an eye' on students who did not seem to have had a good start and ensuring that students "marked with amber are staying focused". For the time point 5 minutes before the end, providing help to students was more important.

Ten respondents gave detailed comments such as: "visual review of everyone's status helps you really check if students are understanding, without asking each of them

---

[1] Although these participants were "trainee" teachers, the session took place towards the end of their training year. PGCE students at the Institute of Education spend 26 weeks of their training year teaching and developing contrasting experiences. Our evaluation participants, therefore, did have experience teaching different student cohorts, working with different maths departments, and were mentored by experienced maths teachers. None the less, we acknowledge that, had it been feasible to conduct a similar evaluation with a group of more experienced teachers, this may have resulted in additional feedback and suggestions.

[2] This session involved 15 students who interacted with the eXpresser in the context of a Mathematics lesson for roughly one hour. All occurrences of indicators generated during this lesson had been stored in the MiGen database: approximately 1,000 indicator occurrences had been generated by the 15 students working with eXpresser over the one-hour lesson.

individually", "allows much faster response to students' queries", "clearly highlights who is active and who isn't", "also shows what they have been doing which helps as sometimes, although the student is active, they may not be on task". Two responses went beyond using the CD tool and recognised that "GA also shows levels of progress" and can be used also for deciding which students need help.

With regards to finding out which students need immediate help without using the TA tools, 14 participants answered that they would use a traditional classroom solution (see Table 1 for examples). Eight felt that they would need innovative solutions like a remote control that would allow students to 'call the teacher'.

US2. Finding out which students are progressing satisfactorily towards completing the task and which ones may be in difficulty

For this usage scenario, for the first time point 19 participants referred to using the GA tool, e.g. "students who have yet to achieve any goals are in difficulty", or a combination of the GA tool with the CD and ST tools. Another five mentioned only the CD tool as a means of finding out at a glance which students are progressing. For the second time point, most referred to their previous answer or did not provide an answer. Two said that they would use the CD tool, and one commented that it would be useful to see this kind of information on a per task basis.

Without using the TA tools, the answers of 23 participants were similar to the previous US, e.g. "periodically ask whole class re. stage of progress, level of understanding, plus constantly circulate to observe them at work and assist as required". Two responses contained a comment about the effort that this would require.

US3: Finding out which students are currently disengaged from the task

All 24 participants provided an answer demonstrating an appreciation of the CD tool, with answers such as "I would be most concerned about amber students who had yet to achieve any tasks". Again, the difference between the two time points related to the level of the teacher's intervention. Some participants recognised that towards the end of the lesson students could be disengaged because they had finished the task, and that one approach could be to get "those who have finished to help those who are struggling".

Without use of the TA tools, 22 participants answered that they would use again some traditional classroom solution. Three commented on the difficulty of doing this and five provided elaborate comments demonstrating an appreciation also of the transformative nature of the information provided by the TA tools, e.g. "The tools provide a way of helping to increase the efficiency of my role as a teacher. […] I can readily identify those pupils that appear to be disengaged and subsequently target them for additional support/ encouragement".

Usage Scenarios 4–6

The answers to the questions relating to these USs were in line with what we expected and further demonstrated participants' appreciation of the information provided by the TA tools and the effort that it would take to achieve a similar level of awareness without them. The difference between the two time points was the level of teachers' intervention: at 30 minutes, their answers revolved around making sure that students were progressing, whereas at 5 minutes before the end of the lesson they wanted to ensure that students had achieved important objectives and to wrap up the lesson. For US 5 (finding out which students have finished the task) we asked participants to give examples of students who had finished the task. 87.5% of them (21 participants) responded with correct answers for both time points, confirming the intuitiveness of the information presented by the tools.

Usage Scenarios 7-8

These USs are more open-ended, so we asked participants to answer questions relating to them only in the end-of-session questionnaire. Here all 14 participants who provided answers appreciated that the different tools can be combined in order to allow teachers to provide support to students both during and after the lesson. Five participants stated that they would look back at the TA tools after the lesson to check which task goals were being accomplished and which were problematic; this would allow them to find out where and how students were struggling. Comments included that having access to the TA tools "could be very helpful after the lesson to assess progress and decide which pupils need more support next lesson, and which pupils need stretching further".

More generally, in relation to our goal to test the full suite of prototype TA tools, we did not identify any major technical or user interface issue. One overarching theme was the lack of adaptability in certain aspects that may be teacher- or school-specific. Examples include being able to configure the length of the time period that would cause a circle to be coloured amber, and the set of indicators that a teacher finds personally important (functionalities that were indeed incorporated into the final version of the tools).

A more critical issue that some participants raised relates to the validity and reliability of the information displayed by the tools, e.g. an amber circle in the CD tool may be showing a slow thinker, or students may be finding ways to "game the system" to avoid an amber indicator. These are valid challenges and raise interesting future areas of work but also helped us identify some subtle points to draw teachers' attention to in terms of the role of the tools. Our view is that the disengagement information provided by the CD tool is a first sign for the teacher to approach such students to find out who is actually disengaged from the task; or to use this information to choose which students' models to observe in detail through the tool.

Usage	General remarks	30 minutes into	5 minutes to end	Without TA tools
US1. Immediate help	Almost all participants said that they would use the CD tool and look for students whose circles were coloured red (23 people or 96%). Two responses recognised that in addition to the CD tool the "GA [tool] shows levels of progress" and can also be used for deciding which students need help.	Responses clustered around "keeping an eye" on those students who did not seem to have had a good start (45%) and ensuring that those "marked with amber are staying focused" (51%)	Responses focused on helping students e.g. "visual review of everyone's status helps you really check if students are understanding, without asking each of them individually"	Would use a traditional classroom solution (59%), e.g. "walk around the classroom viewing students' work and helping any student who may be off task or needing help", or "'students would have to put their hand up for my immediate attention". Other responses (33%) referred to more innovative solutions like a manual 'traffic light' system.
US2. Satisfactory progress or in difficulty	Participants appreciated that a combination of GA tool with the CD and ST tools provides this information. The CD tool in particular seemed more relevant as a means of finding out in a glance which students are progressing.	79% would use the GA tool as "students who have yet to achieve any goals are in difficulty"; 20% would glance the CD tool to find out state of class.	Referred to US1 (or did not provide an answer).	Participants referred to the effort it would take otherwise to take approaches like "periodically ask whole class re. stage of progress, level of understanding, plus constantly circulate to observe them at work and assist as required"
US3 Finding out disengaged students	Participants referred to the need for configuring the length of the time period that would cause a circle to be coloured Amber and advocated a more nuanced definition of disengagement.	100% demonstrated an appreciation of the CD tool (e.g. "I would be most concerned about amber students who had yet to achieve any tasks").	Some participants recognised that disengagement towards the end of the lesson could be because of task completion.	Similar answers to US1 and US2 and further comments on the difficulty of doing this without the TA tools.
US4. Common conceptual and procedural difficulties	White cells show lack of achievement of task goals (37%); ST tool can be used but after class (17%); CD tool (but the 2 responders did not provide any explanation); Most responses praised the time saving potential of the tools and could also see using it as formative assessment tool	Answers revolved around making sure that students were progressing.	Here participants were concerned with ensuring that students have achieved important objectives and with wrapping up the lesson.	Walking around the class and observing or asking students of different abilities to explain what they had done (63%). Initiate a whole-class discussion, referring to critical observations they would have made while walking around the class (8%). Formative assessment to ascertain the level of the class (8%)
US5 & US6. Task completion and goal accomplishments	Participants commented explicitly on the ease with which they could check which students have finished the task (21%). or provided additional ways on taking advantage of the GA tool e.g. identify difficult tasks to modify or choose students to give extension tasks (13%).	Most participants (88%) answered the questionnaires correctly by either referring to the annotations of the number of goals on the CD tool (61%) or to the GA tool (27%).	One participant said that close to the end of the lesson they would consider a whole-class display of GA information to encourage the students that that are behind to catch up.	All respondents acknowledged that they would have to revert to a traditional approach whereby they would ask students who have finished the task to raise their hands.
US7 & US8. Support, guidance; reflection.	For these USs, participants were asked to only answer end-of-session questionnaires. Their answers showed an appreciation of the functionalities of the TA tools and how they can be combined to allow teachers to provide support to students who need immediate help during the lesson, and that for those students who had not received enough help it would allow the teacher to plan to talk to them in the next lesson. Relevant comments: "could be very helpful after the lesson to assess progress and decide which pupils need more support next lesson, and which pupils need stretching further", "useful to identify the most common misconceptions which could then be consolidated in the following lesson".			

Table 1: Summary of participants' answers on evaluation questionnaires per usage scenario

7. DISCUSSION AND CONCLUSION

In this paper, we have described a suite of Teacher Assistance (TA) tools for an exploratory learning context, and their design and formative evaluation with a large number of teachers. We began the paper by highlighting the differences of Exploratory Learning Environments compared with Intelligent Tutoring Systems that are designed mostly for well-structured tasks and procedural learning. We argue that there is value in focusing on ELEs because, from a pedagogical point of view, they provide complementary ways of engaging students and supporting their understanding of difficult concepts. Due to their complexity, ELEs are under-researched and under-used in classrooms, in part because of the additional orchestration challenges (Dillenbourgh, 2013) they pose compared to other digital learning environments (Mavrikis, 2013a).

We advocate that carefully designed learning analytics presented through TA tools can support teachers in their challenging role as facilitators in this context. We have engaged in iterative cycles of co-design with teachers to derive a set of usage scenarios and corresponding tools. The evaluation results show that, in general, teacher participants exposed to the TA tools understand their capabilities and are able to use them quickly and with little training. Teachers perceive the tools as useful for raising their awareness of the classroom status overall and of individual students' progress and needs. Although we were aware of the complexity of the Student Tracking tool, we expected it to be used more for some of the usage scenarios co-designed with teachers, namely for intervention planning and class management (c.f. the framework of Prieto et al., 2015). The teachers' relatively infrequent use of the ST tool, contrasted with their frequent use of the CD and GA tools, point to their appreciation of relatively simple tools and the need for more training to achieve proficiency in using more complex tools. Both our classroom observations and the teachers' reflections confirm the usefulness of simple tools with actionable insights in closing the "regulation loop" (c.f. Dillenbourg et al., 2011).

We reflect below on the lessons learnt from our work and offer a set of design guidelines for Learning Analytics for ELEs, and implications for design and evaluation methodology.

7.1 Design guidelines for Learning Analytics for ELEs

We have shared in this paper our usage scenarios, designs and research findings in the hope that they will inspire others to design similar tools, thus moving away from the design of dashboards that simply show performance on skills or just descriptive statistics, e.g. the number of activities a student has undertaken. Our aim is that TA tools appropriately designed for orchestrating learning with digital technologies will also encourage teachers to adopt constructionist pedagogies whenever they see the need.

Similarly to the need to design carefully the ELE itself so as to help students develop conceptual understanding – in our case algebraic ways of thinking (Mavrikis et al. 2012), we believe that carefully designed TA functionalities have the potential to 'augment' teacher practice and to enable the adoption of ELEs in the classroom. In addition to the usage scenarios that emerged from our work, we can identify the following guidelines that apply more broadly for designing learning analytics relating to orchestration challenges:

Support real-time, ‘at-a-glance’ decision making

A key design decision at early stages of the development of the TA tools was to focus our attention on high-level indicators (milestones and landmarks) of the interaction with the ELE. For example, US 1-3 reflect this goal. The formative evaluation confirmed our early design decisions and teachers’ appreciation of being able to make quick decisions in the classroom. The summative evaluation (Mavrikis et al., 2016) reaffirmed these findings.

More broadly, a recent review (Schwendimann et al., 2017) characterises ‘learning dashboards’ and puts emphasis on ‘at-a-glance’ decision making as a feature that is generally applicable, and we have found this a particularly strong need during exploratory learning activities. Returning to the theme we touched on in the Introduction relating to the criticism of AI applied in education putting more emphasis on instructional pedagogy, the work described here demonstrates that a different approach is possible. By utilising intelligent analysis of student interactions we provide key insights to teachers in line with the constructionist pedagogy underlying exploratory learning. More recent studies support this need for human-AI partnerships, particularly by Holstein et al. (2019) in what they refer to as ‘designing for complementarity’. Their focus has been on classroom use of intelligent tutoring system but the principal need behind the work is the same.

Enable configurability at different layers

Both during the design process and arising from the formative evaluation, the need for configuring the TA tools according to different preferences became clear. This view is reinforced by recent research in implementation of learning analytics, c.f. the principle of customisation (Wise & Vytasek, 2016), and the increased sense of agency in multimodal learning analytics customisation (Rodríguez-Triana et al., 2018). The open nature of exploratory learning activities makes this a particularly strong requirement. While our early versions of the TA tools allowed some level of configuration (e.g. in the placement of circles in the CD tool), the formative evaluation brought out clearly these needs. As such, in later versions we extended the customisation with additional user interface features, e.g. to allow the teacher to choose which indicators they want to focus their attention on in the ST tool. In our classroom trials, the pedagogical strategies for supporting students directly through the ELE itself could also be configured, albeit in collaboration with a learning technologist at this stage and not by the user through a user interface. Therefore, more co-design and evaluation work is needed in this area.

Augment constructionist pedagogical strategies

While the previous guidelines may be applicable and important in other contexts beyond exploratory learning, the successful involvement of teachers early in the design highlights the importance of taking into account the specific context for which learning analytics solutions are being designed (c.f. Wise & Vytasek, 2016). Many of our usage scenarios may be applicable in other blended learning scenarios, e.g. with intelligent tutoring systems such as the ones described in du Boulay (this issue) and Holstein et al. (2017). However, operationalising such tools in the context of exploratory learning brings out the requirement of utilising intelligent analysis of students' interactions and designing visualisations for action. This has the added benefit of enabling constructionist strategies that would otherwise be very consuming or not possible at all.

For example, a common constructionist teaching strategy involves grouping students together during or after a task, for students to work collaboratively or to help each other or, as in our case, to reflect on their constructions. Our Grouping Tool is a case in point: it simplifies an otherwise prohibitively time-consuming process for the teacher and enables meaningful pairings based on intelligent analysis of students' constructions. The formative evaluation helped us to see the Classroom Dynamics tool being unexpectedly appropriated by teachers for getting students who had finished all the task goals to help other students. Future work could combine the functionalities of the Grouping Tool and the CD tool to automatically make such recommendations to the teacher, though again co-design cycles would be required to ensure that the tool's recommendations are accurate, useful and trusted by teachers. Other strategies in constructionist contexts include extending tasks, and holding meaningful plenaries based on reflection and awareness of students' interactions (Foster, 2013). The Student Tracking tool, despite the challenges of using it real-time in the classroom, provides the opportunity for teachers to reflect and take planning decisions. Future work here would again need to extend this tool to provide recommendations to the teacher, e.g. of students who can be set extension tasks or who can be formed into plenary groups for discussion.

7.2 Design and evaluation methodology

Reflecting on our design and evaluation methodology, we noted in Section 3 how, retrospectively, it is similar to the LATUX methodology (Martinez-Maldonado et al., 2015) proposed to address the need to take into account teachers' and pedagogical requirements in learning analytics design through a series of iterative evaluation stages. Despite the field recognising this need, teacher involvement often tends to come rather late in the design process after the design agenda has been set (c.f. Prieto-Alvarez et al. 2017). LATUX, in particular, does not provide *methodological guidance* on how to support teachers' involvement in the early generative design phase – for example, how to account for teachers' lacking experience in a given context. Only recently have techniques from Human Computer Interaction started to be explored in detail for this purpose. For example, we referred earlier to Holstein et al. (2018) adopting a participatory design approach to design a dashboard that answers real-time teachers' needs in the context of an intelligent tutoring system. We argue that if AI-driven Teacher Assistance tools are to effectively support exploratory learning, a participatory design approach from the very early stages of the process is mandatory.

Our approaches, starting from the identification of important landmarks and milestones in students' interactions with the emerging ELE by conducting early Wizard-Of-Oz studies, subsequently co-designing prototype TA tools and Usage Scenarios with teachers, and then continuing with the evaluation methods presented here, are relevant to the design of similar tools for teachers in general and exploratory learning settings in particular. To design the TA tools, we started with collaborative identification of the indicators relating to students' interactions that are useful and meaningful for teachers, continued with developing computational techniques to detect or infer each kind of indicator, and then iteratively co-designed and evaluated visualisations of this information within a set of targeted tools to assess whether they fulfil specific intended purposes in the context of the USs. This methodology has worked successfully in our context and we argue that it is transferrable to other exploratory learning settings.

The formative evaluation described here gave us key insights from a large set of teacher participants. We extended and replicated our formative evaluation methods for the summative evaluation of the TA tools, which we did not report on here due to space limitations. In brief, the summative evaluation comprised two parts: (i) a more detailed trial with one of our teacher collaborators at her school, comprising two lessons, one in which the teacher had access to the TA tools installed on a tablet computer, and the second in which she did not and had to support her students without having access to the tools; and (ii) a lab-based trial with another cohort of PGCE students following a similar methodological approach to that described in Section 6. As discussed in Mavrikis et al. (2016), this summative evaluation resembles the 'validation in the wild' stage of the LATUX workflow, but goes beyond that by obtaining input from a large number of teacher participants. We refer the reader to that paper for a detailed discussion and outcomes, which generally reaffirmed the usability and usefulness of the tools in meeting the USs. Designing and evaluating such tools is classroom is a challenging task. This methodology of reusing real log data from students' interactions to engage a larger set of participants in prototyping and evaluations than would realistically be possible given teacher and research fundings constraints, is emerging as a viable method (Holstein, 2019). In contrast, most evaluation studies in the field have a small number of teacher participants (c.f. Schwendimann et al., 2017). A lab-based method for obtaining input from a large number of participants does not entirely reproduce the conditions of a real classroom of course, where the teacher may know the students and has to manage full classroom orchestration throughout the lesson. So wider classroom deployment remains necessary to gain further validation and insights.

7.3 Concluding remarks and future work

The TA tools presented in this paper are general in their architectural design. By following the participatory iterative methodology laid out here, similar tools could be designed to monitor the activities of students interacting with other digital environments that follow constructionist learning approaches provided that they are able to detect the interaction indicators that are considered meaningful for that setting. Details that would allow others to implement similar tools are given in Pearce-Lazard et al. (2010), Gutierrez Santos et al. (2012), and Noss et al. (2012) for the initial Java and REST-based implementation. Gutierrez Santos et al. (2016) and Mavrikis & Karkalas (2017) expand this work with subsequent cloud-based versions.

Having received more feedback on the TA tools since they were made available, and through our latest work with teachers, anecdotally we see that teachers start requiring such tools as a prerequisite for considering adopting new technologies in classroom. At the same time, more and more platforms are providing a form of learning analytics (c.f. Mangaroska, Giannakos, 2018) This is promising as Holstein at al. (2018) demonstrate that students reap the benefits of the increased teacher awareness in such cases. In addition, with the increased emphasis on evidence-based teaching, such tools empower teachers to provide evidence of learning and to engage in their own inquiry (c.f. Emin-Martnez et al., 2014; Mor et al., 2015) even in a context such as constructionist learning that is less subject to formal assessment. and measurement. Our future work therefore includes investigating how such TA tools could be extended to support teachers in even more complex learning scenarios, e.g. blended or collaborative learning, as well as identifying the training needs and continuous professional development opportunities to embed such tools in teachers' practice (Rodríguez-Triana et al., 2018; Wise & Vytasek, 2016).

To conclude, we have argued here that by taking into account teachers' requirements early on and throughout the design process, a constructionist pedagogy can be supported by Teacher Assistance tools driven by artificial intelligence. Thanks to teachers' early involvement and trials, we quickly moved away from conceptualising AI components that only automate the support that students require in this setting, to a hybrid approach where analysis of students' interactions is used both to provide feedback to students directly whenever possible and for driving the TA tool visualisations. This is consistent with the emerging view of AI as augmenting or assisting human intelligence (Baker, 2016; Gorban, 2018; Pardo et al., 2019; Holstein et al., in press).

Acknowledgements

The authors thank the other members of the original MiGen team for our stimulating collaborative research on the project, and especially the team of teacher collaborators for their insightful feedback. The MiGen project was funded by the ESRC/EPSRC TEL programme (RES-139-25-0381). We would also like to thank the anonymous reviewers and Ken Holstein for comments on earlier versions of this paper, as well as Yvonne Vezzoli and Mina Vasalou from the iRead project that has provided a new context for reflection on these issues. The iRead project has received funding from the European Union's Horizon 2020 innovation programme under grant agreement No 731724.

Statements on ethics and conflict of interest

The formative evaluation presented here and the overall project research had research ethics approval from the Institute of Education committee following BERA guidelines. The student data used for the teacher prototype evaluation was with non-identifiable usage data from previous usage of the system and is not available due to the nature of the consent form at the time. The questionnaire and interview data from teachers is available on request. The authors have no conflicts of interest to declare in relation to this work.

REFERENCES

Amershi, S. and Conati, C. (2009). Combining Unsupervised and Supervised Classification to Build User Models for Exploratory Learning Environments. *Journal of Educational Data Mining*, 1(1), pp.18–71.

Amir, O. and Gal, K. (2013). Plan recognition and visualization in exploratory learning environments. *ACM Transactions on Interactive Intelligent Systems* (TiiS), 3(3), 16.

Baker, R. S. (2016) Stupid tutoring systems, intelligent humans. *International Journal of Artificial Intelligence in Education*, 26 (2), pp. 600-614. https://doi.org/10.1007/s40593-016-0105-0

Ben-Naim, D., Marcus, N., and Bain, M. (2008). Visualization and analysis of student interactions in an exploratory learning environment. In *Proc. of the 1st Int. Workshop on Intelligent Support for Exploratory Environments*, at EC-TEL 2008.

Dillenbourg, P., Zufferey, G., Alavi, H. S., Jermann, P., Do, L. H. S., Bonnard, Q., Cuendet, S., Kaplan, F. (2011). Classroom orchestration: The third circle of usability. In *Connecting Computer-Supported Collaborative Learning to Policy and Practice: CSCL 2011 Conference Proceedings*. Volume I — Long Papers, International Society of the Learning Sciences, 1, pp. 510–517.

Dillenbourg, P. (2013). Design for classroom orchestration. *Computers & Education*, 69, pp. 485 – 492.

Dragon, T., Mavrikis, M., McLaren, B., Harrer, A., Kynigos, C., Wegerif, R., and Yang, Y. (2013). Metafora: A web-based platform for learning to learn together in science and mathematics. *IEEE TLT*, 6(3), pp. 197-207.

Du Boulay, B. (2019). Escape from the Skinner Box: The case for contemporary intelligent learning environments. *British Journal of Educational Technology*. 10.1111/bjet.12860.

Emin-Martnez, V., Hansen, C., Rodrguez-Triana, M. J., Wasson, B., Mor, Y., Dascalu, M., Ferguson, R., and Pernin, J.-P. (2014). Towards teacher-led design inquiry of learning. eLearning Papers.

Foster C. (2013) *The Essential Guide to Secondary Mathematics*: *Successful and enjoyable teaching and learning*. London: Routledge.

Garcia, R. C., Pardo, A., Kloos, C. D., Niemann, K., Scheffel, M., and Wolpers, M. (2012). Peeking into the black box: visualising learning activities. *International Journal on Technology Enhanced Learning*, 4(1/2), pp. 99–120.

Gorban, A. N., Grechuk, B., and Tyukin, I. Y. (2018). Augmented Artificial Intelligence: a Conceptual Framework. *ArXiv:1802.02172 [Cs]*. Retrieved from http://arxiv.org/abs/1802.02172

Grawemeyer, B., Mavrikis, M., Holmes, W., Gutiérrez-Santos, S., Wiedmann, M., & Rummel, N. (2017). Affective learning: Improving engagement and enhancing learning with affect-aware feedback. *User Modeling and User-Adapted Interaction*, 27(1), pp. 119–158. https://doi.org/10.1007/s11257-017-9188-z

Gueraud, V., Lejeune, A., Adam, J.-M., Dubois, M., and Mandran, N. (2009). Supervising Distant Simulation-Based Practical Work: Environment and Experimentation. In *Proceedings of ECTEL 2009*, pp. 602–608.

Gutierrez Rojas, I., Crespo Garcıa, R., and Delgado Kloos, C. (2012). Enhancing Orchestration of Lab Sessions by Means of Awareness Mechanisms. In *Proceedings EC-TEL 2012*, pp. 113–125.

Gutierrez Santos, S., Geraniou, E., Pearce-Lazard, D., and Poulovassilis, A. (2012). Design of Teacher Assistance Tools in an Exploratory Learning Environment for Algebraic Generalization. *IEEE Transactions on Learning Technologies*, 5(4), pp. 366–376.

Gutierrez Santos, S., Mavrikis, M., and Magoulas, G. D. (2012b). A Separation of Concerns for Engineering Intelligent Support for Exploratory Learning Environments. *Journal of Research and Practice in Information Technology*, 44, pp. 347–360.

Gutierrez Santos, S., Mavrikis, M., Geraniou, E., and Poulovassilis, A. (2017). Similarity-based grouping to support teachers on collaborative activities in exploratory learning environments. *IEEE Transactions on Emerging Topics in Computing*, 5(1), pp. 56-68

Healy, L. and Kynigos, C. (2010). Charting the microworld territory over time: design and construction in mathematics education. *ZDM*, 42(1):63–76.

Holstein, K., McLaren, B.M., and Aleven, V. (2017) Intelligent tutors as teachers' aides: exploring teacher needs for real-time analytics in blended classrooms. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference, ACM*, pp. 257–266. https://doi.org/10.1145/3027385.3027451

Holstein, K., McLaren, B. M., & Aleven, V. (2018). Student learning benefits of a mixed-reality teacher awareness tool in AI-enhanced classrooms. In *Proceedings of the 19th International Conference on Artificial Intelligence in Education (AIED 2018).* LNAI 10947, pp. 154-168, Springer: Berlin.

Holstein, K., McLaren, B.M., and Aleven, V. (in press) Co-designing a Real-time Classroom Orchestration Tool to Support Teacher–AI Complementarity. *The Journal of Learning Analytics*.

Holtzblatt, K., & Beyer, H. (1997). Contextual Design: Using Customer Work Models to Drive Systems Design. *CHI '97 Extended Abstracts on Human Factors in Computing Systems*, pp. 184–185. http://doi.acm.org/10.1145/1120212.1120334

Hoyles, C., Noss, R., and Kent, P. (2004). On the Integration of Digital Technologies into Mathematics Classrooms. *International Journal for Computers in Mathematical Learning*, 9(3), pp. 309–326.

Kirscher, P., Sweller, J., and Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiental and inquiry-based learning. *Educational Psychologist*, 41(2), pp. 75–86.

Kynigos, C. (1992). Insights into Pupils' and Teachers' Activities in Pupil-Controlled Problem-Solving Situations. In *Information Technology and Mathematics Problem Solving*, volume 2, pp. 219–238.

Martinez-Maldonado, R., Shum, S. B., Schneider, B., Charleer, S., Klerkx, J., and Duval, E. (2017). Learning Analytics for Natural User Interfaces. *Journal of Learning Analytics,* 4(1)**,** pp 24-57.

Martinez-Maldonado, R., Pardo, A., Mirriahi, N., Yacef, K., Kay, J., and Clayphan, A. (2015). The LATUX workflow: Designing and deploying awareness tools in technology-enabled learning settings. In Proceedings of the Fifth International Conference on *Learning Analytics And Knowledge*, pp. 1–10.

Martinez-Maldonado, R., Clayphan, A., Yacef, K. & Kay, J. (2015b). MTFeedback: Providing Notifications to Enhance Teacher Awareness of Small Group Work in the Classroom. *IEEE Transactions on Learning Technologies,* 8(2)**,** pp. 187-200.

Martinez-Maldonado, R. (2016). Seeing learning analytics tools as orchestration technologies: Towards supporting learning activities across physical and digital spaces. In *Proceedings of the First International Workshop on Learning Analytics Across Physical and Digital Spaces co-located with 6th International Conference on Learning Analytics & Knowledge* (LAK 2016), pp. 70 – 73.

Mavrikis, M., Gutierrez Santos, S., and Poulovassilis, A., 2016. Design and Evaluation of Teacher Assistance Tools for Exploratory Learning Environments, In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge,* (LAK 2016). ACM, New York, NY, USA, pp. 168–172.

Mavrikis, M., Gutierrez Santos, S. Geraniou, E., Hoyles, C., Magoulas, G., Noss R. and Poulovassilis, A. (2013). Iterative context engineering to inform the design of intelligent exploratory learning environments for the classroom in al., Luckin, R. (ed). *Handbook of Design in Educational Technology*, Routledge, pp. 80-92.

Mavrikis, M., Karkalas, S., 2017. Reflective Analytics for Interactive e-books. *IxD&A 33*, pp. 33–53.

Mavrikis, M., Noss, R., Hoyles, C., and Geraniou, E. (2012). Sowing the seeds of algebraic generalization: designing epistemic affordances for an intelligent microworld. Special Issue on Knowledge Transformation, Design and Technology. *Journal of Computer Assisted Learning*, 29(1) pp. 68-84.

Mangaroska, K., and Giannakos, M.N. (2018). Learning Analytics for Learning Design: A Systematic Literature Review of Analytics-Driven Design to Enhance Learning. *IEEE Transactions on Learning Technologies*, https://doi.org/10.1109/TLT.2018.2868673.

Mayer, R. E. (2004). Should There Be a Three-Strikes Rule Against Pure Discovery Learning? - The Case for Guided Methods of Instruction. *American Psychologist*, 59(1), pp. 14–19.

Mor, Y., Ferguson, R., and Wasson, B. (2015). Editorial: Learning design, teacher inquiry into student learning and learning analytics: a call for action. *British Journal of Educational Technology*, 46(2), pp. 221–229.

Noss, R. and Hoyles, C. (1996). *Windows on Mathematical Meanings*. Kluwer Academic Publishers, The Netherlands.

Noss, R., Poulovassilis, A., Geraniou, E., Gutierrez Santos, S., Hoyles, C., Kahn, K., Magoulas, G. D., and Mavrikis, M. (2012). The design of a system to support exploratory learning of algebraic generalisation. *Computers & Education*, 59(1), pp. 63–81.

Papert, S., & Harel, I. (1991). Situating constructionism. In S. Papert & I. Harel (Eds.), *Constructionism*. New York: Ablex Publishing.

Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., and Mirriahi, N. (2019) Using learning analytics to scale the provision of personalised feedback. *British Journal of Educational Technology*, 50(1), pp. 128-138. https://doi.org/10.1111/bjet.12592

Pearce-Lazard, D., Poulovassilis, A., and Geraniou, E. (2010). The Design of Teacher Assistance Tools in an Exploratory Learning Environment for Mathematics Generalisation. In *Proceedings EC-TEL 2010*, Springer LNCS 6383, pp. 260–275.

Prieto-Alvarez, C.G., Martinez-Maldonado, R. and Anderson, T. (2017). Co-designing in learning analytics: tools and techniques. In J.C.H. Lodge, J.C.H., L. Corrin ed. *Learning analytics in the classroom: translating learning analytics research for teachers*, Routledge.

Prieto, L. P., Dimitriadis, Y., Asensio-Pérez, J. I. and Looi, C.-K. (2015). Orchestration in Learning Technology Research: Evaluation of a Conceptual Framework. *Research in Learning Technology,* 23(1) [1]

Rodríguez-Triana, M.J., Prieto, L.P., Martínez-Monés, A., Asensio-Pérez, J.I., Dimitriadis, Y. (2018). “The Teacher in the Loop: Customizing Multimodal Learning Analytics for Blended Learning.” In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (LAK 2018), pp. 417–426. New York, NY, USA: ACM. https://doi.org/10.1145/3170358.3170364.

Roll, I., Aleven, V., Koedinger, K. R. (2010). The invention lab: Using a hybrid of model tracing and constraint-based modeling to offer intelligent support in inquiry environments. In V. Aleven, J. Kay, & J. Mostow (Eds.), *Proceedings of the International Conference on Intelligent Tutoring systems,* pp. 115-24. Berlin: Springer Verlag.

Rummel, N., Mavrikis, M., Wiedmann, M., Loibl, K., Mazziotti, C., Holmes, W., & Hansen, A. (2016). Combining Exploratory Learning with Structured Practice to Foster Conceptual and Procedural Fractions Knowledge. *ICLS*, 12, 58–65. Retrieved from http://discovery.ucl.ac.uk/id/eprint/1475916

Sergis, S., Sampson, D. G., Rodríguez-Triana, M. J., Gillet, D., Pelliccione, L., & de Jong, T. (2019). Using educational data from teaching and learning to inform teachers’ reflective educational design in inquiry-based STEM education. *Computers in Human Behavior*, *92*, 724–738. https://doi.org/10.1016/j.chb.2017.12.014

Schwendimann, B. A., Rodríguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., Gillet, D., and Dillenbourg, P. (2017). Perceiving Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research. *IEEE Transactions on Learning Technologies*, 10(1), pp. 30–41. https://doi.org/10.1109/TLT.2016.2599522

Trouche, L. (2004). Managing the Complexity of Human/Machine Interactions in Computerized Learning Environments: Guiding Students’ Command Process through Instrumental Orchestrations. *International Journal of Computers for Mathematical Learning*, 9(3), pp. 281–307.

Van Leeuwen, A. (2015), Learning analytics to support teachers during synchronous CSCL: Balancing between overview and overload. *Journal of Learning Analytics*, 2(2), pp. 138–162.

Verbert, K., Duval, E., Klerkx, J., Govaerts, S., and Santos, J.L., 2013. Learning Analytics Dashboard Applications. *American Behavioral Scientist* 57, pp. 1500–1509. https://doi.org/10.1177/0002764213479363

Voyiatzaki, E., Polyzos, P., and Avouris, N. (2008). Teacher tools in a networked learning classroom: monitor, view and interpret interaction data. In *Proceedings 6th International Conference on Networked Learning*.

Wise, A., and Vytasek, J. (2016). Learning analytics implementation design. In G. Siemens, Lang, C. (Eds.), *Handbook of Learning Analytics*, pp. 151-159, DOI: 10.18608/hla17.013

Wichmann, A., Giemza, A., Hoppe, U. and Krauß, M. (2009). Effects of awareness support on moderating multiple parallel E-discussions. In *Proceedings of the 8th International Conference on Computer Supported Collaborative Learning,* CSCL '09*,* Volume 1, pp. 646-650. https://doi.org/10.3115/1600053.1600146.

Wilson, C., and Scott, B. (2017). Adaptive systems in education: a review and conceptual unification. *The International Journal of Information and Learning Technology*, 34(1), pp. 2-19. doi:http://dx.doi.org/10.1108/IJILT-09-2016-0040

Zaldivar, V. A. R., Pardo, A., Burgos, D., and Kloos, C. D. (2012). Monitoring student progress using virtual appliances: A case study. Computers & Education, 58(4), pp. 1058–1067.

APPENDICES

Appendix 1. Glossary of terms

Term	Meaning
CD	Classroom Dynamics (tool)
GA	Goal Achievements (tool)
ELE	Exploratory Learning Environment
eGeneraliser	The intelligent component that infers TD indicators, and generalised personalised feedback for students
eXpresser	The mathematical microworld of the MiGen system
MiGen	The overall system designed to support 11-14 year old students' development of algebraic generalisation
ST	Student Tracking (tool)
TA	Teacher Assistance (tools)
TD / TI	Task-dependent / task-independent (indicator)
US	Usage Scenario

Appendix 2. Usage Scenarios and how they can be met by the TA tools

1	*Finding out which students need the teacher's immediate help.*
	The teacher can consult the CD tool and see which students' circles are coloured Red and how many of the task goals they have achieved. The teacher can click on a student's circle to view their current model and rule, to provide some context for the help to be given to the student. The teacher can also open up the ST tool to view the recent indicators relating to the student's actions, to provide additional context. If there is more than one student coloured Red in the CD display, the teacher may select to help first students who have achieved the fewest task goals.
2	*Finding out which students are progressing satisfactorily towards completing the task and which students may be in difficulty.*
	Similarly to US1, the teacher can consult the CD tool to see students in need of help, how many task goals students have achieved, and students' current models and rules; and the ST tool to view students' recent indicators. The teacher can open up the GA tool to view specifically which task goals are being achieved by each student.
3	*Finding out which students are currently disengaged from the task.*
	The teacher can consult the CD tool and see which students' circles are currently coloured Amber. If any of these students has not completed all the task goals, then she/he is likely to be currently disengaged from the task and in need of encouragement from the teacher. If a student has completed all the task goals, then she/he may need to be set additional goals or a new task to work on while waiting for the rest of the class to finish.
4	*Identifying common conceptual and procedural difficulties students are facing in order to provide more explanation to the class as a whole.*
	Consulting the GA tool allows the teacher to see which task goals students are having difficulty completing, so as to inform additional explanation to the class. Consulting the ST tool allows the teacher to see if there are specific Red indicators showing in many of the students' columns, indicating particular procedural difficulties that students may be facing and again informing the provision of additional explanation to the class.
5	*Finding out which students have finished the task.*
	The CD tool can be used to see which students have achieved all the task goals. For these students, the teacher can click on their circles to view their final model and rule, to check if they have achieved a correct solution. The teacher can then go to each student to set them a new task or additional goals relating to the current task, if their solution was correct; or ask them to reflect further on their construction if not.
6	*Finding out which students have achieved which task goals.*
	The GA tool can be used to see which students have achieved which of the task goals.
7	*Providing appropriate support and guidance to individual students (i) during the lesson, and (ii) after the lesson.*
	This can be undertaken during the lesson using a combination of the tools as described for US1, US2, US3 above, and after the lesson by using the GA tool to see which task goals an individual student has not managed to achieve, the CD tool to view the student's final model and rule as produced by the end of the lesson, and the ST tool to view the student's detailed history of interactions during the lesson.
8	*Reflecting on the class' achievements and planning the next lesson.*
	This can be undertaken using a combination of the GA tool, to see which task goals have been largely achieved by the class, the CD tool to view selected students' models and rules, and the ST tool to see a historical record of how students tackled the task during the lesson.
9	*Pairing students for productive discussion of their solution approaches at the end of the task*
	This can be undertaken straightforwardly using the Grouping Tool – the teacher simply requests for a full set of suggested parings relating to the whole class and waits a few seconds for the tool to respond with its recommendations. Without using the Grouping Tool (which was not available for the formative and summative evaluations described here) it would require the teacher loading each student's construction, and a best-effort attempt to put together pairs of students who have taken different construction approaches.

Appendix 3. Evaluation Instruments

Formative Evaluation in the Lab

Below are the questions from the questionnaire form that participants were asked to complete relating to information displayed by the TA tools using the time-stop functionality: (i) 30 minutes into the lesson, and (ii) 5 minutes before the end of the lesson. For each question or task, participants were asked to state how they would answer it using the TA tools. They could also add additional comments, if any:

1) Finding out which students need your immediate help.
2) Which students are progressing satisfactorily towards completing the task goals? Which students may be in difficulty?
3) Which students are currently disengaged from the task?
4) Would this be a time-point that you would give more explanation to the class as a whole? And if so what would you say based on information from the tools?
5) Which students have finished the task?
6) Which task goals have most students achieved?

At the end of the evaluation session, participants were asked to complete a second questionnaire comprising the following questions:

1) Finding out which students need your immediate help.
2) Finding out which students are progressing satisfactorily on an eXpresser activity and which ones may be in difficulty.
3) Finding out which students are disengaged during the activity.
4) Identifying common conceptual and procedural difficulties students are facing in order to provide more explanation to the class as a whole.
5) Providing appropriate support and guidance to individual students during the lesson.
6) Providing appropriate support and guidance to individual students after the lesson.
7) Finding out which students have finished the task.
8) Finding out which students have achieved specific task goals.
9) Pairing students for collaborative tasks.
10) Planning the next lesson using the MiGen system.
11) Planning the next lesson without using the MiGen system.

For each of questions 1)-11), participants were asked: (i) How would you find this out without using the TA tools; (ii) with the TA tools; (iii-a) if the tools helped you achieve this task, how did they do it; (iii-b) if they did not, what additional information would help you?

Questions 1-4 refer to the corresponding US; questions 5 and 6 to US 7(i) and (ii); question 7 to US 5; question 8 to US 6; questions 10 and 11 to US 8; and question 9 to the ninth usage scenario mentioned in the last paragraph of Section 4.