Personalized Visualizations to Promote Young Learners' SRL: The Learning Path App

Inge Molenaar†
Behavioural Science Institute
Radboud University
Nijmegen, The Netherlands
i.molenaar@pwo.ru.nl

Anne Horvers
Behavioural Science Institute
Radboud University
Nijmegen, The Netherlands
a.horvers@pwo.ru.nl

Rick Dijkstra
Behavioural Science Institute
Radboud University
Nijmegen, The Netherlands
r.dijkstra@pwo.ru.nl

Ryan S. Baker
University of Pennsylvania, United States
ryanshaunbaker@gmail.com

ABSTRACT

This paper describes the design and evaluation of personalized visualizations to support young learners' Self-Regulated Learning (SRL) in Adaptive Learning Technologies (ALTs). Our learning path app combines three Personalized Visualizations (PV) that are designed as an external reference to support learners' internal regulation process. The personalized visualizations are based on three pillars: grounding in SRL theory, the usage of trace data and the provision of clear actionable recommendations for learners to improve regulation. This quasi-experimental pre-posttest study finds that learners in the personalized visualization condition improved the regulation of their practice behavior, as indicated by higher accuracy and less complex moment-by-moment learning curves compared to learners in the control group. Learners in the PV condition showed better transfer on learning. Finally, students in the personalized visualizations condition were more likely to under-estimate instead of over-estimate their performance. Overall, these findings indicates that the personalized visualizations improved regulation of practice behavior, transfer of learning and changed the bias in relative monitoring accuracy.

CCS CONCEPTS

• CCS → Human-centered computing → Visualization → Visualization design and evaluation methods
•CCS → Applied computing → Education → E-learning

KEYWORDS

Self-Regulated Learning; Adaptive Learning Technologies; Learner-faced dashboards; Hybrid Human-System Intelligence

ACM Reference format:

Molenaar, I., Horvers, A. Dijkstra, R. Baker, R. 2020. Personalized Visualizations to Promote Young Learners' SRL: The Learning Path App. In *Proceedings of the 10th International Conference on Learning Analytics & Knowledge (LAK'20). ACM, New York, NY, USA, X pages.* https://doi.org/10.1145/1234567890

LAK '20, March 23–27, 2020, Frankfurt, Germany

ACM ISBN 978-1-4503-7712-6/20/03…$15.00
https://doi.org/10.1145/3375462.3375465

1 Introduction

In the field of learning analytics, it has often been suggested that providing learners with a tangible reference to their learning process is a powerful way to support Self-Regulated Learning (SRL)[31, 35]. For example, the Open Learner Model movement proposes to display learners' data to support their SRL [15, 25] and similar propositions have been made from an SRL perspective [31]. Yet few learner-faced dashboards have been developed with a strong foundation in SRL theory [20]. This study describes the design of personalized visualizations, i.e. learner-faced dashboard that are tuned to the needs of individual students, to support young learners' regulation and learning in Adaptive Learning Technologies (ALTs). These personalized visualizations are grounded in SRL theory and based on learners' trace data. The visualizations are derived using the moment-by-moment-learning curve (MbMLC) algorithm [8]. Previous research indicated that MbMLC provides valuable indicators of how learners regulate their learning over time and could potentially be used within personalized visualizations [26]. Moreover, integrating MbMLC with ALT data allowed us to determine learners' SRL support needs [24].

In this paper we describe a study into the effects of personalized visualizations on young learners' regulation and learning in the context of an Adaptive Learning Technology for primary education. We first relate this work to the other work on learner-faced dashboards to support SRL, then we outline the COPES theory on SRL as a theoretical basis for developing personalized visualizations, and finally we describe how the design of the personalized visualizations in the learning app provides an external mirror for students internal SRL process.

1.1 State of the art on learner-faced dashboards

Dashboards have been defined as: "*Single displays that aggregated different indicators about learners, learning processes and or learning contexts into one or multiple visualizations*" [32]. Research around dashboards traditionally has a strong focus on learning analytics and educational data and less attention is paid to the connection to learning theory [20]. Although SRL theory is the most common foundation for learner-faced dashboards [13], most

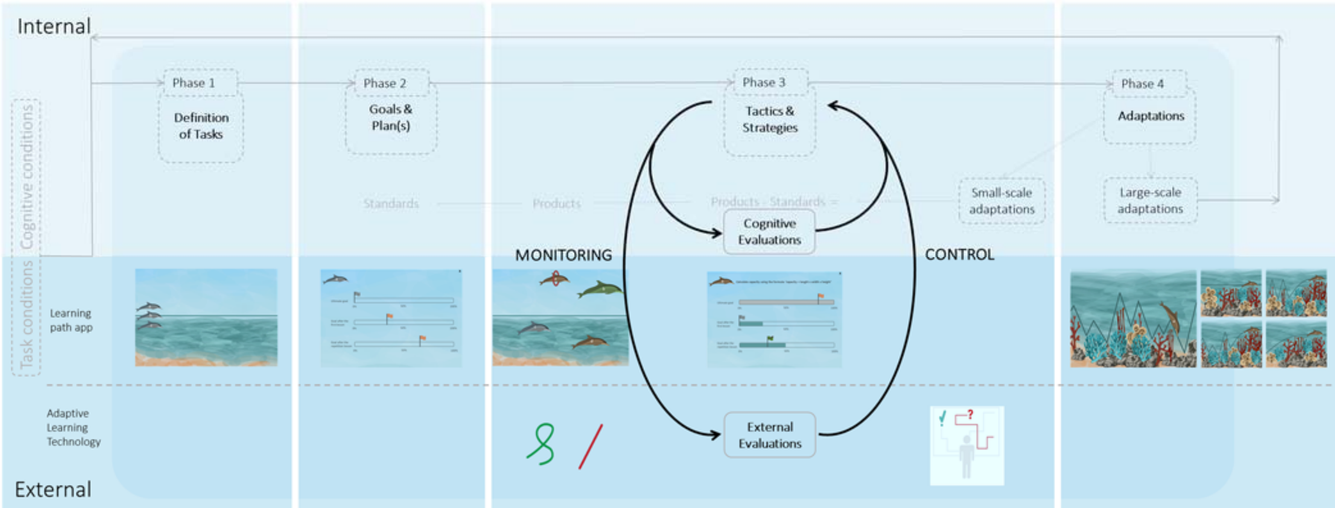


Figure 1: Personalized visualizations related to the COPES model.

of these dashboards visualize indicators of learner performance to support students' regulation [14]. Performance feedback alone is not always enough to help learners to translate progress data into actions that improve regulation[16]. Although there are some good examples of trace data informed progress charts [3] and intelligent tutor systems to support SRL [2, 6], there is limited work within learner-facing dashboards that use trace data to support SRL [3, 13]. Partially this lack of research may stem from challenges in understanding what trace data reveal about SRL [4, 11] as well as finding ways to visualize temporal and sequential characteristics of SRL in a meaningful way for learners [26, 27] . In this paper, we discuss our work to provide learners personalized visualizations to support learners' regulation. Below we elaborate on how SRL theory, specifically the COPES model, informed the design of the personalized visualizations used in the learning path app.

1.2 Designing personalized visualizations based on SRL theory

The COPES model describes the internal regulation processes that learners enact to regulate their learning [33]. The central assumption in this model is that learning is a goal-oriented process in which learners make conscious choices working toward their learning goals [25]. In order to reach these learning goals learners use metacognitive activities to control and monitor their learning and engage in appropriate levels of effort [7]. Regulation in the COPES model unfolds in four loosely coupled phases: i) in the task definition phase, learners develop an understanding of the task, ii) during the goal setting phase, learners set their goals and plan their learning, iii) in the enactment phase, learners execute their plans and control and monitor progress iv) in the adaptation phase, adjustments are made when progress towards the goals is not proceeding as planned. These phases are enacted in the context of task and learner conditions that drive operations, strategies and tactics performed by learners.

The control and monitoring loop are at the heart of COPES model and especially important in the context of ALTs [24]. Learners need to monitor and control their work in order to answer the problems correctly and make progress to their goals. Accuracy can be conceptualized as a function of a student's knowledge and effort [20]. Hence effective self-regulating learners in ALTs adjust their effort to ensure that their planned learning goals are achieved and control their learning to ensure a productive level of accuracy [31].

It is well established that learners often face a utilization deficiency [33], the failure to adequately activate control and monitoring during learning. The internal regulation process as described by the COPES model can be supported by the external regulation in technology. Dashboards are potentially a powerful tool to overcome the utilization deficiency as they can help learners by showing them objective data about their current performance, *how* this performance is related to their learning goals, and *what* progress they have made [19]. In the context of ALTs, personalized visualizations attuned to individual learners needs, can show learners the alignment between learning goals and their actions.. This form of external feedback can consequently drive cognitive evaluation and help learners to optimize their strategies, make adjustments to plans, or choose different actions to reach their learning goals

In designing the learning path app, we follow the four phases of the COPES model to support learners with external cues following internal regulation processes. Learner-faced dashboards function as a visual layer between the internal regulation of the learner and the external regulation support of the ALT. Their primary function is to support learners to explicitly engage in the four phases that are critical for successful self-regulation. As such, the different visualizations in the learning path app function as a reference for learners to better understand their own regulation process. In essence, the app is a mirror for learners to better monitor their progress and recognize the need for control actions and thus drive their internal regulation process, see Figure 1.

First, in the *task definition phase,* learners need to develop an understanding of the task to be able to formulate appropriate learning goals. In this phase, learners are supported by the *overview screen* which summarizes the tasks for the learner. Within the scope of our current study this entails the visualization of three subskills that need to be mastered by the learners. The dolphins each represent one subskill.

Second, in the *goal setting phase,* learners need to translate their task perception into goals. The *goal setting screen* functions as an external trigger to support learners to articulate when learning goals are reached. This is based on the idea of feed-up, which comes from the feedback literature [19]. *Feed-up* represents an external trigger to support learners to articulate *when* learning goals are reached. Feed-up interventions are used to support learners to explicitly set goals and standards for regulation. These standards help learners to formulate criteria that indicate *how* to know that a learning goal is reached. This helps learners to engage in cognitive evaluations in the enactment phase. Consequently, the goal setting screen is developed to ground learners' cognitive evaluations in the enactment phase of the COPES model. In our current study, we ask learners to determine goals at different time scales: during the lesson, during rehearsal and their overall goal so that they have standards throughout their learning process to evaluate their progress.

Third, in the *enactment phase,* learners work towards their learning goals while they monitor their progress and control their actions and strategies in case needed. The need for adaptation is determined by cognitive evaluation in which learners compare their current product to their standard to determine progress. Previous research has shown that young learners' monitoring accuracy is often low [1, 18, 22]. Young learners tend to overestimate their performance, which leads to unjustified regulation actions. For example, when a learner believes he is making strong progress, he may reduce his effort. In case of overestimation this leads to unjustified reduction of effort and could harm further progress.

Also in the context of ALTs, we found that learners tend to overestimate their performance, yet there also is a group of learners that consistently underestimates their performance in this context [23]. Therefore the learning path app includes performance feedback to provide learners with an accurate representation of their actual performance. In the *overview screen* progress is symbolized by the position of the dolphin and the dolphins attributes (hoop, ball) which indicate whether learning goals are reached. One layer deeper, on the *goal setting screen,* detailed information on performance is provided which indicates the exact relation between goals (standards) and current performance (products). This screen can be viewed as an external cue to trigger cognitive evaluation, i.e. to compare the goals set with learning products to evaluate progress. This helps learners realize when their progress is not as expected and they need to adapt (small scale adaptation in figure 1), for instance by re-evaluating their degree of effort. This idea is in line with the notion of feed-forward for the feedback literature. Feed-forward is an external cue to re-evaluate plans and adjust strategies. For example when a learner's verbalizes how to adapt learning strategies and actions to ensure future learning. Hence this screen is a cue for learners to explicitly evaluate progress and determine the need for control actions [19].

Fourth, in *the adaptation phase.* learners enact adaptations. As described above, small-scale adaptations are often embedded in the enactment phase when learners adjust effort or strategies based on cognitive evaluations. Large scale adaptations entail reflection and drive improved regulation in the next learning cycle. The *learning path screen* shows personal progress over time. Here the moment-by-moment-learning curves indicate when learners were likely to have learned during their practice session [8, 9]. These curves show the probability that a learner has learned at each problem-solving opportunity [8]. This information is deducted from learners' data traces in ALTs and specifically highlights the relation between learning actions, in this case problems solved, and progress in the learning path [25]. Previously, student MbMLCs were found to have five characteristic clusters: immediate drop, immediate peak, double spikes, close multiple spikes and separated multiple spikes were found [24]. In the learning path app these clusters function as personalized visualizations to show learners how they enact regulation over time [26]. The names have been adjusted for young learners in "high swimmer (immediate drop), quick swimmer (immediate peak), climber in two steps (double spikes), slow climber (close multiple spikes) and climber and descender (separated multiple spikes),"[1]. Hence, in this study MbMLC curves help learners to explicitly understand the their progress over a lesson to formulate adaptations for the next lesson.

Learners are given support to translate the personalized visualisations into regulation actions. A classroom poster helps learners in deriving meaningful actions from the learning path visualizations. Each of the 5 personalized visualisations are explained on the poster and tips to improve regulation are given. According to the learning analytics process model learners need to translate awareness into action, via reflection and sense making [14]. Therefore, on the poster an explanation and explicit actionable information is given. For example, learners that have a close multiple spikes tend to be dependent on external regulation [25]and are advised to increase effort and pay extra attention to their accuracy.

To summarize, the learning path app contains 3 personalized visualizations (overview, goal setting and learning path) that are designed to support learners' internal regulation. The visualizations are explicitly developed as external feedback to help learners to create a valid reference for their regulation process. Based on this reference learners can optimize their internal regulation process. In the learning path app, trace data from the ALT are used to provide learners with continuous feedback about their performance, progress and how progress towards their learning goal is related to their actions. In this way we extend the role of learner-faced dashboards from discussing *what* learners learned to also incorporate *how* learners have learned. Hence the learning path app is expected to be a first step towards developing a novel way to overcome learners' utilization deficiencies in SRL. In the next section, we compare this app to other SRL support tools.

[1] In the original language it has a more positive sound to it; the terms do not translate well.

1.3 Comparison with other SRL support

The personalized visualizations we present here are distinct from other support technique that have been used to assist learners' regulation such as prompts [10], scaffolding [7, 29] and pedagogical agents providing feedback [5, 30]. These techniques have been effective for improving learning, but they were less successful in developing self-regulated learning skills that sustain effective regulation in absence of support. A drawback of these techniques is that they do not help learners to make explicit inferences about how their actions are related to progress towards learning goals [36]. In the learning path app, we aim to make this relation explicit for learners. Traditional SRL support facilitated local corrections, but it did not provide sufficient information to train monitoring accuracy or teach learners to determine the need for small and large scale adaptations themselves. Tentatively, this support may not trigger learners' own cognitive evaluation, which is essential for learners to develop their SRL skills [7]. In line with this argument , it has been emphasized that in order to engage in accurate cognitive evaluations, learners need reliable, revealing, and relevant data to draw valid inferences about their own regulation process [34]. Learner-faced dashboards have been proposed as a potential external cue to help learners make those inferences and our approach is an initial attempt to design these dashboards.

1.4 This study

This study evaluates the effects of the learning path app on students' regulation, learning and monitoring accuracy while learning in an ALT. Based on earlier research, we expect that the learning app will trigger learners to articulate goals and supports cognitive evaluation. We also hypothesize that the explicit goal setting and performance feedback in the overview screen and goal setting screen will support learners to evaluate their progress, and that the MbMLC in the learning path screen will help learners to better understand their regulation over time and determine what course of action to take. This external support to optimize regulation is expected to improve regulation of practice behavior (effort and accuracy), learning (post-test and transfer) and monitoring accuracy (absolute and relative calibration). We expect that learners in the experimental condition will improve their regulation, leading to more effort (hypothesis 1) and higher accuracy (hypothesis 2), less complex moment-by-moment learning curves (hypothesis 3) and consequently greater learn more (hypothesis 4) and better transfer (hypothesis 5). Finally we hypothesize that learners in the experimental condition will show less deviation in their absolute calibration accuracy (hypothesis 6) and less overestimation in their relative calibration (hypothesis 7).

2 Method

2.1 Participants

The participants in this study were 92 grade 5 learners. The four participating schools were located in the north-east of the Netherlands and had a diverse population. The learners were between 10 and 12 years old with a mean of 10.15 (sd =.45 .46), and 38 boys (42%) and 54 girls (58%) participated in this study. Five classes were randomly assigned, three to the experimental condition (n=60) and the two to the control condition (n=32). The inclusion criterion was that learners had to participate in at least 3 out of 4 lessons. Based on this criterion 16 learners were excluded from the sample. Moreover, 4 learners missed the pre-test and 1 learner did not participate in the post-test.

2.2 Design

This study was conducted with a quasi-experimental pre-test - post-test design, see Figure 2. Learners in the experimental condition (PV condition) worked with the learning path app. They set goals at the beginning of every lesson and evaluated their progress in the learning path app at the start of each lesson. Learners in the control condition completed a puzzle at the start of each lesson to keep total time investment equal over the two conditions. Learners received instruction and practiced the three arithmetic subskills in 3 lessons for 55 minutes each on three consecutive days. The design of the first three lessons followed the direct instruction model including teacher instruction, guided practice, class wide practice and individual practice. In the fourth lesson learners were instructed to practice those skills for which they needed most practice in. The pre-test took place prior to the first lesson and after the completion of all lessons learners took the post-test and the transfer test.

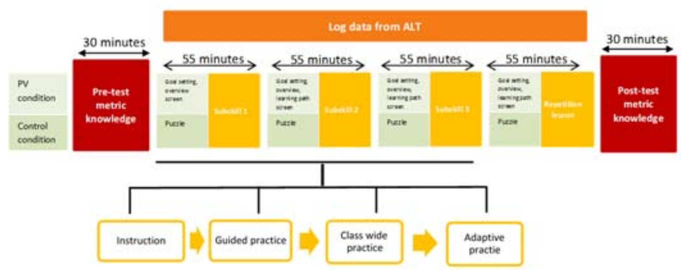


Figure 2: Study design

2.3 Materials

In the experimental condition, learners started their lesson with the learning path app in which they had 3 personalized visualizations to improve their regulation: the *overview* screen, *goal setting* screen and the *learning path* screen (see Figure 3 and 4).

First, in the *overview* screen, learners clicked on the dolphin of a particular arithmetic subskill. Then, learners could set learning goals in the *goal setting* screen, which they did at the start of each lesson. Students were prompted to set goals were for the current lesson, for the rehearsal lesson and for overall proficiency. Learners indicated their learning goals by how proficient they wanted to become at that particular subskill. They represented this by moving the flag on a scale from 0 to 100%. The goal setting screen was designed to act as a feed-up intervention [19] in which learners clearly articulated their learning goal and set their standards to evaluate their progress.

Second, at the beginning of the next lesson, learners could see their progress in the *overview* screen and in the *goal setting* screen. In the overview screen learners, saw their progress on all the three subskills, which was communicated by the position of the dolphin. The placement of the dolphin on the horizontal axis indicated the ability score of the learner as calculated by the ALT. Additionally, the size of the dolphin increased with the number of problems solved. Finally, the dolphin's color provided information about the progress in relation to the overall learning goal set. A grey dolphin indicated that the learning goal was not yet set, an orange dolphin indicated the goal was not yet reached, and a green dolphin showed that the learning goal had been successfully completed. The attributes indicated the progress in relation to the lesson goals. A hoop around the dolphin indicated that the lesson goal was reached and a ball shows the rehearsal lesson goal was obtained.

Third, when learners clicked on a dolphin, they went back to the *goal-setting* screen, which now showed more detailed information on the learner's progress. The blue bars indicated performance based on the ability score as calculated by the ALT. When the ALT did not yet provide an ability score, learners were shown a grey bar. The color of the flag shows how this progress is related to the goals set to support learners. An orange flag indicates that the learner has not yet reached their goal and a green flag indicates that particular goal is reached. The overview screen and the goal setting screen were designed to act as a feed-forward intervention in which learners clearly articulate progress towards their learning goal and engage in cognitive evaluation.

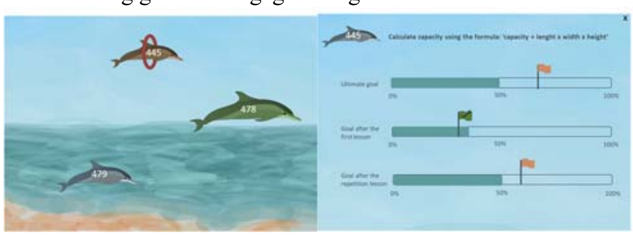


Figure 3: Overview and goal setting screen

Fourth, when learners clicked on the progress bars, they went to the *learning path* screen. Here learners see the learning paths for the selected subskill. The personalized visualizations are based on the Moment-by-Moment Learning Curves calculated based from the ALT data. Learners were shown 5 clusters called high swimmer (immediate drop), quick swimmer (immediate peak), climber in two steps (double spikes), slow climber (close multiple spikes) and climber and descender (separated multiple spikes), see Figure 4. The learning path visualized how learners' actions contribute to their performance and how they made progress towards their goals over time. To make these visualizations understandable, the meaning of the learning paths were explained to learners on posters and by the teachers. On the poster, students were also given recommendations to adapt their regulation in the next lesson. For example, when a learner showed a close multiple spike pattern, this means that he/she learned the skill slowly and that more practice is still needed. Learners were advised to actively monitor their accuracy and increase their effort to ensure they were make progress towards their goals. Hence, these patterns may help learners understand the development of their effort and accuracy during the previous lesson and make adjustments in the next lesson.

Additionally, teachers were given instructions to support learners to understand the learning paths and their implications. A protocol was provided to the teachers that explicitly discussed the function of each step in the intervention. Moreover, teachers were instructed to help learners formulate actions they could take depending on their learning paths.

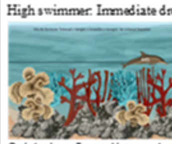
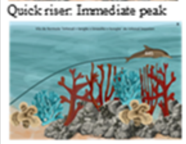
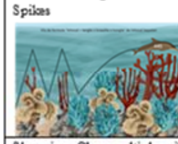
Personalized dashboards	**Planning**	**Monitoring**
High swimmer: Immediate drop 	You already know this skill. → Please practice a different skill.	Your accuracy is high, well done!
Quick riser: Immediate peak 	You have learned this skill quickly after the teacher explained it. → You can practice until you have reached proficiency (green dolphin) and then continue on the next skill.	Your accuracy is high, well done!
Riser in two stages: Double Spikes 	You have learned this skill in two stages during guided instruction and class wide practice. → Please practice until you have reached proficiency.	→ Please monitor your accuracy during practice. → Do you feel that you can put in a little more effort? Try to become a quick riser!
Slow riser: Close multiple spikes 	You are learning this skill somewhat slowly. → Please continue to practice in adaptive mode until you have reached proficiency.	→ Please monitor your accuracy during practicing. → Do you feel that you can put in a little more effort? Try to become a riser in two stages!
Riser and descender: Separate multiple spikes 	You are learning this skill quite slowly. → Please continue to practice in adaptive mode → If you cannot master this skill please notify your teacher	→ Please monitor your accuracy during practicing. → Do you feel that you can put in a little more effort? Try to become a slow riser!

Figure 4: Learning path screens with recommendations

2.3.1 The Adaptive Learning Technology. The adaptive learning technology (ALT) used in this study is widely used for spelling and arithmetic education throughout the Netherlands. This technology is applied in blended classrooms in which the teacher gives instruction, after which learners practice on their tablets. Learners are also given direct feedback (correct or incorrect) after entering an answer to a problem. Teachers can follow learners in teacher dashboards [28]. A lesson has different phases. First, learners practiced in the class-wide practice stage on non-adaptive problems, which were the same for each student in the class. Next, learners worked on adaptive problems, which were selected after each problem solved, based on an estimate of the learner's knowledge: the ability score [21], calculated using a derivative of the ELO algorithm [17]. Based on the learner's ability score, the ALT selected problems with a probability of 75% that the learner will answer the problem correctly. After a learner had answered approximately 25 problems, the system had a reliable estimate of their ability score. This ability score was used as an indicator of performance in the goal setting screen, the bleu bars see figure 3. The difference between the previous ability score and the new score was used as an indicator of progress. In the goal setting screen this is the difference between the second and the third bar.

2.3.2 Subskills learned. The three subskills all included different aspects of measurements of capacity (see Table 1). The Dutch metric system units for measuring capacity were used. The problems related to the first subskill "Calculate capacity using the formula: 'capacity = length x width x height" were relatively easy because learners were given a formula to solve the problem. Also, in this subskill, examples were used to support learners' problem solving. The problems related to the second subskill "Convert from common capacity units to cubic meters" were of medium difficulty. Learners were asked to convert from common capacity units into cubic meters (cm3, dm3, m3). Finally, problems within the third subskill "Convert cubic meters units to liter units" were hard. Learners were asked to convert cubic meters (cm3, dm3, m3) into cubic liter units (cl3, dl3, l3) without a formula.

2.4 Measurements

2.4.1. Pre- and post-test. The pre- and post-test consisted of 24 items, 8 items per subskill. The items in the pre- and post-test were structurally similar, but different numbers were used. The difficulty level of the items, as indicated by the ALT, was used to balance both tests. Figure 5 provides examples of the items for each subskill. The overall Cronbach's alpha for the whole pre-test was .81 with .90 for subskill 1, .85 for subskill 2 and .54 for subskill 3 respectively. The overall Cronbach's alpha for the post-test was .79 with 0.61 for subskill 1, 0.85 for subskill 2 and 0.55 for subskill 3 respectively. Learning gain was calculated as the difference between pre- and post-test. The transfer test consisted of 15 items that tested students' deeper understanding of the relations between meter units and liter units. The Cronbach's alpha for the transfer test was .68

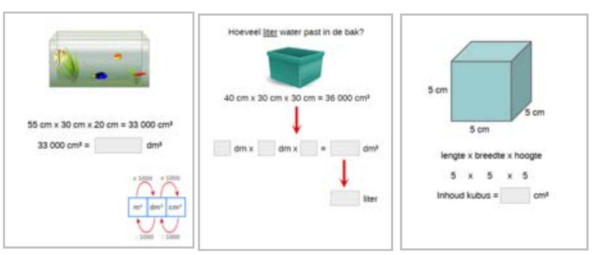


Figure 5: Examples of problems for each subskills

Table 1: Measures and their definition

Learning measures	Definition
Prior knowledge	Pre-test, one per subskill
Post Knowledge	Post-test, one per subskill
Gain	Post-test - pre-test per subskill
Process measures	Log file data
Unique problems	Number of unique problems completed per subskill
Accuracy unique problems	Correct unique problems / total unique problems completed

2.4.3. Measures from the ALT. The knowledge a student has acquired on a subskill is expressed in their ability level as calculated by the ELO algorithm. This score is given by a number between 0 and 600. In order to compare this value to the student's goals we translated the ability score into a percentage. The logs of the ALT stored data on learners' practice activities, including a date and time stamp, student identifier, problem identifier, learning objective identifier, ability score after each problem and correctness of the answer given. Based on this information the following indicators of effort and accuracy were calculated. Effort is measured by one indicator per subskill: the number of unique problems a student completed to practice this subskill. Accuracy is calculated by dividing the number of correctly answered problems by the total number of problems completed. Table 1 provides an overview of all measures calculated and their definition.

2.4.3. Moment by Moment learning curves. The moment-by-moment learning curves were derived based on an algorithm that calculates the probability that the student has just learned the skill [8]. This probability is plotted across the learner's unique problems solved on a single skill over time, to derive the MbMLC. We developed a Python script to automatically classify the form of MbMLC, following the rules in Table 2. Peaks were defined as points that are more than .015 higher than the point before and after.

Table 2: Coding rules for classifying moment-by-moment learning curves.

Curve	Rules
Immediate drop	The curve starts high, drops quickly after solving problems and remains low afterwards.
Immediate peak	The curve starts low, peaks within the first 10 problems and remains low afterwards.
Double spikes	The curve starts low and shows 2 peaks over the course of problem solving.
Close multiple spikes	The curve starts low and shows more than 2 peaks within the first 25 problems and remains low afterwards.
Separated multiple spikes	This curve starts low and continues to show multiple peaks, even after 25 problems

2.4.4. Measures of monitoring accuracy. To measure monitoring accuracy, we asked students to predict how many problems they would solve correctly per subskill on the post-test. We calculated an absolute monitoring accuracy (predicted correct – actual correct) to understand the distance between the expected performance and actual performance, and a relative difference to understand the direction of bias learners have in their monitoring accuracy. We speak of overestimation when a learner's expected score is higher than their actual score and underestimation in the case where a learner's expected score is lower than the actual score.

2.5 Procedure

On the first day learners completed the pre-test (30 minutes) after which the first instruction lesson of 55 minutes was given. The two other instruction lessons and the repetition lesson were given on separate consecutive days following the first lesson. On the fifth day learners completed the post-test (30 minutes) and the transfer test (15 minutes). Each instruction lesson started with the learning path app for the PV condition and the puzzle for the control condition for 10 minutes. In the PV condition, students were asked to look at the overview screen, consult their learning paths and set goals for this lesson. Next, instruction on the subskill was given by the teacher (10 minutes). The instruction was standardized by using an instruction protocol. After the instruction, the teacher and students practiced 6 to 8 problems together in guided practice. Then learners continued to work on problems within that particular subskill. First, learners completed a set of non-adaptive problems (15 problems) which were the same problems for all learners in the class. Next, they continued to work on adaptive problems for the remaining time in the lesson. The total individual practice time was 30 minutes. Finally, 5 minutes were spent for reflection on the lesson. In the fourth lesson the three subskills of the previous lessons were repeated and practiced with adaptive problems. Learners were instructed to select subskills depending on their progress and need for practice.

2.6 Analysis

In order to assess how the personalized visualizations affected effort, accuracy, and transfer, a MANOVA analysis was performed with effort, accuracy and transfer on skill 1, skill 2 and skill 3 as within-subject factors and condition as a between-subject factor. A repeated measures MANOVA was used to assess how learning path app affected learning of the 3 subskills over time. The within-subject factors were the pre and post-test scores (time) on the three skills (subskill 1, subskill 2 and subskill 3). Again condition was used as a between subject factor. In order to investigate how the learning path app affected learners' MbMLC patterns we performed a chi-square analysis. For differences in monitoring accuracy between the conditions an independent samples t-test was used.

3 Results

Table 3: Descriptive statistics per condition

	Subskill 1				Subskill 2			
	PV		Control		PV		Control	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Pre-test	6.32	2.47	5.60	2.85	1.46	2.02	.70	1.23
Post-test	7.39	.73	6.57	1.59	6.32	1.85	5.00	2.44
Gain	1.07	2.55	1.16	2.58	4.86	2.34	4.16	2.75
Effort	53.42	21.61	55.47	17.85	60.33	22.76	57.00	20.54
Accuracy	.81	.08	.76	.10	.64	.16	.49	.15

	Subskill 3			
	PV		Control	
	M	*SD*	*M*	*SD*
Pre-test	1.48	1.61	1.03	1.15
Post-test	4.11	1.87	2.97	1.19
Gain	2.63	2.02	1.93	1.51
Effort	61.58	24.88	69.72	41.11
Accuracy	.63	.11	.55	.12

3.1 Effect on regulation of practice behavior

First, we determined the effects of practice behavior on regulation, looking at effort, accuracy and MbMLCs. For *effort,* there was a significant main effect of skill, $F(2, 85) = 6.31$, $p < 0.002$ indicating that learners showed different effort on the three subskills. There was no significant interaction between skill and condition, $F(2, 85) = 1.62$, $p < 0.05$: Learners did not show more effort in the PV condition than the control condition (Hypothesis 1, rejected). For *accuracy,* there was a significant main effect of skill, $F(2, 85) = 6.31$, $p < 0.002$ indicating that learners showed different accuracy on the three subskills. There was a significant interaction between skill and condition, $F(2, 85) = 4.88$, $p < 0.01$: in the PV condition learners had higher accuracy than learners in the control condition (Hypothesis 2, accepted). For *MbMLCs*, there was no significant difference for subskill 1, but we found a significant difference in the relative occurrence of different MbMLC patterns between the two conditions for skill 2 and 3, chi-square analysis $\chi_2(df = 5, N = 92) = 11.38$, $p < .05$ and $\chi_2(df = 4, N = 92) = 12.38$, $p < .01$, see Figure 6. For subskill 2, the PV condition showed more immediate peaks and double peaks, whereas the control condition showed more close and separate multiple spikes . For subskill 3, the PV condition showed more immediate peaks and double spikes whereas the control condition showed more close and separate multiple spikes (hypothesis 5 accepted).

Table 4: MbMLCs per subskill per condition

Curve	PV			Control		
	S1	S2	S3	S1	S2	S3
Immediate drop	75	2	2	66	6	3
Immediate peak	15	47	33	28	35	9
Double spikes	8	28	30	3	9	19
Close multiple spikes	2	7	22	3	16	31
Separated multiple spikes	0	15	13	0	34	38

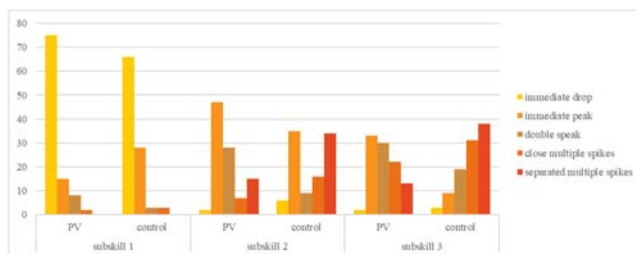


Figure 6: MbMLC per condition and subskill

3.1 Effect on learning

There was a significant main effect of time, $F(1 , 85) = 320.11$, $p < 0.001$: learners scored higher on the post-test than the pre-test. There also was a main effect of condition, $F(1 , 85) = 15.25$, $p < 0.001$: learners in the PV group scored higher then learners in the control group and a main effect of skill, $F(2 , 85) = 247.17$, $p < 0.001$: learners score differently on the three skills. There was an interaction effect between time and skill, $F(2 , 85) = 43.12$, $p < 0.001$, which indicates difference in growth over time between the skills. There were no other interaction effects. Follow up analysis revealed that the PV group scored higher on pre-test for subskill 2 compared to the control condition, but not for subskill 1 and 3. On the post-test the PV group scored higher on all three skills, but progress from pre to post-test was only marginally stronger on skill 3 (see Figure 7). Overall, even though the results are in the anticipated direction, hypothesis 3 is not accepted. Finally with respect to transfer, we find a significant difference between the PV condition (M = 10.68, SD = 2.41) and the control condition (M = 9.16, SD = 3.79), $t(85, 2) = 2.33$, $p < 0.05$ (hypothesis 4, accepted).

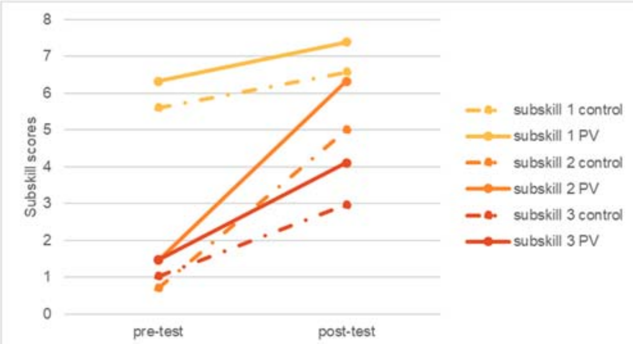


Figure 7. pre- and post-test scores per condition

3.3 Effect on monitoring accuracy

Table 5: absolute and relative calibration per condition

	Absolute calibration				Relative calibration			
	PV		Control		PV		Control	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Subskill 1	1.73	1.20	1.00	1.19	-1.66	1.29	-.87	1.28
Subskill 2	1.63	1.81	1.81	1.51	-1.01	1.91	.62	2.29
Subskill 3	2.10	1.58	2.53	1.26	1.22	2.34	2.47	1.39

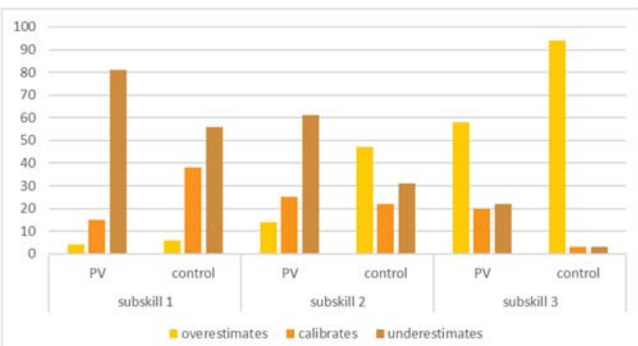


Figure 8: Relative monitoring accuracy per condition and subskill

We only found a significant difference for absolute calibration accuracy for subskill 1 (see Table 5), where the control condition was actually more accurate than the experimental condition t (2, 89) = 2.77, $p < 0.01$. For relative calibration accuracy, we found significant difference for all three subskills. For subskill 1, learners in the PV condition underestimated their performance more than learners in the control condition, t (2, 89) = -2.77, $p < 0.01$. For subskill 2, again learners in the PV condition underestimated their performance, whereas learners in the control condition overestimated their performance, t (2, 89) = -3.64, $p < 0.001$. For subskill 3, learners in both the PV condition and the control condition overestimated their performance, but the effect was larger in the control condition , t (2, 89) = -2.75, $p < 0.01$. Hence even though there are little difference in absolute calibration there are clear difference in the extent to which learners over or under estimate themselves as a result of the working with the learning path app, as shown in Figure 8, although the relationship is complex.

4 Discussion

This paper described the design and evaluation of personalized visualizations to support young learners' SRL in ALTs. The learning path app combined three personalized visualizations (overview, goal setting and learning path screen) that were designed as an external reference to support learners' internal regulation. We found that learners in the personalized visualization condition (PV) improved regulation of their practice behavior as shown by increased accuracy and less complex MbMLCs for subskills 2 and 3 compared to learners in the control group. Although in the PV condition learners scored higher on the post-test they only showed marginally more progress on the most difficult subskill 3. Moreover, initial differences in prior knowledge on subskill 2 prevent us from drawing conclusions with regard to learning outcomes. Learners in the PV group did show enhanced transfer of their knowledge to a structurally different situation. Finally, although both conditions scored equally on absolute monitoring accuracy, there was a difference in relative monitoring accuracy, indicating that students in the personalized visualizations condition were more likely to underestimate their knowledge than students in the control group. Overall, these findings indicates that the personalized visualizations affect learners accuracy during practicing, MbMLC and relative monitoring accuracy. Below we discuss these findings in depth and relate them to learner's regulation.

With regard to improved regulation of practice behavior, we found that the personalized visualizations in the learning path app did improve students' accuracy, but did not improve effort. Effort was measured by the number of problems a students solved. We may have seen an effect, however, if a more advanced measure of effort was used. The MbMLC in the learning path screen had the goal of driving between-lesson adaptation by giving students feedback on their regulation in this lesson to inform their regulation in the next lesson. Indeed, for both subskills 2 and 3 we found an increase of simple curves and a reduction of complex curves (close and separated multiple spikes) in the personalized visualizations condition. Simple curves show more efficient practice behavior, whereas more complex curves indicate that learners had problems adjusting their effort and increasing their accuracy. This showed that an external reference for learning goals, performance and progress over time in different personalized visualizations indeed helped learners to improve their regulation of their practice behavior in the next lesson.

The results on learning outcomes are partially confounded by initial differences in prior knowledge for subskill 2 between the PV and control condition. For subskill 1 and 3 there were no significant initial differences. For all three subskills the PV condition significantly outperformed the control condition on the post-test. A marginally significant difference in progress was found, only for subskill 3, but the trend in the data pointed towards improved progress in the PV condition. This is further reinforced by the improved transfer performance of the PV condition. It has been reported in previous research that improved regulation is more beneficial for transfer compared to immediate learning [12], a finding replicated in this study.

Finally, we expected that goal setting and cognitive evaluations with performance-based feedback would support learners monitoring accuracy during learning. Although we found no difference in absolute monitoring accuracy, interesting differences were found with respect to relative monitoring accuracy. This indicates that while students in both conditions were comparably inaccurate in their perceptions of themselves, the direction of the bias was different. Learners in the PV condition under-estimated their performance more often than learners in the control condition. This indicates that the tendency to overestimate performance which is widely reported on in research [18, 22] was reduced by the learning path app, although it was replaced with a different bias.

Other work has outlined four groups of learners with different SRL support needs, namely a self-regulation group, a teachers regulation group, system regulation group and an advanced system regulation group [25]. This work proposed that personalized visualizations may to have the potential to enhance regulation specifically for two groups of learners namely those in the SRL and Teacher regulation groups. By contrast, this article argued that students in the system regulation and advance system regulation group would need more support to overcome their utilization deficit [25]. In this study we have found a substantial reduction of in close and separated multiple spikes, respectively 57% and 56% for subskill 2, and 30% and 65% for subskill 3. This seems to indicate that even for students with more advanced SRL support needs personalized visualizations did support regulation. Some students appeared to be in need of more advanced support to improve regulation, but this group seems smaller than the system and advanced regulation group. Still the proposal of human-system regulation that takes over parts of the regulation from the learner until he/she is ready to exert more control over learning could be beneficial for these students. Tailoring SRL support in Hybrid human-system regulation to each learner's detected needs could further optimize learning and regulation.

To summarize, we have found evidence for improved regulation of practice behavior, reduction of complex MbMLCs and a reduction of overestimation in monitoring accuracy (albeit at the cost of greater underestimation) as a result using the personal visualizations in the learning path app. A limitation of this study is the initial difference on subskill 2 which prevents strong conclusions with regard to learning outcomes. It is also worth noting that we are unable to derive which PV was responsible for the findings, due to studying a combination of multiple personalized visualizations at once. Although future studies could address this issue through ablating the PV, the existing comprehensive visualization is designed around the COPES model to work in combination; therefore, individual effects may be less than the joint effects. Finally, long-term effects of PV on the development of SRL skills could not be assessed in this study. To address this research question a longer-term intervention with the learning path app is needed.

ACKNOWLEDGMENTS

VENI grant of the first author (451-16-017). We thank participating schools, students teachers and master student Nadine Bril for their contribution.

REFERENCES

[1] Aghababyan, A., Lewkow, N. and Baker, R. 2017. Exploring the asymmetry of metacognition. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (2017), 115–119.

[2] Aleven, V., Roll, I., McLaren, B.M. and Koedinger, K.R. 2016. Help Helps, but only so Much: Research on Help Seeking with Intelligent Tutoring Systems. *International Journal of Artificial Intelligence in Education*. 26, 1 (2016), 205–223. DOI:https://doi.org/10.1007/s40593-015-0089-1.

[3] Arroyo, I., Ferguson, K., Johns, J., Dragon, T., Mehranian, H., Fisher, D., Barto, A., Mahadevan, S. and Woolf, B. 2007. Repairing disengagement with non invasive interventions. *International Conference on Artificial Intelligence in Education*. August 2016 (2007), 195–202.

[4] Azevedo, R. 2015. Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational Psychologist*. 50, 1 (2015), 84–94.

[5] Azevedo, R., Johnson, A., Chauncey, A. and Burkett, C. 2010. Self-regulated learning with MetaTutor: Advancing the science of learning with MetaCognitive tools. *New science of learning*. Springer. 225–247.

[6] Azevedo, R., Martin, S.A., Taub, M., Mudrick, N. V, Millar, G.C. and Grafsgaard, J.F. 2016. Are Pedagogical Agents' External Regulation Effective in Fostering Learning with Intelligent Tutoring Systems? *Intelligent Tutoring Systems* (Cham, 2016), 197–207.

[7] Azevedo, R., Moos, D.C., Greene, J.A., Winters, F.I. and Cromley, J.G. 2008. Why is externally-facilitated regulated learning more effective than self-regulated learning with hypermedia? *Educational Technology Research and Development*. 56, 1 (2008), 45–72. DOI:https://doi.org/10.1007/s11423-007-9067-0.

[8] Baker, R.S., Hershkovitz, A., Rossi, L.M., Goldstein, A.B. and Gowda, S.M. 2013. Predicting Robust Learning With the Visual Form of the Moment-by-Moment Learning Curve. *Journal of the Learning Sciences*. 22, 4 (2013), 639–666. DOI:https://doi.org/10.1080/10508406.2013.836653.

[9] Baker, R.S.J.D., Goldstein, A.B. and Heffernan, N.T. 2011. Detecting learning moment-by-moment. *International Journal of Artificial Intelligence in Education*. 21, 1–2 (2011), 5–25. DOI:https://doi.org/10.3233/JAI-2011-015.

[10] Bannert, M., Hildebrand, M. and Mengelkamp, C. 2009. Effects of a metacognitive support device in learning environments. *Computers in Human Behavior*. 25, 4 (2009), 829–835. DOI:https://doi.org/10.1016/j.chb.2008.07.002.

[11] Bannert, M., Molenaar, I., Azevedo, R., Järvelä, S. and Gašević, D. Relevance of Learning Analytics to Measure and Support Students ' Learning in Adaptive Educational Technologies. DOI:https://doi.org/10.1145/3027385.3029463.

[12] Bannert, M. and Reimann, P. 2012. Supporting self-regulated hypermedia learning through prompts. *Instructional Science*. 40, 1 (2012), 193–211. DOI:https://doi.org/10.1007/s11251-011-9167-4.

[13] Bodily, R., Jivet, I., Kay, J., Davis, D., Verbert, K., Aleven, V. and Xhakaj, F. 2018. Open Learner Models and Learning Analytics Dashboards: A Systematic Review. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (2018).

[14] Bodily, R. and Verbert, K. 2017. Review of research on student-facing learning analytics dashboards and educational recommender systems. *IEEE Transactions on Learning Technologies*. (2017). DOI:https://doi.org/10.1109/TLT.2017.2740172.

[15] Bull, S. and Kay, J. 2013. *Open Learner Models as Drivers for Metacognitive Processes*.

[16] Butler, D.L. and Winne, P.H. 1995. Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research*. (1995). DOI:https://doi.org/10.3102/00346543065003245.

[17] Elo, A. 1978. The rating of chessplayers, past and present. (1978).

[18] Foster, N.L., Was, C.A., Dunlosky, J. and Isaacson, R.M. 2017. Even after thirteen class exams, students are still overconfident: the role of memory for past exam performance in student predictions. *Metacognition and Learning*. 12, 1 (2017), 1–19.

[19] Hattie, J. and Timberley, H. 2007. The power of feedback. *Medical Education*. 44, 1 (2007), 16–17. DOI:https://doi.org/10.1111/j.1365-2923.2009.03542.x.

[20] Jivet, I., Scheffel, M., Specht, M. and Drachsler, H. 2018. License to evaluate: preparing learning analytics dashboards for educational practice. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge - LAK '18*. (2018), 31–40. DOI:https://doi.org/10.1145/3170358.3170421.

[21] Klinkenberg, S., Straatemeier, M. and Van Der Maas, H.L.J. 2011. Computer adaptive practice of Maths ability using a new item response model for on the fly ability and difficulty estimation. *Computers and Education*. 57, 2 (2011), 1813–1824. DOI:https://doi.org/10.1016/j.compedu.2011.02.003.

[22] van Loon, M.H., de Bruin, A.B.H., van Gog, T. and van Merriënboer, J.J.G. 2013. The effect of delayed-JOLs and sentence generation on children's monitoring accuracy and regulation of idiom study. *Metacognition and Learning*. 8, 2 (2013), 173–191. DOI:https://doi.org/10.1007/s11409-013-9100-0.

[23] Molenaar, Inge; Horvers, A.; Dijkstra, R. 2020. Young Learners' Regulation of Practice Behavior in Adaptive Learning Technologies. *Frontiers in Psychology*.

[24] Molenaar, I., Horvers, A. and Baker, R.S. 2019. What can moment-by-moment learning curves tell about students ' self- regulated learning ? *Learning and Instruction*. July 2018 (2019), 101206. DOI:https://doi.org/10.1016/j.learninstruc.2019.05.003.

[25] Molenaar, I., Horvers, A. and Baker, R.S.J.D. 2019. Towards Hybrid Human-System Regulation: Understanding Children' SRL Support Needs in Blended Classrooms. *In proceedings of the 9th International learning analytics & knowledge conference* (2019).

[26] Molenaar, I., Horvers, A., Dijkstra, R. and Baker, R. 2019. Designing Dashboards to support learners' Self-Regulated Learning. *In companion proceedings of the 9th International learning analytics & knowledge conference* (2019), 1–12.

[27] Molenaar, I. and Järvelä, S. 2014. Sequential and temporal characteristics of self and socially regulated learning. *Metacognition and Learning*. (2014).

[28] Molenaar, I. and Knoop-van Campen, C. 2018. How Teachers Make Dashboard Information Actionable. *IEEE Transactions on Learning Technologies*. 1382, c (2018), 1–9. DOI:https://doi.org/10.1109/TLT.2018.2851585.

[29] Molenaar, I., Roda, C., Van Boxtel, C. and Sleegers, P. 2012. Dynamic scaffolding of socially regulated learning in a computer-based learning environment. *Computers and Education*. 59, 2 (2012), 515–523. DOI:https://doi.org/10.1016/j.compedu.2011.12.006.

[30] Roll, I., Aleven, V., McLaren, B.M. and Koedinger, K.R. 2011. Improving students' help-seeking skills using metacognitive feedback in an intelligent tutoring system. *Learning and instruction*. 21, 2 (2011), 267–280.

[31] Roll, I. and Winne, P. 2015. Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics*. (2015).

[32] Schwendimann, B.A., Rodríguez-Triana, M.J., Vozniuk, A., Prieto, L.P., Boroujeni, M.S., Holzer, A., Gillet, D. and Dillenbourg, P. 2016. Understanding learning at a glance. (2016), 532–533. DOI:https://doi.org/10.1145/2883851.2883930.

[33] Winne, P. and Hadwin, A. 1998. Studying as self-regulated learning. *Metacognition in educational theory and practice*.

[34] Winne, P.H. 2010. Improving measurements of self-regulated learning. *Educational Psychologist*. 45, 4 (2010), 267–276. DOI:https://doi.org/10.1080/00461520.2010.517150.

[35] Winne, P.H. and Baker, R.S.J. d. 2013. The Potentials of Educational Data Mining for Researching Metacognition, Motivation and Self-Regulated Learning. *JEDM - Journal of Educational Data Mining*. 5, 1 (2013), 1–8. DOI:https://doi.org/10.1037/1082-989X.2.2.131.

[36] Winne, P.H. and Hadwin, A.F. 2013. nStudy: Tracing and supporting self-regulated learning in the Internet. *International handbook of metacognition and learning technologies*. Springer. 293–308.

[37] Winne, P.H. and Hadwin, A.F. 2017. Studying as self-regulated learning. January 1998 (2017).

[38] Zimmerman, B.J. 2000. Attaining Self-Regulation: A Social Cognitive Perspective. *Handbook of Self-Regulation*. (Jan. 2000), 13–39. DOI:https://doi.org/10.1016/B978-012109890-2/500317.