Do Instrumentation Tools Capture Self-Regulated Learning?

Joep van der Graaf*
Behavioural Science Institute, Radboud University, The Netherlands

Lyn Lim
TUM School of Education, Technical University of Munich, Germany

Yizhou Fan
School of Informatics, University of Edinburgh, United Kingdom

Jonathan Kilgour
School of Informatics, University of Edinburgh, United Kingdom

Johanna Moore
School of Informatics, University of Edinburgh, United Kingdom

Maria Bannert
TUM School of Education, Technical University of Munich, Germany

Dragan Gašević
School of Informatics, University of Edinburgh, United Kingdom

Inge Molenaar
Behavioural Science Institute, Radboud University, The Netherlands

ABSTRACT

Researchers have been struggling with the measurement of Self-Regulated Learning (SRL) for decades. Instrumentation tools have been proposed to help capture SRL processes that are difficult to capture. The aim of the present study was to improve measurement of SRL by embedding instrumentation tools in a learning environment and validating the measurement of SRL with these instrumentation tools using think aloud. Synchronizing log data and concurrent think aloud data helped identify which SRL processes were captured by particular instrumentation tools. One tool was associated with a single SRL process: the timer co-occurred with monitoring. Other tools co-occurred with a number of SRL processes, i.e., the highlighter and note taker captured superficial writing down, organizing, and monitoring, whereas the search and planner tools revealed planning and monitoring. When specific learner actions with the tool were analyzed, a clearer picture emerged of the relation between the highlighter and note taker and SRL processes. By aligning log data with think aloud data, we showed that instrumentation tool use indeed reflects SRL processes. The main contribution is that this paper is the first to show that SRL processes that are difficult to measure by trace data can indeed be captured by instrumentation tools such as high cognition and metacognition. Future challenges are to collect and process log data real time with learning analytic techniques to measure ongoing SRL processes and support learners during learning with personalized SRL scaffolds.

CCS CONCEPTS

• **Education**; • **Interactive learning environments**; • **Measurement**;

*joep.vandergraaf@ru.nl

LAK21, April 12–16, 2021, Irvine, CA, USA

ACM ISBN 978-1-4503-8935-8/21/04…$15.00
https://doi.org/10.1145/3448139.3448181

KEYWORDS

Self-Regulated Learning, Instrumentation tools

ACM Reference Format:
Joep van der Graaf, Lyn Lim, Yizhou Fan, Jonathan Kilgour, Johanna Moore, Maria Bannert, Dragan Gašević, and Inge Molenaar. 2021. Do Instrumentation Tools Capture Self-Regulated Learning?. In *LAK21: 11th International Learning Analytics and Knowledge Conference (LAK21), April 12–16, 2021, Irvine, CA, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3448139.3448181

1 INTRODUCTION

One of the current and future core competences is Self-Regulated Learning (SRL), also referred to as "learning to learn" [14]. Regulation fosters learning outcomes, revealed by the positive relation between SRL processes and measures of learning, as well as future learning, as it increases the learners ability to engage in lifelong learning [32]. SRL consists of three core processes: cognitive processes to acquire knowledge and skills, metacognitive processes to monitor and control learning, and motivational processes to regulate effort for engaging with the materials [39]. Despite the theorized significance and documented empirical benefits of SRL, many students struggle to productively self-regulate their learning [4]. In a growing body of studies [9, 19, 31, 34, 45], learning analytics researchers have attempted to understand SRL processes to determine ways to promote students' SRL. In this endeavor, measuring SRL posed a major challenge to researchers. Initially, in educational psychology, SRL was measured as a trait using self-report questionnaires, but this did not always reflect how students actually regulate when performing a learning task (e.g., [10]). Think aloud data, as an event-based measurement of SRL during learning, has proven to be a good indicator of students' actual regulation [6, 33]. Drawbacks of think aloud are that it can be obtrusive and transcribing and coding is a cumbersome process [15]. Learning analytics have been put forward as a promising way to measure SRL e.g., [34]. However, it is not straightforward to associate specific SRL processes with traces in log data [27], particularly because not all SRL processes are revealed equally well in logs. In fact, the extent to which learning analytics capture SRL processes largely depends on the degree to which students are able to reveal their regulation through the use of learning technology [43]. Instrumentation tools, embedded into learning technology, are proposed as a solution to capture SRL processes that otherwise remain unrevealed [19], such as organization (high cognition) and orientation (metacognition). Learners can use instrumentation tools to interact with information in a learning technology, thereby generating traces. For example, when a learner highlights important sentences, this could indicate the process of organizing information. An example of a learning technology with instrumentation tools is nStudy, proposed by Winne and colleagues [42, 46]. Although tools have been proposed for measurement of SRL before, their ability to improve the measurement of "hard-to-measure" SRL processes has yet to be validated. The contribution of the paper lies in validating the proposition that instrumentation tools can capture SRL processes that otherwise remain uncaptured. Our findings can inform future applications of instrumentation tools and provide a novel methodological approach to validating measurement of SRL with these tools. The present study, therefore, aimed to investigate the measurement potential of instrumentation tools to assess SRL processes. In this exploratory study, learners used an instrumented learning technology while thinking aloud to determine which SRL processes occur during instrumentation tool use.

1.1 Self-Regulated Learning

SRL is a dynamic and cyclical process that unfolds over time with interactions across and within cognitive and metacognitive activities [26]. Cognitive processes deal with understanding and creating meaning [44]. A distinction has been made between low and high cognition [24], where low cognition is processing of information to acquire knowledge, such as reading, and high cognition is constructing meaning, such as elaboration on ideas. Metacognitive processes deal with controlling and monitoring of cognitive processes[43], for example, checking what information is known to make an inference about what is unknown. Four main categories of metacognition can be identified: planning, orientation, monitoring, and evaluation [6, 22]. A final component of SRL is motivation. Students can express positive and negative feelings about the task, situation, or oneself. Motivational processes can interact with cognitive and metacognitive processes to affect learning [13].

1.2 Measuring SRL

Measurement of SRL has advanced from self-report to current approaches in which multiple data channels are used [2]. Students' self-reports were used to assess how students approach learning tasks (trait-based), but later studies showed that this was not a good predictor of learning outcomes. Instead, analysis of students' think aloud during learning proved to be a better predictor [33]. While think aloud improved the validity of SRL measurement, problems are its obtrusiveness during learning and the labor-intensive process of coding it [15]. To further improve event-based measurement of SRL, learning analytics researchers have developed techniques to analyze student log data [30]. Learning analytics provide a way to gather data about SRL unobtrusively. Just as think aloud, processing log data also involves the process of coding and interpreting actions recorded in the events captured by log data[31,34]. Thus, think aloud and log data are data channels that have been used to measure SRL and each has advantages and disadvantages [for an overview see 5]. Currently, think aloud data tend to be more revealing with respect to SRL processes than log data [38]. Therefore, think aloud can be used as a ground truth to evaluate the validity of measurement with other data channels, including log data.

Detection of SRL with log data is a much debated issue; it can be questioned whether logs represent a learner's intended learning actions or actions induced by the design of the environment [8], and whether a logged event consistently represents the same learning action, or are many learning actions subsumed under a single trace in the log data [45]. Instrumentation tools might be a solution to improve measurement of SRL with log data, because they can reveal how a learner operates on particular information at a point in time and in a relatively well-identified context [46]. By using instrumentation tools, students have more opportunities to reveal their SRL and researchers can collect data about SRL [30]. The potential advantage of instrumentation tools is that SRL is more easily measured across all phases of SRL [30], and it may improve the detection of SRL processes that has been hard to measure, such as high cognition and metacognitive processes [7, 23]. Meanwhile, the analysis of log data is still under development with emergence of analytic approaches dealing with temporal and sequential variation to better capture SRL [31]. At present, one way to deal with construct validity of log data is to use different procedures or instruments to record data about SRL [45], such as synchronizing log and think aloud data.

1.3 Instrumentation Tools

Measurement approaches of SRL using log data have been proposed (e.g., [31, 34]). These models help the interpretation of log data in terms of SRL theories by describing ways to label, process, and analyze log data. For example, there are ways to process patterns of actions, such as interpreting a quiz attempt before switching to reading materials, as an indicator of evaluation [31]. It is important to identify log events that represent enactment of a specific SRL process. Such events have been harder to find for SRL processes reflecting high cognition (organization and elaboration) and metacognition (orientation, planning, monitoring, evaluation), compared to low cognitive processes such as reading and processing of information [7, 23]. This may be improved by incorporating instrumentation tools in the learning technology that students can use to reveal these SRL processes [43]. Furthermore, few tools have been developed to capture metacognitive processes [11]. An example of a learning technology that includes instrumentation tools to capture SRL is nStudy, which allows students to take notes, create highlights, add labels to notes and highlight, write essays, search, and check time [42, 46]. Furthermore, because it is unknown which SRL processes are captured by instrumentation tools and to what extent, it is also unknown whether instrumentation tools capture a single or multiple SRL processes.

1.4 Capturing SRL with Instrumentation Tools

A range of instrumentation tools have been proposed and used to measure students' enactment of SRL [35, 46]. While effects of prior knowledge on instrumentation tool use have been investigated [35], as well as effects of instrumentation tool use on learning outcomes [9], little is known about which SRL processes occur during the use of an instrumentation tool. This is especially relevant, because instrumentation tools have been developed to capture hard-to-detect SRL. Instrumentation tools that have been developed include: timer, highlighter, note taker, search tool [46], and planner [35].

First, a timer displays the time left to study and complete the task. This information can be used by learners to monitor their progress in relation to the set goals [42]. It has been shown that time checking is unrelated to students' prior knowledge [35], but whether using the timer is associated with monitoring has not yet been investigated.

Second, a highlighter can be used to create highlights by selecting text and changing its background color. Highlighting has been associated with low cognitive strategies: i.e., to rehearse information, but also with metacognitive monitoring to assess the importance of information [46]. In the first case, creating the highlight has the goal to memorize the highlighted information [12]. In the second case, the highlight is used as an external reference to information [12], which can subsequently be compared with other types of information. An example is that a student creates a highlight after comparing the informative text with his prior knowledge. In this case, the highlight might indicate that the text is new to the learner. Although no studies have indicated that highlights measure memorization nor monitoring, previous research did show that highlighting was positively related to learning [9].

Third, note taking can be used to add written associations to a text. An example is that a learner selects text and adds their interpretation of the text in light of the goal of a task they are working on, such as 'I can use this text to understand one of the core concepts in the task'. Note taking has been linked to two functions: to encode information into long-term memory and to store additional information externally [12]. This means that notes can be used to memorize information (low cognition), to organize information or elaborate on information (high cognition), and to monitor the relevance of information (metacognition) [1]. No studies have associated note taking with these SRL processes yet. However, the highlighting study mentioned above also investigated notes, and found that while note taking was not associated to learning, it was related to highlighting [9].

Fourth, a search function can be used to look for information provided in informative texts. As the learner has to generate search terms, this is an example of goal-directed search; this is in contrast to free search, which is a search without a goal, such as looking at the top of a page to see what is there [15]. In the case of goal-directed search, search would be an indicator of planning behavior. Another reason to use the search function can be to compare products, such as an essay, or understanding about a core topic to specific information in the text, which would indicate monitoring [15]. These relations between use of a search function and co-occurring SRL processes have not been investigated. It has been shown that students, whose SRL was supported by a human tutor, searched less and learned more compared to students, who were not supported [3]. As there were more differences between the conditions, it cannot be concluded that the difference in learning was due to a difference in search behavior.

Fifth, a planner can be used to specify activities a student plans to work on and the time and duration of these activities. Using the planner to create a plan is likely to reflect metacognitive planning [15]. Another reason to use the planner is to monitor progress when a plan has been made, which would indicate monitoring. While some learning environments did incorporate a planner (e.g., [36]), there appear to be no studies addressing the role of a planner to measure this SRL process.

1.5 Relations Between Instrumentation Tools

To measure the use of instrumentation tools, students must use the tools [11]. There are several factors that can affect tools use, such as whether learners are aware of the tool, whether they think the tool can be useful to the given task, and whether they have skills to use it [41]. In addition to determining whether tools are used, it can be informative to investigate which tools are related, as this can provide additional evidence in determining what the tools are used for. For example, highlights and notes share that they can be used to memorize information and/or store it externally [12], and they have been found to be correlated [9]. It has been proposed that tools that share learning processes that occur when using the tool might be correlated [11]. Thus, if tools capture the same SRL processes, a correlation can be expected.

1.6 The Present Study

To summarize, students struggle with employing effective learning strategies and to help them, more information is needed about their use of SRL processes during learning. However, researchers struggle with the measurement of SRL. It has been proposed that instrumentation tools can support the measurement of SRL within log data, especially for SRL processes that are otherwise hard to measure such as higher cognition and metacognition. The present study examined this proposition in detail. We analyzed the frequency of tool use to determine whether the tools were used and to what extent. Next, we aligned think aloud and trace data to identify to what extent tools are related to a particular SRL process and hence, are meaningful for measurement of SRL. Finally, the relations between tools were analyzed to explore if tools that are associated with the similar SRL processes also correlate with respect to their usage.

With respect to the SRL processes co-occurring with tool use, we hypothesized that the use of timer was associated with monitoring [42], highlights with low cognition (superficial writing down) [46], notes with high cognition (organization) and low cognition (superficial writing down) [12], and search and the planner with planning and monitoring [15]. With reference to relations between instrumentation tools, we hypothesized that tools would be related if they capture the same SRL processes. Therefore, the uses of search and planner were expected to be related, and the use of them both might be related to the timer. We also expected the use of the highlighter and note taker to be correlated. We did not hypothesize other relations between tools to be present.

2 METHOD

2.1 Participants

A convenience sample of 46 students from a medium-sized university in the Netherlands with a mean age of 21 years (*SD* = 3 years) participated. One participant was excluded from analyses, due to a technical issue recording think aloud data. Another participant

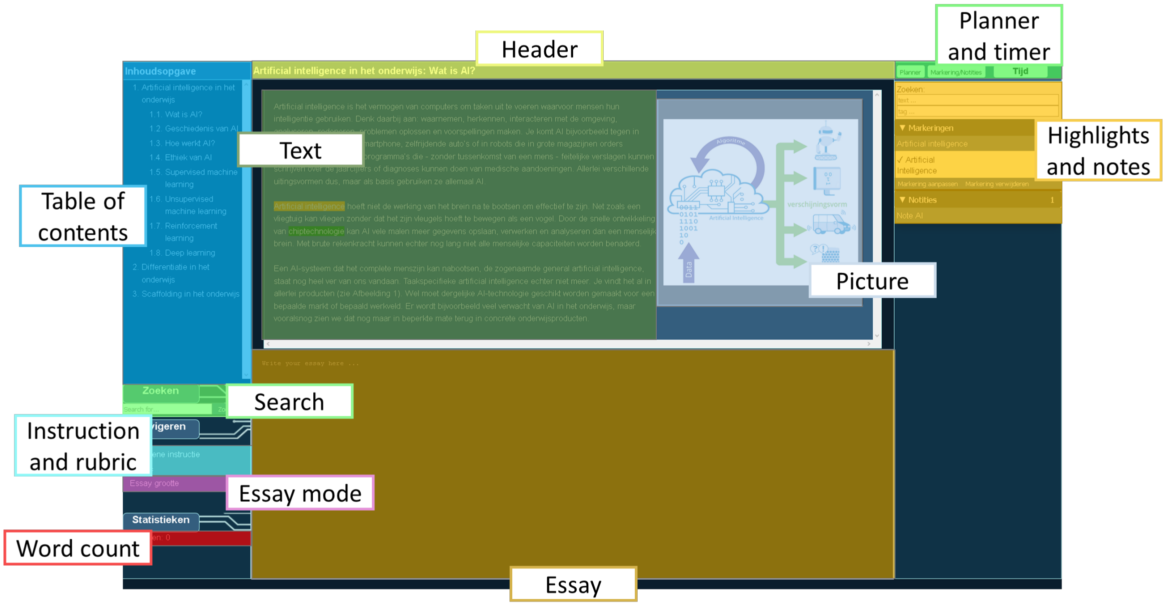

Figure 1: The digital learning environment.

was excluded because Dutch was not her mother tongue, which affected think aloud. Analyses were conducted with 44 participants (34 female and 10 male). Five were enrolled in a master's and 39 in a bachelor's degree program. The participants were students of a wide range of degree programs, but mostly from social sciences. The participants were informed about the study's setup and procedure, and they gave active consent prior to collecting their data. This research was approved by the research lab's ethical committee.

2.2 Procedure

The study used a pretest-posttest design with a learning session of 45 minutes in between. In this paper, we focus on the learning session, because SRL processes were detected with think aloud and tools were used during learning. The task in the learning sessions was to write a vision-essay using the provided informative texts. Before participants started to work on their assignment, there was an introduction to the digital learning environment: how to log in, navigate, and use the tools and essay. There was also a short introduction to thinking out loud. The experimenter verbalized his thoughts during a standardized run through the environment. Then, the participants practiced with the learning environment while thinking out loud. After the introductions, an eye-tracker was calibrated. and the learning sessions started. When 45 minutes had passed, the experimenter made sure the participant stopped. Eye-tracking was used as part of a larger study. In the present study, we did not analyze the eye-tracking data.

2.3 Materials

The experimental setup consisted of a screen-based eye-tracker (Tobii TX300), a laptop with a keyboard and mouse, and a microphone. The laptop was used to run a local PHP-server to present the learning environment to the participants on the screen of the eye-tracker.

2.3.1 Learning environment. The learning environment consisted of a navigation panel on the left, informative texts with an input box to write the essay in the middle, and a tool panel on the right, see Figure 1. The learning environment was used to present instruction (general instruction and a rubric) and informative texts about three topics: artificial intelligence (AI), differentiation, and scaffolds. The landing page contained the general instruction about the task: the participants were instructed to write a vision-essay of 300 to 400 words about education in 2035 using the provided texts. They were also instructed that they should select which texts to use, because 45 minutes is too short to read everything and write an essay. Most of the texts were useful for the essay, but some of them were intentionally not relevant to allow learners exercise their SRL skills. To navigate between the texts, the menu on the left-side panel could be used, which was organized around the three topics, which, when clicked, showed a submenu of sections related to the topic. The text was accompanied with a figure in six sections. The participants could type their essay in the essay area. The size of the essay could be changed to one of three options: small essay (about 10% of the page), medium essay (about 50% of the page, see Figure 1), and large essay (about 90% of the page). The word count was visible on the bottom of the left-side menu.

2.3.2 Instrumentation tools. The tools were a timer, highlighter, note taker, search box, and planner, which were accessible during the whole session.

The timer showed the time left to complete the assignment. By default, the time left was not visible. When a participant clicked on the timer, the time left was visible for 2 seconds. The label *timer* started when clicking on the timer and ended after 2 seconds. Subsequent clicks in 4 seconds were merged.

The highlighter and note taking tool shared functionalities and were presented together in the right-side panel. The highlighter could be used to select text to highlight by changing the color of the background. With the note taker, the participants could also add notes to the selected text. The participants could use the highlighter and note taker to create, read, edit, delete, and search highlights and notes. *Creating* started with selecting text and choosing between creating a highlight or note and ended when clicking on the save button. After choosing to create either a highlight or note, the right-side panel was used to select labels (e.g., "AI"), type a title or content (for notes only), and save the highlight or note. Note that the participants could also create labels themselves. This resulted in the actions: highlight creating, highlight labeling, and note creating. Highlight *creating* is creating a highlight without self-generating labels. Highlight labeling is creating a highlight with a self-generated label. Note creating is creating a note, regardless of a label was self-generated. *Reading* highlights and notes started with opening them in the right-side panel and ended when a new action occurred. Clicking on a highlight showed the labels associated with it. For notes, the title of the note was visible, when the note had no title it showed "Untitled". When clicking on a note, the associated labels and the note content were visible. The click on a highlight or note also caused a navigation to the page, where it was created. To differentiate between clicking to read a highlight/note and clicking to go to a page, a threshold of six seconds was used: when the visit to a new page was shorter than six seconds, it was labelled as highlight reading. *Editing* started with opening a highlight or note and ended when clicking on the save button. In contrast to reading, the information (label, title and/or content) should be edited in between. *Deleting* started with opening the highlight/note and ended when clicking on the delete button. *Searching* started when typing in one of the search boxes (one to search for the selected text and one to search for labels) of the highlights and notes menu and ended when there when there was a click on a highlight or note or when a new action occurred.

Search was a function in the left-side menu to search in the informative texts. *Search* started when typing in the search box started and it ended when there were no search results, when the search was cancelled, or when clicking on one of the search results.

The planner tool consisted of a 45 minute timeline and six blocks. Each block represented a different activity. These blocks could be dragged to the timeline and there their size could be changed from 1 to 45 minutes. The default size was 5 minutes. The label *planner* started when opening the planner and ended when the planner was closed.

2.4 Analysis

2.4.1 Trace data. The actions in the learning environment were logged with the specific action and a timestamp. Mouse-clicks were logged with their coordinates and where possible, their meaning; for example, a click on the planner was logged as planner. Scrolling in the text and essay was also logged. The log was used to identify tool use, as described in Section 2.3.2. In some cases, the end of an action was not recorded. For example, when reading a highlight, there is no clear end. In those cases, the action lasted until the next action, unless this was too long. A maximum duration was chosen for each action so that outliers were set to longest duration of non-outlying values.

2.4.2 Coding think aloud data. The participants' utterances while thinking aloud were recorded and later coded during data analysis. A coding scheme was used to categorize utterances as SRL activities. It was based on previously developed and used coding schemes by Bannert [7] and Molenaar [21]. The main categories were metacognition, cognition, procedural, motivational, and not codable. Main categories: procedural and motivational, had a low frequency (3.52% and 0.52%) and were excluded from further analyses. Utterances that were too unclear to assign any of the previous codes were coded as not codable, such as instances of murmuring, and were excluded from further analyses. Only metacognitive and cognitive processes were used for analyses and Table 1 provides an overview of the subcategories and their frequency. We created segments of the utterances using automatic sound detection. We used the ELAN software [37] to code and modified kappa statistics to correct for the extend of overlap between the timing segments between different coders [16]. Raters coded the utterances and reached acceptable [17] inter-rater reliability: κ = .53-.65 ($\kappa_{\max}$ = .81-.82).

2.4.3 Data analysis. To investigate instrumentation tool use during learning absolute frequencies, calculated based on the log data, were compared using ANOVA with instrumentation tool as within-subjects factor. Huynh-Feldt correction was applied, because sphericity was not assumed, $W = .01$, $p < .001$. Post-hoc comparisons used Bonferroni corrections.

The co-occurrence of instrumentation tool use and SRL processes was analyzed by identifying the time window in which a tool was used. If a think aloud code started and/or ended in this time window, then it was identified as a co-occurring SRL process. If there was a single SRL process in the time window, this process was selected. If there were multiple SRL processes, then the first one was selected. If there was no co-occurring SRL process in the time window, then the first occurring SRL process after tool use was selected. The rational was that tool usage results from the decision that is already made and thus detectable immediately after use.

In follow-up analyses, the type of action when using the highlighter or note taker was taken into account. ANOVA's with action type, SRL process, and their interaction as within-subjects factors and frequency of the SRL processes either during highlighter use or note taker use as dependent variable.

Relations between frequencies of instrumentation tool use were analyzed using Spearman rank correlations, because visual inspection revealed that not all variables were normally distributed, not all relations were linear, and the homoscedasticity assumption was not met. Analyses were conducted in R [28] with the ez [18] and psych [29] packages.

3 RESULTS

3.1 Tool use during learning

The frequencies of tool use revealed that the tools were used to a different extent, as expected, see Table 2. This was confirmed by a main effect of tool on frequency, $F(1.77, 77.78) = 23.82$, $p <$

Table 1: Coding scheme for categorizing students' utterances as SRL processes.

Coding category	Description	Example	Overall usage (%)
Metacognition			30
Orientation	Orientation on learning-related activities.	"Four things are important in the assignment"	4
Planning	Planning of the learning process by arranging activities and determining strategies.	"I will explain these topics in my essay"	6
Monitoring	Monitoring and checking the learning process according to instruction or plan.	"This information does not appear to be relevant"	19
Evaluation	Evaluating content-wise correctness of learning activities.	"I doubt whether Artificial Intelligence will have such an impact"	<1
Low cognition			45
First time reading	Reading the text out loud and superficial describing of pictorial representations.	"Artificial Intelligence is the ability of . . ."	31
Rereading	Rereading of text or figures.	"AI consists of two main components . . ."	1
Superficial repetition	Repeat to memorize information, learn by heart, or use it in, for example, the essay.	". . . a self-learning algorithm and data"	2
Superficial writing down	Copying information by typing it in a note or in the essay, or by creating highlights.	"An algorithm is . . ."	5
Processing	Rereading or editing products, such as notes.	"I should start a new paragraph here"	6
High cognition			25
Elaboration	Connecting content-related comments and concepts; reasoning and association.	"This means that both Artificial Intelligence and humans can learn"	3
Organization	Organizing of content by creating an overview, etc..	"To summarize, . . . are important for future education"	22

Table 2: Descriptive Statistics of Tool use

	M	*SD*	Minimum	Maximum
Timer	6.33	3.59	1	15
Highlighter	9.60	10.35	0	45
Note taker	4.69	5.63	0	22
Search	0.22	0.56	0	2
Planner	0.49	1.01	0	4

.001, generalized $\eta^2 = .30$. As expected, use of the highlighter was relatively frequent. Other tools that were also used frequently were the note taker and timer. The timer was the only tool that was used by everyone. In contrast, the search and planner were used less than once on average. Post-hoc comparisons revealed that the highlights, notes, and timer were used more often than search and the planner, p's $< .001$. Other differences were not significant.

3.2 Co-occurrence of instrumentation tool use and SRL processes

We investigated the proposition that instrumentation tools can be used to measure particular SRL processes. The patterns for the highlighter and note taker were not unambiguous, but our expectations about the timer capturing monitoring, highlighter capturing low cognition, note taker capturing high cognition, and search and planner capturing planning and monitoring were confirmed, see Tables 3 and 4. The use of the timer almost exclusively coincided with monitoring (88%). The highlighter was mostly associated with superficial writing down (32%), but monitoring (26%) and planning (18%) were also relatively frequent. The use of the note taker mostly co-occurred with organization (51%) and, similar to the highlighter, monitoring (17%) and planning (8%). note The use of the search-function coincided with planning (67%) and monitoring (33%). The planner was associated with planning (52%) and monitoring (43%). Comparing these percentages with the overall frequency of SRL processes, see Table 1, indicated that the instrumentation tools captured more of specific SRL processes, especially metacognition and high cognition. Monitoring during use of the highlighter and note taker was the exception, because its occurrence was similar to its occurrence over the whole learning session. Note that the tools captured little and sometimes nothing of the SRL processes: orientation, evaluation, elaboration, superficial repetition, and reading (first time and rereading).

3.2.1 SRL processes during the highlighter and note taker use. Because the overall pattern was somewhat ambiguous when we investigated SRL processes during the highlighter and note taker

Table 3: Frequency of Metacognitive, High Cognitive, and low Cognitive Processes per Tool

	N	Metacognitive	High cognitive	Low cognitive
Timer	264	252 (95%)	4 (2%)	8 (3%)
Highlighter	394	177 (45%)	39 (10%)	178 (45%)
Note taker	172	46 (27%)	95 (55%)	31 (18%)
Search	9	9 (100%)	0 (0%)	0 (0%)
Planner	21	21 (100%)	0 (0%)	0 (0%)

Table 4: Top Three Co-occurring SRL Processes per Tool

	N	Most frequent	Second	Third
Timer	264	Monitoring (88%)	Planning (6%)	Organization (2%)
Highlighter	394	Superficial writing (32%)	Monitoring (26%)	Planning (18%)
Note taker	172	Organization (51%)	Monitoring (17%)	Planning (8%)
Search	9	Planning (67%)	Monitoring (33%)	
Planner	21	Planning (52%)	Monitoring (43%)	Orientation (5%)

Table 5: Top Three Co-occurring SRL Processes per Action Type of the Highlighter and Note taker

	N	Most frequent	Second	Third
Highlighter	394	Superficial writing (32%)	Monitoring (26%)	Planning (18%)
Creating	315	Superficial writing (40%)	Monitoring (23%)	Planning (21%)
Labeling	29	Organization (55%)	Monitoring (21%)	Planning (14%)
Reading	42	Monitoring (48%)	Processing (38%)	First time reading (7%)
Note taker	172	Organization (51%)	Monitoring (17%)	Planning (8%)
Creating	140	Organization (61%)	Monitoring (14%)	Superficial writing (9%)
Reading	28	Processing (46%)	Monitoring (25%)	Planning (14%)

use, we analyzed how students used these tools, i.e., the type of action. Therefore, the effect of action type on the co-occurring SRL process was investigated. Three actions were relatively frequent for highlights (creating, labeling, and reading) and two action for notes (creating and reading).

An ANOVA with action type, SRL process, and their interaction as within-subjects factors and frequency of the SRL processes during highlighter use as a dependent variable, revealed a significant interaction between action type and SRL processes, $F(20, 880) = 12.80$, $p < .001$, $\eta^2 = .14$, and main effects of action type, $F(2, 88) = 34.21$, $p < .001$, $\eta^2 = .06$, and SRL process, $F(10, 440) = 11.98$, $p < .001$, $\eta^2 = .09$. The interaction between action type and SRL processes indicated that the pattern of frequencies across SRL processes differed between highlight creating, highlight labeling, and highlight reading. In other words, highlight creating, highlight labeling, and highlight reading were associated with different SRL processes, see Table 5. All instances of superficial writing down occurred during highlight creating (40% of co-occurring SRL processes). Highlight labeling was mainly related to organization (55%). Highlight reading was associated with monitoring (48%) and processing (38%).

The same pattern of results was found for notes. An ANOVA with action type and SRL processes as within-subjects factors and frequency of the SRL processes during note taker use as a dependent variable, revealed a significant interaction between action type and SRL processes, $F(10, 440) = 20.51$, $p < .001$, $\eta^2 = .18$, and main effects of action type, $F(1, 44) = 28.84$, $p < .001$, $\eta^2 = .04$, and SRL process, $F(10, 440) = 18.89$, $p < .001$, $\eta^2 = .17$. The interaction between action type and SRL process indicated that the pattern of frequencies across SRL processes differed between note creating and note reading. In other words, note creating and note reading were associated with different SRL processes, see Table 5. Almost all instances of organization occurred during note creating (61% of co-occurring SRL processes). Similar to highlight reading, note reading was associated with monitoring (25%) and processing (46%). Thus, the type of action when using the highlighter and note taker matters for which SRL process is captured with these tools. Although there is not a one-to-one mapping of action and SRL process, the picture is clearer when taking the action type into account.

3.3 Relations between tools

The use of the timer, search function, and planner were expected to be related, because they shared the SRL processes they captured. Indeed, we found that frequency of search and planner use was positively correlated. In addition, use of the timer was positively correlated with the planner and its relation with search approached significance. The action types within one tool were positively correlated, see highlighter and note taker. The correlations of note

Table 6: Spearman's Correlations Between Tools

	1	2	2a	2b	2c	3	3a	3b	4	5
1. Timer	1									
2. Highlighter	.117	1								
2a. Creating	.084	.983***	1							
2b. Labeling	.192	.447**	.379**	1						
2c. Reading	.159	.679***	.606***	.399**	1					
3. Note taker	.137	.071	.045	.080	.098	1				
3a. Creating	.142	.062	.041	.067	.073	.994***	1			
3b. Reading	.106	.053	.019	.121	.161	.819***	.777***	1		
4. Search	.284†	.019	.015	-.091	.058	.155	.177	-.005	1	
5. Planner	.310*	.003	.022	.179	-.109	.231	.263†	.165	.326*	1

Note. ***$p < .001$, **$p < .01$, *$p < .05$, † $p < .10$

creating and planner approached significance. Other relations were not significant, see Table 6

4 DISCUSSION

Measuring SRL processes is critical to understand students' SRL and help them engage in a more productive SRL processing which remains a challenge to many students [5]. The aim of the present study was to improve measurement of SRL by embedding instrumentation tools in a learning environment and validating the measurement of SRL with these instrumentation tools using think aloud. Instrumentation tools were used to a different extent, with the highlighter being used most often and the search function least often. We showed that some tools clearly capture a single SRL process (i.e., timer capturing monitoring), while other tools capture a range of SRL processes (e.g., highlighter and note taker capturing superficial writing down, organization, and monitoring). Furthermore, by investigating the action that was performed when using the instrumentation tool (e.g., creating vs reading highlights), we were able to specify which SRL process occurred during which action resulting in a clearer picture of which SRL processes occurred during tool use (e.g., superficial writing down when creating highlights, but monitoring when reading highlights). Finally, frequencies of the timer, search, and planner usage were positively correlated; in line with the shared SRL processes they captured.

4.1 Frequency of Instrumentation Tool use

Tools were used to a different extent. The timer, highlighter, and note taker were used more often than the search and planner. There appears to be no previous studies that recorded trace data about time-checking. Our results showed that the timer was used often and it was used by all students. Regarding the highlighter and note taker, they were used often, which is in line with previous studies [9, 25]. The search function and planner were hardly used, as indicated by a low average and maximum frequency of use. An explanation for the low usage of search is that the informative texts were already structured and they were limited to three topics, which might have reduced the need for goal-directed search behavior. An explanation for the low usage of the planner is that the learning task lasted 45 minutes, which would reduce the need for a planner. In general, planning behavior appears to be relatively infrequent [7].

4.2 SRL During Instrumentation Tool use

We used think aloud to validate SRL processes occurring during instrumentation tool use. Generally, our hypotheses were confirmed about which SRL process was associated with which tool. The instrumentation tools co-occurred with SRL processes more often than would be expected based on the overall occurrence, with one exception. Monitoring was relatively frequent overall, and this was also the case during use of the highlighter and note taker.

The *timer* was almost exclusively used for monitoring, which confirmed our first hypothesis. This means that the timer is an instrumentation tool that can be used to assess monitoring of time left and/or time elapsed. Time left helps to evaluate progress in relation to the set goals [42].

The *highlighter* was associated with superficial writing down (low cognition), which was especially the case when *creating* highlights. However, the process of highlight *labeling*, in which students also created a label for the highlight, was mostly associated with organization (high cognition). Furthermore, highlight *reading* was not associated with either of these two processes, but it was associated with monitoring (metacognition) and processing (low cognition). First, to explain the result of highlight creating and labeling, superficial writing down is an indicator of copying and selecting text without additional processing of the information. This indicates that the highlighter was used for memorizing information (low cognition). However, the result of highlight labeling, revealed that highlights were used to organize information while storing it externally in the highlighter tool. Thus, both memorizing and external storage have been found [12]. The present results add that to be better able to differentiate between these two processes, it helps to take into account how the highlighter was used: creating a highlight without generating a label to be associated with the highlight or labeling (creating a highlight and generating a label to be associated with the highlight). Second, highlight reading, revealed a different pattern of SRL processes, as it was associated with monitoring and processing. Reading of a highlight means the information was already processed at least once, because the highlight was already stored. Monitoring might indicate students compare information from the text with the information in the highlight [46]. This can take various forms, for example, the highlight can serve as knowledge, which is used to create or evaluate a product: "how can I use the information in my highlight to write my essay?" [40]. Processing, on the other hand, is a low cognitive activity, which indicates rereading or editing of the information. Thus, in terms of highlight reading, processing might reflect simply rereading the highlight, as editing was a separate action that was also recorded (but its frequency was too low for further investigation).

The *note taker* was most often used in association with organization of information (high cognition), which was especially the case when *creating* notes. However, the process of note *reading* was associated with monitoring and processing. First, regarding note *creating*, organization was expected, because notes can be used to process information, organize it, and possibly integrate it in long-term memory [20], or use it later in the essay. Second, with respect to note *reading*, the results are the same as for highlight reading, monitoring and processing co-occurred, which thus can be explained in the same way: monitoring qualities of information, such as its relevance in relation to learning goals [46], or processing the note by rereading its content.

The *search* function was mostly used for planning and otherwise for monitoring. Planning was expected, as a search often is goal-directed behavior to reach a (temporarily) set goal [15]. Planning, in this case, is the intention to find (more) information about a specific topic. This can also explain monitoring during search, because whether (more) information is desired depends on the gap between what is present (knowledge) and what is desired (standard) [40].

The *planner* revealed the same pattern as search: it was mostly used for planning and otherwise for monitoring. Planning when using a planner was expected, because a planner helps to create a plan by providing a visual overview and elements to create a planning. Often, planning is conceptualized as establishing subgoals before engaging in a task [40], which likely was the case when using the planner. However, the goal-directed behavior, such as when using the search function, has also been conceptualized as planning [15]. This distinction can explain that the percentage of planning was larger when using search compared to the planner, because it was a different type of planning. A direct comparison would be interesting, but even when aggregating these types of planning, its frequency was low, in line with previous research [7]. While establishing subgoals might affect the whole learning process and its unfolding over time [7], goal-direct behavior might be more relevant as a strategy that can be employed [15]. In addition to planning, monitoring also co-occurred during planner use. Monitoring was expected to monitor the actual progress with the intended progress in the planner [40].

Taken together, instrumentation tools help to measure different SRL processes when student learn in a digital learning environment [30]. The timer captured one SRL process, but other tools were associated with a range of SRL processes, which could be more specifically determined by taking the type of action into account, such as note creating vs reading.

4.3 Instrumentation Tools: Interrelations

The frequency of the use of the timer correlated positively with that of the planner and the correlation with search approached significance. The frequencies of the use of planner and search also correlated positively. An explanation for this finding is that the use of the timer, search, and planner coincided with the same SRL processes: planning and monitoring. This suggests that one of the strategies in this present learning task was to complete the essay task in a goal-directed manner: create a plan, focus on relevant aspects, and monitor progress. Such a strategy might be more common based on the time pressure introduced in the present study (45 minutes to complete the task), but we did not compare different deadlines, such as 60 or 75 minutes, in this study. Other correlations were not significant. We expected the use of highlights and notes to be correlated [9], but the correlation was around zero. An explanation for this insignificant correlation is that a different task is that a different task was used, which requires different learning processes. In the present study, the use of highlights was mainly associated with low cognition, the use of notes was mainly associated with high cognition. This suggests that highlights and notes were used differently and for different purposes. This can explain the lack of correlation.

4.4 Limitations and Design Implications

Overall, our instrumentation tools were fairly specific and consistent in which SRL processes were captured: meaning that one or two SRL processes were sufficient to explain at least 63% of the tool usage. Although the total sample constitutes of 44 participants, we have analyzed 860 interactions with the tools and a total of 38,701 utterances in think aloud data, which contained 36870 SRL processes. This revealed medium-sized effects and findings in line with theoretical assumptions and hypotheses, which justifies our sample size. It can, therefore, be recommended to add instrumentation tools to a learning environment to capture relevant SRL processes in log data. To further improve capturing of SRL with instrumentation tools, we analyzed the type of action in addition to which tool was used. We suggest to take into account how students use a tool to be better able to interpret the associated learning process. There still is room for improvement in capturing SRL with instrumentation tools. One suggestion is to analyze the text that was selected to create a highlight or note. This can provide information into a student's knowledge and beliefs, to study the role of knowledge and beliefs in SRL [40]. The affordances of a task may affect how tools are used. In the present study, the assignment was to write an essay, but tool use has been studied in other contexts, e.g., [9]. Therefore, interpretations about learning processes co-occurring during tool use can be improved if effects of task affordances on tool use are known. It can be speculated, for example, that the role of integrating information into long-term memory [20] when using the note taker would be more prominent in learning-focused assignments compared to production-focused assignments. However, it can be hypothesized that our findings can be generalized to other situations, because by capturing SRL processes, we captured processes that allow transfer of knowledge and skills across domains of subject matter [40]. To collect more specific trace data about the search and choose whether the aim of using the search function aligned with planning or monitoring, a suggestion is to take into account prior knowledge [36] and what was read so far. If there is little knowledge about a topic, then a search might be a better indicator of planning than monitoring.

Each data source has its advantages and disadvantages in capturing SRL [2] and this applies to the present study as well. Think aloud does not capture everything, for example for automatized processes, people struggling with think aloud, or timing effects: stating what you were doing instead of what you are doing. For trace data other limitations apply, such as the limitation to capture SRL as interaction with the learning environment. However, as our expectations about SRL processes during tool use were largely confirmed, the combination of different data sources can be used to validate inferences from different data sources [45]. More specifically, our assumptions about co-occurring SRL processes during tool use were confirmed by using think aloud, which supports the proposition that think aloud can be used as a ground truth to investigate other data sources. It would be interesting to follow the development of log data analysis [31] and think about how log data might be used to replace or complement think aloud [2]. Furthermore, process mining might be used to detect patterns of co-occurring SRL processes during instrumentation tool use. For such studies, it would help to identify a time window in which multiple SRL processes occur, because, in some instances, we did not detect any co-occurring SRL process.

5 CONCLUSION

To conclude, instrumentation tools help to capture SRL and this was validated with think aloud data in this study. Although SRL can be hard to detect [30], we detected relevant SRL processes, including planning and high cognition, which are examples of SRL processes that are hard to detect and that appear to be relatively infrequent. A timer can be very specific (monitoring only), but other instrumentation tools show a variety of SRL processes: planning and monitoring for search and planner. The highlighter mainly captured low cognition and the note taker mainly captured high cognition. Instrumentation tools that were associated with the same SRL processes, were also correlated. Thus, by aligning think aloud, we showed that tool use indeed reflects SRL processes. Future challenges are to collect and process log data real time with learning analytic techniques to measure ongoing SRL processes and support learners during learning with personalized SRL scaffolds.

ACKNOWLEDGMENTS

This work is a result of the FLoRA research project funded by DFG (Germany), NWO (The Netherlands), and ESRC (United Kingdom) as part of the Open Research Area (ORA) BA20144/10-1, NWO464.18.104, ES/S015701/1.

REFERENCES

[1] Roger Azevedo, Jennifer G Cromley, and Diane Seibert. 2004. Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology* 29, 3 (July 2004), 344–370. DOI:https://doi.org/10.1016/j.cedpsych.2003.09.002

[2] Roger Azevedo and Dragan Gašević. 2019. Analyzing Multimodal Multichannel Data about Self-Regulated Learning with Advanced Learning Technologies: Issues and Challenges. *Computers in Human Behavior* 96, (July 2019), 207–210. DOI:https://doi.org/10.1016/j.chb.2019.03.025

[3] Roger Azevedo, Daniel C. Moos, Jeffrey A. Greene, Fielding I. Winters, and Jennifer G. Cromley. 2008. Why is externally-facilitated regulated learning more effective than self-regulated learning with hypermedia? *Education Tech Research Dev* 56, 1 (February 2008), 45–72. DOI:https://doi.org/10.1007/s11423-007-9067-0

[4] Roger Azevedo, Daniel C. Moos, Amy M. Johnson, and Amber D. Chauncey. 2010. Measuring Cognitive and Metacognitive Regulatory Processes During Hypermedia Learning: Issues and Challenges. *Educational Psychologist* 45, 4 (October 2010), 210–223. DOI:https://doi.org/10.1080/00461520.2010.515934

[5] Roger Azevedo, Michelle Taub, and Nicholas V. Mudrick. 2018. Understanding and Reasoning About Real-Time Cognitive, Affective, and Metacognitive Processes to Foster Self-Regulation With Advanced Learning Technologies. In *Handbook of Self-Regulation of Learning and Performance* (Second edition), Dale H. Schunk and Jeffrey Alan Greene (eds.). Routledge, Taylor & Francis Group, New York, NY.

[6] Maria Bannert. 2007. Metakognition beim Lernen mit Hypermedia. Erfassung, Beschreibung und Vermittlung wirksamer metakognitiver Lernstrategien und Regulationsaktivitäten. Waxmann, Münster, Germany.

[7] Maria Bannert, Peter Reimann, and Christoph Sonnenberg. 2014. Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition Learning* 9, 2 (August 2014), 161–185. DOI:https://doi.org/10.1007/s11409-013-9107-6

[8] Matthew L. Bernacki. 2018. Examining the Cyclical, Loosely Sequenced, and Contingent Features of Self-Regulated Learning: Trace Data and Their Analysis. In *Handbook of Self-Regulation of Learning and Performance* (Second edition), Dale H. Schunk and Jeffrey Alan Greene (eds.). Routledge, Taylor & Francis Group, New York, NY.

[9] Matthew L. Bernacki, James P. Byrnes, and Jennifer G. Cromley. 2012. The effects of achievement goals and self-regulated learning behaviors on reading comprehension in technology-enhanced learning environments. *Contemporary Educational Psychology* 37, 2 (April 2012), 148–161. DOI:https://doi.org/10.1016/j.cedpsych.2011.12.001

[10] Ivar Bråten and Marit S. Samuelstuen. 2007. Measuring strategic processing: comparing task-specific self-reports to traces. *Metacognition Learning* 2, 1 (August 2007), 1–20. DOI:https://doi.org/10.1007/s11409-007-9004-y

[11] Geraldine Clarebout and Jan Elen. 2006. Tool use in computer-based learning environments: towards a research framework. *Computers in Human Behavior* 22, 3 (May 2006), 389–411. DOI:https://doi.org/10.1016/j.chb.2004.09.007

[12] Francis J. Di Vesta and G. Susan Gray. 1972. Listening and note taking. *Journal of Educational Psychology* 63, 1 (1972), 8–14. DOI:https://doi.org/10.1037/h0032243

[13] Anastasia Efklides, Bennett L. Schwartz, and Victoria Brown. 2018. Motivation and Affect in Self-Regulated Learning: Does Metacognition Play a Role? In *Handbook of Self-Regulation of Learning and Performance* (Second edition), Dale H. Schunk and Jeffrey Alan Greene (eds.). Routledge, Taylor & Francis Group, New York, NY.

[14] Publications Office of the European European Union. 2019. Key competences for lifelong learning. Retrieved October 22, 2020 from http://op.europa.eu/nl/publication-detail/-/publication/297a33c8-a1f3-11e9-9d01-01aa75ed71a1

[15] Jeffrey Alan Greene and Roger Azevedo. 2009. A macro-level analysis of SRL processes and their relations to the acquisition of a sophisticated mental model of a complex system. *Contemporary Educational Psychology* 34, 1 (January 2009), 18–29. DOI:https://doi.org/10.1016/j.cedpsych.2008.05.006

[16] Henning Holle and Robert Rein. 2015. EasyDIAg: A tool for easy determination of interrater agreement. *Behav Res* 47, 3 (September 2015), 837–847. DOI:https://doi.org/10.3758/s13428-014-0506-7

[17] J. Richard Landis and Gary G. Koch. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33, 1 (March 1977), 159. DOI:https://doi.org/10.2307/2529310

[18] Micheal A. Lawrence. 2016. *ez: Easy Analysis and Visualization of Factorial Experiments.*

[19] Zahia Marzouk, Mladen Rakovic, Amna Liaqat, Jovita Vytasek, Donya Samadi, Jason Stewart-Alonso, Ilana Ram, Sonya Woloshen, Philip H Winne, and John C Nesbit. 2016. What if learning analytics were based on learning science? *AJET* 32, 6 (December 2016). DOI:https://doi.org/10.14742/ajet.3058

[20] Richard E. Mayer. 1996. Learning strategies for making sense out of expository text: The SOI model for guiding three cognitive processes in knowledge construction. *Educ Psychol Rev* 8, 4 (December 1996), 357–371. DOI:https://doi.org/10.1007/BF01463939

[21] Inge Molenaar, Carla A. M. van Boxtel, and Peter J. C. Sleegers. 2011. Metacognitive scaffolding in an innovative learning arrangement. *Instr Sci* 39, 6 (November 2011), 785–803. DOI:https://doi.org/10.1007/s11251-010-9154-1

[22] Inge Molenaar, Carla A.M. van Boxtel, and Peter J.C. Sleegers. 2010. The effects of scaffolding metacognitive activities in small groups. *Computers in Human Behavior* 26, 6 (November 2010), 1727–1738. DOI:https://doi.org/10.1016/j.chb.2010.06.022

[23] Inge Molenaar and Ming Ming Chiu. 2014. Dissecting sequences of regulation and cognition: statistical discourse analysis of primary school children's collaborative learning. *Metacognition Learning* 9, 2 (August 2014), 137–160. DOI:https://doi.org/10.1007/s11409-013-9105-8

[24] Inge Molenaar and Ming Ming Chiu. 2017. Effects of Sequences of Cognitions on Group Performance Over Time. *Small Group Research* 48, 2 (April 2017), 131–164. DOI:https://doi.org/10.1177/1046496416689710

[25] John C. Nesbit, Philip H. Winne, Dianne Jamieson-Noel, Jillianne Code, Mingming Zhou, Ken Mac Allister, Sharon Bratt, Wei Wang, and Allyson Hadwin. 2006. Using Cognitive Tools in Gstudy to Investigate How Study Activities Covary with Achievement Goals. *Journal of Educational Computing Research* 35, 4 (December 2006), 339–358. DOI:https://doi.org/10.2190/H3W1-8321-1260-1443

[26] Ernesto Panadero. 2017. A Review of Self-regulated Learning: Six Models and Four Directions for Research. *Front. Psychol.* 8, (April 2017), 422. DOI:https://doi.org/10.3389/fpsyg.2017.00422

[27] Zacharoula Papamitsiou and Anastasios A. Economides. 2014. Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence. *Journal of Educational Technology & Society* 17, 4 (2014), 49–64.

[28] R Core Team. 2020. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.R-project.org/

[29] William Revelle. 2020. psych: Procedures for Psychological, Psychometric, and Personality Research.

[30] Ido Roll and Philip H. Winne. 2015. Understanding, evaluating, and supporting self-regulated learning using learning analytics. *JLA* 2, 1 (2015), 7–12. DOI:https://doi.org/10.18608/jla.2015.21.2

[31] J. Saint, A. Whitelock-Wainwright, D. Gasevic, and A. Pardo. 2020. Trace-SRL: A Framework for Analysis of Micro-Level Processes of Self-Regulated Learning from Trace Data. *IEEE Transactions on Learning Technologies* (2020), 1–1. DOI:https://doi.org/10.1109/TLT.2020.3027496

[32] Dale H. Schunk and Jeffrey Alan Greene (Eds.). 2018. *Handbook of Self-Regulation of Learning and Performance* (Second edition ed.). Routledge, Taylor & Francis Group, New York, NY.

[33] Dale H. Schunk and Jeffrey Alan Greene. 2018. Historical, Contemporary, and Future Perspectives on Self-Regulated Learning and Performance. In *Handbook of Self-Regulation of Learning and Performance* (Second edition), Dale H. Schunk and Jeffrey Alan Greene (eds.). Routledge, Taylor & Francis Group, New York, NY.

[34] Melody Siadaty, Dragan Gašević, and Marek Hatala. 2016. Trace-Based Micro-analytic Measurement of Self-Regulated Learning Processes. *JLA* (April 2016), 183–214. DOI:https://doi.org/10.18608/jla.2016.31.11

[35] Michelle Taub and Roger Azevedo. 2019. How Does Prior Knowledge Influence Eye Fixations and Sequences of Cognitive and Metacognitive SRL Processes during Learning with an Intelligent Tutoring System? *Int J Artif Intell Educ* 29, 1 (March 2019), 1–28. DOI:https://doi.org/10.1007/s40593-018-0165-4

[36] Michelle Taub, Roger Azevedo, François Bouchet, and Babak Khosravifar. 2014. Can the use of cognitive and metacognitive self-regulated learning strategies be predicted by learners' levels of prior knowledge in hypermedia-learning environments? *Computers in Human Behavior* 39, (October 2014), 356–367. DOI:https://doi.org/10.1016/j.chb.2014.07.018

[37] The Language Archive. 2020. *ELAN.* Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands. Retrieved from https://archive.mpi.nl/tla/elan

[38] Marcel V.J. Veenman. 2013. Assessing Metacognitive Skills in Computerized Learning Environments. In *International Handbook of Metacognition and Learning Technologies*, Roger Azevedo and Vincent Aleven (eds.). Springer New York, New York, NY.

[39] Philip H. Winne. 1996. A metacognitive view of individual differences in self-regulated learning. *Learning and Individual Differences* 8, 4 (January 1996), 327–353. DOI:https://doi.org/10.1016/S1041-6080(96)90022-9

[40] Philip H Winne. 1997. Experimenting to Bootstrap Self-Regulated Learning. *Journal of Educational Psychology* 89, 3 (1997), 14. DOI:https://doi.org/10.1037/0022-0663.89.3.397

[41] Philip H. Winne. 2006. How Software Technologies Can Improve Research on Learning and Bolster School Reform. *Educational Psychologist* 41, 1 (March 2006), 5–17. DOI:https://doi.org/10.1207/s15326985ep4101_3

[42] Philip H. Winne. 2017. Learning Analytics for Self-Regulated Learning. In *Handbook of Learning Analytics* (First), Charles Lang, George Siemens, Alyssa Wise and Dragan Gasevic (eds.). Society for Learning Analytics Research (SoLAR), 241–249. DOI:https://doi.org/10.18608/hla17.021

[43] Philip H. Winne. 2018. Theorizing and researching levels of processing in self-regulated learning. *British Journal of Educational Psychology* 88, 1 (2018), 9–20. DOI:https://doi.org/10.1111/bjep.12173

[44] Philip H. Winne. 2018. Cognition and Metacognition Within Self-Regulated Learning. In *Handbook of Self-Regulation of Learning and Performance* (Second edition), Dale H. Schunk and Jeffrey Alan Greene (eds.). Routledge, Taylor & Francis Group, New York, NY.

[45] Philip H. Winne. 2020. Construct and consequential validity for learning analytics based on trace data. *Computers in Human Behavior* 112, (November 2020), 106457. DOI:https://doi.org/10.1016/j.chb.2020.106457

[46] Philip H. Winne, Kenny Teng, Daniel Chang, Michael Pin-Chuan Lin, Zahia Marzouk, John C. Nesbit, Alexandra Patzak, Mladen Rakovic, Donya Samadi, and Jovita Vytasek. 2019. nStudy: Software for Learning Analytics about Learning Processes and Self-Regulated Learning. *JLA* 6, 2 (August 2019), 95–106. DOI:https://doi.org/10.18608/jla.2019.62.7