Received: 13 January 2022 | Revised: 7 February 2023 | Accepted: 14 February 2023

DOI: 10.1111/jcal.12801

ARTICLE

Journal of Computer Assisted Learning WILEY

Towards a fuller picture: Triangulation and integration of the measurement of self-regulated learning based on trace and think aloud data

Yizhou Fan[1] | Mladen Rakovic[2] | Joep van der Graaf[3] | Lyn Lim[4] | Shaveen Singh[2] | Johanna Moore[5] | Inge Molenaar[3] | Maria Bannert[4] | Dragan Gašević[2,5,6]

[1]Graduate School of Education, Peking University, Beijing, China

[2]Faculty of Information Technology, Monash University, Clayton, Victoria, Australia

[3]Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands

[4]TUM School of Education, Technical University of Munich, Munich, Germany

[5]School of Informatics, University of Edinburgh, Edinburgh, United Kingdom

[6]Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia

Correspondence
Yizhou Fan, Graduate School of Education, Peking University, Beijing 100871, China.
Email: fyz@pku.edu.cn

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Number: BA20144/10-1; Nederlandse Organisatie voor Wetenschappelijk Onderzoek, Grant/Award Number: NWO464.18.104; the Economic and Social Research Council of the United Kingdom, Grant/Award Number: ES/S015701/1

Abstract

Background: Many learners struggle to productively self-regulate their learning. To support the learners' self-regulated learning (SRL) and boost their achievement, it is essential to understand the cognitive and metacognitive processes that underlie SRL. To measure these processes, contemporary SRL researchers have largely utilized think aloud or trace data, however, not without challenges.

Objectives: In this paper, we present the findings of a study that investigated how concurrent analysis and integration of think aloud and trace data could advance the measurement of SRL and assist in better understanding the mechanisms of SRL processes, especially those details that remain obscured by observing each data channel individually.

Methods: We concurrently collected think aloud and trace data generated by 44 university students in a laboratory setting and analysed those data relative to the same timeline.

Results: We found that the two data channels could be interchangeably used to measure SRL processes for only 17.18% of all the time segments identified in a learning task. Moreover, SRL processes for around 45% of all the time segments could be detected via either trace data or think aloud data. For another 27.17% of all the time segments, different SRL processes were detected in both data channels.

Conclusions: Our results largely suggest that the two data collection methods can be used to complement each other in measuring SRL. In particular, we found that think aloud and trace data could provide different perspectives on SRL. The integration of the two methods further allowed us to reveal a more complex and more comprehensive temporal associations among SRL processes compared to using a single data collection method. In future research, the integrated measurement of SRL can be used to improve the detection of SRL processes and provide a fuller picture of SRL.

KEYWORDS
eye-tracking data, learning analytic, process mining, self-regulated learning, think aloud, trace data

1 | INTRODUCTION

To adapt to dynamic changes in modern society, students and professionals are commonly required to maintain and extend their knowledge bases and skill sets. To this end, they often need to study on their own and quickly acquire new knowledge and skills that emerge as requirements for jobs. Their ability to productively self-regulate learning, that is, engaging in conscious and goal-oriented learning processes (Winne & Hadwin, 1998), is critical to ensure success in different learning tasks. Researchers have consistently shown that productive engagement in self-regulated learning (SRL) improves learning performance and motivation (Broadbent & Poon, 2015; Dent & Koenka, 2016; Rienties et al., 2019). For this reason, the development of SRL skills represents a critical aim for all learners involved in contemporary education.

Regardless of the theorized significance and documented benefits of SRL, many learners struggle to productively self-regulate their learning (Bjork et al., 2013). To provide these learners with the required support for SRL and boost their achievement, it is essential to understand the cognitive and metacognitive processes that underlie SRL (Winne & Jamieson-Noel, 2002). Studying these processes is, however, considered a challenging endeavour as they influence each other and dynamically change throughout a learning session, that is, SRL processes are sequential and temporal by nature (Bannert et al., 2014; Molenaar & Järvelä, 2014).

Over the last three decades, different methods for the analysis of the processes involved in SRL have been proposed. These methods can be grouped into categories such as self-report inventories, think aloud protocols, and traces of learning behaviours (Winne, 2010). However, each method comprises a considerable number of challenges. For example, data collected via self-report inventories asking learners to recall their prior learning activities and experiences are not consistently associated with actual SRL processes (Hadwin et al., 2007; Veenman, 2007). One of the many reasons is that a learner may not be able to retrieve all the details from memory after completing a learning session. Another method, think aloud protocols, was designed to document SRL processes from learners in verbal form and in real-time. Despite the considerable results that researchers have achieved using this approach, for example, in identifying SRL processes that differentiate between low- and high-achieving learners (Azevedo et al., 2004b; Bannert et al., 2014) and in evaluating the effects of external SRL support (Azevedo et al., 2004a, 2007; Sonnenberg & Bannert, 2016), think aloud protocols are deemed limited and challenging, with abundant data being generated that is difficult and time-consuming to analyse (Aleven et al., 2010; Eccles & Arsal, 2017; Greene & Azevedo, 2009; Ramey et al., 2006; Veenman, 2013; Winne, 2010; Young, 2005).

Recent advances in computational methods for data collection and analysis opened up new opportunities to improve the measurement of SRL, namely by collecting and analysing trace data, that is, data that includes navigational logs, eye and mouse movements, and keystrokes that learners generate when they learn online. Trace data are collected unobtrusively in computer-based learning environments and analysed to approximate cognitive and metacognitive processes learners engage in (Fan et al., 2021; Saint et al., 2020; Siadaty et al., 2016a, 2016c). For example, based on MOOC learners' navigational logs, Matcha et al. (2019) detected different learning tactics and learning strategies by identifying clusters of similar SRL behaviours of learners; Rakovic et al. (2022) unobtrusively measured metacognitive processes using learners' trace data during a multi-source writing task (e.g., mouse clicks and keyboard strokes) and used features extracted from these trace data to investigate how SRL processes are associated with the quality of a written product; Fan et al. (2022) used eye-tracking, mouse movement, and keystroke data as additional data channels to improve the recording of learning actions (e.g., learner's interactions with certain learning tools) and detect SRL processes (e.g., monitoring and evaluating processes during learning) while enhancing the granularity of the measurement; Fan et al. (2022) concluded that eye-tracking data is particularly valuable for measuring and extracting SRL processes.

However, associating specific SRL processes with raw trace data represents a formidable challenge (Bernacki, 2018). Hence, data collected using other methods, in particular, think aloud protocols that have been shown to more accurately reflect SRL processes than survey instruments (Beishuizen et al., 1999; De Backer et al., 2012; Rovers et al., 2019; Veenman, 2013), should be used to support analysis of trace data. The triangulation and integration of multiple data channels and methods have important implications for the measurement and cross-validation of SRL (Saint et al., 2022), however, studies entertaining this approach are very rare. Only a few studies combined or integrated multiple data channels. For instance, Sobocinski et al. (2017) categorized learners' log files and then coded video data from each learning session to explore temporal sequences of SRL phases; Munshi et al. (2018) combined observational data of student affect with log files of student interactions in the Betty's Brain environment to model learners' cognitive and affective states; Järvelä et al. (2019) triangulated heart rate, electrodermal and facial expression data to advance understanding of self-regulation in collaborative learning; and Mudrick et al. (2019) integrated and aligned eye-tracking data with learners' self-reported metacognitive judgements to understand learning processes and meta-comprehension. However, to our knowledge, no study has yet been conducted to examine how think aloud and behavioural trace data can be analysed relative to the same theoretical framework, and also how these data channels can further be aligned and integrated relative to the same timeline, towards improved understanding of SRL processes.

In this study, we set out to investigate how the triangulation and integration of think aloud and trace data can advance the measurement of SRL and assist in a better understanding of the mechanisms of SRL processes, especially those details that remain obscured by observing each data channel individually. We concurrently collected think aloud and trace data generated by university students in a laboratory setting and analysed those data relative to the same timeline. Overall, our results suggested that using a single method to measure SRL processes can often reveal SRL processes partially and the integration of two methods can be a more helpful approach to obtaining a more comprehensive picture of SRL. More specifically, our study proposed a novel approach to align and integrate think aloud and trace data, and revealed what each data channel could contribute to measuring SRL processes. In order to deepen our understanding of the comprehensive models of SRL measured using integrated data channels, we also used a process mining technique to analyse and compare the sequential and temporal associations among SRL processes. We demonstrated that the SRL process map based upon integrated think aloud and trace data channels is a more complex, more complete, and more informative SRL process map compared to process maps that are constructed based on individual data channels.

2 | RELATED WORK

2.1 | Measuring SRL with think aloud data

Studies using think aloud protocols ask participants to verbalize their thoughts while remaining engaged in the experimental task (Ericsson & Simon, 1984). These thoughts can be generated from immediate cognitive and metacognitive processes (e.g., comprehending a definition in a book chapter, or specifying learning goals) or recalled from long term memory (e.g., restating a previously learned definition). Information verbalized in such a manner should provide a window into the learner's thought process and more accurately capture dynamically changing learning processes compared to information collected via self-report questionnaires (De Backer et al., 2012; Rovers et al., 2019). For this reason, many researchers who study learning processes, including those interested in SRL, have developed different think aloud protocols and analysed collected utterances on a number of learning constructs, for example, cognitive, metacognitive, affective, to empirically test theoretical assumptions about learning (Greene & Azevedo, 2009; Pollard et al., 2019; Veenman, 2007).

Think aloud protocols have been adopted in a group of studies to overcome limitations of self-report surveys in assessing learners' self-regulatory processes. Azevedo et al. (2004b) analysed lengthy think aloud protocols to examine SRL processes that distinguished learners who demonstrated a deep conceptual understanding of a science topic they studied from learners who did not. Azevedo et al. (2004b) also found that high-achievers more extensively used the key SRL processes (task understanding, planning, monitoring, effective strategy use) throughout the learning session when compared to their lower-achieving counterparts. The think aloud data collected in the Azevedo, Cromley, and Seibert (2004) study indicated that learners who received SRL training starting to learn about complex scientific topics enacted the key SRL processes more often throughout the learning session, and also demonstrated deeper conceptual understanding at post-test compared to their colleagues in the control condition. Another study examined differences between learners who regulated their learning all by themselves and learners who were provided with adaptive external support on their self-regulation in an online learning environment (Azevedo et al., 2007). By analysing think aloud data, the authors found that externally supported learners exercised the key SRL processes more frequently and achieved higher learning gains than their unsupported counterparts. Bannert et al. (2014) examined learners' think aloud protocols not only to assess differences in frequencies of SRL processes between high and low achieving learners, but also to understand the temporal patterns of SRL processes. Their findings suggested that compared to their less productive counterparts, productive self-regulated learners spent more time in preparing activities (e.g., orientation and planning) before they engaged with the learning material. Furthermore, productive self-regulated learners elaborated on information more thoroughly and constantly monitored and evaluated their learning activities throughout learning sessions. Sonnenberg and Bannert (2016) utilized a think aloud protocol to reveal learners' learning activities after receiving metacognitive prompts designed to elicit SRL processing. The think aloud protocol used in their study was not only useful in revealing differences between effective and non-effective metacognitive prompts, but also in shading light on prompt-induced learning activities that positively correlated with performance.

In summary, these previous studies demonstrated that think aloud protocols can reliably reflect learners' SRL processes and can accurately predict learning performance. Importantly, think aloud protocols can provide a deeper insight into the effects of instructional support SRL. However, three major challenges remain when using a think aloud method. First, the procedure required when using think aloud protocols might affect the validity of certain measured constructs. For example, if the procedures of think aloud “entail describing or explaining thoughts and actions are significantly reactive” (Fox et al., 2011, p. 316), the think aloud approach might lead to higher learning performance compared to silent control conditions (Fox et al., 2011). Engagement in think aloud can therefore influence processes learners enact, that is, asking learners to explain and reflect on what they think might lead them to exercise some processes (e.g., metacognitive monitoring) more than usual (Bannert & Mengelkamp, 2008; Eccles & Arsal, 2017; Greene & Azevedo, 2009; Ramey et al., 2006). Second, ahead of a learning task, researchers often need to train participants how to think aloud and during the learning task, researchers need to carefully observe the think aloud process to avoid interruptions when collecting data (e.g., by reminding participants to say their thoughts), as the moments of silence are hard to analyse and interpret from think aloud protocols. Even though researchers do their best to ensure the participant's think aloud process is not interrupted, it is difficult to completely avoid moments of silence in think aloud studies (McDonald & Petrie, 2013). Third, data collected using think aloud protocols is typically extensive and linguistically diverse; the analysis not only requires substantial human labour to segment and label verbal protocols, but also the results can be subject to a coder's bias, as participants articulate their thoughts in different ways, often in a form that cannot straightforwardly be labelled with one distinct category from the code book (Young, 2005).

2.2 | Measuring SRL with trace data

Researchers also collect data as digital traces of learners' interactions with online learning environments. Researchers often link these data to different theoretical constructs, including those of SRL (Bernacki, 2018; Winne, 2011, 2020), and perform analyses to deepen their understanding of learning processes. Over the past decades, several studies have shown that trace data can be successfully utilized to examine different elements of SRL, for example, effective and ineffective help-seeking strategies (Aleven et al., 2006), frequency and interdependence of study tactics (Hadwin et al., 2007), time learners spent on different learning strategies (Azevedo et al., 2009) and, more recently, fine-grained characteristics of learning tactics for example, (Saint et al., 2020b) and impact of instrumentation tools and computer-based scaffolds on SRL (van der Graaf et al., 2021; Siadaty, Gašević, & Hatala, 2016b). In most of these studies, trace data have been extracted from computer logs in a form of navigation behaviour (Kinnebrew et al., 2013a), for example, moving between pages and accessing digital learning resources. Recently, new trace data channels, for example, peripheral data originated from mouse and keyboard use and eye-tracking data (Di Mitri et al., 2017; Giannakos et al., 2019; Mudrick et al., 2019; Song et al., 2021; Stark et al., 2018), have been used to capture learner activity at finer levels of granularity, aiming to provide a more detailed account of enacted SRL processes (Azevedo & Gašević, 2019; Järvelä & Bannert, 2019; Paans et al., 2019).To make trace data useful for better understanding of SRL processes, it is essential to clearly operationalize and map these data to theoretical models of SRL (Siadaty et al., 2016a). Trace data, therefore, can be enriched with the theoretical meta-information to approximate cognitive and metacognitive processes of SRL. To this purpose, several protocols for measuring SRL processes with trace data have been developed to date, for example, (Fan et al. 2021; Fan, Lim et al. 2022 Greene & Azevedo, 2009; Saint et al., 2021; Saint, Gašević, et al., 2020; Siadaty et al., 2016a, 2016c). Typically, the protocols include a theoretical framework as a reference, a coding scheme to identify and label SRL processes (e.g., orientation and monitoring), and a trace parser to convert raw log data into learning events and create more meaningful sequences of actions and SRL processes, that is, action and process libraries. For example, the process library developed by Fan et al. (2021) proposes 15 sequences of actions, and these sequences can be mapped to five SRL processes. Further, a sequence "Reading -> Annotation_Confusing" resembles a small learning episode that involves reading over the textbook chapter and then annotating information in that chapter as confusing. This pattern indicates that a learner monitored their prior knowledge of a domain and identified a gap, the process associated with the task understanding stage of Winne and Hadwin's model of SRL (Winne & Hadwin, 1998).

The trace-based method is generally considered a more accurate approach in measuring SRL when compared to the self-report methods (Gasevic et al., 2017; Zhou & Winne, 2012), because trace data are collected dynamically and unobtrusively while learners naturally interact with the learning environment, that is, the data collection procedure typically does not influence the learning processes. Despite all the promise of the trace data approach, accomplishing accurate mapping between trace data and SRL processes (Winne, 2010) is often a challenge for researchers. Inaccuracies are mostly the result of researchers' multiple inferences about why learners interact with objects in online learning platforms the way they do (Winne, 2010). For example, the previously mentioned "Reading -> Annotation_Confusing" sequence in the Fan et al. (2021) study can be categorized as monitoring for prior knowledge; however, it is possible that a learner annotates a chunk of text as confusing for a different reason, for example, multiple grammatical errors in the text, which would not be monitoring for prior knowledge.

2.3 | The current study

According to Bannert (2007), SRL processes can be grouped into metacognitive, cognitive, and other (motivation and procedure-related). This theoretical work informed the development of the coding scheme that has been used in several studies to measure SRL based on think aloud data (e.g., Bannert, 2007; Engelmann and Bannert, 2019; Sonnenberg and Bannert, 2015) or based on trace data (e.g., Fan, Lim et al. 2022; Fan, van der Graaf et al. 2022). As shown in Table 1, the Metacognition category includes the *Orientation*, *Planning*, *Monitoring*, and *Evaluation* subcategories; the Cognition category includes the *First-reading* and *Re-reading* and *Elaboration/Organization* subcategories; and Motivational and Procedural, In this study, we also used Bannert's (2007) theoretical framework to guide our measurement, interpretation and integration of SRL since it (Table 1) can describe and define specific SRL processes which could be captured in different data channels. For example, in a reading and writing task, learners may: (i) read or re-read specific learning content (data: read aloud the content or eye gaze captured in the text reading zone) to operate on information; (ii) then, write down and elaborate on how to use this information (data: keyboard strokes and verbal expressions) to solve a problem; during which, (iii) learners may also monitor the learning process (data: mouse click on the timer) or plan what to do next (data: glance the catalogue or think aloud); and sometimes, all these processes can be (iv) accompanied by motivational expressions, such as "the task is too challenging for me to perform well".

More research is needed to validate inferences obtained from think aloud and trace data, given the methodological challenges associated with both methods (Bernacki et al., 2012; Greene & Azevedo, 2009; Winne, 2010). An earlier paper (Winne, 2010) already suggested that think aloud and trace data can complement each other and, when analysed concurrently, can provide a more valid and fuller

TABLE 1 Theoretical framework used in the current study.

Main categories	Subcategories	Codes	Definitions
Metacognition	Orientation	MC.O	Orientation on the task and learning activities; Reading of general instructions and rubrics.
	Planning	MC.P	Planning of the learning process by arranging activities and determining strategies. Proceeding to the next topic.
	Monitoring	MC.M	Monitoring and checking the learning process; checking of progress according to the instruction or plan.
	Evaluation	MC.E	Evaluation of the learning process; checking of content-wise correctness of learning activities. Saying that one's own work is correct.
Cognition	First-reading	LC.F	Reading information from the text and superficial descrition of pictorial representations.
	Re-reading	LC.R	Rereading of information in the text or figures.
	Elaboration/ Organization	HC.E/O	Elaborate by connecting content-related comments and concepts; reasoning and association. Organizing of content by creating an overview; write down information point by point; summarizing; adding information generated by oneself; and editing information by rephrasing or integrating information with prior knowledge.
Motivation/ Procedural	Motivational and Procedural issues	Other	Learners' positive/negative expressions about the task, situation, or the ability; Learners ask a researcher whether they can begin working on the task

picture of SRL than when analysed separately, which is typically the case in the SRL research to date. Moreover, Winne et al. (2010) suggested investigation into whether think-aloud and trace data are interchangeable when studying SRL, and also whether these two data channels should be integrated to complement each other and provide a fuller account of SRL. Here, by interchangeable, we refer to cases where think-aloud and trace data reveal the same SRL processes from the same time period in a learning session. By data channels that complement each other, on the other hand, we refer to cases where the SRL processes can only be revealed in one data channel, that is, in either think aloud or trace data, during the observed time period in a learning session.

Although Winne made this suggestion over a decade ago, there has been little research since then that attempted to empirically investigate the value of combining and concurrently analysing these two measurement methods following the same theoretical framework. We, therefore, set out to examine how think aloud protocols and digital trace data generated by undergraduate and graduate learners during computer-based learning can be combined and jointly analysed to deepen insight into SRL processes. Accordingly, we aim to answer the following two research questions:

> **RQ1.** To what extent can think aloud and trace data be integrated when measuring SRL processes? Specifically, to what extent can think aloud and trace data can be used interchangeably or complement each other when measuring SRL processes?

> **RQ2.** What are the differences in temporal and sequential associations between SRL processes when the SRL processes are observed based on integrated think aloud and trace data versus based on each of these two data channels individually?

3 | METHODS

3.1 | Research design and learning environment

The laboratory study was conducted at a university in the Netherlands and involved 44 participants, including 39 undergraduate and 5 graduate students, with an average age of 21.70 years ($SD = 2.99$ years). The participants declared diverse majors (e.g., psychology and communication science). During the learning session (45 min), the participants were asked to read materials (text and pictures) on three topics: (1) artificial intelligence (e.g., general information about artificial intelligence, 11 pages and about 2300 words), (2) differentiation in the classroom (e.g., how teachers deal with differences among learners, 9 pages and about 1400 words), and (3) scaffolding (e.g., providing learners with external and structured support during the learning process, 11 pages and about 1900 words). The learning session also included a multi-source writing task, that is, learners were required to integrate the three topics into a vision essay (300–400 words) that describes learning in school in 2035.We developed a technology-enhanced learning environment (TEL) for the purposes of this study. The TEL consisted of a catalogue and navigation area on the left; an area for reading and writing in the middle; instrumentation tools including annotation, planner and search on the right (see Figure 1). The learners could use the navigation area to navigate to task instructions and the essay scoring rubric. The catalogue area and search tool could be used to navigate through learning materials. The learners could also use the planner tool[1] to plan their learning session and the timer tool to monitor time left for the

[1]The planner tool can be used to specify activities a learner plans to work on and the time and duration of these activities. This tool provides learners with a timeline (e.g., 45 min) and several activity blocks, including recommended activity blocks (e.g., read the AI section or write an essay) and customizable activity blocks by learners (e.g., learners can create a new block as "evaluate the essay quality"). Learners can drag and drop these activity blocks onto the timeline and allocate corresponding time to different modules (e.g., assign 10 min at the end for essay writing).

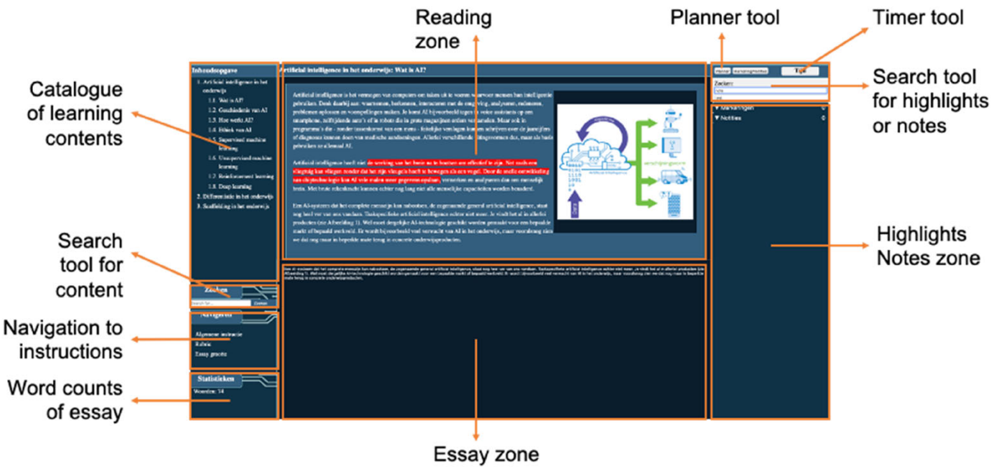


FIGURE 1 Learning environment and different functional zones.

task. The annotation tool afforded the learners the opportunity to highlight and tag parts of text, take notes or search for highlights, tags and notes they created earlier. More detailed introductions about these tools are provided by van der Graaf et al. (2021).

To collect think aloud data we utilized a webcam and a microphone. To collect trace data, we utilized an Internet capable computer with keyboard and mouse, and a screen-based eye-tracker. We presented the TEL to the participants on a 23-inch monitor (1920 × 1080 pixels) via a laptop, running the Windows 10 operating system. We used the Tobii TX300 Eye Tracker (with a sampling rate of 300 Hz) to capture eye movement during the learning session. At the outset of the study, researchers introduced the study requirements to participants and provided a short training session (15–20 min) of thinking aloud in which the participants had a chance to practice thinking aloud. The participants then engaged in the learning task using the TEL, while, at the same time, verbalizing their thoughts. Throughout the study, the researchers ensured the think aloud process unfolded with no major interruptions, for example, should they notice an extended period of silence they prompted the participants to continue talking. The **trace data** collected in this study included *navigation logs*, *keyboard strokes*, *mouse traces* (*move, click and scroll*), and *eye-tracking data* (*eye fixations*), which were stored via a local PHP-server. The **think aloud** data were collected in audio formats which were analysed and coded for SRL processes. The screen recordings were occasionally used to resolve ambiguities in audio recordings. Importantly, the trace and think aloud data channels were synchronized within the same timeline.

3.2 | Measuring SRL

3.2.1 | Theoretical framework and measuring approach

As explained in the Introduction section, we use Bannert's (2007) theoretical framework (as shown in Table 1) to guide the measurement of SRL processes in this paper. The measurement and integration approach in this study (Figure 2) involved three major groups of activities: **measuring SRL using trace data** with an **action library** and a **process library** that was used to parse the raw trace data into **SRL processes (V1)**; **measuring SRL using think aloud data** with the **think aloud coding protocol** that was used to code raw think aloud data into **SRL processes (V2)**; and the **integrated results** that included the **integration rules** we utilized to obtain the **integrated SRL processes (V3)** from the two data channels. In the following subsections, we provide more details about our measurement and integration approach.

3.2.2 | Measuring SRL using trace data

Building upon the existing literature (Fan et al., 2021, Fan, Lim et al. 2022; Saint et al., 2020a, 2020b; Siadaty et al., 2016a, 2016c), we created a trace parser to identify SRL processes from trace data. This trace parser included two components: the **action library** – responsible for labelling raw log data with meaningful learning actions; and the **process library** – responsible for detecting SRL processes from the sequences of actions. In the action library, we defined 18 action labels (Table 2) that could be mapped to individual learning actions. For example, when a learner created a note (mouse clicks and keyboard strokes) with the annotation tool, this action was labelled as NOTE_EDITING; when the learner's eye fixations were fixed on the catalogue and navigation zone, this action was labelled as NAVIGATION. We also combined different data channels to validate our labelling of learning actions. For example, if the learner's eye gaze moved within the note taking zone with the mouse cursor when typing their notes, these trace data (mouse clicks, keyboard strokes and eye fixations) were all labelled as NOTE_EDITING; however, if the learner's eye gaze moved back and forth between the note zone and the reading zone during their typing of a long note, the eye fixations back to the reading zone were labelled as READING actions. In this way, the multi-channel trace data improved the granularity and validity of action labelling.

To build a reliable and valid process library, we had initially harnessed the SRL processes based on Bannert's SRL coding protocol (Bannert, 2007), and then included additional relevant processes from

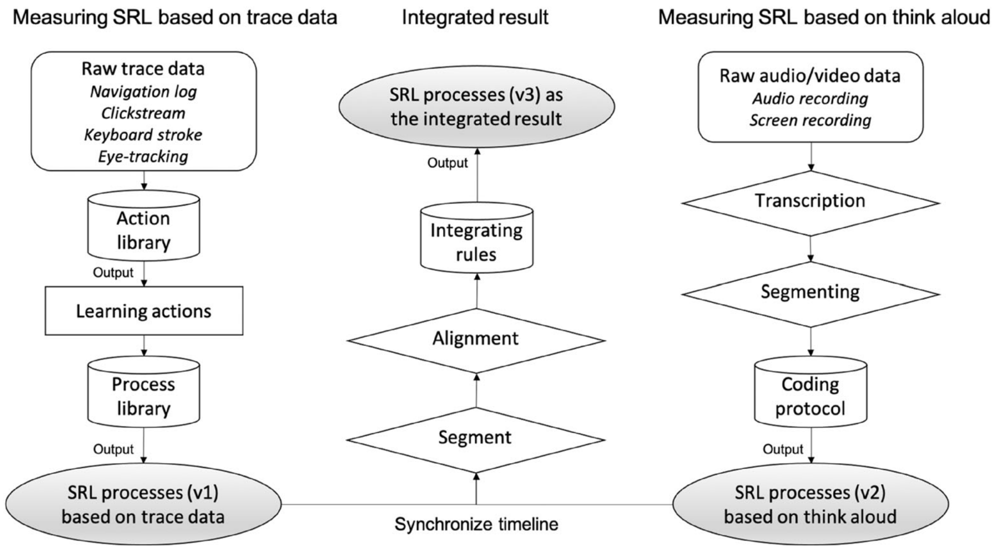


FIGURE 2 Measurement and integration approach of SRL processes.

several related studies (e.g., Fan et al., 2021; Fan, Lim et al. 2022; Kizilcec et al., 2017; Saint et al., 2021, 2020; Siadaty et al., 2016). As a result, we developed the process library shown in Table 3. The process library included three main categories and seven subcategories, encompassing 31 different sequences of actions. The three main categories were metacognition (with subcategories Orientation, Planning, Evaluation, and Monitoring), cognition (with subcategories First-reading, Re-reading and Elaboration and Organization). The sequences of actions that we interpreted as SRL processes were mapped to these categories. For example, when a learner took a note while reading the general instruction page for the task, we labelled this sequence of learning actions as GENERAL_INSTRUCTION <-> NOTE_EDITING based on the action library. Further, this sequence of actions was mapped to the Orientation category, based on our process library. In this way, we were able to parse the raw trace data into SRL processes. The segment of the above Orientation process starts with the starting timestamp of action GENERAL_INSTRUCTION and ends with the ending timestamp of action NOTE_EDITING. In cases when actions recorded in the trace data could not be mapped to any of the proposed sequences in Table 3, we labelled those actions as No_Process and did not include them into further analysis to construct SRL process maps. We provide additional details about action and process libraries in the supplemental document.

3.2.3 | Measuring SRL using think aloud data

At this stage, we identified SRL processes (V2) from think aloud data based on our theoretical framework in Table 1. We recorded learners' utterances during the learning sessions and these utterances were segmented and coded as SRL processes using a think aloud coding protocol. We first segmented the audio recordings using an automatic sound detection tool (plug-in in Audacity) into "segments with sound" that can be coded. If audio was below 26 dB, then it was considered silence, and if there was 0.30 s of silence or longer, then a segment was created. Then, three coders coded these utterances as SRL processes using our coding scheme which was based on previously developed coding schemes (Bannert, 2007; Molenaar et al., 2011). The main categories in this scheme are metacognition, cognition and motivational and procedural utterances, and No_process (non-codable utterances, for example, moments of silence or murmuring, see Table 1). We used the ELAN software (Max Planck Institute for Psycholinguistics, 2021) to code and deal with the timing of codes, because codes from different coders could overlap to different extents. For example, the coders were instructed to discuss and modify the automatically detected segments when in doubt, such as in the situations when one coder heard speech but there was no automatically detected segment and the other coder did not notice this issue. The coders reached an acceptable inter-rater reliability of $\kappa = .53–.65$ (kmax $= .81–.82$). In Table 4, we present our coding protocol with examples of learners' utterances and their corresponding think aloud codes.

3.2.4 | Aligning and integrating trace data and think aloud data

The third step of our analytical approach was to align and integrate SRL processes based on trace (V1) and SRL processes based on think aloud data (V2). First, we aligned the two sets of results on the same timeline. Next, we segmented the timeline (time unit in milliseconds) and defined five alignment situations (S1–S5). As shown in Figure 3, all starting timestamps and ending timestamps of all segments in trace data results and think aloud results were aligned in the same timeline and segmented the timeline into fine-grained segments (grey boxes). Then we assigned the alignment situations (S1-S5) to these segments according to the corresponding integration rules:

Situation 1: No_measurement occurrences; if No_process was detected in both V1 (trace data results) and V2 (think aloud results) for one segment, then assign No_process to that segment, see S1 in Figure 3;

TABLE 2 The action library for labelling learning actions.

Action labels	Action descriptions
GENERAL_INSTRUCTION	Learners read or re-read general instructions and learning goals
RUBRIC	Learners read or re-read the rubric for essay writing
RELEVANT_READING	Learners read and learn learning content for the first time
RELEVANT_RE-READING	Learners re-read and review learning content which they have read before
IRRELEVANT_READING	Learners read the pages which are not relevant to the learning goal or the task
IRRELEVANT_RE-READING	Learners re-read the irrelevant pages
NAVIGATION	Learners navigate through pages or scroll at catalogue zone
WRITE_ESSAY	Learners write, edit or stay in the essay zone
COPY_PASTE	Learners copy and paste content from reading materials into the essay or notes
NOTE_EDITING	Learners create, delete, edit or label the notes
NOTE_READING	Learners click to open and read or re-read the notes
HIGHLIGHT_EDITING	Learners create, delete or edit the highlights
HIGHLIGHT_READING	Learners click to open and read or re-read the highlights
HIGHLIGHT_LABELLING	Learners create tags for highlights
TIMER	Learners click to check timer during learning
SEARCH_CONTENT	Learners use search tool on the left to search learning contents
SEARCH_HIGHLIGHT_NOTE	Learners use search tool on the right to search notes or highlights
PLANNER	Learners click to open planner tool, and create or edit their plans

Situation 2: Only think aloud occurrences; if No_process was detected in V1 and an SRL process was detected in V2 for one segment, then assign the think aloud result (e.g., MC.P) to that segment, see S2 in Figure 3;

Situation 3: Only trace occurrences; if No_process was detected in V2 and an SRL process was detected in V1 for one segment, then assign the trace data result to that segment, see S3 in Figure 3;

Situation 4: Matched co-occurrences; if same SRL process was detected in both V1 and V2 for one segment, then assign the process to that segment, see S4 in Figure 3;

Situation 5: Unmatched co-occurrences; if different SRL processes were detected in both V1 and V2 for one segment, for example, MC.O in V1 and LC.F in V2, then assign both processes (e.g., MC.O and LC.F) to that segment, see S5 in Figure 3;

Here, it is worthwhile to mention an important assumption of our study: we adopted a neutral position towards two methods and considered the different measurement results based on trace data and think aloud data as equally valid. Therefore, in situation 5, we considered two unmatched SRL processes co-occurred without presuming only one process as the correct result and the other one as the incorrect result. We are also fully aware of the complexity of explaining these co-occurrence situations of SRL processes; therefore, in the discussion section, we further discuss multiple possible interpretations about the matched and unmatched co-occurrences together with other essential dimensions of measuring SRL, such as the validity issue. There are also cases of partial overlaps (see the MC.P example in Figure 3) and different parts were treated differently. For example, the matched part of MC.P (first half) was considered situation 4 (both trace and think aloud data detected MC.P) and the remaining part was considered situation 2 (only think aloud measured MC.P). Different segments had variable lengths of duration, which depended on the actual alignment situations, which could be relatively short or long.

It is also worth noting that, the above integration rules created new segments which resulted in the number of SRL processes detected in the integrated data streams is not equal to the simple addition of the number of SRL processed in the trace data and think aloud data. For example, in Figure 3, there are six processes in the results based on trace data and six processes in the results based on think aloud data, but there are 14 processes in the results based on integrated data. Because the think aloud data are very fine-grained and a large number of gaps (very short and silent segments in the audio being labelled as No_process) exist in the think aloud results, a significant amount of new segments were created when integrating two data channels. Additional technical explanations can be found in the supplemental document.

Based on the five integrating rules above, we integrated SRL processes obtained via each channel separately into the third set of processes, **Integrated SRL processes (V3)**. In the next subsection, we explain how we analysed these three sets of results to answer our research questions.

3.3 | Data analysis

To answer the *first research question on whether think aloud and trace data be used interchangeably and complement each other in measuring SRL processes*, we conducted a descriptive and inferential statistical analysis to calculate the duration of the five alignment situations. We unpacked these alignment situations by providing the detailed alignment results for each SRL process, and then further unpacked the co-occurrences between trace data and think aloud using a cross table. We also performed a descriptive statistical analysis to calculate the duration of each SRL process in the three sets of results.

TABLE 3 Process library for detection of SRL processes from action labels.

Categories	No.	SRL processes
Orientation	1	GENERAL_INSTRUCTION*/RUBRIC* -> NAVIGATION -> RELEVANT_READING
(MC.O)	2	GENERAL_INSTRUCTION/RUBRIC -> GENERAL_INSTRUCTION/RUBRIC
	3	GENERAL_INSTRUCTION/RUBRIC <-> HIGHLIGHT_EDITING/NOTE_EDITING/NAVIGATION
	4	GENERAL_INSTRUCTION*/RUBRIC*
Planning	5	PLANNER -> NAVIGATION -> RELEVANT_READING
(MC.P)	6	GENERAL_INSTRUCTION/RUBRIC <-> PLANNER* (during first 15 min)
	7	PLANNER* (during first 15 min)
	8	SEARCH_CONTENT*
Evaluation (MC.E)	9	IRRELEVANT_(RE-)READING -> (NAVIGATION) -> GENERAL_INSTRUCTION*/ RUBRIC* -> (NAVIGATION) -> RELEVANT_(RE-)READING
Monitoring	10	NAVIGATION <-> NOTE_READING
(MC.M)	11	GENERAL_INSTRUCTION/RUBRIC <-> PLANNER* (after the first 15 min)
	12	WRITE_ESSAY <-> PLANNER*
	13	TIMER*
	14	PLANNER* (after the first 15 min)
	15	SEARCH_HIGHLIGHT_NOTE*
	16	HIGHLIGHT_READING/NOTE_READING*
First-reading (LC.F)	17	(IR)RELEVANT_READING->HIGHLIGHT_EDITING/NOTE_EDITING-> (IR)RELEVANT_READING
	18	(IR)RELEVANT_READING -> NAVIGATION -> (IR)RELEVANT_READING
	19	RELEVANT_READING -> IRRELEVANT_READING -> IRRELEVANT_READING
	20	(IR)RELEVANT_READING<->HIGHLIGHT_EDITING/NOTE_EDITING/ HIGHLIGHT_READING/NOTE_READING
	21	(IR)RELEVANT_READING <-> (IR)RELEVANT_READING
	22	IRRELEVANT_READING*
	23	RELEVANT_READING*
Re-reading	24	RELEVANT_RE-READING*
(LC.R)	25	IRRELEVANT_RE-READING*
Elaboration/	26	(IR)RELEVANT_RE-READING -> (NAVIGATION) -> WRITE_ESSAY
Organization	27	GENERAL_INSTRUCTION*/RUBRIC* -> (NAVIGATION) -> WRITE_ESSAY
(HC.E/O)	28	WRITE_ESSAY -> WRITE_ESSAY
	29	WRITE_ESSAY <-> HIGHLIGHT_READING/NOTE_READING
	30	HIGHLIGHT_LABELLING*
	31	NOTE_EDITING*

Note: "->" means a transition from action A to action B; "<->" means a transition from action A to action B or the other way around; "()" means optional; "*" means one or more consecutive instances of the same action; "/" means either action A or action B.

As the data were not normally distributed, we report the 25th, median, and 75th percentile values. To compare the measurement results based on three different data channels, we conducted a Friedman test followed by post hoc Wilcoxon signed rank tests (with the Bonferroni correction) for the pairwise comparison.

To answer our *second research question on sequential and temporal analysis of SRL based on the different data channels*, we utilized pMineR, the process mining analytical technique (Gatta et al., 2017) that has recently been increasingly applied in research on SRL (Ahmad Uzir et al., 2020; Fan et al., 2021; Matcha et al., 2019; Saint et al., 2021; Saint, Gašević, et al., 2020). By using this technique, we were able to unveil sequential and temporal characteristics of SRL processes. We thus computed the temporal transitions between SRL processes and created **SRL process maps** using (i) the trace data, (ii) the think aloud data, and (iii) the integrated data. Specifically, first-order Markov models (FOMMs) using the above three data channels were first deployed to train and present the process maps based on the transition matrix between all SRL processes (e.g., containing the probabilities of transition from MC.O to MC.O, MC.M, MC.P, MC.E and all the other processes). Then, in these process maps, each node represented an SRL process, whereas the edges between nodes indicated the probabilities of transition between SRL processes, and the thickness of the edges indicated the relative probabilities of transitions (e.g., an edge with a 20% probability of transition is two times thicker than an edge with 10% probability of transition). We used 5% as the edge threshold, which means all the edges showing transition probabilities below 5% were omitted from the SRL process maps, to keep the process maps concise and interpretable. This threshold was also used in several previous studies which used process mining, for example, Fan, van der Graaf et al. (2022) and Lim et al. (2022).

Furthermore, we compared the results to identify differences in transition probabilities across the three process maps to answer RQ2 in that way. The black edge with a single transition probability value indicated that the corresponding pairwise transition probabilities in the compared FOMM models were similar to each other, that is, the difference between the two values was below 10%. In cases where differences exceeded 10%, both probabilities have been shown, and the edge was either green or red, green indicating that the second transition probability was greater and red indicating that second transition probability was lower. By comparing the transition probabilities between processes across process maps in this way, we unveiled differences between the three data channels in measuring SRL processes.

TABLE 4 Think aloud coding protocol with examples of learners' utterances.

Metacognition	Code	Example
Orientation	MC.O	*Four things are important in the assignment*
Planning	MC.P	*I will explain these topics in my essay*
Evaluation	MC.E	*I doubt whether Artificial Intelligence will have such an impact*
Monitoring	MC.M	*I am checking how much time is left*
Cognition	**Code**	**Example**
First-reading	LC.F	*Artificial Intelligence is the ability of …*
Re-reading	LC.R	*Artificial Intelligence is the ability of … (re-read out aloud again)*
Elaboration/ Organization	HC.E/O	*This means that both Artificial Intelligence and humans can learn …* *To summarize, … are important for future education*
Motivation/ Procedural	**Code**	**Example**
Motivation	Other	*I find it difficult to perform well in this task …*
Procedural	Other	*Oh, this is how I save my highlights*
No_Process	Code	Example
Uncodable	No_Process	*Instances of murmuring or completely silent*

4 | RESULTS

4.1 | Five alignment situations between trace data and think aloud data

4.1.1 | Situation 1: No_measurement occurrences

To answer **RQ1**, we defined five alignment situations (see Section 4.2.4) and aligned SRL processes measured from trace and think aloud data. When only using the think aloud data, we detected SRL processes for 57.82% (median) of the whole learning session; and when only using the trace data. we detected SRL processes for 80.21% (median) of the whole learning session. However, by aligning these two data channels, our results (see Table 5) indicate that only in 6.37% of segments, no SRL processes in both trace data and think aloud were observed (No_measurement occurrences).

4.1.2 | Situation 2 and 3: Only think aloud occurrences and only trace occurrences

Next, we detected SRL processes either from think aloud (11.34% of the time segments, only think aloud occurrences) or from trace data (34.48% of time segments, Only trace occurrences). In the only think aloud occurrences, learners probably verbally articulated SRL processes that do not result in any specific observable interactions by trace data with the learning environment at that time (e.g., a mouse click or keyboard stroke), and therefore, relevant SRL processes can not be captured by trace data. In Table 6, which unpacked Table 5 by showing the details of the alignment results for each SRL process, we

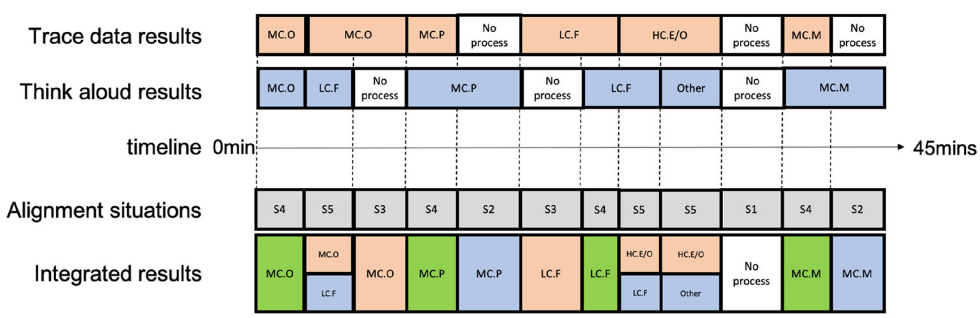


FIGURE 3 Integrating trace data and think aloud data.

TABLE 5 Alignment between trace data and think aloud: Median (25th, 75th) duration (%).

	Example		
Alignment situation	Trace data	Think aloud data	Median (25th, 75th)
Situation 1: No_measurement occurrences	No_Process	No_Process	6.37 (4.18, 11.68)
Situation 2: Only think aloud occurrences	No_Process	MC.O	11.34 (4.65, 16.84)
Situation 3: Only trace occurrences	MC.O	No_Process	34.48 (26.10, 39.91)
Situation 4: Matched co-occurrences	MC.O	MC.O	17.18 (13.71, 23.63)
Situation 5: Unmatched co-occurrences	MC.O	MC.M	27.17 (21.47, 33.72)

TABLE 6 Detailed alignment results for each SRL process: Median (25th, 75th) duration (%).

Processes	S2: Only think aloud	S3: Only trace	S4: Matched	S5:Unmatched
MC.O	1.96 (0.88, 4.97)	43.76 (34.22, 53.17)	6.24 (1.64, 8.63)	40.37 (31.90, 50.94)
MC.P	12.65 (7.68, 28.33)	0.00 (0.00, 6.29)	0.00 (0.00, 0.00)	75.98 (60.63, 87.94)
MC.E	0.00 (0.00, 19.82)	5.70 (0.00, 33.52)	0.00 (0.00, 0.00)	66.62 (56.00, 85.47)
MC.M	11.14 (5.89, 20.00)	12.89 (8.91, 21.20)	1.11 (0.60, 3.21)	68.05 (60.23, 76.95)
LC.F	9.07 (5.34, 18.75)	29.27 (23.22, 35.53)	26.73 (18.17, 34.01)	30.84 (24.90, 37.09)
LC.R	0.02 (0.00, 3.31)	38.44 (23.86, 50.47)	0.00 (0.00, 0.16)	56.38 (48.32, 68.33)
HC.E/O	4.93 (1.91, 12.94)	30.02 (20.61, 37.85)	11.30 (4.59, 17.90)	45.25 (38.43, 58.67)
Other	11.69 (7.66, 25.16)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	88.31 (74.84, 92.34)

Note: When analysing each SRL process in detail, there were no occurrences of situation 1, therefore this table only contains four situations (S2-S5). Take LC.F as an example, 9.07% (by median) of all LC.F processes were detected only in think aloud (and trace data detected No_Process), 29.27% (by median) were detected only in trace data (and think aloud detected No_Process), 26.73% (by median) were detected synchronously in both think aloud and trace data, and the other 30.84% (by median) were detected either in think aloud or trace data (and a different SRL process was detected in the other data channel).

found 12.65% of the *Planning* (MC.P) processes and 11.14% of the *Monitoring* (MC.M) processes fall into the situation 2 (Only think aloud occurrences). These findings indicate that, compared to other SRL processes, learners sometimes verbally articulated *Planning* (e.g., "I am gonna read the definition of AI") and *Monitoring* (e.g., "this is not relevant") without actual observable interaction with the learning environment. It is worth noting that the absence of observable interactions with the learning environment does not necessarily imply that no interaction was taking place, and our results indicated that learners did interact with the learning material as recorded by think aloud data only.

In contrast, in the only trace occurrences, learners engaged with SRL processes that could be detected using the trace data and did not or could not verbally articulate any SRL processes at the same time segments. As shown in Table 6, we found 43.76% of the *Orientation* (MC.O) processes, 38.44% of the *Re-reading* (LC.R) processes and 30.02% of the *Elaboration/Organization* (HC.E/O) processes fall into situation 3 (Only trace occurrences). These findings indicate that, compared to other SRL processes, a relatively large proportion of learners' *Orientation*, *Re-reading* and *Elaboration/Organization* processes were only detected in trace data. There were many possible scenarios can cause such findings. One possible scenario is that learners tended not to verbally articulate certain SRL processes. For example, when a learner re-opened a reading page they might not read out loud again for the same content; however, we labelled such learning action as *Re-reading* using the trace data. Another possible scenario is that for the same SRL process, learners could only think aloud for a short time period and kept silent in the rest time period. For example, when a learner made highlights or took notes while reading the general instruction page, such an action pattern was interpreted in *Orientation* process based on our process library (Table 3) and may take 30 s; however, during these 30 s, this learner might only verbally stated one sentence (e.g., "this is the main task requirement, I better note it down") which might only take less than 5 s and the other 25 s fall into the situation 3 (Only trace occurrences).

These above examples demonstrated possible scenarios of both situation 2 (Only think aloud occurrences) and situation 3 (Only trace occurrences). For these around 45% of all time segments, the two data channels could be used to complement each other. However, we also found two SRL processes co-occurred in around another 45% of all time segments, and these co-occurrences situations are more complicated.

4.1.3 | Situation 4 and 5: Matched co-occurrences and unmatched co-occurrences

As shown in Table 5, the same SRL processes were detected in trace and think aloud data channels in about 17% of time segments

TABLE 7 Cross table showing the matched and unmatched details of each SRL process: Duration (%).

Trace data → Think aloud ↓	MC.O	MC.P	MC.E	MC.M	LC.F	LC.R	HC.E/O
MC.O	45.27	4.23	3.51	3.41	25.14	1.76	16.69
MC.P	14.02	4.45	0.80	2.53	43.99	5.29	28.93
MC.E	10.68	0.00	0.00	2.48	57.85	12.93	16.05
MC.M	13.29	1.10	0.42	4.70	40.54	8.64	31.32
LC.F	2.33	0.34	0.59	2.31	75.01	6.15	13.28
LC.R	4.58	0.54	0.36	4.24	25.35	9.66	55.27
HC.E/O	11.97	0.15	0.20	3.68	34.92	12.51	36.56
OTHER	14.89	0.62	0.25	3.55	32.44	9.84	38.42

Note: This table shows in the time period corresponding to eight think aloud codes of all 44 learners, the respective cumulative time distributions of different SRL processes detected by the trace data. For example, in all occurrences that think aloud coded as MC.O, 45.27% (based on the duration of all MC.O processes in think aloud) was also detected as MC.O in trace data, and 4.23% was detected as MC.P in trace data.

TABLE 8 Descriptive statistics of the duration of SRL processes detected from the multi-channel data: median (25th, 75th) duration (%).

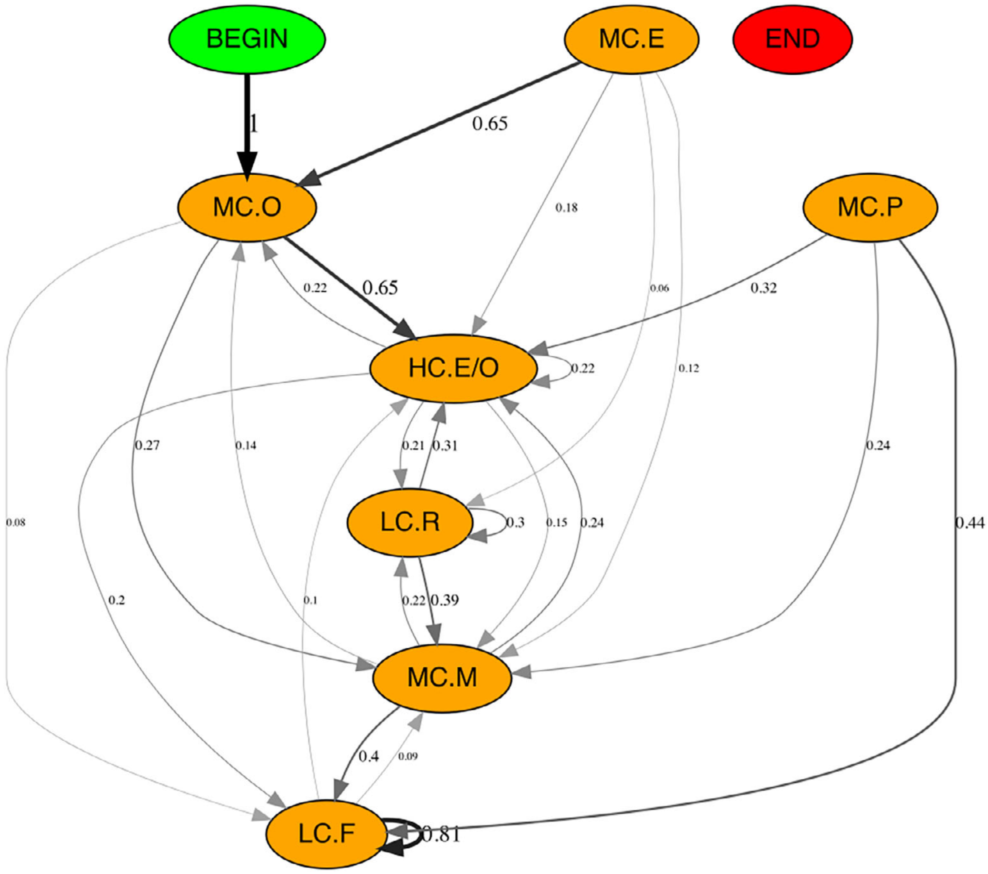
Processes	Trace data M (25th, 75th)	Think aloud data M (25th, 75th)	Integrated data M (25th, 75th)	Pairwise comparisons
MC.O	7.31 (3.87, 12.55)	1.97 (1.11, 3.75)	7.23 (3.92, 10.64)	b***
MC.P	0.00 (0.00, 0.38)	2.36 (1.55, 3.81)	2.21 (1.25, 2.98)	a***
MC.E	0.00 (0.00, 0.30)	0.00 (0.00, 0.17)	0.08 (0.00, 0.39)	b*
MC.M	2.20 (1.09, 3.25)	5.20 (3.51, 7.31)	6.06 (4.56, 7.17)	a***
LC.F	33.78 (25.58, 46.29)	21.78 (16.39, 27.39)	35.75 (30.25, 44.54)	b***
LC.R	5.07 (1.80, 9.67)	0.45 (0.11, 0.98)	4.91 (1.66, 7.94)	b***
HC.E/O	25.33 (10.22, 29.71)	16.40 (9.95, 20.99)	26.22 (22.19, 31.21)	b***
Other	0.00 (0.00, 0.00)	7.84 (5.95, 11.55)	5.84 (4.44, 8.35)	a***; b*

Note: Statistical comparison was done with the Friedman test followed by Wilcoxon signed rank tests for pairwise comparison (with the Bonferroni correction). Legend: a – significant difference between Trace data and Integrated results; b – significant difference between Think aloud and Integrated results; * – $p < 0.05$; ** – $p < 0.01$; and *** – $p < 0.001$. The complete results generated from all statistical tests including r -values, Z -values and p-values can be found in the supplemental document.

(Matched co-occurrences). As well, we were able to identify SRL processes that differed across two data channels during the same segment (Unmatched co-occurrences) in about 27% of all time segments. For example, when learners read the general instructions page and then navigated reading materials, the *Orientation* process (MC.O.1 in Table 3) was detected from trace data. At the same time, learners could also verbally express their SRL processing, for example, “These four things are important in the assignment …”, the utterance coded as *Orientation* (MC.O) (Table 4; or “This information does not appear to be relevant …”, which was coded as *Monitoring* (MC.M), as per the think aloud coding protocol. The former utterance detected the same SRL process (MC.O) as the trace data, that is the matched co-occurrences as shown in Table 5; the later utterance, on the other hand, detected different SRL process (MC.M) as the trace data (MC.O), that is the unmatched co-occurrences in Table 5.

In order to better understand the matched and unmatched details of each SRL process, we created a cross table of co-occurrences situations. As shown in Table 7, we found trace data and think aloud data more matched at *Orientation* (MC.O), *First-reading* (LC.F) and *Elaboration/Organization* (HC.E/O) processes, but much less matched at other processes. For example, 75.01% of all co-occurrences that think aloud coded as LC.F was also detected as LC.F in trace data, which shows these two methods aligned very well at measuring *First-reading* processes. However, for example, 0% of MC.E, only 4.45% of MC.P and 4.70% of MC.M (coded in think aloud) was also detected as the matched processes in trace data, and most of these metacognition processes coded in think aloud were detected as LC.F or HC.E/O processes in trace data. These three SRL processes are usually very low frequency (total less than 8% as shown in Table 8), and the duration of them is usually very short (e.g., “I now check timer” which is *Monitoring* process and can be less than 2 s), so it is very difficult to detect these processes using trace data at exactly the same time. Because learners verbally articulated these SRL processes occasionally during their reading and writing which mainly labelled as LC.F and HC.E/O in trace data, we found bigger proportion of these co-occurrences as unmatched situations. These findings indicate that, trace data and

FIGURE 4 FOMM of the temporal transition probabilities between SRL processes detected based on the trace.



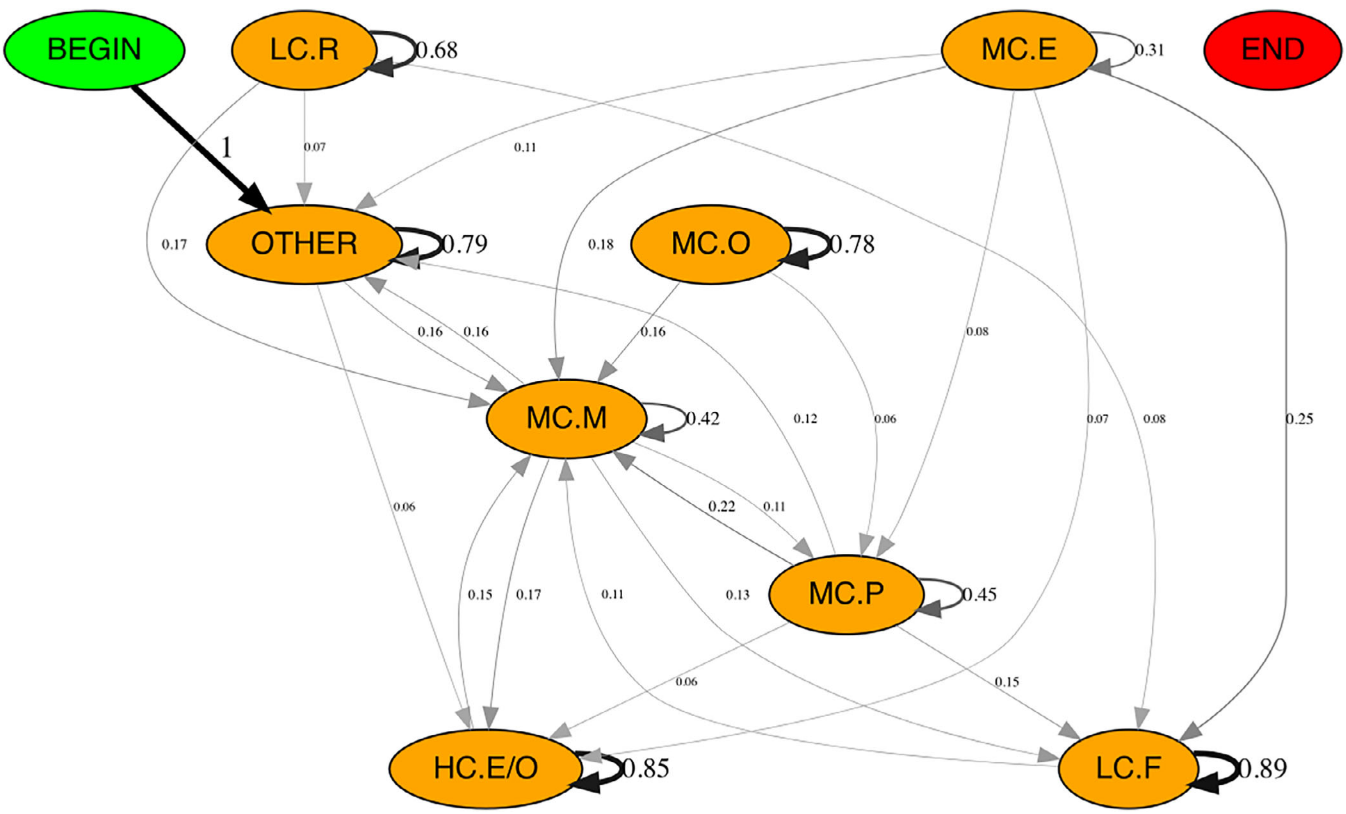


FIGURE 5 FOMM of the temporal transition probabilities between SRL processes detected based on the think aloud.

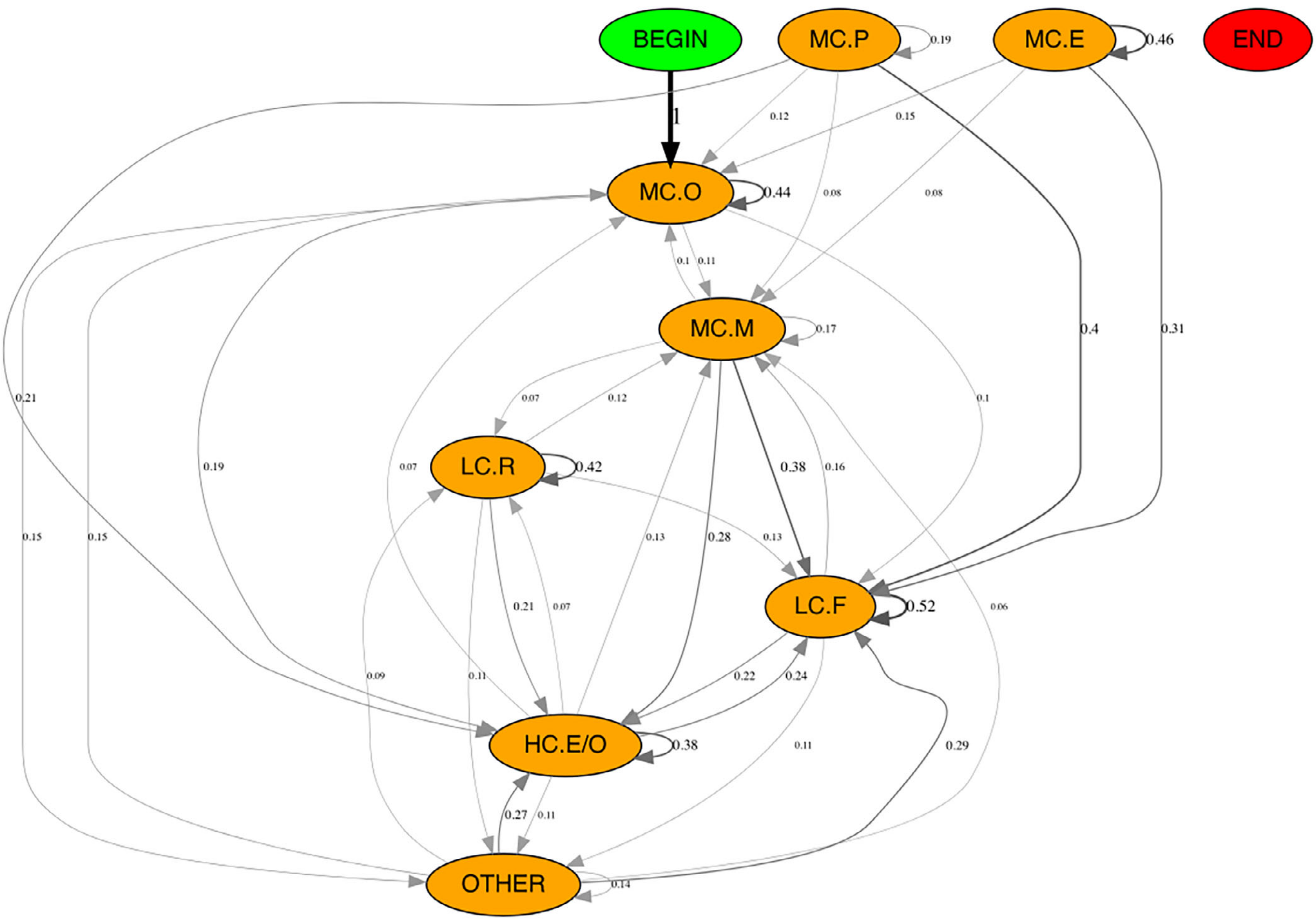


FIGURE 6 FOMM of the temporal transition probabilities between SRL processes detected based on integrated data.

think aloud data could be interchangeably used to measure SRL processes for only a small percent (17.18%) of all time segments in a learning task, and also should be limited used in measuring certain SRL processes. Therefore, integrating the trace data and think aloud data to provide a fuller measurement of SRL becomes promising.

4.1.4 | The integrated SRL processes

Analysing data generated by 44 learners, we obtained 9,993 SRL processes from their trace data, 38,856 SRL processes from their think aloud data and 83,121 SRL processes from integrated trace and think aloud data channels. We note that the integrated SRL processes are not a sum of processes obtained in the two data channels individually. Instead, the integrated SRL processes have been obtained through a fine-grained analytical approach described earlier (see Subsection 3.2.4). For example, when observing a 20-s time segment, researchers could notice that the LC.F process was generated from trace data over this entire segment. However, think aloud protocols may indicate more diverse processes during this same segment: LC.F in the first 10 s, MC.M in the next 5 s and No_Process in the last 5 s. Therefore, upon integrating the two data channels, four SRL processes would be generated from the observed time segment: (1) LC.F in the first 10 s; (2) LC.F and MC.M in the period between 10 and 15 s; and (3) LC.F in the last 5 s.

Moreover, the duration of SRL processes in the three sets of results was different, as indicated by the pairwise statistical comparison presented in Table 8. Of all the processes detected in the three data channels, cognitive processes were the most prominent. As shown in Table 8, the integrated data revealed significantly more cognitive processes (LC.F, LC.R and HC.E/O) than the think aloud data. Metacognitive processes, on the other hand, were detected about 10% of time, both in trace and think aloud data channels. However, the proportion of metacognitive processes increased to 15.58% in the integrated results. Compared to trace data results, the integrated results revealed significantly more *Planning* (MC.P) and *Monitoring* (MC.M), according to pairwise comparisons (Table 8). Compared to think aloud results, the integrated results revealed significantly more *Orientation* (MC.O) and *Monitoring* (MC.M).

4.2 | SRL process maps from three data channels

To answer the research question RQ2, we analysed temporal and sequential differences between SRL processes detected in trace, think aloud and integrated data channels.

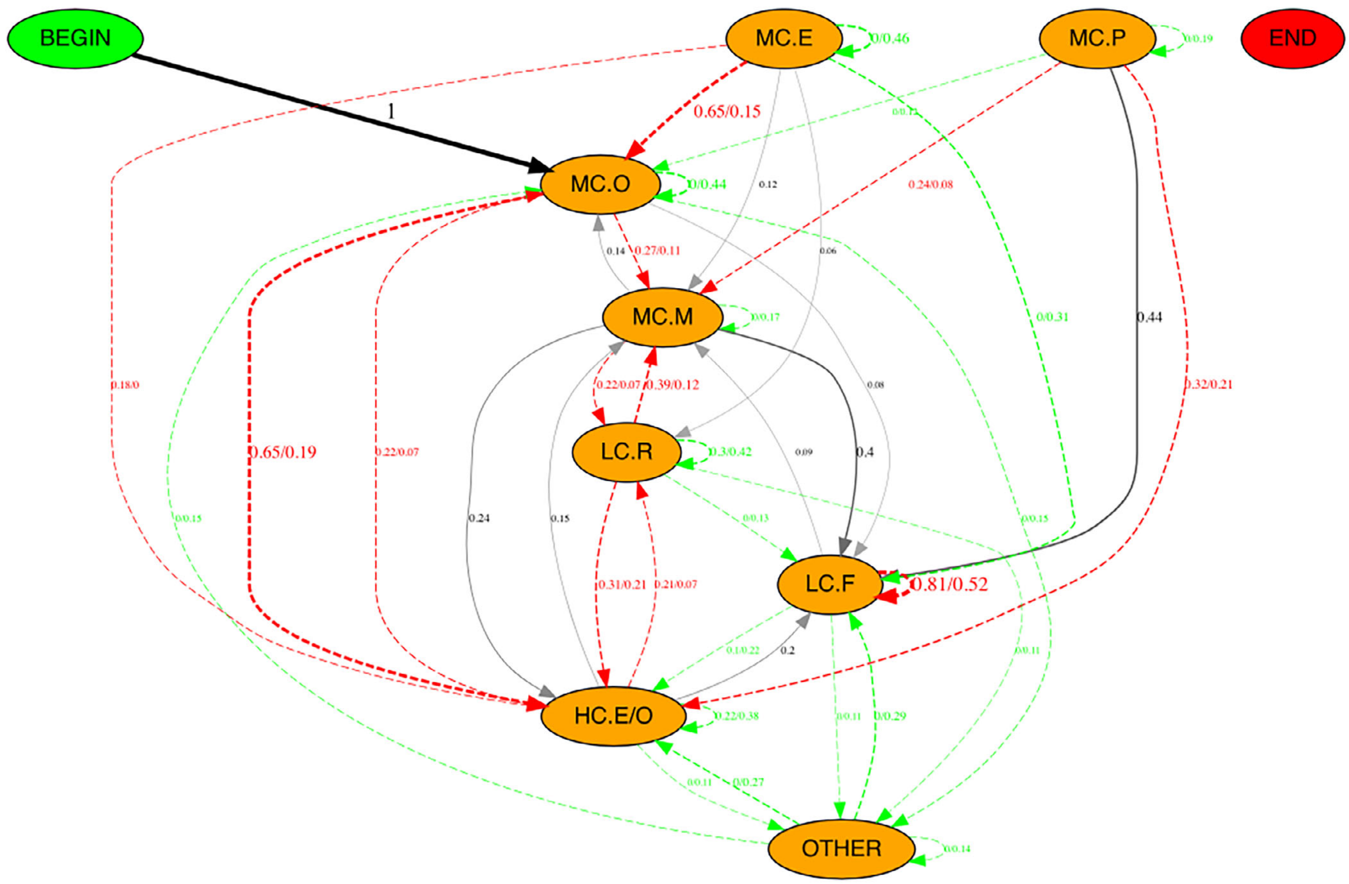


FIGURE 7 Comparison the trace data process map and the integrated data.

4.2.1 | SRL process map (V1) based on trace data

As shown in Figure 4, the learners engaged in *Orientation* (MC.O) at the outset of learning session. Further, either simultaneously with or after *Orientation* was done, the learners engaged in *Elaboration and Organization* (HC.E/O), *Monitoring* (MC.M), and *First-reading* (LC.F) as the SRL processes that the learners most frequently and most continuously engaged in. For instance, the transition probability from LC.F to LC.F (self-loop) in the process map was 81%. Occasionally, *Monitoring* and *Elaboration and Organization* interspersed during *First-reading*, as indicated by the transition probabilities of 10% from LC.F to HC.E/O and 9% from LC.F to MC.M. This finding, in turn, indicates that learners monitored their reading (e.g., by looking at the timer) and also elaborated or organized the information they read (e.g., by writing down what they read in the essay). The process map generated from the trace data also shows that *Elaboration and Organization* was a very prominent process, as its corresponding node in the map was connected to many other nodes (i.e., processes), such as MC.O, MC.M and LC.R. As well, the learners tended to go back to HC.E/O from other SRL processes, such as MC.E. This finding was expected, since the task in this study was to write an essay from multiple sources of information. We also note that *Re-reading* was more closely connected to *Elaboration and Organization* than the *First-reading* process, which indicates that many learners did not begin writing the essay after they read sources for the first time; rather, some of them typically re-read parts of source information before they decided to engage in writing.

4.2.2 | SRL process map (V2) based on think aloud data

As shown in Figure 5, *OTHER* was the most prominent process at the outset of a learning session. In most cases, this process was related to *Procedural* issues, according to the think aloud protocol. For example, learners sometimes said "Now, I entered the task" or "Oh, the page is loaded and I now start reading". The process map generated from think aloud data further shows that self-loops of many SRL processes were more prominent compared to their self-loops in the trace data process map. For example, the transition probability from MC.O to MC.O was 78% in the think aloud process map. We also note that pair-wise transition probabilities between many SRL processes in the think aloud map were relatively small (less than 5%), and, for that reason, many edges were omitted in this process map, for example, the edges between HC.E/O and LC.F, or HC.E/O and LC.R. Still, the edges that remained in the map unveiled important information about SRL processing.

As shown in Figure 5, *Monitoring* and *Planning* (MC.P) are two central nodes in the think aloud process map, formed multiple edges with other nodes, such as LC.F and HC.E/O. The think aloud process

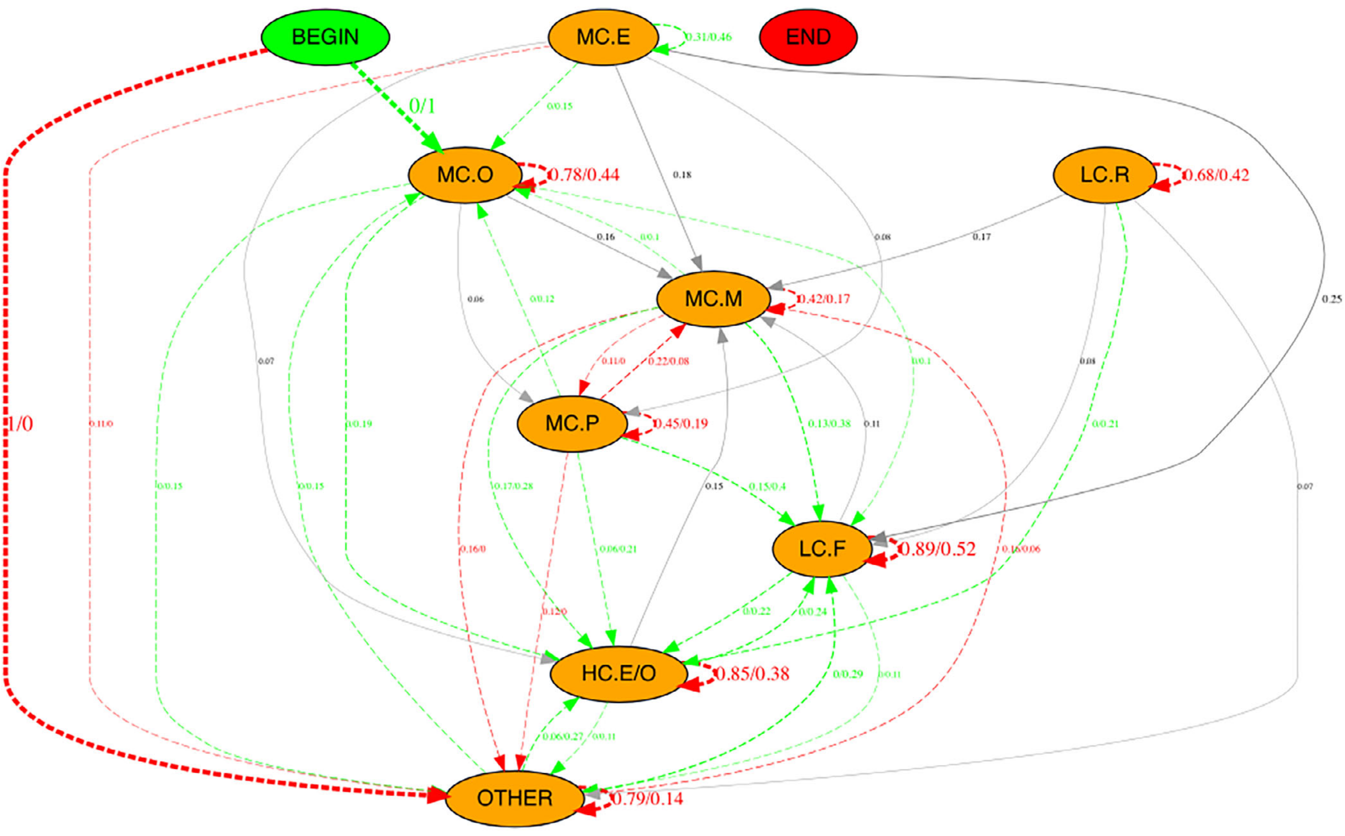


FIGURE 8 Comparison of the think aloud process maps and the integrated data.

map further indicates that the transition probability from *Evaluation* (MC.E) to *First-reading* (LC.F) was 25%, which was much higher than the transition probability between the same pair of nodes in the trace data process map (0%). However, it must be pointed out here that MC.P and MC.E are very low frequency SRL processes: that is, even though the transition probabilities are high in think aloud, the actual frequency of occurrences are very low. Overall, the patterns detected in the think aloud process map suggest that many learners verbally expressed metacognitive processes (e.g., MC.M) before cognitive processes (e.g., LC.F and HC.E/O). Moreover, these transitions from metacognitive to cognitive processes were also detected in the trace data process map (Figure 4). Importantly, these findings indicate that the results based on think aloud data corroborated the results based on trace data, for example, we observed the same transitions from metacognitive to cognitive processes in both process maps. The findings also showed that think aloud provided new insights into learners' SRL processes that were not obvious based on the analysis of the trace data only, for example, the transition from Evaluation to First-reading was only identified in the process map of think aloud results.

4.2.3 | Integrated SRL process map (V3) based on integrated data

We further created the integrated process map that involved both think aloud and trace data transition probabilities (Figure 6). The integrated process map showed more transitions (i.e., edges) among the SRL processes than the two process maps individually. As indicated in the map, learners engaged in *Orientation* at the outset of a learning session, and then mainly engaged in the *Monitoring*, *First-reading*, *Re-reading*, and *Elaboration and Organization* processes. These four SRL processes formed transitions between each other, mostly with transition probabilities of more than 5%. Most of these transitions were also captured in the trace data process map (Figure 5) with more emphasis on HC.E/O) and the think aloud based process map (Figure 4 with more emphasis on MC.M and less emphasis on LC.R). Some processes, however, appeared to be more prominent in each individual process map, for example, *Elaboration and Organization* in the trace data process map and *Monitoring* in the think aloud process map. In the integrated process map, the relationships among these four processes were more balanced. For example, the transition probabilities from *Rereading* to *Elaboration and Organization* and from *Monitoring* to *First-reading* were prominent in the integrated map. Below, we show the overlay between the integrated and trace data process maps (Figure 7), and integrated and think aloud process maps (Figure 8). A green edge between two processes in Figure 7 depicts the transition probability in the integrated process map that was greater than the transition probability between the same processes in the trace data process map, relative to difference threshold of 10%. Similarly, a red edge between two processes depicts the transition probability in the integrated process map that was lower than the transition probability between the same processes in the trace data process map.

For example, all the edges to/from the OTHER process were green, as motivation and procedural issues were not captured by the trace data channel; the transition probability from *Re-reading* to *Re-reading* (self-loop) was 12% higher in the integrated process map than in that in the trace data process map, and the transition probability from *Elaboration and Organization* to *Elaboration and Organization* (self-loop) was 16% higher in the integrated process map than that in the trace data process map, two findings that may indicate that more information about the continuity of certain SRL process in a learning session can be obtained by integrating trace and think aloud data, than by looking at trace data only. We also compared the integrated process map and the think aloud process map (Figure 8) and found that the incoming edges to *Orientation*, *First-reading* and *Elaboration and Organization* were mostly coloured in green (e.g., the four edges from *Monitoring* and *Planning* to *First-reading* and *Elaboration and Organization*). On the other hand, the edges between *Monitoring* and *Planning*, including the self-loops of these processes, were red, that is, they indicate higher transition probabilities in the think aloud than those in the integrated process map. This finding, in turn, may suggest think aloud data highlighted the continuity of metacognitive processes of planning and monitoring during a learning session, but at the same time, led to an imbalance process model compared to the integrated result.

5 | DISCUSSION

5.1 | Triangulation between think aloud and trace data when measuring SRL

SRL processes have mainly been measured and studied using either think aloud or traced methods, but rarely using a combination of the two. The purpose of our research was to juxtapose learners' think aloud and trace data and, through the analysis of data from both channels, gain a deeper insight into SRL processes. Specifically, we simultaneously collected think aloud and trace data generated by university students in a laboratory setting, and then synchronized the two data sets relative to the same timeline (Figure 2). Our results indicate that the same SRL processes were observed by both think aloud and trace data methods in about 17% of all time segments. This evidence resonates with previous research identifying the same learning processes between what students said they did and what was captured by their trace data (Bråten & Samuelstuen, 2007; Rogiers et al., 2020). This finding suggests that think aloud and trace data could be interchangeably used to measure only a small amount of SRL processes in this learning task.

SRL processes in more than 30% of all time segments in this study were captured by trace data (Only trace occurrences as described in Section 4.2.4). These processes were not captured during the same time segments in the think aloud protocols, that is, they were detected as No_process in the think aloud data. Mainly, these processes were *Orientation*, *First-reading*, and *Elaboration and Organization* (Bannert, 2007). They often cannot be straightforwardly inferred from think aloud protocols due to differences in learners' ability to articulate their thoughts in the form of useful data and cognitive demands that many students face when studying and verbalizing their thoughts at the same time (Winne, 2020; Young, 2005). Instead, our results suggest that *Orientation*, *First-reading* and *Elaboration and Organization* processes can be inferred from traces of learners' interactions with online learning resources, conforming the results of previous studies that found trace data useful in obtaining fine-grained insights into cognitive and metacognitive processes of SRL, for example, Azevedo et al. (2013); Bondareva et al. (2013); Chen and Su (2019); Kinnebrew et al. (2013b); Taub et al. (2016); Trevors et al. (2016); Winne et al. (2017). For example, a learner could take several notes on what they have read on the general instruction page. These notes would be all interpreted as occurrences of the *Orientation* process, based on the process library (Table 3). The learner, however, may only verbally express the occurrence of *Orientation* before or after all the notes were taken in the general instructions page, for example, “These topics are important in the task”. As a result, the *Orientation* process would be identified in the think aloud data over only a short period of time.

Under certain circumstances, think aloud protocols can also be more insightful into SRL processing than trace data. Our study showed that SRL processes in about 11% of time segments were only detected in think aloud data and were not identified as SRL processes in trace data (Only think aloud occurrences defined in Section 4.2.4). Those were *Planning* and *Monitoring*, metacognitive processes that are often difficult to reveal with trace data and rich think aloud protocols can be more helpful to this end (Deekens et al., 2018; Moos & Azevedo, 2008). For instance, we note that few learners verbally expressed occurrences of the *Planning* process when developing their plans using the Planner tool. The *Planning* process was hence captured both in think aloud and trace data. However, in most cases, learners did not use the planner tool to make specific plans for their learning, but they still verbally expressed *Planning*, for example, "I will re-read this page when I start writing the essay". In those cases, *Planning* remained undetected in trace data.

We also found that in 27.17% of the time segments the observed SRL processes did not match with each other across the two data channels (Unmatched co-occurrences, Section 4.2.4). Therefore, think aloud and trace methods are not directly interchangeable in measuring SRL for these segments in timeline. Especially by unpacking the matched and unmatched co-occurrences, we found these two methods should not be used interchangeably for certain SRL processes such as *Planning*, *Evaluation* and *Monitoring*.

These unmatched co-occurrences reflect the complexity of the learner's self-regulation process. For example, *Planning* can be performed not only at the beginning of the task when learners use the planner tool, but also can occur intertwining with the execution of the task which in line with the cyclical nature of SRL (Panadero, 2017; Pintrich, 2000; Winne & Hadwin, 1998; Zimmerman, 2000). At the beginning of the task, learners may verbally articulate their understanding of the task (think aloud: MC.O) and simultaneously open the planner tool to arrange their overall learning plan (trace data: MC.P);

during the execution of the task, learners may read the content page with informational text (trace data: LC.F) which occasionally may trigger their articulation of more specific plans about what to read next (think aloud: MC.P). Therefore, these unmatched co-occurrences can be constituted by equal valid measurement results in both trace-based method and think aloud method, but when aligned and integrated, two data channels can reveal a more complex, complete, and comprehensive picture of SRL.

However, these unmatched co-occurrences may also cause the shortcomings of the methods themselves (Greene & Azevedo, 2009; Winne, 2010) such as coders' bias in interpreting linguistically diverse think aloud utterances, learners' inability to timely think aloud about all their SRL processes, or incorrect SRL processes inferred from some action patterns in trace data. For example, learners may have verbally expressed *Planning* only after actually using the Planner tool, which in this case, caused the unmatch between think aloud and trace data. We hence believe that the triangulation of think aloud and trace data does not necessarily need to follow the strict temporal matching across segments in a learning session (although it is valuable), but should further emphasize the triangulation of SRL measurement in a more flexible and macroscopic way.

We, therefore, further evaluated these two methods from the high-level perspective, that is, by looking at how overall SRL processes unfolded over time for the whole task, and the evaluation indicated that the two methods showed certain levels of consistency and uniqueness. That is, several important process-to-process temporal transitions were confirmed in both data channels, for example, the same transitions from metacognitive processes of *Planning* and *Monitoring* to cognitive processes of *Elaboration and Organization* and *First-reading* were prominent in both process maps (Figures 4 and 5). Therefore, when measuring SRL to holistically explain how students regulate their learning throughout learning task, think aloud and trace data revealed several similar temporal transitions between SRL processes. However, think aloud and trace data also revealed different results when measuring SRL and, for that reason, we further examined the value of integrating the two data channels.

5.2 | Integration of think aloud and trace data when measuring SRL

We further investigated whether the integration of think aloud and trace data can provide researchers with a fuller picture of SRL processes, to answer our RQ2. To this end, we proposed an approach to integrating SRL processes captured using the two methods. In particular, we developed a set of alignment rules (Figure 3) and applied those rules to synchronize the two data channels while keeping the information from both of them. The results obtained through this integration between think aloud and trace data provided a more elaborate insight into SRL processing compared to insights gained from each method separately. For example, while reading the content, a learner could verbalize their thoughts as "This information does not appear to be relevant" and "I will include this in my essay", with the 5 s of silence between the two utterances. Accordingly, the utterances would be labelled in think aloud protocols as *Monitoring* and *Planning*, respectively, and the 5 s period between them will be labelled as No_Process. However, upon integrating think aloud and trace data, this 5 s period would be labelled as *First-reading*, because trace data (e.g., learner's eye fixations in the reading zone) indicated the reading activity during that time.

In order to deepen the understanding of temporal processes of SRL using the integrated data, we created the integrated process map that involved both think aloud and trace data with transition probabilities. The integrated process map revealed a more comprehensive SRL processing compared to a single method map. For instance, the think aloud process map highlighted the self-loop transition probabilities of processes such as MC.M and LC.F, and therefore relatively weakened the transitions between these SRL processes. For example, by just looking at Figure 5, researcher might conclude that learners rarely transit from LC.F to HC.E/O; however, the integrated process map (Figure 6) revealed this important transition between LC.F and HC.E/O as a result of re-balancing.The integrated results also showed that two different SRL processes labelled by the two different methods may occur at the same time. This implies an important assumption of our study: we considered the measurement results based on trace data and think aloud data as equally valid, because neither of the two methods can be considered as fully "truth" in measuring SRL processes (Winne, 2019). Therefore, instead of considering the unmatched cooccurrences as contradictions where researchers need to choose one, we adopt that such unmatched co-occurrences may reflect SRL processes that a learner simultaneously has engaged in (Dresel et al., 2015; Kim et al., 2020; Nodoushan, 2012; Pintrich, 2000; Schoor & Bannert, 2012; Schuitema et al., 2012; Schunk, 2005; Winne & Hadwin, 1998). For example, the *First-reading* and *Elaboration and Organization* processes appeared to simultaneously occur near the middle of the learning session (Figure 3), resonating with previous research that suggested strategy use (e.g., reading) in SRL learning episodes can often have an overlapping relation with a learner's domain-general executive functions (e.g., organization) (Garner, 2009). For instance, many learners may look for organizational linguistic cues (e.g., "To sum up …") in the text they read.

Moreover, Kim et al. (2020) pointed out the need to investigate these simultaneous processes of SRL and findings from our study can further motivate this line of research. It is also worth noting that there are other possible interpretations of the simultaneous SRL processes we found in this study, for example: (i) these co-occurred SRL processes, in fact, happened one after the other at very fine-grained level, but due to the limited accuracy of the measurement method, they overlapped on the time axis; (ii) these co-occurred SRL processes, in fact, should be considered as a new SRL process that we have not yet been able to accurately define using the current theoretical framework. We acknowledge that the conceptualisation of the simultaneous SRL processes is still a controversial topic and researchers may argue that only a very small number of information elements can be simultaneously active due to the limited working memory (Shipstead et al., 2014). Therefore, we believe this requires follow-up empirical research and in-depth theoretical exploration. In particular, future research should focus on the conditions under which the simultaneous SRL processes can exist as pointed out by Winne and Hadwin (1998), for example, the nature of the task, student metacognitive knowledge or interest in subject, and the effects such an overlap may have on goal attainment, strategy use, and subsequent learning performance.

5.3 | Limitations

The findings in this study need to be interpreted with a few limitations in mind. First, we focused on cognition and metacognition processes in this study and did not infer learners' motivational and affective states, which are also important components of SRL. In the future, new technologies such as facial recognition and physiological sensors can also be used to collect richer trace data and further improve the measurement of motivational and affective processes. Second, we did not elaborate much about the validity issue of our measurement protocols in this paper due to the space limitation, but more detailed information can be found in (Fan, van der Graaf et al., 2022). The validity of both think aloud based and trace based measurement protocols could be further evaluated and improved in the future, by using a larger data set or applying in new learning tasks and contexts. In particular, it is necessary to explore further the relative validity of the two measurement methods in different conditions, for example, whether think aloud method is more or less valid than trace based method in more challenging tasks that require higher-level metacognitive engagement. Also, how the SRL processes operationalized in the think aloud coding scheme (Table 4) and the trace-based measurement protocol (Table 3) heavily influenced the alignment results. For example, we found it difficult to construct more meaningful and valid action patterns for the Evaluation processes using trace data. Therefore, future work should investigate other think aloud coding schemes and trace-based measurement protocols to further examine the integration results of these two data channels.

6 | IMPLICATIONS AND CONCLUSION

First, our results suggest that using a single measurement method can often reveal SRL processes only partially. An integration of the two methods improves detection of SRL processes and gives a fuller picture of SRL, including the identification of different SRL processes that can occur simultaneously. For this purpose, we proposed a set of data integration rules. Second, our findings indicated that the integration of the two measurement methods could not address all their methodological shortcomings and more research is needed towards new integrative approaches that can further reduce the number of misaligned results. Finally, we note that replication of our study may pose challenges to many researchers as they may not be able to collect both think aloud and trace data at the same time, due to practical constraints.

It is also important to notice that the measurement of SRL is context-sensitive (Winne, 2017), such as how learning materials are arranged, how the learning environment is designed, which learning tools are available. For example, as discussed above, the planner tool in our learning environment provided an opportunity to measure and triangulate the *Planning* process based on think aloud and trace data. Therefore, when researchers consider implicating the theoretical framework (Table 1) and SRL measurement protocols (Tables 2, 3, 4) presented in this paper, context-related elements deserve special attention. However, we believe the triangulation and integration approach proposed in this paper has a certain degree of generality because even if other theoretical models or measurement protocols to be used, a similar alignment approach could still be valid to test various measurement results as long as they follow the same theoretical model.In summary, this paper proposed a novel approach to measuring SRL processes by triangulating and integrating the results obtained from two methods: think aloud protocols and trace data. In this way, we attempted to advance the research on SRL measurement methods (Winne, 2010). More importantly, our study suggested what each data channel could contribute differently to measuring SRL processes with their own advantages, and the integrated results revealed more comprehensive insights into what SRL looks like compared to analyses that use only a single measurement method. The integrated measurement of SRL can be used in the future to (1) test the effects of instructional SRL interventions, for example, scaffolding; and (2) evaluate how learners use specific learning tools.

FUNDING INFORMATION

This study was funded through Call 5 of the Open Research Area (ORA) which is jointly supported by Deutsche Forschungsgemeinschaft (BA20144/10-1), Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO464.18.104), and the Economic and Social Research Council of the United Kingdom (ES/S015701/1).

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/jcal.12801.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ETHICS STATEMENT

The ethical committee in the Faculty of Social Sciences of Radboud University approved the present research.

ORCID

Yizhou Fan https://orcid.org/0000-0003-2777-1705
Mladen Rakovic https://orcid.org/0000-0002-1413-1103

REFERENCES

Ahmad Uzir, N., Gašević, D., Matcha, W., Jovanović, J., & Pardo, A. (2020). Analytics of time management strategies in a flipped classroom. *Journal of Computer Assisted Learning*, *36*(1), 70–88 eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jcal.12392

Aleven, V., McLaren, B., Roll, I., & Koedinger, K. (2006). Toward metacognitive tutoring: A model of help seeking with a cognitive tutor. *International Journal of Artificial Intelligence in Education*, *16*(2), 101–128.

Aleven, V., Roll, I. D. O., McLaren, B. M., & Koedinger, K. R. (2010). Automated, unobtrusive, action-by-action assessment of self-regulation during learning with an intelligent tutoring system. *Educational Psychologist*, *45*(4), 224–233.

Azevedo, R., Cromley, J. G., & Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology*, *29*(3), 344–370.

Azevedo, R., & Gašević, D. (2019). Analyzing multimodal multichannel data about self-regulated learning with advanced learning technologies: Issues and challenges. *Computers in Human Behavior*, *96*, 207–210.

Azevedo, R., Greene, J. A., & Moos, D. C. (2007). The effect of a human agent's external regulation upon college students' hypermedia learning. *Metacognition and Learning*, *2*(2–3), 67–87.

Azevedo, R., Guthrie, J. T., & Seibert, D. (2004). The role of self-regulated learning in fostering Students' conceptual understanding of complex systems with hypermedia. *Journal of Educational Computing Research*, *30*(1-2), 87–111.

Azevedo, R., Harley, J., Trevors, G., Duffy, M., Feyzi-Behnagh, R., Bouchet, F., & Landis, R. (2013). Using trace data to examine the complex roles of cognitive, metacognitive, and emotional self-regulatory processes during learning with multi-agent systems. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies. Springer International Handbooks of Education* (pp. 427–449). Springer.

Azevedo, R., Witherspoon, A., Chauncey, A., Burkett, C., & Fike, A. (2009). MetaTutor: A metacognitive tool for enhancing self-regulated learning. p. 6.

Bannert, M. (2007). *Metakognition beim lernen mit hypermedien*. Waxmann Verlag.

Bannert, M., & Mengelkamp, C. (2008). Assessment of metacognitive skills by means of instruction to think aloud and reflect when prompted. Does the verbalisation method affect learning? *Metacognition and Learning*, *3*(1), 39–58.

Bannert, M., Reimann, P., & Sonnenberg, C. (2014). Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition and Learning*, *9*(2), 161–185.

Beishuizen, J. J., & Stoutjesdijk, T. (1999). Study strategies in a computer assisted study environment. *Learning and Instruction*, *9*(3), 281–301.

Bernacki, M. L. (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* (2nd ed., Educational psychology handbook series, pp. 370–387). Routledge/Taylor & Francis Group.

Bernacki, M. L., Byrnes, J. P., & Cromley, J. G. (2012). The effects of achievement goals and self-regulated learning behaviors on reading comprehension in technology-enhanced learning environments. *Contemporary Educational Psychology*, *37*(2), 148–161.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*(1), 417–444 eprint: https://doi.org/10.1146/annurev-psych-113011-143823

Bondareva, D., Conati, C., Feyzi-Behnagh, R., Harley, J. M., Azevedo, R., & Bouchet, F. (2013). Inferring learning from gaze data during interaction with an environment to support self-regulated learning. In H. C. Lane, K. Yacef, J. Mostow, & P. Pavlik (Eds.), *Artificial intelligence in education, lecture notes in computer science* (pp. 229–238). Heidelberg. Springer.

Bråten, I., & Samuelstuen, M. S. (2007). Measuring strategic processing: Comparing task-specific self-reports to traces. *Metacognition and Learning*, *2*(1), 1–20.

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, *27*, 1–13.

Chen, C.-H., & Su, C.-Y. (2019). Using the BookRoll E-book system to promote self-regulated learning, self-efficacy and academic achievement for university students. *Journal of Educational Technology & Society*, *22*(4), 33–46.

De Backer, L., Van Keer, H., & Valcke, M. (2012). Exploring the potential impact of reciprocal peer tutoring on higher education students' metacognitive knowledge and regulation. *Instructional Science*, *40*(3), 559–588.

Deekens, V. M., Greene, J. A., & Lobczowski, N. G. (2018). Monitoring and depth of strategy use in computerbased learning environments for science and history. *British Journal of Educational Psychology*, *88*(1), 63–79 eprint: https://bpspsychub.onlinelibrary.wiley.com/doi/pdf/10.1111/bjep.12174

Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review*, *28*(3), 425–474.

Di Mitri, D., Scheffel, M., Drachsler, H., Börner, D., Ternier, S., & Specht, M. (2017). Learning pulse: A machine learning approach for predicting performance in self-regulated learning using multimodal data. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference, LAK '17* (pp. 188–197). Association for Computing Machinery.

Dresel, M., Schmitz, B., Schober, B., Spiel, C., Ziegler, A., Engelschalk, T., Jöstl, G., Klug, J., Roth, A., Wimmer, B., & Steuer, G. (2015). Competencies for successful self-regulated learning in higher education: Structural model and indications drawn from expert interviews. *Studies in Higher Education*, *40*(3), 454–470 eprint: https://doi.org/10.1080/03075079.2015.1004236

Eccles, D. W., & Arsal, G. (2017). The think aloud method: What is it and how do I use it? *Qualitative Research in Sport, Exercise and Health*, *9*(4), 514–531 eprint: https://doi.org/10.1080/2159676X.2017.1331501

Engelmann, K., & Bannert, M. (2019). Analyzing temporal data for understanding the learning process induced by metacognitive prompts. *Learning and Instruction*, 72, 101205.

Ericsson, K. A., & Simon, H. A. (1984). Protocol analysis: Verbal reports as data. In *Protocol analysis: Verbal reports as data* (p. 426). The MIT Press.

Fan, Y., Tan, Y., Raković, M., Wang, Y., Cai, Z., Shaffer, D. W., & Gašević, D. (2022). Dissecting learning tactics in MOOC using ordered network analysis. *Journal of Computer Assisted Learning*, 39(1), 154-166.

Fan, Y., Saint, J., Singh, S., Jovanovic, J., & Gašević, D. (2021). A learning analytic approach to unveiling self-regulatory processes in learning tactics. LAK21: 11th International Learning Analytics and Knowledge Conference. https://doi.org/10.1145/3448139.3448211

Fan, Y., Lim, L., van der Graaf, J., Kilgour, J., Raković, M., Moore, J., Molenaar, I., Bannert, M., & Gašević, D. (2022). Improving the measurement of self-regulated learning using multi-channel data. *Metacognition and Learning*, *17*(3), 1025–1055. https://doi.org/10.1007/s11409-022-09304-z

Fan, Y., van der Graaf, J., Lim, L., Raković, M., Singh, S., Kilgour, J., Moore, J., Molenaar, I., Bannert, M., & Gašević, D. (2022). Towards investigating the validity of measurement of self-regulated learning based on trace data. *Metacognition and Learning*, *17*(3), 949–987. https://doi.org/10.1007/s11409-022-09291-1

Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin*, *137*(2), 316–344.

Garner, J. K. (2009). Conceptualizing the relations between executive functions and self-regulated learning. *The Journal of Psychology*, *143*(4), 405–426 eprint: https://doi.org/10.3200/JRLP.143.4.405-426

Gasevic, D., Jovanovic, J., Pardo, A., & Dawson, S. (2017). Detecting learning strategies with analytics: Links with selfreported measures and academic performance. *Journal of Learning Analytics*, *4*(2), 113–128.

Gatta, R., Lenkowicz, J., Vallati, M., Rojas, E., Damiani, A., Sacchi, L., De Bari, B., Dagliati, A., Fernandez-Llatas, C., Montesi, M., Marchetti, A., Castellano, M., & Valentini, V. (2017). pMineR: An innovative R library for performing process Mining in Medicine. In A. ten Teije, C. Popow, J. H. Holmes, & L. Sacchi (Eds.), *Artificial Intelligence in Medicine, Lecture Notes in Computer Science* (pp. 351–355). Springer International Publishing.

Giannakos, M. N., Sharma, K., Pappas, I. O., Kostakos, V., & Velloso, E. (2019). Multimodal data as a means to understand the learning experience. *International Journal of Information Management*, *48*, 108–119.

Greene, J. A., & Azevedo, R. (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), 18–29.

Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning*, *2*(2), 107–124.

Järvelä, S., & Bannert, M. (2019). Temporal and adaptive processes of regulated learning-what can multimodal data tell. *Learning and Instruction*, *72*, 101268.

Järvelä, S., Malmberg, J., Haataja, E., Sobocinski, M., & Kirschner, P. A. (2019). What multimodal data can tell us about the students' regulation of their learning process. *Learning and Instruction*, *72*(7), 4.

Kim, Y.-e., Brady, A. C., & Wolters, C. A. (2020). College students' regulation of cognition, motivation, behavior, and context: Distinct or overlapping processes? *Learning and Individual Differences*, *80*, 101872.

Kinnebrew, J. S., Biswas, G., Sulcer, B., & Taylor, R. S. (2013). Investigating self-regulated learning in teachable agent environments. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies, Springer International Handbooks of Education* (pp. 451–470). Springer.

Kinnebrew, J. S., Loretz, K. M., & Biswas, G. (2013). A contextualized, differential sequence mining method to derive Students' learning behavior patterns. *JEDM | Journal of Educational Data Mining*, *5*(1), 190–219.

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in massive open online courses. *Computers & Education*, *104*, 18–33.

Lim, L., Bannert, M., van der Graaf, J., Singh, S., Fan, Y., Surendrannair, S., Rakovic, M., Molenaar, I., Moore, J., & Gašević, D. (2022). Effects of real-time analytics-based personalized scaffolds on students' self-regulated learning. *Computers in Human Behavior*, *139*, 107547.

Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J., Pardo, A., Maldonado-Mahauad, J., & Pérez-Sanagustín, M. (2019). Detection of learning strategies: A comparison of process, sequence and network analytic approaches. In M. Scheffel, J. Broisin, V. Pammer-Schindler, A. Ioannou, & J. Schneider (Eds.), *Transforming learning with meaningful technologies, Lecture Notes in Computer Science* (pp. 525–540). Springer International Publishing.

Max Planck Institute for Psycholinguistics, T. L. A. (2021). ELAN.

McDonald, S., & Petrie, H. (2013). The effect of global instructions on think-aloud testing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '13* (pp. 2941–2944). Association for Computing Machinery.

Molenaar, I., & Järvelä, S. (2014). Sequential and temporal characteristics of self and socially regulated learning. *Metacognition and Learning*, *9*(2), 75–85.

Molenaar, I., van Boxtel, C. A., & Sleegers, P. J. (2011). Metacognitive scaffolding in an innovative learning arrangement. *Instructional Science*, *39*(6), 785–803.

Moos, D. C., & Azevedo, R. (2008). Monitoring, planning, and self-efficacy during learning with hypermedia: The impact of conceptual scaffolds. *Computers in Human Behavior*, *24*(4), 1686–1706.

Mudrick, N. V., Azevedo, R., & Taub, M. (2019). Integrating metacognitive judgments and eye movements using sequential pattern mining to understand processes underlying multimedia learning. *Computers in Human Behavior*, *96*, 223–234.

Munshi, A., Rajendran, R., Ocumpaugh, J., Biswas, G., Baker, R. S., & Paquette, L. (2018). Modeling learners' cognitive and affective states to scaffold SRL in open-ended learning environments. In Proceedings of the 26th conference on user modeling, adaptation and personalization (pp. 131–138).

Nodoushan, M. A. S. (2012). Self-regulated learning (SRL): Emergence of the RSRLM model, vol. 6. Issue: 3; pp. 1–16. Online Submission.

Paans, C., Molenaar, I., Segers, E., & Verhoeven, L. (2019). Temporal variation in children's self-regulated hypermedia learning. *Computers in Human Behavior*, *96*, 246–258.

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*, 422.

Pintrich, P. R. (2000). Chapter 14 - the role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). Academic Press.

Pollard, B., Dounas-Frazer, D. R., Lewandowski, H., & Ríos, L. (2019). Using think-aloud interviews to characterize model-based reasoning in electronics for a laboratory course assessment. *Physical Review Physics Education Research*, *15*(1), 10140.

Rakovic, M., Fan, Y., Van Der Graaf, J., Singh, S., Kilgour, J., Lim, L., Moore, J., Bannert, M., Molenaar, I., & Gasevic, D. (2022). Using learner trace data to understand metacognitive processes in writing from multiple sources. In LAK22: 12th International Learning Analytics and Knowledge Conference (pp. 130–141).

Ramey, J., Boren, T., Cuddihy, E., Dumas, J., Guan, Z., van den Haak, M. J., & De Jong, M. D. T. (2006). Does think aloud work? How do we know? In CHI '06 extended abstracts on human factors in computing systems, CHI EA '06, pp. 45–48, New York, NY, USA. Association for Computing Machinery.

Rienties, B., Tempelaar, D., Nguyen, Q., & Littlejohn, A. (2019). Unpacking the intertemporal impact of self-regulation in a blended mathematics environment. *Computers in Human Behavior*, *100*, 345–357.

Rogiers, A., Merchie, E., & Van Keer, H. (2020). What they say is what they do? Comparing task-specific self-reports, thinkaloud protocols, and study traces for measuring secondary school students' text-learning strategies. *European Journal of Psychology of Education*, *35*(2), 315–332.

Rovers, S. F. E., Clarebout, G., Savelberg, H. H. C. M., de Bruin, A. B. H., & van Merriënboer, J. J. G. (2019). Granularity matters: Comparing different ways of measuring self-regulated learning. *Metacognition and Learning*, *14*(1), 1–19.

Saint, J., Fan, Y., Gašević, D., & Pardo, A. (2022). *Temporally-focused analytics of self-regulated learning: A systematic review of literature* (p. 100060). Artificial Intelligence.

Saint, J., Fan, Y., Singh, S., Gasevic, D., & Pardo, A. (2021). Using process mining to analyse self-regulated learning: A systematic analysis of four algorithms. In *LAK21: 11th international learning analytics and knowledge conference, LAK21* (pp. 333–343). Association for Computing Machinery.

Saint, J., Gašević, D., Matcha, W., Ahmad Uzir, N., & Pardo, A. (2020). Combining analytic methods to unlock sequential and temporal patterns of self-regulated learning. In *Proceedings of the 10th International Conference on Learning Analytics & Knowledge* (p. 10). Frankfurt.

Saint, J., Whitelock-Wainwright, A., Gasevic, D., & Pardo, A. (2020). Trace-SRL: A framework for analysis of micro-level processes of self-regulated learning from trace data. In *IEEE transactions on learning technologies* (p. 1). IEEE Transactions on Learning Technologies.

Schoor, C., & Bannert, M. (2012). Exploring regulatory processes during a computer-supported collaborative learning task using process mining. *Computers in Human Behavior*, *28*(4), 1321–1331.

Schuitema, J., Peetsma, T., & Veen, I. V. D. (2012). Self-regulated learning and students' perceptions of innovative and traditional learning environments: A longitudinal study in secondary education. *Educational Studies*, *38*(4), 397–413 eprint: https://doi.org/10.1080/03055698.2011.643105

Schunk, D. H. (2005). Self-regulated learning: The educational legacy of Paul R. Pintrich. *Educational Psychologist*, *40*(2), 85–94 eprint: https://doi.org/10.1207/s15326985ep4002_3

Shipstead, Z., Lindsey, D. R. B., Marshall, R. L., & Engle, R. W. (2014). The mechanisms of working memory capacity: Primary memory, secondary memory, and attention control. *Journal of Memory and Language*, *72*, 116–141.

Siadaty, M., Gasevic, D., & Hatala, M. (2016). Trace-based micro-analytic measurement of self-regulated learning processes. *Journal of Learning Analytics*, *3*(1), 183–214.

Siadaty, M., Gašević, D., & Hatala, M. (2016b). Associations between technological scaffolding and micro-level processes of self-regulated learning: A workplace study. *Computers in Human Behavior*, *55*, 1007–1019.

Siadaty, M., Gašević, D., & Hatala, M. (2016c). Measuring the impact of technological scaffolding interventions on microlevel processes of self-regulated workplace learning. *Computers in Human Behavior*, *59*, 469–482.

Sobocinski, M., Malmberg, J., & Järvelä, S. (2017). Exploring temporal sequences of regulatory phases and associated interactions in low-and high-challenge collaborative learning sessions. *Metacognition and Learning*, *12*(2), 275–294.

Song, D., Hong, H., & Oh, E. Y. (2021). Applying computational analysis of novice learners' computer programming patterns to reveal self-regulated learning, computational thinking, and learning performance. *Computers in Human Behavior*, *120*, 106746.

Sonnenberg, C., & Bannert, M. (2015). Discovering the effects of metacognitive prompts on the sequential structure of SRL-processes using process mining techniques. *Journal of Learning Analytics*, *2*(1), 72–100.

Sonnenberg, C., & Bannert, M. (2016). Evaluating the impact of instructional support using data mining and process mining: A micro-level analysis of the effectiveness of metacognitive prompts. *Journal of Educational Data Mining*, *8*(2), 51–83.

Stark, L., Brünken, R., & Park, B. (2018). Emotional text design in multimedia learning: A mixed-methods study using eye tracking. *Computers & Education*, *120*, 185–196.

Taub, M., Mudrick, N. V., Azevedo, R., Millar, G. C., Rowe, J., & Lester, J. (2016). Using multi-level modeling with eye-tracking data to predict metacognitive monitoring and self-regulated learning with Crystal Island. In A. Micarelli, J. Stamper, & K. Panourgia (Eds.), *Intelligent tutoring systems, lecture notes in computer science* (pp. 240–246). Springer International Publishing.

Trevors, G., Feyzi-Behnagh, R., Azevedo, R., & Bouchet, F. (2016). Self-regulated learning processes vary as a function of epistemic beliefs and contexts: Mixed method evidence from eye tracking and concurrent and retrospective reports. *Learning and Instruction*, *42*, 31–46.

van der Graaf, J., Lim, L., Fan, Y., Kilgour, J., Moore, J., Bannert, M., Gasevic, D., & Molenaar, I. (2021). Do Instrumentation Tools Capture Self-Regulated Learning? LAK21: 11th International Learning Analytics and Knowledge Conference. https://doi.org/10.1145/3448139.3448181

Veenman, M. V. J. (2007). The assessment and instruction of self-regulation in computer-based environments: A discussion. *Metacognition and Learning*, *2*(2), 177–183.

Veenman, M. V. J. (2013). Assessing metacognitive skills in computerized learning environments. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies, springer international handbooks of education* (pp. 157–168). Springer.

Winne, P. H. (2010). Improving measurements of self-regulated learning. *Educational Psychologist*, *45*(4), 267–276.

Winne, P. H. (2011). *A cognitive and metacognitive analysis of self-regulated learning: Faculty of education*. Simon Fraser University.

Winne, P. H. (2017). The trajectory of scholarship about self-regulated learning. *Teachers College Record*, *119*(13), 1–16.

Winne, P. H. (2019). Paradigmatic dimensions of instrumentation and analytic methods in research on self-regulated learning. *Computers in Human Behavior*, *96*, 285–289.

Winne, P. H. (2020). Construct and consequential validity for learning analytics based on trace data. Computers and Human Behavior, (In press).

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated engagement in learning. in metacognition in educational theory and practice. *Metacognition in Educational Theory and Practice*, 277–304.

Winne, P. H., Hadwin, A. F., & Gress, C. (2010). The learning kit project: Software tools for supporting and researching regulation of collaborative learning. *Computers in Human Behavior*, *26*(5), 787–793.

Winne, P. H., & Jamieson-Noel, D. (2002). Exploring students' calibration of self reports about study tactics and achievement. *Contemporary Educational Psychology*, *27*(4), 551–572.

Winne, P. H., Nesbit, J. C., Ram, I., Marzouk, Z., Vytasek, J., Samadi, D., & Stewart, J. (2017). Tracing metacognition by highlighting and tagging to predict recall and transfer. AERA Online Paper Repository. ISSN: AERA Online Paper Repository.

Young, K. A. (2005). Direct from the source: The value of 'think-aloud' data in understanding learning.

Zhou, M., & Winne, P. H. (2012). Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction*, *22*(6), 413–419.

Zimmerman, B. J. (2000). Chapter 2 - attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Fan, Y., Rakovic, M., van der Graaf, J., Lim, L., Singh, S., Moore, J., Molenaar, I., Bannert, M., & Gašević, D. (2023). Towards a fuller picture: Triangulation and integration of the measurement of self-regulated learning based on trace and think aloud data. *Journal of Computer Assisted Learning*, 1–22. https://doi.org/10.1111/jcal.12801