



OPEN ACCESS

EDITED BY

Manuel Gentile,
Istituto per le Tecnologie Didattiche
ITD—Consiglio Nazionale delle
Ricerche, Italy

REVIEWED BY

Wahyu Nanda Eka Saputra,
Ahmad Dahlan University, Indonesia
Alexander Nussbaumer,
Graz University of Technology, Austria

*CORRESPONDENCE

Elizabeth Brooke Cloude
ecloude94@gmail.com

SPECIALTY SECTION

This article was submitted to
Digital Learning Innovations,
a section of the journal
Frontiers in Education

RECEIVED 26 April 2022

ACCEPTED 22 July 2022

PUBLISHED 11 August 2022

CITATION

Cloude EB, Azevedo R, Winne PH,
Biswas G and Jang EE (2022) System
design for using multimodal trace data
in modeling self-regulated learning.
Front. Educ. 7:928632.
doi: 10.3389/feduc.2022.928632

COPYRIGHT

© 2022 Cloude, Azevedo, Winne,
Biswas and Jang. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

System design for using multimodal trace data in modeling self-regulated learning

Elizabeth Brooke Cloude^{1*}, Roger Azevedo², Philip H. Winne³,
Gautam Biswas⁴ and Eunice E. Jang⁵

¹Penn Center for Learning Analytics, University of Pennsylvania, Philadelphia, PA, United States,

²School of Modeling, Simulation, and Training, University of Central Florida, Orlando, FL,

United States, ³Faculty of Education, Simon Fraser University, Burnaby, BC, Canada, ⁴Department of
Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN, United States,

⁵Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada

Self-regulated learning (SRL) integrates monitoring and controlling of cognitive, affective, metacognitive, and motivational processes during learning in pursuit of goals. Researchers have begun using multimodal data (e.g., concurrent verbalizations, eye movements, on-line behavioral traces, facial expressions, screen recordings of learner-system interactions, and physiological sensors) to investigate triggers and temporal dynamics of SRL and how such data relate to learning and performance. Analyzing and interpreting multimodal data about learners' SRL processes as they work in real-time is conceptually and computationally challenging for researchers. In this paper, we discuss recommendations for building a multimodal learning analytics architecture for advancing research on how researchers or instructors can standardize, process, analyze, recognize and conceptualize (SPARC) multimodal data in the service of understanding learners' real-time SRL and productively intervening learning activities with significant implications for artificial intelligence capabilities. Our overall goals are to (a) advance the science of learning by creating links between multimodal trace data and theoretical models of SRL, and (b) aid researchers or instructors in developing effective instructional interventions to assist learners in developing more productive SRL processes. As initial steps toward these goals, this paper (1) discusses theoretical, conceptual, methodological, and analytical issues researchers or instructors face when using learners' multimodal data generated from emerging technologies; (2) provide an elaboration of theoretical and empirical psychological, cognitive science, and SRL aspects related to the sketch of the visionary system called SPARC that supports analyzing and improving a learner-instructor or learner-researcher setting using multimodal data; and (3) discuss implications for building valid artificial intelligence algorithms constructed from insights gained from researchers and SRL experts, instructors, and learners SRL *via* multimodal trace data.

KEYWORDS

multimodal trace data, self-regulated learning, emerging technologies, system architecture, artificial intelligence

1. Introduction

Technology is woven into the fabric of the twenty-first century. Exacerbated by the pandemic of COVID-19, these emerging technologies have the capacity to increase accessibility, inclusivity, and quality of education across the globe (UNESCO, 2017). Emerging technologies include serious games, immersive virtual environments, simulations, and intelligent tutoring systems that have assisted learners in developing self-regulated learning (SRL) and problem-solving skills (Azevedo et al., 2019) across multiple domains (Biswas et al., 2016; Azevedo et al., 2018; Winne, 2018a; Lajoie et al., 2021), populations, languages, and cultures (Chango et al., 2021). Empirical evidence shows that SRL with emerging technology results in better learning gains compared to conventional methods (Azevedo et al., 2022). These technology-rich learning environments can record learners' multimodal trace data (e.g., logfiles, concurrent verbalizations, eye movements, facial expressions, screen recordings of learner-system interactions, and physiological signals) that instructors and education researchers can use to systematically monitor, analyze, and model SRL processes, and study their interactions with other latent constructs and performance with overall goals to augment teaching and learning (Azevedo and Gašević, 2019; Hadwin, 2021; Reimann, 2021).

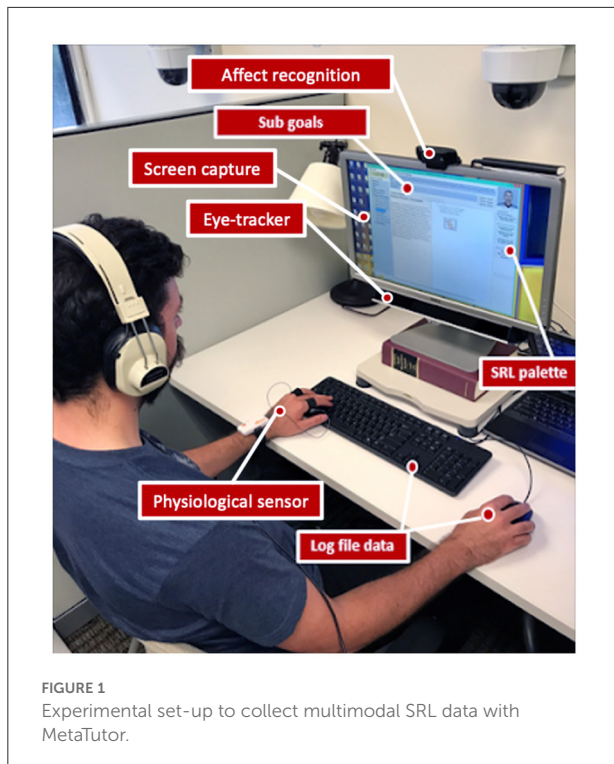
Emerging evidence points to key roles that multimodal data can play in this context (Jang et al., 2017; Taub et al., 2021) and has sparked promising data-driven techniques for discovering insights into SRL processes (Cloude et al., 2021a; Wiedbusch et al., 2021). Yet, major issues remain regarding roles for various SRL processes (e.g., cognitive and metacognitive; Mayer, 2019) and their properties: evolution or recursive nature over time, frequency and duration, interdependence, quantity vs. quality (e.g., accuracy in metacognitive monitoring), and methods for fusing multimodal trace data to link SRL processes to learning task performance. As research using multimodal trace data unfolds, we perceive an increased need to understand how instructional decisions can be forged by modeling regulatory patterns reflected by multimodal data in both learners and their instructors. We pose a fundamental question: Has the field developed the knowledge and the supporting processes to help researchers and instructors interpret and exploit multimodal data to design productive and effective instructional decisions?

In this paper, we provide an elaboration on psychological aspects related to the design of a teaching and learning architecture called SPARC that allows researchers or instructors to standardize, process, analyze, recognize, and conceptualize (SPARC) SRL signals from multimodal data. The goal is to help researchers or instructors represent and strive to understand learners' real-time SRL processes, with the aim to intervene and support ongoing learning activities. We envision a SPARC system to reach this goal. Specifically, we recommend

that the design of SPARC should embody a framework grounded in (1) conceptual and theoretical models of SRL (e.g., Winne, 2018a); (2) methodological approaches to measuring, processing, and modeling SRL using real-time multimodal data (Molenaar and Järvelä, 2014; Segedy et al., 2015; Bernacki, 2018; Azevedo and Gašević, 2019; Winne, 2019), and (3) analytical approaches that coalesce etic (researchers/instructors) and emic (learners) trace data to achieve optimal instructional support. We first discuss previous studies using learners' multimodal trace data to measure SRL during learning activities with emerging technologies. Next, we describe challenges in using these data to capture, analyze, and understand SRL by considering recent developments in analytical tools designed to handle challenges associated with multimodal learning analytics. Lastly, we recommend a hierarchical learning analytics framework and discuss theoretical and empirical guidelines for designing a system architecture that measures (1) learners' SRL alongside (2) researchers'/instructors' monitoring, analyzing, and understanding of learners' SRL grounded in multimodal data to forge instructional decisions. Implications of this research could pave the way for training artificial intelligence (AI) using data insights gained from researchers, instructors, and experts within the field of SRL that vary by individual characteristics including training background/experience, country, culture, gender, and many other diversity aspects. Algorithms trained using data collected on a diverse sample of interdisciplinary and international (1) SRL experts and researchers, (2) instructors, and (3) learners has the potential to automatically detect and classify SRL constructs across a range of data channels and modalities could serve to mitigate the extensive challenges associated with using multimodal data and assist educators in making effective instructional decisions guided by both theory and empirical evidence.

1.1. Characteristics of multimodal data used to reflect SRL

To gather multimodal data about SRL processes during learning, learners are instrumented with multiple sensors. Examples include electro-dermal bracelets (Lane and D'Mello, 2019), eye tracking devices (Rajendran et al., 2018), and face tracking cameras to capture facial expressions representing emotions (Taub et al., 2021). These channels may be supplemented by concurrent think-aloud data (Greene et al., 2018), online behavioral traces of learners using features in a software interface (Winne et al., 2019) and gestures and body movements in an immersive virtual environment (Raca and Dillenbourg, 2014; Johnson-Glenberg, 2018). This wide array of data can reflect when, what, how, and how long learners interact with specific elements in a

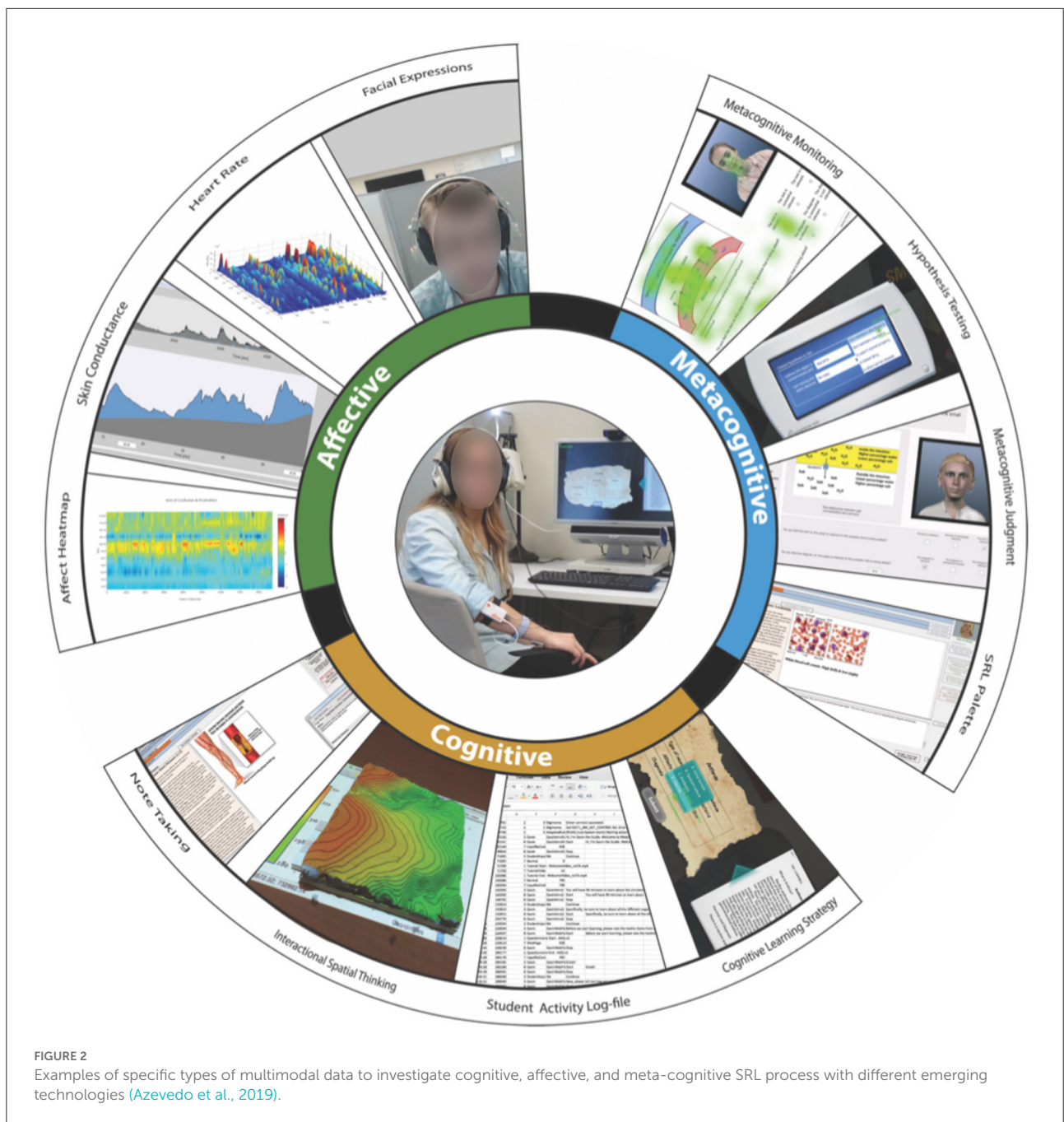


learning environment—e.g., reading and highlighting specific text, inspecting diagrams, annotating particular content, manipulating variables in simulations, recording, and analyzing data in problem-solving tasks, and interacting with pedagogical agents (Azevedo et al., 2018; Winne, 2019). Multimodal data gathered across these channels offer advantages in representing latent cognitive, affective, metacognitive, and motivational processes that are otherwise weakly signaled in any single data channel (Greene and Azevedo, 2010; Azevedo et al., 2018) (Figure 4).

A typical laboratory experimental set-up shown in Figure 1 illustrates a college student instrumented during learning with MetaTutor, a hypermedia-based intelligent tutoring system designed to teach about the human circulatory system (Azevedo et al., 2018). In addition to pre- and post-measures of achievement and self-report questionnaires not represented in the figure, multimodal instrumentation gathers a wide range of data about learning and SRL processes. Mouse-click data indicate when, how long, and how often the learner selects a page to study. Features and tools available for to the learner in a palette, such as self-quizzing and typing a summary, identify when the learner makes metacognitive judgments about knowledge (Azevedo et al., 2018) and how that might change across different learning goals (Cloude et al., 2021b). An electro-dermal bracelet records signals documenting changes in skin conductance produced by sympathetic innervation of sweat glands, a signal for arousal that can be matched to

the presence of external sensory stimuli (Lane and D'Mello, 2019; Messi and Adrario, 2021; Dindar et al., 2022). Eye movements operationalize what, when, where, and how long the learner attends to, scans, revisits, and reads (or rereads) content and consults displays, such as a meter showing progress toward goals (Taub and Azevedo, 2019; Cloude et al., 2020). Dialogue recorded between the learner and any of the four pedagogical agents embedded in MetaTutor identify system-provided scaffolding and feedback. Screen-capture software records and time stamps how and for how long the learner interacts with all these components and provides valuable contextual information supplementing multimodal data. A webcam samples facial features used to map the sequence, duration, and transitions between affective states (e.g., anger, joy) and learning-centered emotions (e.g., confusion).

Figure 2 illustrates examples of multimodal data used to study SRL across several emerging technologies including MetaTutor, Crystal Island, and MetaTutor IVH (Azevedo et al., 2019). The figure omits motivational beliefs because motivation has been measured almost exclusively using self-reports (Renninger and Hidi, 2019). Multimodal data structures are wide in scope, complexly structured and richly textured. For example, a learner reading about the anaphase stage of cell division may have metacognitively elected to apply particular cognitive tactics (e.g., selecting key information while reading, then assembling those selections across the text and diagram as a summary). At that point, eye-gaze data show repeated saccades and fixations between text and diagrams as the learner utters a metacognitive judgment captured *via* think-aloud, “I do not understand the structures of the heart presented in the diagram.” Concurrently, physiological data reveal a spike in heart rate and analysis of the learner’s facial expressions indicate frustration. Inspecting and interpreting this array of time-stamped data sampled across multiple scales of measurement and spanning several durations pose significant challenges for modeling cognition, affect, metacognition, and motivation. Which data channels relate to the different SRL features (cognition, affect, metacognition, and motivation)? Is one channel better at operationalizing a specific SRL feature? How should the different data channels be configured so that researchers can accurately monitor, analyze, and interpret SRL processes in real-time? What is (are) the appropriate temporal interval(s) for sampling each data channel, and how are characterizations across data channels used to support accurate and valid interpretations of latent SRL processes? Assuming these questions are answered, how can researchers be guided to make instructional decisions that support and enhance learners’ SRL processes? We suggest guidelines to address these questions in the form of a SPARC system. Our paper is based in theoretical and empirical literature from the science of learning, and evolving understanding about multimodal trace data.



1.2. Challenges in representing SRL using multimodal data

Time is a necessary yet perplexing feature needing careful attention in analyses of multimodal data sampled over multiple channels. How should data with differing frequencies be synchronized and aligned when modeling processes? To blend multisynchronous data, time samples need to be rescaled to a uniform metric (e.g., minutes or seconds). Multimodal data

may require filtering to dampen noise and lessen measurement errors. Decisions about these adjustments can be made usually only after learners have completed segments in or an entire study session. Judgments demand intense vigilance as researchers and instructors scan multimodal data and update interpretations grounded in multimodal data. If researchers or instructors attempt to monitor and process multimodal data in real-time to intervene during learning—e.g., prompting learners to avoid or correct unproductive studying tactics—vigilance will be one key.

In the presence of dense and high-velocity data, critical signals in multimodal data that should steer instructional decision-making may be missed as demonstrated in Claypoole et al.'s (2019) study. Their findings showed increases in stimuli per minute decreased participants' sensitivity (discriminating hits and false alarms) and increased time needed to detect pivotal details (Claypoole et al., 2019). As well, because vigilance declines over time and tasks (Hancock, 2013; Greenlee et al., 2019), counters need to be developed if multimodal data are to be useful inputs for real-time instructional decision-making to support learners' SRL. Furthermore, a particular and pressing challenge for moving this research forward is determining what information can be used from learners, such as *who* will be allowed to access potentially personal data, and *how* might such users obtain permissions to ethically use the data (Ifenthaler and Schumacher, 2019), meanwhile maintaining confidentiality, reliability, security, privacy, among many others that align with security and privacy policies that may vary across international lines (Ifenthaler and Tracey, 2016).

International researchers have begun to engineer systems to manage challenges associated with processing, analyzing, and understanding multimodal data. For example, SensePath (Nguyen et al., 2015) was built in the United Kingdom and designed to reduce demand for vigilance by providing visual tools that support articulating multichannel qualitative information unfolding in real-time, such as transcribed audio mapped onto video recordings. Blascheck et al. (2016) developed a similar visual-analytics tool in Germany to support coding and aligning mixed-method multimodal data gathered over a learning session in the form of video and audio recordings, eye-gaze tracks, and behavioral-interactions. Their system was designed to support researchers in (1) identifying patterns, (2) annotating higher-level codes, (3) monitoring data quality for errors, and (4) visually juxtaposing codes across researchers to foster discussions and contribute to inter-rater reliability (Blascheck et al., 2016). These systems illustrate progress in engineering tools researchers and instructors need to work with complex multimodal data, such as those required to reflect learning and SRL. But two gaps need filling. First, systems developed so far could further mine and apply research on how humans make sense of information derived from complex multimodal data. Second, systems have not yet been equipped to gather and mine data about how researchers and instructors use system features using data representations and visual tools. Furthermore, how might researchers and instructors use system features differently as their goals and intentions, training, and beliefs about phenomenon may vary? In other words, developing models that represent how diverse users leverage the system and its features need to be considered in future work to build a multi-angled view of the total system.

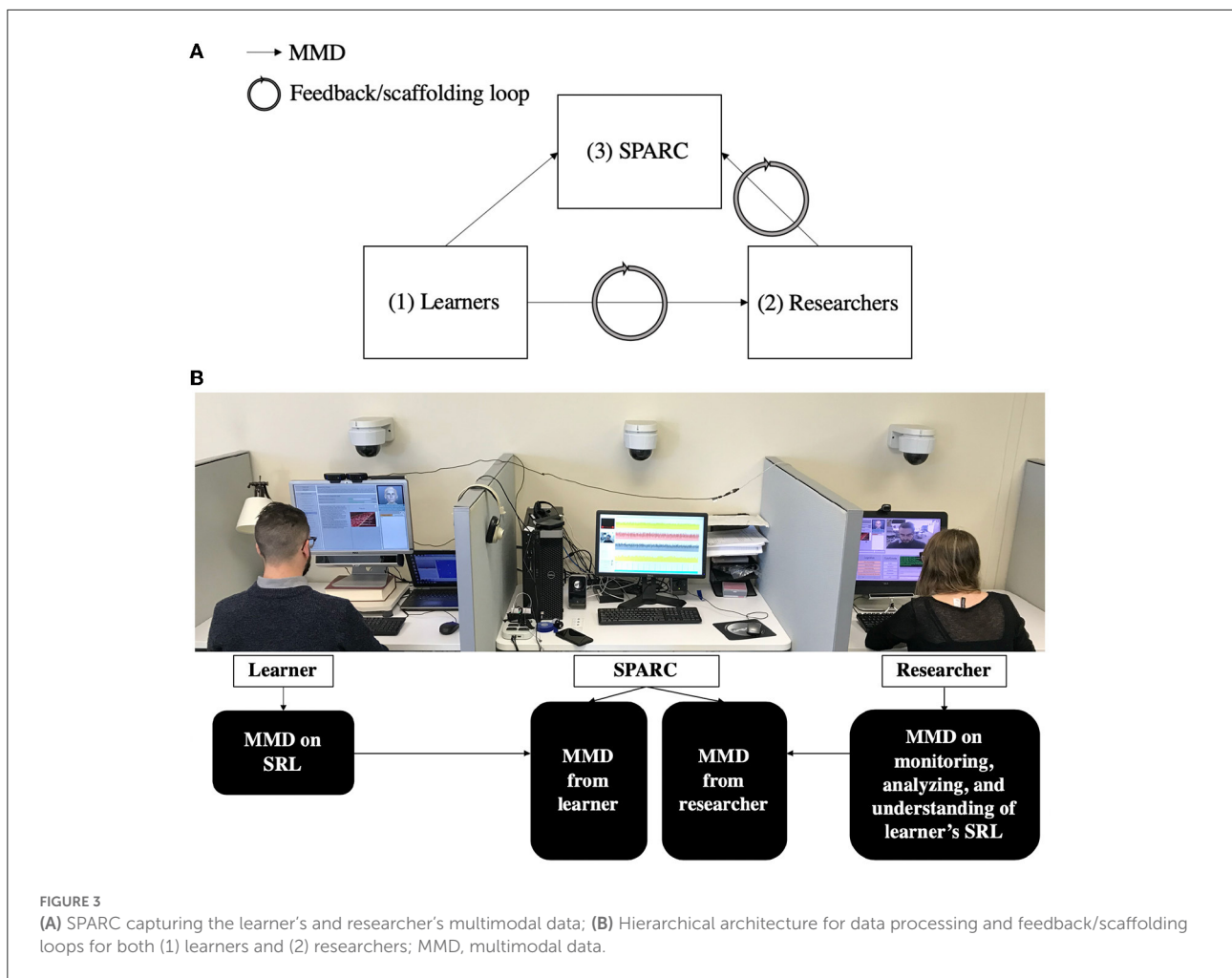
One notable system designed for multimodal signal processing and pattern recognition in real-time is the Social Signal Interpretation framework (SSI) developed in Germany

by Wagner et al. (2013). SSI was engineered to simultaneously process data ranging from physiological sensors and video recordings to Microsoft's Kinect. A machine-learning (ML) pipeline automatically aligns, processes, and filters multimodal data in real-time as it is collected. Once data are processed, automated recognition routines detect and classify learners' activities (Wagner et al., 2013). Another multimodal data tool, SLAM-KIT (Noroozi et al., 2019), was built in the Netherlands and designed to study SRL in collaborative contexts. It reduced the volume and variety in multimodal data to allow teachers, researchers, or learners to easily navigate in and across data streams, analyze key features of learners' engagements, and annotate and visualize variables or processes that analysts identify as signals of SRL. Notwithstanding the advances these systems represent, issues remain. One is how to coordinate (a) data across multiple channels with multiple metrics alongside (b) static and unfolding contextual features upon which learners pivot when they regulate their learning (Kabudi et al., 2021). Another target for improvement is supporting researchers to monitor, analyze and accurately interpret matrices of the multimodal data for tracking SRL processes. Factors that may affect such interpretations include choosing and perhaps varying optimal rates to sample data, synchronizing and temporally aligning data in forms that support searching for patterns, and articulating online data with contextual data describing tasks, domains learners study, and characteristics of settings that differentiate the lab from the classroom from home, and individual vs. collaborative work. All these issues have bearing on opportunities to test theoretical models (e.g., Winne, 2018a) and positively influence learning.

1.3. Overview of the SPARC architecture

Making effective just-in-time and just-in-case instructional decisions demands expertise in monitoring, analyzing, and modeling SRL processes. Are researchers and/or instructors equipped to meet these challenges when delivered fast-evolving multimodal data? To address these issues, we discuss a hierarchical learning analytics framework and guidelines for designing a theoretically and empirically-based suite of analytical tools to help researchers and/or instructors solve challenges associated with receiving real-time multimodal data, monitoring SRL, assembling interventions, and tracking dynamically unfolding trajectories as learners work in emerging technologies. The SPARC system we recommend is a dynamic data processing framework in which multimodal data are generated by (1) learners, (2) researchers and/or instructors, and (3) the system itself. These streams of data are automatically¹ processed in real-time negative feedback loops. Two features

¹ We cannot elaborate on this process due to space limitations. However, there are several tools currently used by interdisciplinary

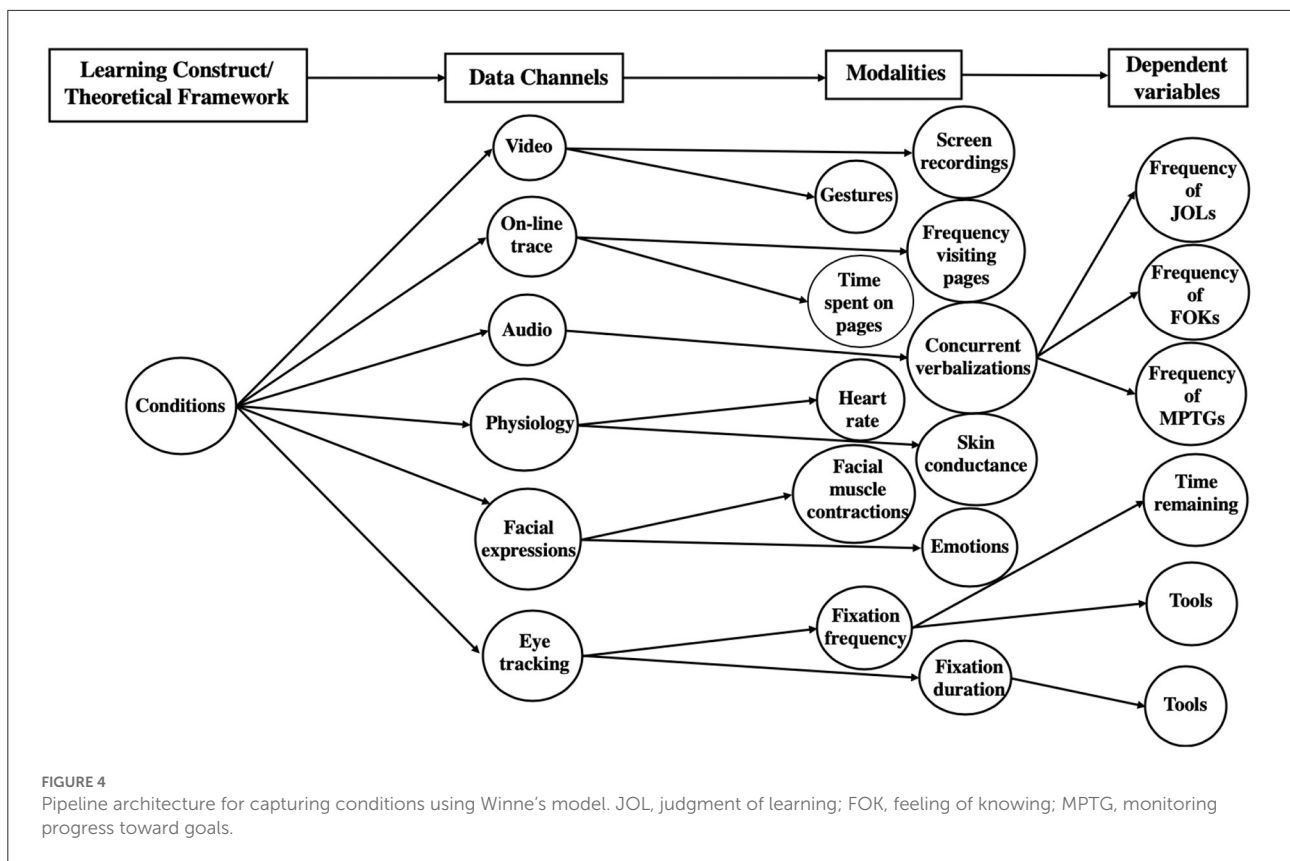


distinguish SPARC from other tools. First, researchers and/or instructors are positioned in three roles: data generators, data processors, and instructional decision makers. Second, the overall system has a triifb-partite structure designed to record series of multimodal data for real-time processing across the timeline of instructional episodes. The target SPARC aims for is iteratively tuning data gathering, data processing, and scaffolding for both learners and researchers/instructors, thereby helping both players more productively self-regulate their respective and interactive engagements (see Figure 3).

Imagine an instructor and learner are about the engage in a learning session with an emerging technology. See Figure 3—both the users (researchers and learners) interact with content while their data are recorded on such interaction. For example, both instructor and learner are instrumented with multiple sensors, including a high-resolution eye tracker and physiological bracelet, meanwhile, both users' video, audio, and

researchers to view, process, and analyze learners' multimodal data in real-time such as the iMotions's research platform.

screens are being recorded. Further, once the learner begins interacting with content, their data are recorded and SRL variables are generated in real-time. These data are displayed to the instructor so they can see what the learner is doing such as where their eyes are attending to specific text and/or diagrams, including the sequence and amount of time they are engaging with content. The instructor can also see the learner's physiology spikes, facial expressions of emotions, screen recording, and speech. Meanwhile, the instructor is also being recorded with sensors. Once the learner begins engaging with materials, data on the instructor measure the degree to which they attend—i.e., monitor, analyze, and understand the learner's SRL *via* data channels and modalities over the learning activity. From these data, SPARC can calculate the degree to which the instructor is biased to oversample a specific channel of learner data, say, eye-gaze behaviors. For instance, SPARC can detect this bias when the instructor's eye-gaze and logfiles data show they infrequently sample other data channels carrying critical SRL signals. Here, SPARC should take three steps. First, alert the instructor to shift attention, e.g., by posting a notification, "Is



variance in the learner's eye-gaze data indicating a change in standards used for metacognitive monitoring?" Second, SPARC varies illumination levels of its panels to cue the instructor to shift attention to the panel displaying learner eye movements. Third, a pop-up panel shows the instructor a menu of alternative interventions. In this panel, each intervention is described using a 4-spoke radar chart grounded in prior data gathered from other learners: the probability of learner uptake e.g., Bayesian knowledge tracing (Hawkins et al., 2014), the cognitive load associated with the intervention, negative impact on other study tactics such as note-taking, and learner frustration triggered upon receiving SPARC's recommendation to adapt standards for metacognitively monitoring understanding. Then, SPARC monitors the instructor's inspection of display elements to update its model of the instructor's biases for particular learner variables—e.g., a preference to limit learner frustration when selecting an intervention to be suggested to the learner. And later, as SPARC assembles data about how the learner reacts to the instructor's chosen intervention, the model of this learner is updated to sharpen a forecast about the probability of intervention uptake and impact of the instructor's chosen intervention on the profile of study tactics this learner uses. SPARC'S complex and hierarchical approach to recording, analyzing, and interacting with both agents in

instructional decision-making and self-updating models both pushes SPARC past the boundaries of other multimodal systems. By dynamically updating models of learners and instructors (or researchers), instructional decisions are grounded and iteratively better grounded on the history of all three players—self-regulating learners, self-regulating instructors, and the self-regulating system itself. If widely distributed to create genuinely diverse big data, SPARC would significantly advance research in learning science and mobilize research based on expanding empirical evidence about SRL and interventions that affect it (e.g., Azevedo and Gašević, 2019; Winne, 2019).

1.3.1. Information processing theory of self-regulated learning

The first step toward a SPARC system is building a pipeline architecture that operationalizes a potent theoretical framework to identify, operationalize, and estimate values for parameters of key variables (Figure 4). From a computer-science perspective, a pipeline architecture captures a sequence of processing and analysis routines such that outputs of one routine (e.g., capturing multimodal data on researchers and learners) can be fed directly into the next routine without human intervention (e.g., processing learners' and researchers'

multimodal data separately). Overall, an ideal system should support valid interpretations of latent causal constructs. SPARC adopts an information processing view of self-regulated learning along with assumptions fundamental to this perspective (Winne and Hadwin, 1998; Winne, 2018a, 2019).

According to the Winne-Hadwin model of SRL, human learning is an agentic, cyclic, and multi-faceted process centered on monitoring and regulating information in a context of physical and internal conditions bearing on cognitive, affective, metacognitive, and motivational (CMMM) processes during learning (Malmberg et al., 2017; Azevedo et al., 2018; Schunk and Greene, 2018; Winne, 2018a). Individual differences, such as prior knowledge about a domain and self-efficacy for a particular task, and contextual resources (e.g., tools available in a learning environment) set the stage for a cycle of learning activity (Winne, 2018a). Consequently, to fully represent learning as SRL requires gathering data to represent cognitive, affective, metacognitive, and motivational processes while learners (and researchers) learn, reason, problem solve, and perform. Also, to ensure that just-in-case instructional decisions can be grounded in this dynamic process and assumptions of SRL (Winne, 2018a), we propose capturing multimodal data about the instructional decision maker (i.e., researcher) is just as relevant and important as capturing multimodal data about the learner. Thus, the pipeline architecture should be fed data across channels and modalities tapping cognitive, affective, metacognitive, and motivational processes separately for both learners and researchers (see Figure 3). The Winne-Hadwin model of agentic SRL (Winne and Hadwin, 1998; Winne, 2018a) describes learning in terms of four interconnected and potentially nonsequential phases. In Phase 1, the learner surveys the task environment to identify internal and external conditions perceived to have bearing on the task. Often, this will include explicit instructional objectives set by an instructor. In phase 2, based on the learner's current (or updated) understanding of the task environment, the learner sets goals and develops plans to approach them. In phase 3, tactics and strategies set out in the plan are enacted and features of execution are monitored. Primary among these features is progress toward goals and subgoals the learner framed in phase 2; and emergent characteristics of carrying out the plan, such as effort spent, pace, or progress. In phase 3, the learner may make minor adaptations as judged appropriate. In phase 4, which is optional, the learner reviews work on the task writ large. This may lead to adaptive re-engagement with any of the preceding phases as well as forward reaching transfer (Salomon and Perkins, 1989) to shape SRL in similar future tasks.

For the physical set-up² illustrated in Figure 3, entries in Table 1 demonstrate a complex coordination between learners'

² This figure is for illustration purposes only since, ideally, we would physically separate the researcher and learner to avoid bias, social desirability, etc.

and researchers' multimodal data facilitated by a SPARC system. In this table, we provide two examples that map assumptions based on Winne's phases to the learners' and researchers' multimodal SRL data (Winne, 2018a). Included are a researcher's monitoring, analyzing, and understanding of a learner's SRL based on the learner's multimodal data and instructional interventions arising from the researcher's inferences about the learner's SRL. Two contrasting cases are provided. The first is a straightforward example of a learner's multimodal data that is easy to monitor, analyze, and understand. This leads, subsequently, to an accurate inference about SRL by the researcher who does not require SPARC to intervene in supporting the researcher's instructional intervention. A second scenario is more complex. The learner presents several signals in multimodal data, which could reflect multiple and diverse issues related to their motivation, affect, and cognition. The researcher must intervene to scaffold and prompt the learner but it is not clear where to begin given multiple instructional concerns. So, SPARC intervenes to scaffold the researcher to optimize instructional decision-making based on pooled knowledge about the learner's SRL and the researcher's past successful interventions.

Throughout each phase of SRL, five facets characterize information and metacognitive events, encapsulated in the acronym COPES. Conditions are resources and constraints affecting learning. Time available to complete the task, interest, and free or restricted access to just-in-time information resources are examples. Operations are cognitive processes learners choose for manipulating information as they address the task. Winne models five processes: searching, monitoring, assembling, rehearsing, and translating (SMART). Products refer to information developed by operations. These may include knowledge recalled, inferences constructed to build comprehension, judgments of learning, and recognition of an arising affect. Evaluations characterize the degree to which products match standards, criteria the learner set or adopted from external sources (e.g., an avatar) to operationalize success in work on the task (e.g., pace), and its results. Since SPARC's pipeline architecture is intended to capture learner's SRL processes, we argue it is critical to investigate how to map multimodal data from specific data channels to the theoretically-referenced constructs in the Winne and Hadwin model of SRL and the cognate models for tasks (COPES) and operations (SMART) within tasks. For instance, what data channels or combinations of data channels and modalities best represent a cognitive strategy? Do these data also indicate elements of metacognition? A system such as SPARC should help answer these questions by examining how researchers' and learners' multimodal data might reflect these processes and how those processes impact performance and instructional decision-making, respectively.

We outline a potentially useful start for mapping data channels and modalities to theoretical constructs based on

TABLE 1 Learner's and researcher's multimodal SRL data aligned with phase 2 of Winne's model of SRL and corresponding instructional strategies based on unambiguous signals in data.

Learning context	Learner's MMD	Researcher's MMD	SPARC
1. Learner engages with biology content and sets	<ul style="list-style-type: none"> • Concurrent verbalizations via audio recording (e.g., "my goal is to learn about how blood flows through the four chambers of the heart."); • Screen recordings showing learner-system interactions; • Eye gaze illustrating where learner is searching for relevant information on the IF 	<ul style="list-style-type: none"> • Concurrent verbalizations via audio recording (e.g., "I will review learner's data within the context of their goal.... They are not monitoring the right information."); • Screen recordings of system interactions with learner's MMD; • Eye gaze on learner's interaction with content on IF 	<p>SPARC observes and updates its user model based on</p> <ol style="list-style-type: none"> 1. Researcher's MMD, including eye gaze, utterances, and screen recordings of researcher-system interactions 2. Learner's MMD, including eye gaze, utterances, and screen recordings of learner-system interactions
• Goals and			
• Plans based on their current (or updated) understanding			
2. Researcher observes learner's MMD on their IF			
3. SPARC observes both the researcher's and learner's MMD			

the COPES model (Winne, 2018a, 2019; Winne P., 2018), specific to conditions³ (Figure 3). As specified in Figure 3, some criterion variables refer to information captured as a learner verbalizes monitoring of engagement in a task (e.g., frequency of judgment of learning, feeling of knowing). Other data obtained from eye tracking instrumentation represents learners' assessing conditions, such as time left for completing the task signaled by viewing a countdown timer in the interface. These data expand information on conditions beyond records of how frequently learners visit pages in MetaTutor (Azevedo et al., 2018), edit a causal map in Betty's Brain (Biswas et al., 2016), or highlight text in nStudy (Winne et al., 2019). Together, these multimodal data characterize how, when, and with what the learner is proceeding with the task and engaging in SRL. A pipeline architecture for SPARC affords modularity as illustrated in Figure 3. The pipeline can customize data cleaning, pre-processing, and analysis routines for each one of the sensing modalities (e.g., think-alouds, eye-gaze and on-line behavioral traces). It also provides separate analysis routines for each of the constructs (e.g., conditions vs. operations vs. products), data channels, modalities, and criterion variables of which can work with the time series (or event sequence data) generated by the previous module. In sum, supporting researchers in constructing meaningful and valid inferences about SRL from multimodal signals in learners' data requires building a pipeline architecture aligned to a theoretical model of SRL. But this begs a key question. After variables are mapped onto a model of SRL, how can researchers' inferences be reasonably adjudicated? How valid are they?

³ Due to space limitations, we will not go into depth on how the SPARC pipeline will be structured to capture, operationalize, and process variables across settings, tasks, and domains that are aligned with the information-processing theory of SRL, including COPES and SMART.

2. Methods

2.1. Empirical synthesis on monitoring, analyzing, and understanding of multimodal data

Setting aside for the moment issues of alignment between SPARC and the Winne and Hadwin model of SRL, it is prudent to synthesize empirical research related to what, when, and how researchers might examine learners' multimodal data to model and understand dynamically unfolding SRL processes. Consequently, we next examine research on humans (a) monitoring information for patterns, (b) analyzing signals detected in patterns, and (c) constructing understanding(s) of this information matrix by monitoring and analyzing stimuli. Further, we emphasize previous methods and findings in literature as potential directions for leveraging trace data to define cognitive and metacognitive aspects of SRL constructs such as monitoring, analyzing, and understanding of SRL in researchers, instructors, and learners. Finally, we discuss challenges and future directions for the field to consider in ways to leverage multimodal data to advance the design of emerging technologies in modeling SRL.

2.1.1. Monitoring real-time multimodal data

Cognitive psychological research on information processing and visual perception—specifically, selective attention, and bottom-up/top-down attentional mechanisms (Desimone and Duncan, 1995; Desimone, 1996; Duncan and Nimmo-Smith, 1996)—is a fruitful starting point to examine factors that bear on how approaches for defining monitoring of SRL signals in multimodal data. When instructors or researchers encounter multimodal data, only a select partition of the full information matrix can be attended to at a time. One factor governing what

can be inspected is the size of the retina which determines how much visual information is available for processing. Where humans look typically reveals foci of attention and, thus, what information is available for processing (van Zoest et al., 2017). Multimodal data are typically presented across multiple displays and, often, as temporal streams of data. Attending to displays, each representing a particular modality, precludes attending to other data channels. This gives rise to two key questions: (1) What is selected? (2) What is screened out? (Desimone, 1996). One theory describing a mechanism for controlling attention is the biased-competition theory of selective attention (Desimone, 1996; Duncan and Nimmo-Smith, 1996). It proposes a biasing system driven by bottom-up and top-down attentional control. Bottom-up attentional control is driven by stimuli, e.g., peaks in an otherwise relatively flat progression of values in the learners' data that SPARC supplies to researchers and instructors. Bottom-up visual attention is skewed to sample information in displays based on shapes, sizes, and colors, and motion, while top-down attentional control is influenced by a researcher's or instructors goals and knowledge—declarative, episodic, and procedural—both of which are moderated by their beliefs and attitudes (Anderson and Yantis, 2013; Anderson, 2016). In the context of multimodal data SPARC displays, attention is directed in part by a researcher's knowledge about data in a particular channel, e.g., the relative predictive validity of facial expressions compared to physiological signals as indicators of learners' arousal. Another factor affecting the researcher's or instructors attention is the degree of training or expertise in drawing grounded inferences about an aspect of learners' SRL—e.g., recognizing facial expressions of frustration. A third factor is the researchers' or instructors preferred model of learning (e.g., this is what I believe SRL looks like). In the case of SPARC, this is familiarity with and commitment to the 4-phase model of SRL and the COPES schema within each phase. Thus, a key aspect of designing a system like SPARC required situating multimodal data around the goal of the (a) session (e.g., detect SRL in a learner's multimodal data) and (b) the user's goals, beliefs, training, education, and familiarity with and commitment to the 4-phase model of SRL and COPES schema within each phase. Variables defining monitoring behaviors need to be contextualized or evaluated against these criteria or set of standards.

2.1.2. Data channels that capture monitoring behaviors

Eye-tracking methodologies have opened a window into capturing implicit monitoring processes (Scheiter and Eitel, 2017). Mudrick et al. (2019) studied pairs of fixations to identify implicit metacognitive processing. Participants' fixations across text and a diagram were examined for dyads where the information was experimentally manipulated to be consistent

or inconsistent (e.g., the text described blood flow but a diagram illustrated lung gas exchange). For each dyad, participants metacognitively judged how relevant information in one medium was to information in the other medium. When information was consistent across dyads, participants more frequently traversed sources and made more accurate metacognitive judgments on the relevance of information in each medium. Eye-gaze data were a strong indicator of metacognitive monitoring and accuracy of judgments. Eye-gaze data also signal other properties of metacognition. Participants in Franco-Watkins et al.'s (2016) study were required to make a decision in a context of relatively little information. In this case, they fixated longer on fewer varieties of information. As variety of information increased, fixations settled on more topics for shorter periods of time. Variety and density of information affected metacognitive choices about sampling information in their complex information displays (Franco-Watkins et al., 2016).

These findings forecast how researchers or instructors may attend to multimodal data with SPARC. For example, if less information is available—i.e., a learner is not thinking aloud and displays a facial expression signaling confusion, will researchers or instructors bias sampling of data in classifying the learner's state by fixating longer on a panel displaying facial expression data, or will they suspend classification to seek data in another channel? A SPARC system would need to collect information on if, when, where, and for how long the user attends to a specific modality or channel, and then prompt the researcher or instructors to introduce data from another channel before classifying learner behavior and recommending a shift in learner behavior. The value of eye-gaze data as proxies for implicit processes such as attention and metacognitive monitoring lead us to suggest that SPARC measures researchers' or instructors' eye-gaze behaviors. Sequences of saccades, fixations, and regressions while monitoring multi-panel displays of learners' multimodal SRL data during a learning activity may reveal how, when, and what researchers or instructors are monitoring in the learners' multimodal data as they strive to synthesize information across modalities. Furthermore, information on what the user is attending to would reveal what the user is *not* attending to that may be potentially important. It would be important for SPARC to also define lack of attention to potentially operationalize the users' goal or intention and whether they are aligned with detecting SRL processes across the data. These data can track whether, when, and for how long users attend to discriminating or non-discriminating signals, sequences, and patterns in learners' multimodal data. Logged across learners and over study sessions, SPARC's data could be mined to model a researcher's or instructors' biases for particular channels in particular learning situations. Beyond eye-tracking data, can other methodologies reveal how researchers analyze learners' engagements during a learning session?

2.2. Analyzing real-time multimodal data

Analyzing and reasoning are complex forms of cognition (Laird, 2012). They dynamically combine knowledge and critical-thinking skills such as inductive and deductive reasoning and may involve episodically-encoded experiences (Blanchette and Richards, 2010). Theoretically, after researchers or instructors allocate attention to multimodal data, they must analyze and then reason about patterns and their sequences in relation to SRL phases and processes. As such, SPARC needs to measure how researchers or instructors search and exploit patterns of multimodal data (relative to other patterns) to make inferences about learners' SRL and recommend interventions. A fundamental issue here is to operationally define a pattern in a manner that achieves consensus among researchers or instructors and can be reliably identified when multimodal data range across data channels. SPARC's capabilities should address questions such as is there a pattern in eye-gaze data that is indicative of SRL? What patterns of gaze data does a researcher use to infer a learner's use and adaptation of tactics, or occurrences of metacognitive monitoring prompted by changing task conditions? In what ways do researchers' or instructors' eye-gaze patterns change over time (e.g., are they focusing on one data channel or more than one? Does their degree of attention change to other data channels?) and, if paired with other modalities, does this change reflect limits or key features in learners' SRL? Do changes in one modality of data indicate changes in other multimodal data and are these changes related to a user's instructional decision-making? We establish a ground truth regarding the validity and reliability of learners', instructors', and researchers' multimodal data patterns (see Winne, 2020). SPARC should be capable of detecting when users are analyzing and reasoning based on multimodal data describing how the user examined learners' multimodal eye-gaze behavior, interactions with the content (logfiles), physiology profile, facial expressions, and other channels. Again, researchers' or instructors' multimodal data play a key role in successively tuning the overall system.

2.2.1. Data channels that capture analyzing behaviors

Some research capturing data to infer implicit processing, such as analyzing and reasoning, used online behavioral traces (Spires et al., 2011; Kinnebrew et al., 2015; Taub et al., 2016); other studies used eye movements data (Catrysse et al., 2018) or concurrent think-aloud verbalizations (Greene et al., 2018). Taub et al. (2017) analyzed learners' clickstream behavior as relevant or irrelevant to the learning objective (e.g., learn about biology), and then applied sequential pattern mining analyses. Two distinct patterns of reasoning differed in efficiency, defined as fewer attempts toward successfully meeting the objective of the learning session. Defining logged learner actions based

on relevance to a learning objective is useful for capturing and measuring analyzing and reasoning behaviors (Taub et al., 2017). This technique could also be applied to the researchers' multimodal data as well. For example, do the researchers' mouse clicks, keyboard strokes, etc. reflect their analysis of the learners' multimodal data in relation to meeting the objective of the learning session—e.g., learning about the circulatory system. Is the researcher selecting modalities to evaluate whether the learner is working toward this objective (e.g., the learner is reading through content that is unrelated to the circulatory system and so, for instance, the researcher examining what content the learner is reviewing?) to guide their instructional decision making? Eye movements may also indicate the how extensively information is processed during learning (Catrysse et al., 2018). For instance, when participants reported both deep and surface-level information processing, they tended to fixate longer and revisit content more often than participants who reported only surface-level processing (Catrysse et al., 2018); but see (Winne, 2018b, 2020) for a critique of the “depth” construct).

Other studies have used think-aloud protocols for data mining to seek emic descriptions of information processing and reasoning (Greene et al., 2018). Muldner et al. (2010) drew inferences from concurrent verbalizations representing self-explaining, describing connections between problems or examples, and other key cognitive processes (e.g., summarizing content) in Physics during learning with an intelligent tutoring system. Similar analytic approaches were used to understand clinicians' diagnostic reasoning (Kassab and Hussain, 2010). Si et al. (2019) used a rubric to quantify the quality of their participants' reasoning about a diagnosis. Quality of reasoning was positively related to clinical-reasoning skills and accurate diagnoses. These findings indicate that think-aloud methods can quantify how and when researchers analyze and reason about learners' multimodal data (Si et al., 2019). Multimodal data about researchers' engagements with learners' multimodal data can inform where researchers' monitoring and analyzing behaviors about deciding if, when, and how to scaffold the learners' SRL. Negative feedback loops built into SPARC (see Figure 3) offer pathways for efficiently examining researchers' understanding of learners' SRL, and how the researchers' biases related to their beliefs about SRL and the effectiveness of their instructional decision-making. Overall, the studies reviewed here illustrate compounding of value by coordinating think-aloud protocols, eye-gaze data, and online behavioral traces to capture implicit processes such as analyzing and reasoning. Therefore, the SPARC system should be engineered to capture and mine patterns within researchers' concurrent verbalizations, eye movements, and clickstream data to mark with what, when, how, and how long researchers reason and analyze learners' multimodal data as they forge inferences about learners' SRL. However, data streams sampling researchers' activation of monitoring processes and marking instances of analyzing and reasoning merely set a stage for inquiring whether researchers

understand how learners' multimodal data represent SRL. Simply tracking researchers' metacognition is insufficient to guide instructional decision-making that optimizes scaffolding learners' SRL. Researchers' understanding is also necessary.

2.3. Understanding real-time multimodal data

People acquire conceptual knowledge by coordinating schema and semantic networks to encode conceptual and propositional knowledge (Anderson, 2000). Therefore, researchers' understanding of learners' SRL represented by multimodal data depends on access to valid schemas and slots within them, and a well-formed structure of networked information about learners' SRL. For example, (Mudrick et al., 2019) results indicate a learner's eye fixations oscillating between text and a diagram (i.e., saccades) should fill a slot in a schema for metacognitive monitoring within a schema describing motivation to build comprehension by, for this slot, resolving confusion. SPARC should detect whether an instructional decision-maker activates and instantiates schemas like this. Then, merging that information and other data about learners and researchers into negative feedback within a pipeline architecture, the system can iteratively scaffold researchers toward successively improved decisions about interventions that optimize learners' performance and self-regulation.

2.3.1. Data channels capturing understanding

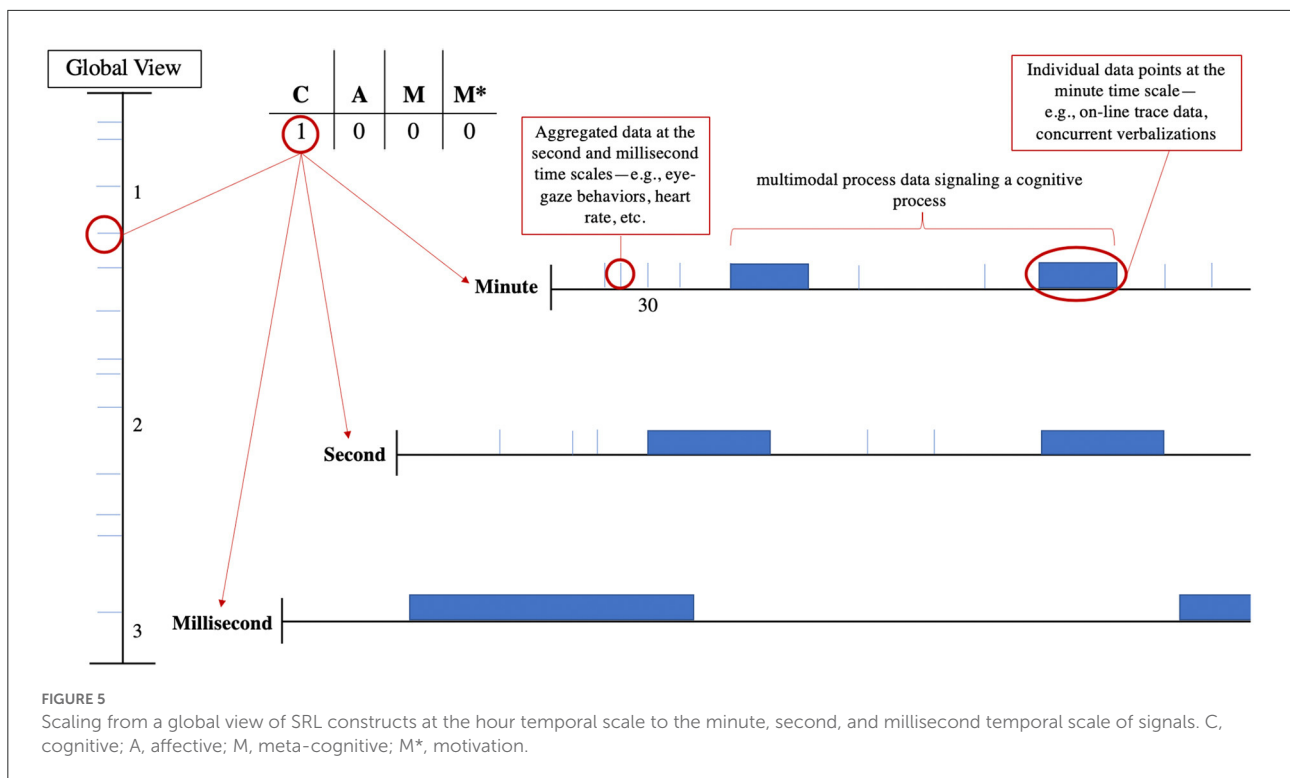
Traditionally, comprehension has been assessed using aggregated total gain scores drawn from selected-response, paper, and pencil tests before and after domain-specific instruction (Makransky et al., 2019). However, process-oriented and performance-based methods using multimodal data offer promising alternatives (James et al., 2016). Liu et al. (2019) sampled multiple data streams over a learning session using video and audio recordings, physiological sensors, eye tracking, and online behavioral traces. Their model formed from multimodal process-based data more strongly predicted learners' understanding than models based on a single modality of data such as online behavioral traces (Liu et al., 2019). Similarly, Makransky et al. (2019) amalgamated multimodal data across online behavioral traces, eye tracking, electrophysiological signals, and heart rate to build models predicting variance in learners' understanding of information taught during a learning session. A unimodal model using just online behavioral traces explained 57% of the variance ($p < 0.05$) in learners' understanding. A model incorporating multiple data streams explained 75% of the variance ($p < 0.05$) in learners' understanding (Makransky et al., 2019). Thus, in order for SPARC to capture researchers' understanding of learners' SRL using multimodal data, the system would need

to sample various data channels to learn data indices that indicate if, when, and how the researcher is understanding the learners' SRL. Depending on the researcher's multimodal data and their accuracy in understanding a learner's SRL, over time SPARC would learn each researcher's understanding of learners' data so that the system may make accurate inferences about when to scaffold researchers to optimize their understanding of learners' multimodal data. We suggest that in order to capture researchers' understanding of information, SPARC could be built to automatically capture understanding based on using eye-gaze behaviors, concurrent verbalizations, mouse-clicks (e.g., what is the researcher and/or learner attending to, and is the action related to the objective of the session—for instance, is the researcher attending to data channels or modalities signaling an SRL process that needs scaffolding, such that of a learner attending to text and diagrams that are irrelevant to the learning objective in which they are studying. Does the researcher monitor these data channels to guide their instructional decision-making, and if so, how does the instructional decision impact the learners' subsequent SRL during the session? Using SPARC to sample and model the multimodal data generated by both researchers and learners could help answer these questions and advance research on the science of learning.

3. Results

3.1. Theoretically- and empirically-based system guidelines for SPARC

SPARC should be engineered to offload tasks that overload researchers' attention and working memory to key analytics describing learners' SRL and integration of those analytics in forming productive instructional decisions. We describe here how SPARC can address these critical needs. Some of SPARC's functionality can be adapted from existing multimodal analytical tools such as SSI (Wagner et al., 2013; Noroozi et al., 2019). SSI's machine-learning pipeline automatically aligns, processes, and filters multimodal data as they are generated in real-time. SPARC will incorporate this functionality and augment it according to theoretical frameworks and empirical findings mined from learning science. Specifically, SPARC's pipeline will be calibrated to weight data channels (e.g., eye gaze, concurrent verbalizations), modalities (e.g., fixation vs. saccades in eye gaze, frequency vs. duration of fixations, etc.), and combinations of data to reflect meaningful and critical learning and SRL processes. For example, facial expressions and physiological sensor data would be assigned greater weight in modeling affect and affective state change, such as frustration, while screen recordings, concurrent verbalizations, and eye-movement data would be assigned greater weights to model cognitive strategy use).



Moreover, the SPARC system will use automated-recognition routines to detect and classify learners’ and researchers’ SRL activities separately while analyzing data from the activities concurrently to guide the scaffolding of the researcher and assessment of how the instructional prescriptions of the researcher are impacting the learners’ SRL and performance. For example, when capturing conditions marked by the COPES model (Winne, 2018a), eye-gaze, and think-aloud data may best indicate conditions learners perceive about a learning task and the learning environment. SPARC’s algorithms would assign these data greater weight compared to clickstream and physiological sensor data to represent conditions from the learner’s perspective. It is important to note algorithms should reflect a full scan of conditions regarding signals about conditions present and conditions absent (Winne, 2019). Theory plays a key role here because it is the source for considering potential roles for a construct that has zero value in the data vector. Temporal dimensions (see Figure 5) are a critical feature in SPARC’s approach to modeling a learner’s multimodal data considering that different learning processes may unfold across varying time scales. Data within and across channels collected over time helps to ensure adequate sampling (e.g., how long does an affective state last?) and multiple contextual cues (e.g., what did the learner do before and after onset of an affective state?). This wider context enhances interpretability beyond single-channel, single time-point data. For example, a 250 Hz eye tracker supplying 250 data points per second may be insufficient to infer learner processing in the one-second

sample. Other data, e.g., sequence of previewing headings, reading, and re-reading indicating multiple metacognitive judgments augmented by screen recording and concurrent verbalizations across several minutes provide a more complete structure for a researcher to draw inferences about the learner’s engagement in a task (Mudrick et al., 2019; Taub and Azevedo, 2019).

As such, SPARC features will allow a researcher to scale up—i.e., scale upsampling rates to a uniform temporal scale such as from milliseconds to seconds, or seconds to minutes, or down—i.e., scale downsampling rates to a uniform temporal scale such as from seconds to milliseconds, to pinpoint how, when, why, and what learning processes were occurring (Figure 5). When researchers scale up or down, it also captured critical information revealing how the researcher is selecting, monitoring, analyzing, and understanding learners’ multimodal data representing operations in the COPES model and modulations of operations that represent SRL. The opportunity for the researcher to explore the learner’s temporal learning progression is a critical feature that researchers need to guide their instructional decision making related to adaptive scaffolding and feedback to the learners (Kinnebrew et al., 2014, 2015; Basu et al., 2017). For SPARC to continuously capture data and update its models of learners and researchers, it should apply predictive models to track the learners’ trajectories and project future learning events prompted by researcher intervention. For instance, if the researcher gave learners feedback and redirected a learner to another section content more relevant

to learning objectives, SPARC should forecast the probability of learner uptake of that recommendation and patterns of multimodal data that confirm uptake. Iterating across learning sessions, this allows SPARC to dynamically converge models to more accurately predict both behavior by the learner and the researcher.

4. Discussion

Emerging research on SRL sets the stage for using temporally sequenced multimodal data to examine the dynamics of multiple processes and interventions to adapt those processes in emerging technologies. Using large volumes of multimodal data to analyze and interpret learners' SRL processes in near real-time is theoretically and algorithmically challenging (Cloude et al., 2020; Emerson et al., 2020). We crafted a theoretical, conceptual, and empirically grounded framework for designing a system that guides researchers and instructors in analyzing and understanding the complex nature of SRL. A novel aspect of the SPARC system is modeling all players in instruction—learner and instructional decision maker—to dynamically upgrade capabilities to enhance learning, SRL, and the empirical foundations for understanding those processes. Further, including insights gained from data collected on researchers and SRL experts could potentially contribute to enhancing our understanding of how to automatically build detectors of SRL processes on both instructors and learners. Emerging research using multimodal data shows promise in approaching this goal, but this research stream has not yet tackled major challenges facing interdisciplinary and international researchers and instructors in monitoring, analyzing, and understanding learners' SRL multimodal data based on what, where, when, how, and with what learners self-regulate to understand content. In particular, the SPARC system we outline defines and sets a framework for addressing a new and fundamental question. How do researchers and instructors monitor, analyze, and understand learners' and groups of learners' multimodal data; and, how can data about those processes be merged with data about learners to bootstrap the full system involving learners, instructional decision makers, and interventions? The SPARC system we suggest takes the first steps toward addressing major conceptual, theoretical, methodological, and analytical issues associated with using real-time multimodal data (Winne, 2022).

4.1. Implications and future directions

Implications of this research are threefold. First, leveraging insights gained from researchers' and SRL experts' multimodal data based on their understanding of both (1) instructors' and (2) learners' multimodal data could be used to build

valid algorithms for SRL detection. For example, the SRL expert could potentially tag whether the instructor identified the learners' misuse of SRL while viewing materials? If the instructor did identify this behavior, did they intervene accordingly based on their own SRL and understanding of the learner's processes to make an informed instructional decision? Utilizing the information that the SRL expert or researcher referenced could be used to build SRL detectors. Training AI on how researchers, instructors, and learners monitor, infer, and understand information across multiple data sources has the potential to build valid algorithms that are empirically and theoretically based derived. Building valid AI is a current challenge for the field, where most AI is built by experts that have little knowledge about SRL theory. Instead, AI algorithms are data-driven such that the steps are built to maximize the detection of significant findings with the highest accuracy. This approach slows progress on deriving meaningful insights from relationships present in multiple data sources. Through utilizing SPARC, it would ensure that the best algorithms/data channels/modalities/dependent variables are selected based on a combination of the researchers', instructors', and learners' information as a whole. Furthermore, this would also spark researchers, instructors, and learners to think critically about what the algorithm should be doing to facilitate understanding of SRL for supporting informed instructional decision-making. This research may highlight areas for teaching training, such as integrating data science and visualizations courses in the curriculum since data are being increasingly used in the classroom to enhance the quality of education. Was this monitoring or behavior? If not, why did the algorithm fire to suggest it was so? It could provide a world of information about where the researcher is doing quality control on the algorithm to assess if they are working properly in all contexts. This could generate a library of open-source algorithms/production rules for a range of contexts, domains, users, countries, theories, and many others. Another important area is leveraging SPARC to reveal user biases. For example, is an instructor focusing on specific data sources or all data sources? Is the instructor supporting all learners in the same way? SPARC would allow us to compare and contrast where users could be biased toward certain data channels relative to others, and potential shed light on these behaviors to mitigate bias and draw awareness to our perspectives when we are not using SPARC, thus potentially enhancing our objectivity as scientists and instructors.

One area of future research that could advance this work is moving away from solely relying on a linear paradigm to define SRL such as linear regression. It is imperative that we utilize sophisticated statistical techniques to model the complexity and dynamics that emerge within multimodal data across varying system levels such as multimodal data collected from the researcher or instructor in their understanding of analytics presented back to the user. An interdisciplinary approach toward data processing and analysis may provide the analytical

tools needed to exploit meaningful relationships and insights within the data. Specifically, we need to go beyond information-processing theory which assumes that self-regulated learning results from linear sequences of learning processes as assumed in linear models. We challenge research to make a paradigmatic shift toward dynamic systems thinking (Van Gelder and Port, 1995) to investigate researchers', instructors', and learners' SRL processes as self-organizing, dynamically emergent, and non-linear phenomena. Leaning on nonlinear dynamical analyses to study SRL is starting to gain momentum (Dever et al., 2022; Li et al., 2022). This interdisciplinary approach would allow us to study SRL across multiple levels and nonlinear dynamical analyses offer more flexibility in utilizing multiple data sources that do not need to adhere to rigid assumptions of normality, equal variation, and independence of observation. Finally, SPARC offers implications for building AI-enabled adaptive learning systems that repurpose information back to instructors, learners, and research to augment both teaching and learning (Kabudi et al., 2021). As outlined by Kabudi et al. (2021), AI-enabled adaptive learning systems could detect and select the appropriate learning intervention using evidence from SRL experts, researchers, instructors, and learners. data collected during learning activities should not only be predictive analytics but rather leverages data in various ways depending on the (a) user and (b) objective of the session. Specifically, analytics fed back to users such as instructors should include both prescriptive and descriptive analytics. (go into prescriptive, descriptive, and predictive; Kabudi et al., 2021) as these all hold different implications for teaching and learning.

5. Conclusions

Much work remains to realize SPARC. New research is needed to widen and deepen understandings of (1) how to map researchers' and learners' multimodal data onto COPES constructs, (2) how differences in these mappings suggest interventions and the degree to which this may vary by country, and (3) how a pipeline architecture should be designed to iterate over these results to optimize both learner's SRL, instructor's SRL, and instructional decision making. Our next steps in building and implementing a prototype system like SPARC are to begin collecting real-time data about how researchers examine and use learners' multimodal data in specific learning and problem-solving scenarios, e.g., learning with serious games, intelligent tutoring systems, and virtual reality. This requires recruiting a number of leading experts within the field of self-regulated learning across a range of emerging technologies, but also a range of SRL theories including socially-shared self-regulated learning and co-regulation. Each scenario presented to participants (i.e., researchers/experts) will encompass gathering their multimodal data while they review learners' multimodal data and provide annotations that classify various SRL processes and strategies. This study will allow further understanding

of how researchers monitor, analyze, and make inferences about SRL using multimodal data. Results will guide system architecture and design for a theoretically- and empirically-based system that supports researchers and instructors in monitoring, analyzing, and understanding learners' multimodal data to make effective instructional decisions and foster self-regulation (Hwang et al., 2020). Through training AI using the multiple data channels collected from leading researchers and experts within the field of self-regulated learning, in conjunction with instructors and learners, it opens opportunities to build valid and reliable AI that goes beyond data-driven techniques to determine when theoretically-relevant constructs emerge across a range of data channels and modalities.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Ethics statement

Ethical review and approval was not required for the current study in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

EC conceptualized, developed, synthesized, and constructed the manuscript. RA assisted in conceptualizing and developing the SPARC vision. PW provided extensive edits and revisions to the manuscript and assisted in conceptualizing the SPARC vision. GB provided edits and revisions to the manuscript. EJ provided edits and revisions to the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was partially supported by a grant from the National Science Foundation (DRL#1916417) awarded to RV.

Acknowledgments

The authors would like to thank team members of the SMART Lab at the University

of Central Florida for their assistance and contributions.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Anderson, B. A. (2016). The attention habit: How reward learning shapes attentional selection. *Ann. N. Y. Acad. Sci.* 1369, 24–39. doi: 10.1111/nyas.12957
- Anderson, B. A., and Yantis, S. (2013). Persistence of value-driven attentional capture. *J. Exp. Psychol.* 39, 6. doi: 10.1037/a0030860
- Anderson, J. R. (2000). *Cognitive Psychology and Its Implications*. New York, NY: Worth Publishers.
- Azevedo, R., Bouchet, F., Duffy, M., Harley, J., Taub, M., Trevors, G., et al. (2022). Lessons learned and future directions of metatutor: leveraging multichannel data to scaffold self-regulated learning with an intelligent tutoring system. *Front. Psychol.* 13, 813632. doi: 10.3389/fpsyg.2022.813632
- Azevedo, R., and Gašević, D. (2019). Analyzing multimodal multichannel data about self-regulated learning with advanced learning technologies: issues and challenges. *Comput. Hum. Behav.* 96, 207–210. doi: 10.1016/j.chb.2019.03.025
- Azevedo, R., Mudrick, N. V., Taub, M., and Bradbury, A. E. (2019). “Self-regulation in computer-assisted learning systems,” in *The Cambridge Handbook of Cognition and Education*, eds J. Dunlosky and K. A. Rawson (Cambridge, UK: Cambridge University Press), 587–618. doi: 10.1017/9781108235631.024
- Azevedo, R., Taub, M., and Mudrick, N. V. (2018). “Understanding and reasoning about real-time cognitive, affective, and metacognitive processes to foster self-regulation with advanced learning technologies,” in *Handbook on Self-Regulation of Learning and Performance*, eds D. H. Schunk and J. A. Greene (New York, NY: Routledge). doi: 10.4324/9781315697048-17
- Basu, S., Biswas, G., and Kinnebrew, J. S. (2017). Learner modeling for adaptive scaffolding in a computational thinking-based science learning environment. *User Model. User Adapt. Interact.* 27, 5–53. doi: 10.1007/s11257-017-9187-0
- 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*, eds D. H. Schunk and J. A. Greene (New York, NY: Routledge). doi: 10.4324/9781315697048-24
- Biswas, G., Segedy, J. R., and Bunchongchit, K. (2016). From design to implementation to practice a learning by teaching system: Betty’s brain. *Int. J. Artif. Intell. Educ.* 26, 350–364. doi: 10.1007/s40593-015-0057-9
- Blanchette, I., and Richards, A. (2010). The influence of affect on higher level cognition: a review of research on interpretation, judgement, decision making and reasoning. *Cogn. Emot.* 24, 561–595. doi: 10.1080/02699930903132496
- Blaschek, T., Beck, F., Baltes, S., Ertl, T., and Weiskopf, D. (2016). “Visual analysis and coding of data-rich user behavior,” in *Proceedings of the 2016 IEEE Conference on Visual Analytics Science and Technology, VAST ’16* (New York, NY), 141–150. doi: 10.1109/VAST.2016.7883520
- Catrysse, L., Gijbels, D., Donche, V., De Maeyer, S., Lesterhuis, M., and Van den Bossche, P. (2018). How are learning strategies reflected in the eyes? Combining results from self-reports and eye-tracking. *Br. J. Educ. Psychol.* 88, 118–137. doi: 10.1111/bjep.12181
- Chango, W., Cerezo, R., Sanchez-Santillan, M., Azevedo, R., and Romero, C. (2021). Improving prediction of students’ performance in intelligent tutoring systems using attribute selection and ensembles of different multimodal data sources. *J. Comput. High. Educ.* 33, 614–634. doi: 10.1007/s12528-021-09298-8
- Claypoole, V. L., Dever, D. A., Denues, K. L., and Szalma, J. L. (2019). The effects of event rate on a cognitive vigilance task. *Hum. Factors* 61, 440–450. doi: 10.1177/0018720818790840
- Cloude, E. B., Dever, D. A., Wiedbusch, M. D., and Azevedo, R. (2020). Quantifying scientific thinking using multichannel data with crystal island: implications for individualized game-learning analytics. *Front. Educ.* 5, 217. doi: 10.3389/feduc.2020.572546
- Cloude, E. B., Wortha, F., Dever, D. A., and Azevedo, R. (2021a). “Negative emotional dynamics shape cognition and performance with metatutor: toward building affect-aware systems,” in *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)* (New York, NY), 1–8. doi: 10.1109/ACII52823.2021.9597462
- Cloude, E. B., Wortha, F., Wiedbusch, M. D., and Azevedo, R. (2021b). “Goals matter: changes in metacognitive judgments and their relation to motivation and learning with an intelligent tutoring system,” in *International Conference on Human-Computer Interaction* (Cham: Springer), 224–238. doi: 10.1007/978-3-030-77889-7_15
- Desimone, R. (1996). “Neural mechanisms for visual memory and their role in attention,” in *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 93 of *PANAS ’96* (Washington, DC: National Academy of Sciences), 13494–13499. doi: 10.1073/pnas.93.24.13494
- Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222. doi: 10.1146/annurev.ne.18.030195.001205
- Dever, D. A., Amon, M. J., Vrzáková, H., Wiedbusch, M. D., Cloude, E. B., and Azevedo, R. (2022). Capturing sequences of learners’ self-regulatory interactions with instructional material during game-based learning using auto-recurrence quantification analysis. *Front. Psychol.* 13, 813677. doi: 10.3389/fpsyg.2022.813677
- Dindar, M., Järvelä, S., Nguyen, A., Haataja, E., and Çini, A. (2022). Detecting shared physiological arousal events in collaborative problem solving. *Contemp. Educ. Psychol.* 69, 102050. doi: 10.1016/j.cedpsych.2022.102050
- Duncan, J., and Nimmo-Smith, I. (1996). Objects and attributes in divided attention: Surface and boundary systems. *Percept. Psychophys.* 58, 1076–1084.
- Emerson, A., Cloude, E. B., Azevedo, R., and Lester, J. (2020). Multimodal learning analytics for game-based learning. *Br. J. Educ. Technol.* 51, 1505–1526. doi: 10.1111/bjet.12992
- Franco-Watkins, A. M., Mattson, R. E., and Jackson, M. D. (2016). Now or later? Attentional processing and intertemporal choice. *J. Behav. Decis. Making* 29, 206–217. doi: 10.1002/bdm.1895
- Greene, J. A., and Azevedo, R. (2010). The measurement of learners’ self-regulated cognitive and metacognitive processes while using computer-based learning environments. *Educ. Psychol.* 45, 203–209. doi: 10.1080/00461520.2010.515935
- Greene, J. A., Deekens, V. M., Copeland, D. Z., and Yu, S. (2018). “Capturing and modeling self-regulated learning using think-aloud protocols,” in *Handbook of Self-Regulation of Learning and Performance*, eds D. H. Schunk and J. A. Greene (New York, NY: Routledge). doi: 10.4324/9781315697048-21
- Greenlee, E. T., DeLucia, P. R., and Newton, D. C. (2019). Driver vigilance in automated vehicles: effects of demands on hazard detection performance. *Hum. Factors* 61, 474–487. doi: 10.1177/0018720818802095
- Hadwin, A. F. (2021). Commentary and future directions: What can multimodal data reveal about temporal and adaptive processes in self-regulated learning? *Learn. Instruct.* 72, 101287. doi: 10.1016/j.learninstruc.2019.101287
- Hancock, P. A. (2013). In search of vigilance: the problem of iatrogenically created psychological phenomena. *Am. Psychol.* 68, 97. doi: 10.1037/a0030214
- Hawkins, W. J., Heffernan, N. T., and Baker, R. S. (2014). “Learning Bayesian knowledge tracing parameters with a knowledge heuristic and empirical probabilities,” in *International Conference on Intelligent Tutoring Systems* (Cham: Springer), 150–155. doi: 10.1007/978-3-319-07221-0_18
- Hwang, G.-J., Xie, H., Wah, B. W., and Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Comput. Educ.* 1, 100001. doi: 10.1016/j.caeai.2020.100001

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Ifenthaler, D., and Schumacher, C. (2019). "Releasing personal information within learning analytics systems" in *Learning Technologies for Transforming Large-Scale Teaching, Learning, and Assessment*, eds D. Sampson, J. Spector, D. Ifenthaler, P. Isaias, and S. Sergis (Cham: Springer), 3–18. doi: 10.1007/978-3-030-15130-0_1
- Ifenthaler, D., and Tracey, M. W. (2016). Exploring the relationship of ethics and privacy in learning analytics and design: implications for the field of educational technology. *Educ. Technol. Res. Dev.* 64, 877–880. doi: 10.1007/s11423-016-9480-3
- James, P., Antonova, L., Martel, M., and Barkun, A. N. (2016). Measures of trainee performance in advanced endoscopy: a systematic review. *342. Am. J. Gastroenterol.* 111, S159. doi: 10.14309/00000434-201610001-00342
- Jang, E. E., Lajoie, S. P., Wagner, M., Xu, Z., Poitras, E., and Naismith, L. (2017). Person-oriented approaches to profiling learners in technology-rich learning environments for ecological learner modeling. *J. Educ. Comput. Res.* 55, 552–597. doi: 10.1177/0735633116678995
- Johnson-Glenberg, M. C. (2018). Immersive VR and education: embodied design principles that include gesture and hand controls. *Front. Robot. AI* 5, 81. doi: 10.3389/frobot.2018.00081
- Kabudi, T., Pappas, I., and Olsen, D. H. (2021). Ai-enabled adaptive learning systems: a systematic mapping of the literature. *Comput. Educ.* 2, 100017. doi: 10.1016/j.caeai.2021.100017
- Kassab, S. E., and Hussain, S. (2010). Concept mapping assessment in a problem-based medical curriculum. *Med. Teach.* 32, 926–931. doi: 10.3109/0142159X.2010.497824
- Kinnebrew, J. S., Mack, D. L., Biswas, G., and Chang, C.-K. (2014). "A differential approach for identifying important student learning behavior patterns with evolving usage over time," in *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD '14*, eds W. C. Peng, et al. (Cham: Springer), 281–292. doi: 10.1007/978-3-319-13186-3_27
- Kinnebrew, J. S., Segedy, J. R., and Biswas, G. (2015). Integrating model-driven and data-driven techniques for analyzing learning behaviors in open-ended learning environments. *IEEE Trans. Learn. Technol.* 10, 140–153. doi: 10.1109/TLT.2015.2513387
- Laird, J. E. (2012). *The Soar Cognitive Architecture*. Cambridge, MA: The MIT Press. doi: 10.7551/mitpress/7688.001.0001
- Lajoie, S. P., Zheng, J., Li, S., Jarrell, A., and Gube, M. (2021). Examining the interplay of affect and self-regulation in the context of clinical reasoning. *Learn. Instruct.* 72, 101219. doi: 10.1016/j.learninstruct.2019.101219
- Lane, H. C., and D'Mello, S. K. (2019). "Uses of physiological monitoring in intelligent learning environments: a review of research, evidence, and technologies," in *Mind, Brain and Technology: Learning in the Age of Emerging Technologies*, eds T. D. Parsons, L. Lin, and D. Cockerham (Cham: Springer International Publishing), 67–86. doi: 10.1007/978-3-030-02631-8_5
- Li, S., Zheng, J., Huang, X., and Xie, C. (2022). Self-regulated learning as a complex dynamical system: examining students' stem learning in a simulation environment. *Learn. Individ. Differ.* 95, 102144. doi: 10.1016/j.lindif.2022.102144
- Liu, S., Ji, H., and Wang, M. C. (2019). Nonpooling convolutional neural network forecasting for seasonal time series with trends. *IEEE Trans. Neural Netw. Learn. Syst.* 95. doi: 10.1109/TNNLS.2019.2934110
- Makransky, G., Terkildsen, T. S., and Mayer, R. E. (2019). Role of subjective and objective measures of cognitive processing during learning in explaining the spatial contiguity effect. *Learn. Instruct.* 61, 23–34. doi: 10.1016/j.learninstruct.2018.12.001
- Malmberg, J., Järvelä, S., and Järvenoja, H. (2017). Capturing temporal and sequential patterns of self-, co-, and socially shared regulation in the context of collaborative learning. *Contemp. Educ. Psychol.* 49, 160–174. doi: 10.1016/j.cedpsych.2017.01.009
- Mayer, R. E. (2019). Thirty years of research on online learning. *Appl. Cogn. Psychol.* 33, 152–159. doi: 10.1002/acp.3482
- Messi, D. and Adrario, E. (2021). "Mixed reality simulation for medical training: how it affects learners; cognitive state," in *Advances in Simulation and Digital Human Modeling: Proceedings of the AHFE 2021 Virtual Conferences on Human Factors and Simulation, and Digital Human Modeling and Applied Optimization, July 25-29, 2021, USA, Vol. 264* (Cham: Springer Nature), 339. doi: 10.1007/978-3-030-79763-8_41
- Molenaar, I. and Järvelä, S. (2014). Sequential and temporal characteristics of self and socially regulated learning. *Metacogn. Learn.* 9, 75–85. doi: 10.1007/s11409-014-9114-2
- Mudrick, N. V., Azevedo, R., and Taub, M. (2019). Integrating metacognitive judgments and eye movements using sequential pattern mining to understand processes underlying multimedia learning. *Comput. Hum. Behav.* 96, 223–234. doi: 10.1016/j.chb.2018.06.028
- Muldner, K., Burleson, W., and VanLehn, K. (2010). "yes!: using tutor and sensor data to predict moments of delight during instructional activities," in *Proceedings of the International Conference on User Modeling, Adaptation, and Personalization, UMAP '10*, eds P. De Bra, A. Kobsa, and D. Chin (New York, NY: Springer), 159–170. doi: 10.1007/978-3-642-13470-8_16
- Nguyen, P. H., Xu, K., Wheat, A., Wong, B. W., Attfield, S., and Fields, B. (2015). Sensepath: understanding the sensemaking process through analytic provenance. *IEEE Trans. Visual. Comput. Graph.* 22, 41–50. doi: 10.1109/TVCG.2015.2467611
- Noroozi, O., Alikhani, I., Järvelä, S., Kirschner, P. A., Juuso, I., and Seppänen, T. (2019). Multimodal data to design visual learning analytics for understanding regulation of learning. *Comput. Hum. Behav.* 100, 298–304. doi: 10.1016/j.chb.2018.12.019
- Raca, M., and Dillenbourg, P. (2014). "Holistic analysis of the classroom," in *Proceedings of the 2014 ACM Workshop on Multimodal Learning Analytics Workshop and Grand Challenge, MLA '14* (New York, NY: ACM), 13–20. doi: 10.1145/2666633.2666636
- Rajendran, R., Kumar, A., Carter, K. E., Levin, D. T., and Biswas, G. (2018). "Predicting learning by analyzing eye-gaze data of reading behavior," in *Proceedings of the 11th International Conference on Educational Data Mining* (New York, NY).
- Reimann, P. (2021). Methodological progress in the study of self-regulated learning enables theory advancement. *Learn. Instruct.* 72, 101269. doi: 10.1016/j.learninstruct.2019.101269
- Renninger, K. A., and Hidi, S. E. (2019). "Interest development and learning," in *The Cambridge Handbook of Motivation and Learning*, eds K. A. Renninger and S. E. Hidi (Cambridge, UK: Cambridge University Press). doi: 10.1017/9781316823279.013
- Salomon, G., and Perkins, D. N. (1989). Rocky roads to transfer: rethinking mechanism of a neglected phenomenon. *Educ. Psychol.* 24, 113–142. doi: 10.1207/s15326985ep2402_1
- Scheiter, K., and Eitel, A. (2017). "The use of eye tracking as a research and instructional tool in multimedia learning," in *Eye-Tracking Technology Applications in Educational Research* (Hershey, PA: IGI Global), 143–164. doi: 10.4018/978-1-5225-1005-5.ch008
- Shunk, D. H., and Greene, J. A. (2018). *Handbook of Self-Regulation of Learning and Performance*. New York, NY: Routledge. doi: 10.4324/9781315697048
- Segedy, J. R., Kinnebrew, J. S., and Biswas, G. (2015). Using coherence analysis to characterize self-regulated learning behaviours in open-ended learning environments. *J. Learn. Analyt.* 2, 13–48. doi: 10.18608/jla.2015.21.3
- Si, J., Kong, H.-H., and Lee, S.-H. (2019). Developing clinical reasoning skills through argumentation with the concept map method in medical problem-based learning. *Interdiscip. J. Probl. Based Learn.* 13, 1–16. doi: 10.7771/1541-5015.1776
- Spires, H. A., Rowe, J. P., Mott, B. W., and Lester, J. C. (2011). Problem solving and game-based learning: effects of middle grade students' hypothesis testing strategies on learning outcomes. *J. Educ. Comput. Res.* 44, 453–472. doi: 10.2190/EC.44.4.e
- Taub, M., and Azevedo, R. (2019). How does prior knowledge influence eye fixations and sequences of cognitive and metacognitive SRL processes during learning with an intelligent tutoring system? *Int. J. Artif. Intell. Educ.* 29, 1–28. doi: 10.1007/s40593-018-0165-4
- Taub, M., Azevedo, R., Rajendran, R., Cloude, E. B., Biswas, G., and Price, M. J. (2021). How are students' emotions related to the accuracy of cognitive and metacognitive processes during learning with an intelligent tutoring system?. *Learn. Instruct.* 72, 101200. doi: 10.1016/j.learninstruct.2019.04.001
- Taub, M., Mudrick, N. V., Azevedo, R., Millar, G. C., Rowe, J., and Lester, J. (2016). "Using multi-level modeling with eye-tracking data to predict metacognitive monitoring and self-regulated learning with crystal island," in *Proceedings of the International Conference on Intelligent Tutoring Systems, ITS '16* (New York, NY: Springer), 240–246. doi: 10.1007/978-3-319-39583-8_24
- Taub, M., Mudrick, N. V., Azevedo, R., Millar, G. C., Rowe, J., and Lester, J. (2017). Using multi-channel data with multi-level modeling to assess in-game performance during gameplay with crystal island. *Comput. Hum. Behav.* 76, 641–655. doi: 10.1016/j.chb.2017.01.038
- UNESCO (2017). *More Than One-Half of Children and Adolescents Are Not Learning Worldwide. UIS Fact Sheet No. 46*. UNESCO.
- Van Gelder, T., and Port, R. F. (eds.). (1995). "It's about time: an overview of the dynamical approach to cognition," in *Mind as Motion: Explorations in the Dynamics of Cognition, Vol. 1* (Cambridge, MA: The MIT Press), 1–44.
- van Zoest, W., Van der Stigchel, S., and Donk, M. (2017). Conditional control in visual selection. *Attent. Percept. Psychophys.* 79, 1555–1572. doi: 10.3758/s13414-017-1352-3

- Wagner, J., Lingenfelter, F., Baur, T., Damian, I., Kistler, F., and André, E. (2013). "The social signal interpretation (ssi) framework: multimodal signal processing and recognition in real-time," in *Proceedings of the 21st ACM international conference on Multimedia*, 831–834. doi: 10.1145/2502081.2502223
- Wiedbusch, M., Dever, D., Wortha, F., Cloude, E. B., and Azevedo, R. (2021). "Revealing data feature differences between system-and learner-initiated self-regulated learning processes within hypermedia," in *International Conference on Human-Computer Interaction* (Cham: Springer), 481–495. doi: 10.1007/978-3-030-77857-6_34
- Winne, P. (2018). Paradigmatic issues in state-of-the-art research using process data. *Frontl. Learn. Res.* 6, 250–258. doi: 10.14786/flr.v6i3.551
- Winne, P. H. (2018a). "Cognition and metacognition within self-regulated learning," in *Handbook of Self-Regulation of Learning and Performance*, eds D. H. Schunk and J. A. Greene (New York, NY: Routledge), 52–64. doi: 10.4324/9781315697048-3
- Winne, P. H. (2018b). Theorizing and researching levels of processing in self-regulated learning. *Br. J. Educ. Psychol.* 88, 9–20. doi: 10.1111/bjep.12173
- Winne, P. H. (2019). Paradigmatic dimensions of instrumentation and analytic methods in research on self-regulated learning. *Comput. Hum. Behav.* 96, 285–289. doi: 10.1016/j.chb.2019.03.026
- Winne, P. H. (2020). Construct and consequential validity for learning analytics based on trace data. *Comput. Hum. Behav.* 112. doi: 10.1016/j.chb.2020.106457
- Winne, P. H. (2022). Modeling self-regulated learning as learners doing learning science: how trace data and learning analytics help develop skills for self-regulated learning. *Metacogn. Learn.* 1–19. doi: 10.1007/s11409-022-09305-y
- Winne, P. H., and Hadwin, A. F. (1998). "Studying as self-regulated learning," in *Metacognition in Educational Theory and Practice*, eds D. J. Hacker, J. Dunlosky, and A. C. Graesser (Mahwah, NJ: Lawrence Erlbaum Associates), 277–304.
- Winne, P. H., Teng, K., Chang, D., Lin, M. P.-C., Marzouk, Z., Nesbit, J. C., et al. (2019). nStudy: software for learning analytics about processes for self-regulated learning. *J. Learn. Anal.* 6, 95–106. doi: 10.18608/jla.2019.6.2.7