



REVIEW ARTICLE



<https://doi.org/10.1057/s41599-024-02987-6>

OPEN

A systematic review of Stimulated Recall (SR) in educational research from 2012 to 2022

Xuesong Zhai^{1,2,7}, Xiaoyan Chu^{1,7}, Minjuan Wang^{3,4}, Chin-Chung Tsai⁵, Jyh-Chong Liang⁵ & Jonathan Michael Spector⁶

Stimulated Recall (SR) has long been used in educational settings as an approach of retrospection. However, with the fast growing of digital learning and advanced technologies in educational settings over the past decade, the extent to which stimulated recall has been effectively implemented by researchers remains minimal. This systematic review reveals that SR has been primarily employed to probe the patterns of participants' thinking, to examine the effects of instructional strategies, and to promote metacognitive level. Notably, SR video stimuli have advanced, and the sources of stimuli have become more diverse, including the incorporation of physiological data. Additionally, researchers have applied various strategies, such as flexible intervals and questioning techniques, in SR interviews. Furthermore, this article discusses the relationships between different SR research items, including stimuli and learning contexts. The review and analysis also demonstrate that stimulated recall may be further enhanced by integrating multiple data sources, applying intelligent algorithms, and incorporating conversational agents enabled by generative artificial intelligence such as ChatGPT. This article provides a comprehensive analysis of SR studies in the realm of education and proposes a promising avenue for researchers to proactively apply stimulated recall in investigating educational issues in the digital era.

Introduction

Stimulated Recall (SR) is an approach commonly used to prompt participants' retrospection by employing diverse stimuli and interview strategies. This method is frequently applied to examine instructors' and students' reflections on their cognitive and affective responses during or after specific educational events or activities (Calderhead, 1981; Gass and Mackey, 2016). This type of SR represents an effective qualitative method for educational researchers to gather implicit data and has been broadly practiced to investigate various teaching and learning occurrences, including teacher cognition, study strategies, and language learning (Meade and McMeniman, 1992; Van der Kleij et al., 2017; White et al., 2016; Sundberg et al., 2018; Martinelle, 2018; Cao et al., 2019; Martinelle, 2020). Moreover, in addition to serving as a

¹ College of Education, Zhejiang University, Hangzhou, China. ² Hangzhou International Urbanology Research Center & Zhejiang Urban Governance Studies Center, Hangzhou, China. ³ Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong SAR, China.

⁴ Learning Design and Technology, San Diego State University, San Diego, CA, USA. ⁵ Program of Learning Sciences and Institute for Research Excellence in Learning Sciences, National Taiwan Normal University, Taipei, Taiwan. ⁶ Department of Learning Technologies, University of North Texas, Denton, TX, USA.

⁷ These authors contributed equally: Xuesong Zhai, Xiaoyan Chu. ✉email: xiaoyanchu@zju.edu.cn; tit-eic@ieee.org

research tool to explore instructors’ and learners’ internal thoughts, several studies have innovatively implemented SR as a teaching and learning strategy to foster students’ metacognition (Zhai et al., 2018; Jensen, 2019). Nevertheless, although the purposes of SR-enabled research appear to be diverse, there are reasons to extend its use in more educational and research settings.

The vast technology integration into education has ushered in changes in the selection of stimuli and technologies for adopting SR in educational research (Gazdag et al., 2019). Technological advancement applied in teaching and learning settings have also expanded the sources of stimuli beyond traditional written notes, classroom photographs, and video recordings. Participants’ learning records on digital platforms and mobile devices can also be used as stimuli to evoke the memory of their own learning path (Koltovskaia, 2020; Lindfors et al., 2020). Furthermore, the shift of instructional environments from offline to online has rendered educational activities in physical scenes more static, lacking observable interactivity to generate an effective stimulus (Duo and Song, 2012; Gijssels et al., 2016; Tan et al., 2021). Some studies have leveraged physiological feedback signals such as eye movement, setting position, and EEG data to provide valuable cues about changes in learners’ inner thoughts (Zhai et al., 2018; El and Windeatt, 2019). However, owing to the constantly evolving technological landscape in learning environments and pedagogical strategies, the question of whether traditional stimuli need to be improved and how to choose new stimuli remains unresolved (Wijayasundara, 2020).

When practicing interview strategies, researchers have exhibited distinctive tendencies in time arrangement and questioning techniques (Gass and Mackey, 2016). Even when the same stimuli were selected, the adoption of interview strategies varied across studies. Concerning the time arrangement of the interview, most researchers contend that participants should be presented with the stimuli and interviewed immediately after the instructional activity, while some researchers intentionally introduce an interval before further interviewing (Gass and Mackey, 2000; Kurki et al., 2016). In terms of questioning techniques, interviewers’ questions can be either entirely open-ended or focused, depending on the research design and educational settings. For instance, Heikonen et al., (2017) commenced with general questions and subsequently narrowed the question scope to explore student and instructors’ reflections on classroom incidents. In contrast, Hu and Gao (2020) posed rather specific questions on students’ responses to linguistic challenges in learning science through English. These disparities may be attributed to the

distinct subjects and research questions that SR measures aim to address (Jackson and Cho, 2018; Tiainen et al., 2018).

In light of the ongoing developments in education and technology, it is worthwhile to conduct a meticulous review of the latest research on applying SR methods in education. Previous reviews were either outdated or narrow in scope. For instance, Keith’s (1988) review centered on studies that applied SR to investigate instructors’ cognitive processes, which, although valuable at the time, can only provide limited guide for current applications of SR in education. More recently, Gazdag et al., (2019) reviewed 35 articles on the use of Video Stimulated Recall (VSR) to enhance instructors’ reflective thinking. However, this study’s scope was confined to implementing VSR in teacher training and excluded studies in broader educational settings. Therefore, further studies are needed to comprehensively examine the application of SR across diverse contexts.

The present study offers a comprehensive review of research using SR in manifold teaching and learning contexts over the past decade. The investigation scrutinizes the characteristics of these studies, such as their research aims, stimuli, and interview strategies. It examines the interplay among these elements, including variations in the purposes of SR employment across disciplines. The ultimate goal of our study is to provide valuable insights for future applications of SR in education and also to aid researchers in exploring the external behaviors and internal thought processes of both instructors and students in a more effective manner.

Literature review

The theoretical foundation of SR in education. SR is a research technique inspired by Dewey’s (1933) reflective thinking concept, which involves presenting participants with vivid prompts to evoke their memories of an original scenario (Bloom, 1953). Since its inception by researchers at Stanford University in 1970, SR has been an essential tool in pedagogical research and widely adopted to investigate various teaching and learning activities in educational research (Stough, 2001). Typically, SR comprises two stages: presenting stimuli and proposing recall questions (see Fig. 1) (Chu and Zhai, 2023). Researchers select specific artifacts, such as notes, audio or video recordings, that exhibit participants’ behavior or cognitive tasks as stimuli, followed by interviews that prompt participants to articulate their intrinsic thoughts, mental processes, or individual feelings at the moment when the stimuli were generated (Calderhead, 1981; Lyle, 2003).

The theoretical basis of SR in educational research draws on the Retrocue Effect and the Cognitive Theory of Multimedia Learning (CTML) (Mayer and Moreno, 1998; Moreno and

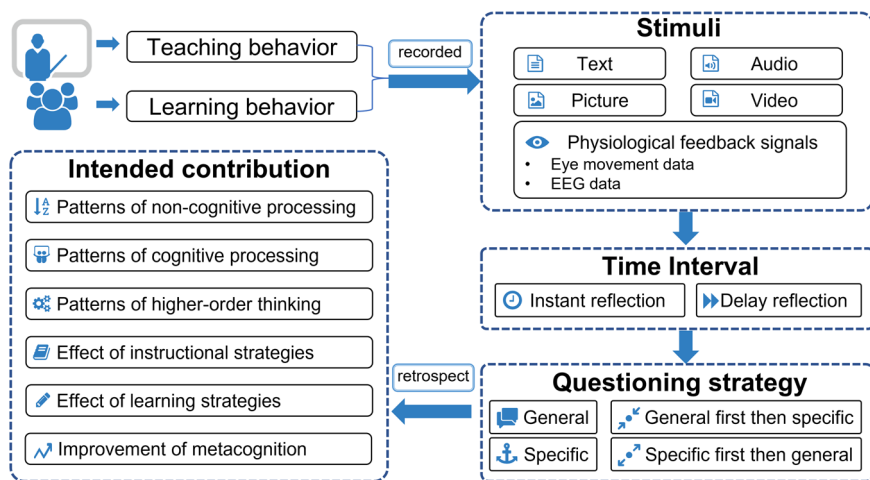


Fig. 1 The flow chart of the SR procedure. The figure shows the main stages of presenting stimuli and proposing recall questions when applying SR.

Mayer, 1999; Souza and Oberauer, 2016; Shepherdson et al., 2018). The Retrocue Effect, a cognitive psychology theory, suggests that an individual's visual working memory is enhanced when their attention is directed toward prior information, even after a delay or distraction (Souza and Oberauer, 2016). Neuroscientific and biopsychological research both provide evidence supporting the protective effect of retroactive attentional focusing on working memory (Duarte et al., 2013; Schneider et al., 2017). According to this theory, retro cues, such as visual stimuli, improve the quality of retrieval and cognitive processes while also reducing cognitive load effects (Shepherdson et al., 2018). Based on this mechanism, SR can offer accurate and specific insights into an instructor or a learner's thoughts and attitudes towards educational tasks.

In addition, the Cognitive Theory of Multimedia Learning (CTML) suggests that multimedia learning is most effective when information is presented in both visual and auditory formats, as learners are actively engaged in the learning process (Mayer and Moreno, 1998; Moreno and Mayer, 1999). As described in the CTML, learners have two separate channels for processing information: visual and verbal (Mayer, 2002; Mayer and Moreno, 2003). When multiple forms of stimuli are presented during the SR interview, instructors and learners become more cognizant of their prior experiences in each channel, which helps them articulate their thought processes in greater detail and enhances their retrospection of previous knowledge and cognition. In conclusion, the application of SR in educational research is rooted in the principles of the ICT and the CTML. Implementing SR provides researchers and practitioners with a valuable tool to gain insight into learners' and instructors' cognitive processes, ultimately leading to more effective teaching and learning.

The educational application using SR. The SR method is an effective technique used in qualitative educational research to gather data on instructors' and learners' thought patterns related to specific events. This method allows researchers to explore instructors' and learners' thinking and decision-making processes, making it a valuable tool for data collection (Nguyen et al., 2013; Bowles, 2018). The use of SR in educational research is critical for maintaining internal validity, as it provides introspective data. Additionally, SR has broad applicability and can be employed in various disciplines for a range of research aims (Meade and McMeniman, 1992; Kurki et al., 2016; Yu and Hu, 2017; Rietdijk et al., 2018; Martinelle, 2020). For instance, Yu and Hu (2017) used SR to probe second language learners' intrinsic and personalized perceptions of peer feedback in collaborative writing assessment, by exploring students' learning behaviors through interviews. Similarly, Kurki et al. (2016) and Rietdijk et al. (2018) tapped into SR to explore how instructors use various teaching strategies and their underlying beliefs, particularly concerning non-cognitive dimensions such as social and emotional factors.

Aside from its application as an educational research method, SR can also serve as an effective teaching and learning strategy. Instructors can use SR to assess what learners remember or may have overlooked to determine learning reinforcement strategies. SR enables learners to examine and articulate their thoughts through memory retrieval and thus elevating their thinking to a new level of expression. Therefore, SR can enhance learning rather than solely serving as a research approach (Smagorinsky, 1998; Egi, 2008). In addition, VSR is a valuable teacher training and development tool that includes video-supported reflection and questioning. This approach motivates instructors to reflect on themselves and their practice consciously, facilitates metacognitive reflection among preservice teachers, and provides reflective

prompts for educational interactions (Geiger, Muir, and Lamb, 2016; Endacott, 2016).

While linguistics and teacher education are the primary application areas of SR, it is also used in other subjects, such as STEM education (Gass and Mackey, 2016; Al Mamun, Lawrie, and Wright, 2020; Schindler and Lilienthal, 2019). However, the purpose of applying SR varies depending on the subject and learning environment. Recent advances in instructional technologies have transformed teaching strategies and educational settings, yet the relationship between these elements and the principles of SR application in distinct contexts is still ambiguous.

Stimuli and interview strategies in SR. The rapid diffusion of Information and Communication Technology (ICT) and the exponential growth of online learning have brought new challenges and opportunities for using SR in educational research and in teaching and learning. Integrating ICT into education requires a careful selection of stimuli that can adapt to the constantly evolving learning environments. When applied in physical environments, audio or video stimuli respond favorably to interactions between teachers and students, enabling subsequent interviews to investigate their inner impressions or perceptions (Chu and Zhai, 2023). In contrast, educational activities incorporating digital technologies are not easily observable, with instructional behaviors conducted through electronic devices and in video or audio conference systems. It is often challenging to find informative stimuli reflecting teacher-student interactions in digital settings.

Nevertheless, technological breakthroughs have enriched stimuli by expanding data collection channels and capacities at the same time. Through the integration of additional stimuli sources such as weblogs, computer screen captures, and biofeedback data, researchers are able to unearth information about learners' inner workings. For example, Révész et al., (2017) gained a comprehensive picture of the L2 writing process and acquired a deeper understanding of implicit thinking using eye movement data-based stimuli. Overall, considering the diverse data collection methods and changing learning contexts, stimuli selection in SR in technology-enabled schooling still requires further clarification.

The interview stage is another crucial aspect of SR that distinguishes it from conventional memory recall. This stage emphasizes estimating internal thinking processes and determining how the method can encourage instructors' and learners' reflection and delve into their internal ideas. During the interview stage, researchers mainly acquire tacit data. Previous studies are inclined to perform interviews promptly after class and employ standard open-ended questions to encourage participants' agency in reflecting on their experiences (Gass and Mackey, 2000; Chu and Zhai, 2023). However, some studies have chosen different approaches. For example, when investigating early childhood teachers' instructional activities, behavior, and emotions, Kurki and his colleagues (2016) delayed inviting teachers to take part in the interview by two weeks. Additionally, researchers argue that, apart from using generic questions, incorporating specific follow-up questions that closely align with the research aim is equally crucial (Heikonen et al., 2017; Hu and Wu, 2020). Despite the significance and disparities in interview strategies, few studies have specifically analyzed this issue, and well-developed principles of organizing interviews and questions is absent.

SR has become an essential technique for examining cognition and behavioral patterns in education by activating instructors' and students' retrospection through stimuli and interviews. As SR has evolved and educational paradigms have transformed, the research purpose and critical steps, such as stimuli selection and

Table 1 Research objectives and coding items.

Research objectives	RQ	Coding items
Contributions of SR to educational research	RQ1	According to the selected sample pool, what were the research aims of applying SR in educational research from 2012 to 2022?
Implementation of SR in educational research	RQ2	How have the stimuli evolved during the implementation of SR in the last decade?
	RQ3	What time factors should or could be considered with regard to the efficacy of stimulated recall? Are they limited to the duration of a single course?
	RQ4	What interview strategies have SR research reviewed in this paper employing?
	RQ5	What is the relationship between different components (research aim, stimuli, question strategies, learning subject, etc.) of SR?
Challenges and potentials of SR in future educational research	RQ6	How can educational research adopting SR be improved in the future?

interview strategies, of applying SR in educational research require further discussion. Education is a complex system with intertwined intrinsic elements such as discipline differences and learning environments (Jacobson and Wilensky, 2006; Jacobson et al., 2019), which can influence stimuli preference and the conduct of interviews.

Therefore, in order to provide insights for learning from past educational applications of SR and enhancing future developments, the present systematic review scrutinized the evolution of the SR method in educational research from 2012 to 2022. It aimed to elucidate what the contributions SR has made, how SR has been implemented, and the challenges and potentials it presents. To fulfill these objectives, we further proposed six specific coding items (see Table 1) to guide our content analysis coding procedure and decoding interpretation.

Methods

Guided by the aforementioned research questions, we systematically analyzed and interpreted studies related to SR from 2012 to 2022. Given the significant changes in teaching environments and research methods associated with the rapid development in educational technology, we believe that a 10-year time span can provide sufficient coverage of research in a variety of disciplines. We used qualitative content analysis to examine these studies, which consists of two steps: selecting papers for review and coding these papers by using an established coding scheme.

Paper selection. To guarantee the quality of selected, our research team reviewed well-recognized peer-reviewed articles in the Web of Science (WOS) core collection, Scopus, and IEEE Xplore. These databases contain reputable journals with recognized impact factors. The articles retrieved in WOS and Scopus can be further refined into social science or educational categories, allowing for more precise retrieval. Additionally, given the focus of this research on the use of technology in education, the IEEE database provides focused research in scientific and technical disciplines.

Two stages comprise the processes used to identify the research papers. In the first stage, the keyword “SR” was selected, and the discipline was refined to “education and educational research”. This process yielded 309 articles. In the second stage, the abstracts and full text of the chosen articles were manually and systematically screened by researchers to confirmed that they: (1) included the SR protocols, (2) prompted participants to reflect on their thinking process, (3) focused on research issues in the field of education, and (4) provided empirical evidence or evaluation rather than solely summarizing previous findings. For example, some articles merely reviewed others’ research on the SR method employed in teaching settings or using painting-based stimuli to spark students’ prior knowledge did not meet the inclusion criteria (Gazdag et al., 2019; Walan and Enochsson, 2019).

According to Golhasany and Harvey’s (2023) study, the coder should pose doctoral degrees or professorships in the relevant field, and each identified papers should be individually scrutinized by experts. Finally, three experts were selected to examine the sample pool: two of whom have doctoral degrees and professorships on learning technology, while the third have a doctoral degree in educational management and post-doctoral experience on learning technology. Moreover, to ensure there are no conflicts of interest, only one coder is involved in the authorship. The inter-coder reliability was assessed following a specific schedule: first, the coders independently examine the selected samples and provided their judgment. Then we use the Fleiss Kappa test in SPSS 26 to test the reliability. The results ranged from 0.874 to 0.973 indicating satisfactory inter-rater reliability and consistent coding for each item. Finally, we adopted the final coding results if all the experts or at least two of them agreed. Finally, the research team identified 257 representative papers as the research sample of this study. As recommended by Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA, Page et al. 2021), we conducted the systematic review with a strong and thorough methodology. Figure 2 depicts the flow of this screening process, which is in accordance with Moher et al. (2009).

Coding procedure. The identified articles were systematically coded to carry out a precise and thorough examination of the utilization of SR in education. By adopting Gass and Mackey’s (2000) definition of SR, this study identifies its key components. They established a coding framework, including the research aim, stimuli, questioning technique, and questioning interval. Additionally, coding the learning subject and educational context helped clarify how to implement SR effectively in various situations. Table 2 illustrates the background information of SR research, such as the learning subject and educational context. Despite reviewing research involving instructors and students as subjects, this study did not differentiate between these two groups as the primary focus of SR is to investigate the participants’ consciousness and thinking behind their behavior, regardless of their roles. The table included in supplementary information described our data collection process.

The coding process involved identifying and extracting relevant data from the selected papers. Any discrepancies were resolved through discussion and consensus-reaching among the research team. We then analyzed the coded data and identified patterns and themes abiding by the content analysis method. The findings of this review are presented in the following sections, addressing the research questions outlined earlier.

Results and discussion

In accordance with the content analysis and coding criteria mentioned above, 257 papers were thoroughly examined. The

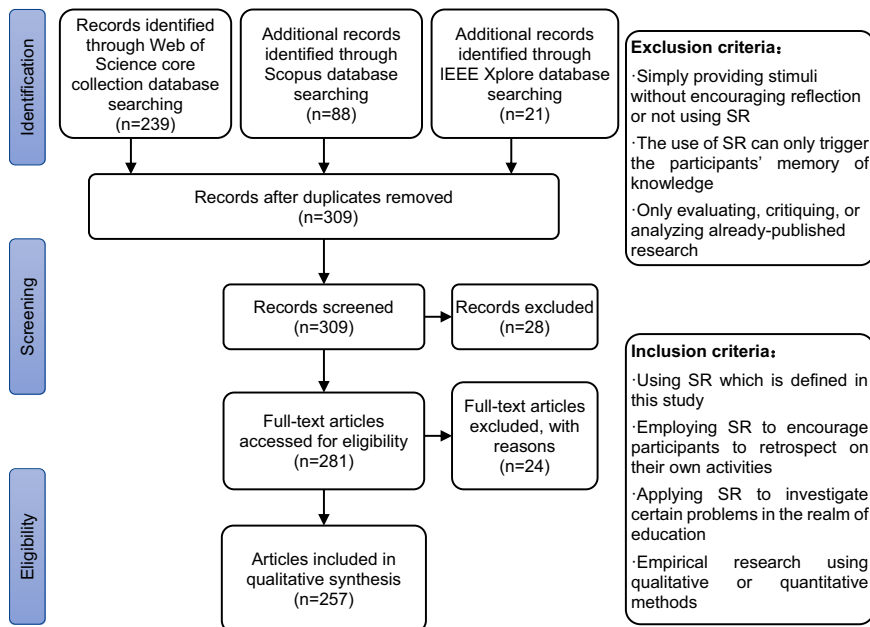


Fig. 2 Research flow. The diagram presents the systematic review flow according to PRISMA.

Table 2 Research foci.

Research foci	Description
Research aim	The expected purposes of using SR by the researcher
Stimuli	The materials used to trigger reflection among participants on prior activities
Questioning technique	Techniques implemented for questioning to stimulate retrospection during SR
Questioning interval	The time duration between the in-class activities and the SR interview
Learning subject	Disciplines where SR was employed to conduct research
Educational context	The learning environment in which SR was employed to conduct research

following sections present the results and provide a corresponding discussion of the research questions.

RQ1: Research aims. The current literature reveal that SR is often applied to studying instructors’ and learners’ inner thoughts and ideas, prompting them to recall and comment on their thinking process. Furthermore, this approach can also examine the effect of teaching and learning strategies and to improve participants’ metacognitive skills. Because SR has long been used in educational settings, it is surprising that more substantial research has yet to be conducted on how it might be expanded and how to overcome its limitations such as time factors and distractions. Therefore, our work focuses on promoting the effectiveness of widespread application of SR in teaching and learning.

Exploring thought patterns. Exploring thought patterns is the primary focus of most educational research that uses SR. This includes investigating non-cognitive and cognitive processes, as well as higher-order thinking. As shown in Table 3, over half (157 in total) of the research reviewed employed SR to achieve this objective. Eighty-two of the reviewed studies explored patterns of non-cognitive processing, such as motivation, emotions, and cultural orientation (Lichtinger and Kaplan, 2015; Ucan and Webb, 2015; Wilby et al., 2017). This method can also investigate the variables influencing the willingness to communicate or the ethical considerations of instructional practices in language learning (Rissanen et al., 2018; Chichon, 2019; Peng, 2020).

In addition, fifty-five studies explored patterns of cognitive processing, probing the epistemic thinking of diverse participants in various subjects, including language learning, STEM, and the arts (Bogard et al., 2013; dos Santos and Loveridge, 2014; Révész et al., 2017). Furthermore, some studies based on SR also obtained insight into both cognitive as well as non-cognitive processes through the integration of multimodal data (Lambert and Zhang, 2019). It is worth mentioning that a total of 11 papers explored both cognitive and non-cognitive thought processes with SR.

Finally, nine studies applied SR to investigate patterns of higher-order thinking, such as creative thinking and critical thinking, as well as the collaborative process (Rissanen et al., 2019; Schindler and Lilienthal, 2020; Łuczniak and May, 2021). The application of SR in these studies allowed for a more comprehensive understanding of participants’ thinking processes and the factors that contribute to effective collaboration and higher-order thinking.

Investigating the effect of educational strategies. Another purpose for research employing SR is to investigate how participants’ learning processes and experiences are affected by instructional design or learning models. Specifically, 35 articles used SR to investigate the impact of specific learning strategies in educational settings, yet 54 studies examined instructional techniques. It seems that SR can facilitate investigating how instructors and students understand and apply newly adopted teaching or learning strategies.

Table 3 Descriptive results of included articles (n = 257).

Description	n	%	Description	n	%
Research aim			Questioning interval		
Patterns of cognitive processing	55	21.40	Instant reflection	126	49.03
Patterns of non-cognitive processing	82	31.91	Delay reflection	23	8.95
Patterns of higher-order thinking	9	3.50	Not mentioned	108	42.02
Effect of instructional strategies	54	21.01	Learning subject		
Effect of learning strategies	35	13.62	Linguistics	115	44.75
Improvement of metacognition	9	3.50	Teacher education	48	18.68
Patterns of cognitive & non-cognitive process	11	4.28	STEM	35	13.62
Patterns of cognitive process & Effect of instructional strategies	1	0.39	Physical Education (PE)	10	3.89
Patterns of cognitive process & Effect of learning strategies	1	0.39	Art	9	3.50
Stimuli			Other subjects	5	1.95
Text	31	12.06	No specific domain	35	13.62
Pictures	5	1.95	Educational context		
Video recordings	173	67.32	Physical settings	194	75.49
Audio recordings	13	5.06	Online platform	48	18.68
Physiological data	14	5.45	OMO environment	6	2.33
Not mentioned	5	1.95	Virtual Reality	1	0.39
Text & pictures	4	1.56	Not mentioned	8	3.11
Text & audio/video recordings	10	3.89			
Pictures & audio/video recordings	2	0.78			
Questioning technique					
General	63	24.51			
Specific	73	28.40			
General first and then specific	25	9.73			
Specific first and then general	1	0.39			
Not mentioned	95	36.96			

A few papers may have more than one research aims and make use of several sources of stimuli.

For instance, SR has revealed the pedagogical knowledge base related to the use of dashboards and the provision of feedback by novice teachers (Karimi and Norouzi, 2017; Molenaar and Knoop-van Campen, 2018; Yu, 2021). In terms of the effectiveness of learning techniques, such as computer-enhanced self-directed learning, SR offers a more precise and comprehensive understanding from students' viewpoints (White et al., 2016; Deng, 2020; Chu and Zhai, 2023).

Extensive empirical studies have shown that data acquired through SR can enhance the interpretability of single-outcome data, such as test results, and also produce more insightful information to evaluate and enhance strategies for improved learning outcomes for both instructors and students (Van der Kleij et al., 2017). In these studies, SR provided a deeper comprehension of how instructional design or learning strategies impact participants' learning experiences and outcomes.

Improving metacognition. Nine articles took advantage of SR to improve participants' metacognition. One example is using SR in online language learning, where learners can compare feedforward and eye-movement data to develop their metacognitive skills (Zhai et al., 2022). Metacognition refers to an individual's awareness of their thinking processes and understanding of the underlying patterns (Flavell, 1979). Educational psychologists have widely acknowledged the significance of metacognition due to its substantial correlation with learners' academic achievements.

Metacognitive activities usually occur during the self-reflection phase and involve the participants' evaluation of their own cognition, understanding, and task performance. Using recorded learning processes as stimuli, participants are prompted to explain or evaluate their past behavior instead of simply recalling knowledge. Encouraging student participation in video-stimulated recall conversations enhances their self-reflection and improves their metacognitive skills by giving them a scaffold (Van der Kleij et al., 2017).

RQ2: Stimuli. Regarding the stimuli used to arouse recall, video recordings of the learning process have gained tremendous popularity. While some changes have occurred during the evolution of SR methods, such as the improvement of video stimuli and the integration of physiological feedback data.

Optimizing video stimuli. Video recordings are a widely popular type of stimuli in educational SR research, as evidenced by nearly 70% of the reviewed studies (173 articles) that utilized them. This prevalence of video stimuli has been noted in previous review research by Gazdag et al. (2019), which to some extent, explains the exclusive focus on video stimuli in his study. These recordings commonly consist of real-life scenes from classrooms, laboratories, and screen captures of technology-mediated learning settings. To serve as effective incentives for participants, video recordings should reflect the interactions between instructors and learners, and researchers ought to regulate their length to prevent participant weariness (Lee, 2020).

A number of enhancement options have been suggested by researchers. One technique is to use multiple cameras to record and display split-screen videos, providing various perspectives of the learning environment. For instance, Jackson and Cho (2018) produced a split-screen video recording of teachers' and students' simultaneous behaviors, enabling a more potent stimulus for supporting event recall, contextual and situational recall.

Additionally, some researchers have used head-mounted video cameras to record video from the participant's perspective, visually reproducing original learning scenarios. Such an approach is beneficial in studies examining interpersonal communication, such as those exploring teacher-student interactions or teacher interventions in early childhood peer conflict (Agricola et al., 2021; Myrtil et al., 2021).

Utilizing biofeedback data. With more accurate and detailed data, biofeedback data (14 articles) is increasingly considered a

stimulus choice for self-reflection. Currently, eye-tracking technology is the most commonly used physiological feedback technique. The Eye-Mind hypothesis suggests that eye movements correspond to mental operations, allowing researchers to infer cognitive processes from gaze patterns (Obersteiner and Tumpek, 2016, p. 257). By combining eye-tracking data with self-reflection, potential ambiguity and uncertainty in eye-tracking techniques are reduced, giving a more thorough overview of the educational process for reflection (Schindler and Lilienthal, 2019; El and Windeatt, 2019; Chu and Zhai, 2023).

Moreover, other physiological indicators have served as informative stimuli in self-reflection. Zhai et al. (2018) found that online learners' reading comprehension and cognitive abilities were significantly improved by using both eye-movement and EEG physiological signals as stimuli. Multiple physiological indicators can be included to provide a more thorough and accurate picture of the cognitive and affective states of learners during the learning process.

RQ3: Time factors. Considering time factors is crucial for the utilizing SR method in educational research. This is because time not only influences the selection and processing of stimuli but also has implications for the subsequent interviews. Specifically, enhancing the temporal properties involves both reducing the presentation time and increasing the span of information provided by the stimuli source. Moreover, it is essential to set appropriate time intervals to schedule the interviews. The reviewed literature suggests that the interview schedule may vary depending on the study.

Enhancing the temporal properties of stimuli. Presenting learners with stimuli is intended to assist them in reflecting on their previous learning activities. Nevertheless, if the presentation of stimulus sources persists for too long, it can also impose a heavier cognitive load on learners (Pratt and Martin, 2017). In general, stimuli sources in textual, image, and other static formats are more convenient due to their controllable presentation duration for participants. However, for classroom video recordings stimuli, direct video replay may take a considerable amount of time. Considering the time spent, one such approach involves selecting clips from full-length video footage, reducing the duration of the stimuli, and enabling participants to concentrate specifically on behaviors that are pertinent to the research aim (Määttä et al., 2016).

The time span of stimuli is also not limited to the classroom. As demonstrated in the 16 studies reviewed, combining multiple data sources has proved more effective. The integration of various materials, including text, video, and pictures, enhances the information capacity and authenticity of the recorded details. For instance, in limited interaction scenarios, researchers often use think-aloud methods, allowing participants to verbalize their thoughts, along with the videotapes, to augment the information provided (Kang and Pyun, 2013).

In addition, incorporating stimuli from multiple sources can encompass both subjective and objective aspects. Video recordings only capture a limited timeframe, while learning behaviors extend beyond the confines of the classroom. Cues to stimulate participants' recall can also come from guide sheets, teacher preparation notes, and student class notes (Chu and Zhai, 2023). In an investigation on the use of metacognitive interventions in twenty-first century writing pedagogies, stimuli included a classroom tour video, a literacy autobiography, a teaching plan, and other instructional materials (Jensen, 2019).

Setting up flexible intervals. The time interval between in-class instructional activities and SR interviews generally varies across

researchers. Among the 149 reviewed studies where the time interval was specified, the majority of the study (126) preferred instant reflection. Instant reflection involves conducting SR interviews as soon as participants finish their learning tasks, typically with only a 5- or 10-min interval or a slight delay according to the timetable for curriculum (White et al., 2016; Rassaei, 2015; Shintani, 2016; Fernandez, 2018).

A shorter time span makes sure that participants recollect the task's cognitive processes more precisely, which improves the accuracy of the interview results (Gass and Mackey, 2000). Instant reflection is particularly valuable in studies that require precise information about the learners' cognitive processes and strategies during the learning task.

However, some researchers (23 studies) purposefully extended the time interval between instruction and recall, for example, 2–4 weeks after the task was completed (Harvey et al., 2014). This design may optimize the study by allowing more time for the process of previously recorded raw data and footage (Nurmu-khamedov and Kim, 2010; Kurki et al., 2016). Delayed interviews can also reduce research impact on participants by avoiding interference with subsequent teaching and learning activities (Dos Santos and Hentschke, 2011).

RQ4: Interview strategies. During the interview phase of SR, to better guide participants in autonomously reflecting on the teaching and learning process, researchers also need to pay attention to the use of strategies, including the openness and value-oriented nature of the questions.

Posing appropriate questions. SR interviews are utilized to encourage reflective thinking in participants within an open and dialogic environment through questioning strategies. Typically, this kind of interview consists of a succession of open-ended, introspective, and generic questions that do not require pre-determined answers. This pattern has been observed in 88 reviewed studies, including research merely posing general questions, as well as those starting with general questions and then progressively narrows down the focus. During these interviews, researchers should take on the role of listeners, serving to train, facilitate, and illuminate while avoiding asking leading questions that could result in biased responses (Ramnarain and Modiba, 2013; Egi, 2008; Gass and Mackey, 2000; Sato, 2019). For instance, researchers should avoid yes-no questions that could encourage participants to react a specific way or provide presentational responses. This approach ensures that the purpose of the SR interview is maintained and that the risk of biased responses is minimized (Thararuedee and Wette, 2020; Rassaei, 2020;).

While questioning in SR interview should leave ample room for participants to retrospect, it must also address the research questions. Thus, 25 papers suggest that questions should be open-ended at the beginning but become increasingly specific as the interview progresses (Stolpe and Björklund, 2013). Researchers can use supplementary queries as prompts to ensure that the interview stays on topic and delves deeper into the research questions, depending on participants' responses (Qiu and Lo, 2017; Qiu, 2020; Chu and Zhai, 2023).

Staying value-neutral in guiding. In addition to the scope of questioning, the neutrality and guidance provided by the interviewer are crucial. Participants receive training before the interview on how to reflect on previous cognitive processes, and the interviewer should remain as neutral as possible during the interview to capture retrospective thinking solely supported by the stimuli (Consuegra et al., 2016). If respondents feel that the

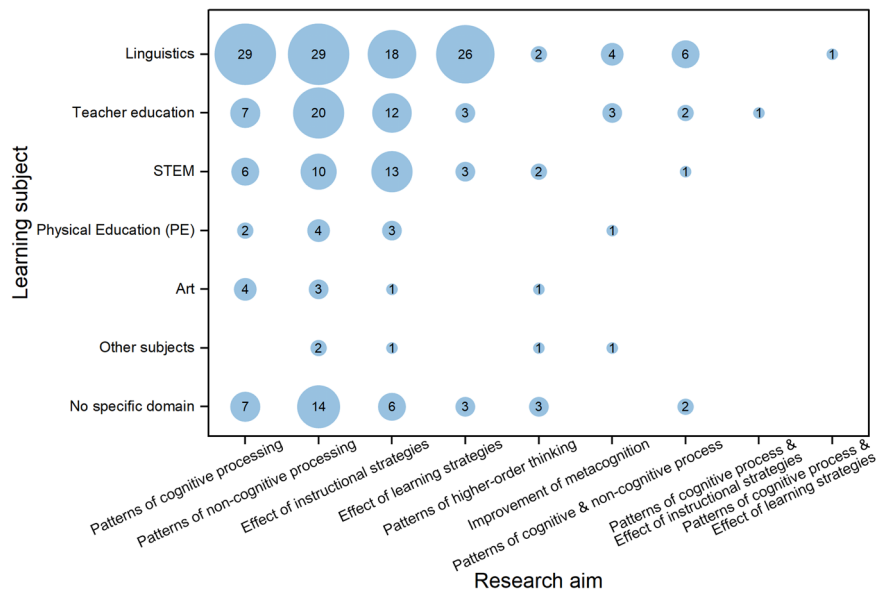


Fig. 3 Bubble chart involving research aim and learning subject. The relationship between research aims and learning subjects is depicted in this bubble chart, where the size of the bubbles represents the quantity of relevant studies.

questions are biased or contain value judgments, they may feel pressured to rationalize or make up explanations, leading to inaccurate reporting of their thoughts. Therefore, the interviewer must carefully design questions wording and adopt a supportive attitude that indicates curiosity in the descriptions provided by participants rather than making judgments (Wu, 2019; Schindler and Lilienthal, 2019).

RQ5: The relationship between different coding items. In addition to key application procedures such as research aims, stimuli, time factors, and interview strategies, the implementation of SR in educational research is also influenced by intrinsic factors within the educational context, such as learning subjects and educational context. The results of the review (see Table 3) indicate that SR is primarily employed within the fields of linguistics (115 articles) and teacher education (48 articles), with relatively fewer studies in areas such as the arts (9 articles). Over 75% of the articles still apply SR in physics learning environments, while nearly 20% explore the use of SR in online platforms.

To enhance the exploratory nature of the research discussion, the current study delved deeper into the intricate relationship between coding items. It is important to note that only outcomes warranting further exploration and discussion are presented in the subsequent section.

The relationship between research aims and learning subject. This bubble chart (Fig. 3) illustrates the connection between research aims and learning subjects, with the size of the bubbles indicating the number of relevant papers reviewed. Our current analysis aimed to explore whether SR is more suitable for investigating specific research questions in different disciplines.

Regarding research aims, SR was primarily used to investigate patterns of non-cognitive processing and the effect of instructional strategies across all subjects. In linguistics, researchers most frequently utilized SR to explore patterns of cognitive processing (29 articles), non-cognitive processing (29 articles), and learning strategies (26 articles). Another six studies focused on both cognitive and non-cognitive occurrences in linguistic teaching and learning. These findings are consistent with prior research highlighting the importance of non-cognitive factors (e.g.,

motivation) and learning strategies in language learning (Lin et al., 2017). Furthermore, 20 studies using SR investigated non-cognitive elements in teacher education contexts where teachers' non-cognitive factors, such as intrinsic motivation, are strongly associated with their professional development (Maaranen et al., 2019).

In the realm of educational subjects, SR has also occupied a pivotal within the domain of STEM and art education research. Within the STEM disciplines, researchers have employed this methodology to probe the impact of pedagogical strategies (13 articles), non-cognitive processing (10 articles), and cognitive processing (6 articles). Intriguingly, SR has been invoked more frequently to investigate cognitive rather than non-cognitive factors within the sphere of art education (four articles versus three). This inclination may stem from the intricate nature that cognitive processing exhibits in artistic creation (Révész et al., 2017). Nonetheless, SR has demonstrated its utility as an effective facilitator, enabling arts educators to acquire profound insights into the cognitive aspects of art instruction and learning. For instance, dos Santos (2018) documented music teachers' approaches to the instruction of rhythmic skills as stimuli, facilitating their reflection upon their cognitive processes and their utilization of their didactic content knowledge.

Linguistics and teacher education are two fields that more frequently took advantage of SR as a teaching strategy beyond research methods (Meade and McMeniman, 1992; Geiger et al., 2016; Sanchez and Grimshaw, 2019). Four articles in linguistics and three in teacher education explore using SR to improve participants' metacognition. In particular, teacher's professional development and language learning emphasize reflective practice and metacognition (Belvis et al., 2013; Zahid and Khanam, 2019). For example, research on teachers' noticing highlights the importance of their cognition and behavior in classroom situations (Chan et al., 2021; Amador et al., 2021). In language learning, metacognitive awareness has been found to enhance foreign language writing ability, emphasizing the need for metacognitive strategies to improve learners' skills (Ramadhanti and Yanda, 2021; Farahian, 2015). These requirements for introspective behavior and metacognition in language learning and teachers' professional development align with the essential steps and reflective characteristics of SR.

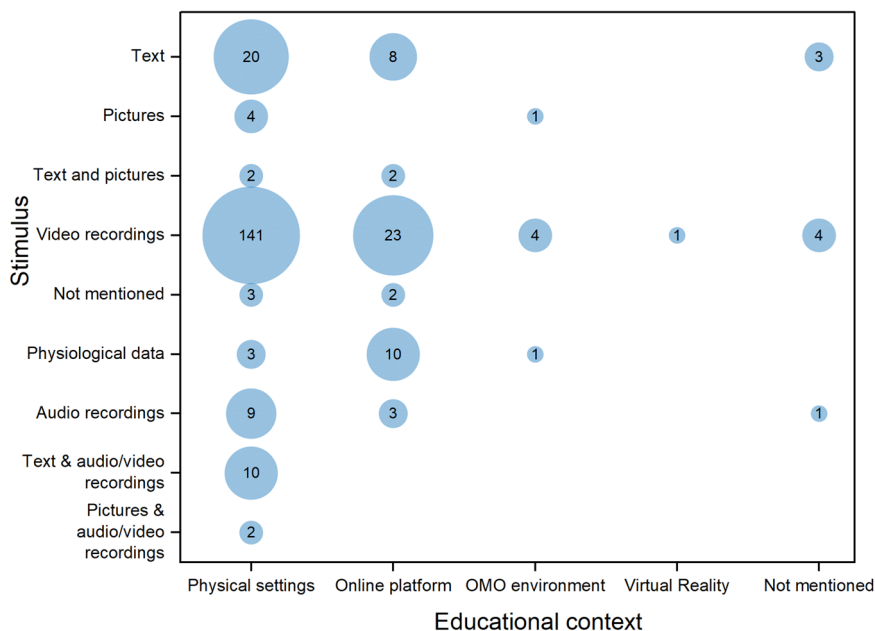


Fig. 4 Bubble chart involving stimuli and educational context. With the size of the bubbles indicating the number of articles in the review, the bubble diagram illustrates connection between learning environment and stimuli.

The relationship between stimuli and educational context. The bubble diagram (Fig. 4) displays the stimuli and learning environment, with the size of the bubbles corresponding to the number of articles in the review. Our analysis aimed to scrutinize which stimuli are commonly adopted in different learning environments.

Firstly, video footage remains the dominant stimulus across various scenarios, with over half of the studies (141 articles) utilizing video in physical learning settings and 23 studies using video footage to stimulate reflection in digital learning platforms and OMO settings. However, there is a disparity in the type of videos used in these settings. Physical learning environments mostly relied on live-action videos that authentically recorded participants’ behaviors and interactions (Chan and Yung, 2015; Nyberg and Larsson, 2017), whereas digital learning environments utilized device screen recordings that captured participants’ operations on computer-supported learning platforms (Rassaei, 2013; Lee, 2020).

Secondly, alongside video recordings, physiological data of participants is most frequently used as a stimulus for SR research based on online platforms (10 articles). This trend is reasonable as screen recordings alone may not fully reflect students’ behavior, mainly if they do not perform mouse manipulation or keyboard input. For instance, eye-gaze behaviors provide direct and objective evidence, including fixation duration, fixation count, and scanning path, allowing for stronger conclusions about learners’ cognitive processes and learning strategies (Lai et al., 2013; Michel et al., 2020).

Thirdly, the diagram indicates that studies utilizing SR in physical environments are more mature and inclined to utilize multimodal stimuli. However, studies in online platforms, OMO environments, and VR environments are still limited, with predominantly homogeneous stimuli. Only one study explored students’ learning strategies in an English task using video footage as stimuli in a virtual reality setting (Park, 2018). Thus, more than relying on text, images, or audio and video alone as stimuli is required, and more physiological and multimodal stimuli should be employed in future teaching and learning environments.

The relationship between questioning strategy and research aim. The diagram presented here (Fig. 5) displays the questioning

strategy and purpose of the study using the SR method, with the size of the bubbles representing the number of articles. Based on this information, our analysis aimed to explore whether the purpose of the study influenced questioning strategies.

They were excluding articles that did not mention questioning techniques, 23 studies exploring non-cognitive processing utilized general questions during SR, outnumbering studies that implemented more focused questioning strategies (20 articles). This preference for general questioning may stem from the diverse and individualized nature of non-cognitive skills, which include motivation, responsibility, and perseverance (Smithers et al., 2018). Consequently, general questions are more appropriate as they allow participants to autonomously recall non-cognitive processing activities with the aid of stimulus materials. Moreover, reflection on non-cognitive processing is prone to interference from external factors. If interview questions are too directed towards the research objectives, they may interfere with the results.

In contrast, studies focusing on cognitive processing patterns predominantly utilized questioning sessions centered on research questions (21 articles), nearly double the number of studies using general questioning strategies (14 articles). Cognitive processing is often intimately related to the teaching or learning activity. Thus, researchers tend to focus their questioning on the learning activity that concerns the research goals. Notably, one article exploring cognitive patterns adopted a different questioning technique: focused first and then general. This article investigated what musicians learned when teaching older adults (Perkins et al., 2015). On the one hand, the research questions themselves were exploratory, and the researcher expected participants to provide more cognitive information. On the other hand, this phenomenon may also reflect the divergent and creative thinking of art learning, requiring questions that encourage participants to reflect freely after satisfying the research objectives.

RQ6: Potential improvements. In addition to exploring how effectively employ SR in education, the current review also points to possible future directions on SR research with existing models of computer-supported learning and technology-assisted instruction.

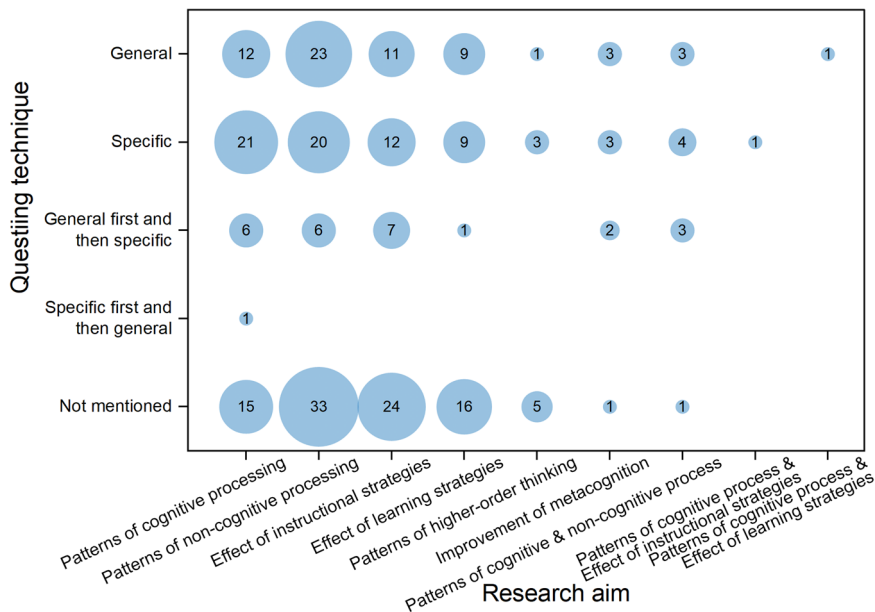


Fig. 5 Bubble chart involving questioning strategy and research aim. With the size of the bubbles signifying the number of articles, the diagram illustrates the SR questioning strategy and the research aim.

Enhancing the dependability of outcomes through the synthesis of multifaceted data sources. Table 3 shows that a mere fraction under 10% of the studies (16 articles) utilized multi-source stimuli. Indeed, the amalgamation of data derived from disparate stimuli can provide a complementary and robust scaffold for the outcomes of SR. This is attributable to the fact that the integration of heterogeneous types of stimuli broadens the information spectrum, providing participants with supplementary prompts that facilitate the recollection of cognitive processes. Such an approach diminishes the cognitive load on subjects, assists them in articulating more accurate reflections, and arguments the reliability of SR. Furthermore, this practice contributes to the transparency of educational research (D’Oca and Hrynaszkiewicz, 2015). For instance, combining video, audio, and text stimuli can offer a more comprehensive and nuanced understanding of learners’ cognitive processes and behaviors.

Additionally, using multimodal stimuli can help address the limitations of using a single type of stimuli and enhance the ecological validity of the study by better replicating real-world learning environments. Some researchers (e.g., Rankanen et al., 2022) conducted a study that employed both quantitative and qualitative methods to investigate the impact of non-instructional clay-making in art education on learners’ creative thinking and positive emotion stimulation. By combining multiple data sources, including physiological data such as heart rate variability (HRV) and electrocardiogram signals, this study provides a more detailed understanding of the art experience and the mechanisms at work in different art forms. Unlike previous research that relied solely on questionnaires, this study includes more objective and in-depth quantitative data analyses of art learning tasks. Additionally, the researchers used video-stimulated recall in addition to HRV data to provide a comprehensive perspective on the learners’ experience of non-instructional clay-making in art education. Including qualitative data can reveal the positive or negative value of the emotional experience of artmakers and provide a more nuanced understanding of the emotional complexity of art.

Strengthening the acquiring and processing of stimuli by adopting intelligent algorithms. The synthesis of the review indicates that

over 70% of the studies (185 articles) employ video recordings as a singular or combined source of stimuli as depicted in Table 3. Therefore, the employment of AI algorithms to refine the processing of video stimuli could markedly enhance the application of SR in education research.

Firstly, algorithm-supported techniques can assist in selecting the relevant learner interaction portion of video stimuli, thus shortening the length of SR and automatically extracting key information. Wass and Moskal (2017) proposed an automatic video annotation tool, which scaffolds more profound reflections and reduces the cognitive load in participating instructors and students. This intelligent excerpting and annotation process saves time and reduces labor, thus improving the efficiency of SR. Furthermore, algorithm-supported techniques can help to automate the coding and analysis of the data, reducing the potential for human error and increasing the reliability of the findings.

Intelligent algorithms can effectively address the challenge of identifying specific moments or events in classroom videos that are relevant to research questions and require meticulous observation, particularly in cases where the video playback duration is extended. Recent advances in computer vision and machine learning have made it possible to automatically extract valuable information from classroom videos, such as the head pose, gaze direction, and facial expressions of instructors and learners, as well as synchronous behaviors between neighboring students (Goldberg et al., 2021). Furthermore, the use of Convolutional Neural Networks (CNN) and Deep Neural Network (DNN) enables the analysis of audiovisual data to identify and annotate class environments, such as the teacher’s instructional strategies, student engagement, and classroom management (Ramakrishnan et al., 2023). The application of these intelligent algorithms has significant implications for using video recordings in SR, as they provide an accurate and comprehensive depiction of the classroom experience, enabling a more efficient analysis of video recordings in SR. By integrating intelligent algorithms, the effectiveness of retrospection can be enhanced, as algorithm-supported stimuli playback offers a reflective cognitive scaffolding beyond the mere recollection of the learning process.

Considering that many researchers have begun to incorporate physiological data as a source of stimuli (14 articles, as indicated

in Table 3), the application of computer vision or machine learning algorithms could also be instrumental in capturing and analyzing learners' physiological data in a lightweight manner, such as recognizing and analyzing their gestures and micro-expressions via webcam, which enriches the informativeness of stimuli (Zhai et al., 2022). Machine learning algorithms can now identify and analyze patterns in learner behavior that may not be apparent to human observers, providing a more nuanced understanding of cognitive processes. Moreover, intelligent algorithms enhance the reliability of findings and can also prevent the potential for the Hawthorne effect resulting from direct observation and data collection.

Facilitating the implementation interviews by using virtual agents powered by generative AI. Interviews are instrumental in the process of SR, with the majority of researchers opting to pose not merely general enquiries, but targeted ones (99 articles, as referenced in Table 3). This necessitated the undertaking of comprehensive interviews with each participant, a process that is notably time-intensive and requires substantial human endeavor. Future research could explore the use of virtual agents supported by generative AI technologies as an alternative approach to completing the questioning process of SR. Educational research has shown that intelligent agents positively affect learner motivation, academic performance, and cognitive load, making them ideal for training learners' metacognitive abilities (Dinçer and Doğanay, 2017; Kautzmann and Jaques, 2019).

Intelligent conversational agents powered by natural language processing (NLP) and large language models (LLM) can replace researchers in providing participants with questions that prompt their recall and offer adaptive feedback based on their responses (Bozkurt, 2023). Employing intelligent agents to conduct interviews increases the number of subjects without increasing labor costs. For instance, OpenAI has developed several cutting-edge AI technologies, including the GPT series of language models such as ChatGPT, which can presumably be applied to provide personalized intelligent tutoring services in which feedback-enabled iterative learning occurs (Qadir, 2022). Furthermore, the LangChain architecture makes it easier to develop domain-specific agents. Such technologies provide tailored feedback to learners, enhancing their metacognitive awareness and learning outcomes. Additionally, recent advancements in generative AI have shown promising results in producing various forms of multimedia content, including text, images, videos, and 3D models (Gozalo-Brizuela and Garrido-Merchan, 2023). This ability to generate multimodal content aligns with the Cognitive Theory of Multimedia Learning, which emphasizes using multiple sensory channels to facilitate learning experiences (Mayer, 2002; Mayer and Moreno, 2003). By providing learners with diverse visual and auditory information, this technology can enhance the effectiveness of educational activities and promote reflection among instructors and students.

Nonetheless, using virtual agents in educational SR research raises ethical and privacy concerns that require attention in future studies. Firstly, the employment of generative AI in SR interviews involves communicating students' sensitive data, including grades or personal information. Secondly, conversational virtual agents are trained on specific data, leading to the possibility of biased and discriminatory responses when posing SR questions. Therefore, SR researchers must utilize generative AI responsibly and ethically (Mhlanga, 2023).

Conclusions

This study reviews 257 empirical articles on using SR in education research from 2012 to 2022. The paper examines the changes and

adaptations of the SR method in the present educational landscape, where virtual and online spaces are prevalent, and technological tools are increasingly involved in the teaching and learning process.

The revealed that researchers frequently employed SR to investigate participants' internal viewpoints and thoughts, improving their metacognitive abilities. In terms of stimuli selection and processing, the commonly employed video stimuli undergo continuous advancements. Moreover, the sources of stimuli are becoming diverse, with the inclusion of physiological feedback data. Additionally, providing participants with space to respond to interview questions is crucial. Researchers should ensure the discussion does not deviate from the research questions and avoid influencing participants' thoughts.

Furthermore, using technologies such as generative AI can enhance the reliability and generalizability of SR, and the study proposes suggestions for future research in result enhancement, stimuli optimization, and interview implementation. This study provides theoretical supplementation to manifesting the Retro-cue Effect in educational settings. It strengthens the Cognitive Theory of Multimedia Learning (CTML) with specific pedagogical strategies by combining it with SR. From a practical perspective, the current research synthesizes current findings and can serve as a valuable reference for educators and researchers in this field.

As with any systematic review, the current research has limitations inherent to the selection and filtering process. Primarily, the sample size is restricted to articles available through the Web of Science, IEEE, and Scopus databases. There might be relevant and high-quality studies published outside these three databases and are worthy studying. We hope future researchers can build on our research and offer a more comprehensive review of the use of SR in education.

In addition, education has now entered the era of Metaverse and artificial intelligence (Wang et al., 2022). How can instructors effectively apply SR in 3D virtual learning environments and in learning setting empowered by AI and AIGC (AI-generated content) remains a new territory for our continued research.

Data availability

All data generated or analyzed during this study are included in this published article.

Received: 1 June 2023; Accepted: 25 March 2024;

Published online: 05 April 2024

References

- Agricola BT, Prins FJ, van der Schaaf MF et al. (2021) Supervisor and student perspectives on undergraduate thesis supervision in higher education. *Scand J Educ Res* 65(5):877–897
- Al Mamun MA, Lawrie G, Wright T (2020) Instructional design of scaffolded online learning modules for self-directed and inquiry-based learning environments. *Comput Educ* 144:103695
- Amador JM, Bragelman J, Superfine AC (2021) Prospective teachers' noticing: a literature review of methodological approaches to support and analyze noticing. *Teach Teach Educ* 99:103256
- Belvis E, Pineda P, Armengol C et al. (2013) Evaluation of reflective practice in teacher education. *Eur J Teach Educ* 36(3):279–292
- Bloom BS (1953) Thought-processes in lectures and discussions. *J Gen Educ* 7(3):160–169
- Bogard T, Liu M, Chiang YHV (2013) Thresholds of knowledge development in complex problem solving: a multiple-case study of advanced learners' cognitive processes. *Educ Tech Res Dev* 61(3):465–503
- Bowles MA (2018) Introspective verbal reports: think-alouds and stimulated recall. In: Phakiti A, De Costa P, Plonsky L et al. (eds.) *The Palgrave handbook of*

- applied linguistics research methodology. Palgrave Macmillan, London, p 423-457
- Bozkurt A (2023) Generative artificial intelligence (AI) powered conversational educational agents: the inevitable paradigm shift. *Asian J Distance Educ* 18(1):198–204
- Calderhead J (1981) Stimulated recall: a method for research on teaching. *Br J Educ Psychol* 51(2):211–217
- Cao Z, Yu S, Huang J (2019) A qualitative inquiry into undergraduates' learning from giving and receiving peer feedback in L2 writing: Insights from a case study. *Stud Educ Eval* 63:102–112
- Chan KKH, Yung BHW (2015) On-site pedagogical content knowledge development. *Int J Sci Educ* 37(8):1246–1278
- Chan KKH, Xu L, Cooper R et al. (2021) Teacher noticing in science education: do you see what I see? *Stud Sci Educ* 57(1):1–44
- Chu X, Zhai X (2023) A Systematic Review of Stimulated Recall (SR) in Education from 2012 to 2021. In: Shih JL et al. (eds.), Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023) (pp. 100-108). China: Beijing Normal University. <https://aic-fe.bnu.edu.cn/fj/lunwenji2023/EnglishPaper.pdf>
- Consuegra E, Engels N, Willeghem V (2016) Using video-stimulated recall to investigate teacher awareness of explicit and implicit gendered thoughts on classroom interactions. *Teach Teach* 22(6):683–699
- D'Oca G, Hrynaskiewicz I (2015) Palgrave Communications' commitment to promoting transparency and reproducibility in research. *Palgr Commun*. 1(1):1–3
- Deng L (2020) Laptops and mobile phones at self-study time: examining the mechanism behind interruption and multitasking. *Australas J Educ Technol* 36(1):55–67
- Dewey J (1933) *How we think: a restatement of the relation of reflective thinking to the educative process*. DC Heath
- Dinçer S, Doğanay A (2017) The effects of multiple-pedagogical agents on learners' academic success, motivation, and cognitive load. *Comput Educ* 111:74–100
- dos Santos RAT (2018) Ways of using musical knowledge to think about one's piano repertoire learning: three case studies. *Music Educ Res* 20(4):427–445
- Dos Santos RAT, Hentschke L (2011) Praxis and poesis in piano repertoire preparation. *Music Educ Res* 13(3):273–292
- dos Santos S, Loveridge J (2014) Using video to promote early childhood teachers' thinking and reflection. *Teach Teach Educ* 41:42–51
- Duarte A, Hearons P, Jiang Y et al. (2013) Retrospective attention enhances visual working memory in the young but not the old: an ERP study. *Psychophysiology* 50(5):465–476
- Duo S, Song LX (2012) An e-learning system based on affective computing. *Phys Procedia* 24:1893–1898
- Egi T (2008) Investigating stimulated recall as a cognitive measure: Reactivity and verbal reports in SLA research methodology. *Lang Aware* 17(3):212–228
- El E, Windatt S (2019) Eye tracking analysis of EAP Students' regions of interest in computer-based feedback on grammar, usage, mechanics, style and organization and development. *System* 83:36–49
- Endacott JL (2016) Using video-stimulated recall to enhance preservice-teacher reflection. *N Educ* 12(1):28–47
- Farahian M (2015) Assessing EFL learners' writing metacognitive awareness. *J Lang Linguist Stud* 11(2):39–51
- Fernandez CJ (2018) Behind a spoken performance: test takers' strategic reactions in a simulated part 3 of the IELTS speaking test. *Lang Test Asia* 8(1):18
- Flavell JH (1979) Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. *Am Psychol* 34(10):906
- Gass SM, Mackey A (2000) *Stimulated recall methodology in second language research*. Routledge, New York
- Gass SM, Mackey A (2016) *Stimulated recall methodology in applied linguistics and L2 research*. Routledge, New York
- Gazdag E, Nagy K, Szivák J (2019) "I Spy with My Little Eyes..." The use of video stimulated recall methodology in teacher training—The exploration of aims, goals and methodological characteristics of VSR methodology through systematic literature review. *Int J Educ Res* 95:60–75
- Geiger V, Muir T, Lamb J (2016) Video-stimulated recall as a catalyst for teacher professional learning. *J Math Teach Educ* 19:457–475
- Gijsselaers HJ, Kirschner PA, Verboon P et al. (2016) Sedentary behavior and not physical activity predicts study progress in distance education. *Learn Individ Differ* 49:224–229
- Goldberg P, Sümer Ö, Stürmer K et al. (2021) Attentive or not? Toward a machine learning approach to assessing students' visible engagement in classroom instruction. *Educ Psychol Rev* 33:27–49
- Golhasany H, Harvey B (2023) Capacity development for knowledge mobilization: a scoping review of the concepts and practices. *Hum Soc Sci Commun* 10(1):1–12
- Gozalo-Brizuela R, Garrido-Merchan EC (2023) ChatGPT is not all you need. A State-of-the-Art Review of large Generative AI models. arXiv. <https://arxiv.org/abs/2301.04655>
- Harvey W, Wilkinson S, Pressé C et al. (2014) Children say the darndest things: physical activity and children with attention-deficit hyperactivity disorder. *Phys Educ Sport Pedag* 19(2):205–220
- Heikonen L, Toom A, Pyhältö K et al. (2017) Student-teachers' strategies in classroom interaction in the context of the teaching practicum. *J Educ Teach* 43(5):534–549
- Hu J, Gao X (2020) Appropriation of resources by bilingual students for self-regulated learning of science. *Int J Biling Educ Biling* 23(5):567–583
- Hu J, Wu P (2020) Understanding English language learning in tertiary English-medium instruction contexts in China. *System* 93:102305
- Jackson DO, Cho M (2018) Language teacher noticing: a socio-cognitive window on classroom realities. *Lang Teach Res* 22(1):29–46
- Jacobson MJ, Wilensky U (2006) Complex systems in education: scientific and educational importance and implications for the learning sciences. *J Learn Sci* 15(1):11–34
- Jacobson MJ, Levin JA, Kapur M (2019) Education as a complex system: conceptual and methodological implications. *Educ Res* 48(2):112–119
- Jensen A (2019) Fostering preservice teacher agency in 21st century writing instruction. *Engl Teach Pract Crit* 18(3):298–311
- Kang YS, Pyun DO (2013) Mediation strategies in L2 writing processes: a case study of two Korean language learners. *Lang Cult Curric* 26(1):52–67
- Karimi MN, Norouzi M (2017) Scaffolding teacher cognition: changes in novice L2 teachers' pedagogical knowledge base through expert mentoring initiatives. *System* 65:38–48
- Kautzmann TR, Jaques PA (2019) Effects of adaptive training on metacognitive knowledge monitoring ability in computer-based learning. *Comput Educ* 129:92–105
- Keith MJ (1988) Stimulated recall and teachers' thought processes: a critical review of the methodology and an alternative perspective. Paper presented at the 17th Annual Meeting of the Mid-South Educational Research Association, Louisville, KY, 9–11
- Koltovskaia S (2020) Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: a multiple case study. *Assess* 44:100450
- Kurki K, Järvenoja H, Järvelä S et al. (2016) How teachers co-regulate children's emotions and behaviour in socio-emotionally challenging situations in day-care settings. *Int J Educ Res* 76:76–88
- Lai ML, Tsai MJ, Yang FY et al. (2013) A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educ Res Rev* 10:90–115
- Lambert C, Zhang G (2019) Engagement in the use of English and Chinese as foreign languages: the role of learner-generated content in instructional task design. *Mod Lang J* 103(2):391–411
- Lee C (2020) A study of adolescent English learners' cognitive engagement in writing while using an automated content feedback system. *Comput Assist Lang. Learn* 33(1-2):26–57
- Lichtinger E, Kaplan A (2015) Employing a case study approach to capture motivation and self-regulation of young students with learning disabilities in authentic educational contexts. *Metacogn Learn* 10(1):119–149
- Lin CH, Zhang Y, Zheng B (2017) The roles of learning strategies and motivation in online language learning: a structural equation modeling analysis. *Comput Educ* 113:75–85
- Lindfors M, Bodin M, Simon S (2020) Unpacking students' epistemic cognition in a physics problem-solving environment. *J Res Sci Teach* 57(5):695–732
- Łuczniak K, May J (2021) Measuring individual and group flow in collaborative improvisational dance. *Think Skills Creat* 40:100847
- Lyle J (2003) Stimulated recall: a report on its use in naturalistic research. *Br Educ Res J* 29(6):861–878
- Maaranen K, Kynäslähti H, Byman R et al. (2019) Teacher education matters: Finnish teacher educators' concerns, beliefs, and values. *Eur J Teach Educ* 42(2):211–227
- Määttä E, Mykkänen A, Järvelä S (2016) Elementary schoolchildren's self- and social perceptions of success. *J Res Child Educ* 30(2):170–184
- Martinelle R (2018) Video-stimulated recall: aiding teacher practice. *Educ Leadersh* 76(3):55
- Martinelle R (2020) Using video-stimulated recall to understand the reflections of ambitious social studies teachers. *J Soc Stud Res* 44(3):307–322
- Mayer RE, Moreno R (1998) A cognitive theory of multimedia learning: Implications for design principles. *J Educ Psychol* 91(2):358–368
- Mayer RE, Moreno R (2003) Nine ways to reduce cognitive load in multimedia learning. *Educ Psychol* 38(1):43–52
- Mayer RE (2002) Multimedia learning. In: Brian HR (ed.) *Psychology of learning and motivation*, vol 41. Academic Press, Cambridge, p 85-139
- Meade P, McMeniman M (1992) Stimulated recall—An effective methodology for examining successful teaching in science. *Aust Educ Res* 19(3):1–18
- Mhlanga D (2023) Open AI in Education, the Responsible and Ethical Use of ChatGPT Towards Lifelong Learning. In: Mhlanga D (ed.) *FinTech and Artificial Intelligence for Sustainable Development*. Sustainable Development Goals Series. Palgrave Macmillan, Cham, pp 387–409

- Michel M, Révész A, Lu X et al. (2020) Investigating L2 writing processes across independent and integrated tasks: a mixed-methods study. *Second Lang Res* 36(3):307–334
- Moher D, Liberati A, Tetzlaff J et al. (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 151(4):264–269
- Molenaar I, Knoop-van Campen CA (2018) How teachers make dashboard information actionable. *IEEE Trans Learn Technol* 12(3):347–355
- Moreno R, Mayer RE (1999) Cognitive principles of multimedia learning: the role of modality and contiguity. *J Educ Psychol* 91(2):358
- Myrttil MJ, Lin TJ, Chen J et al. (2021) Pros and (con) flict: using head-mounted cameras to identify teachers' roles in intervening in conflict among preschool children. *Early Child Res Q* 55:230–241
- Nguyen NT, McFadden A, Tangen D et al. (2013) Video-stimulated recall interviews in qualitative research. Paper presented at the Annual Meeting of the Australian Association for Research in Education (AARE), Adelaide, Australia, 1-5 December 2013
- Nurmukhamedov U, Kim SH (2010) 'Would you perhaps consider...': hedged comments in ESL writing. *ELT J* 64(3):272–282
- Nyberg G, Larsson H (2017) Physical education teachers' content knowledge of movement capability. *J Teach Phys Educ* 36(1):61–69
- Obersteiner A, Tumpek C (2016) Measuring fraction comparison strategies with eye-tracking. *ZDM* 48(3):255–266
- Page MJ, McKenzie JE, Bossuyt PM et al. (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Int J Surg* 88:105906
- Park M (2018) Innovative assessment of aviation English in a virtual world: windows into cognitive and metacognitive strategies. *ReCALL* 30(2):196–213
- Peng JE (2020) Teacher interaction strategies and situated willingness to communicate. *ELT J* 74(3):307–317
- Perkins R, Aufegger L, Williamon A (2015) Learning through teaching: exploring what conservatoire students learn from teaching beginner older adults. *Int J Music Educ* 33(1):80–90
- Pratt SM, Martin AM (2017) The differential impact of video-stimulated recall and concurrent questioning methods on beginning readers' verbalization about self-monitoring during oral reading. *Read Psychol* 38(5):439–485
- Qadir J (2022) Engineering education in the era of ChatGPT: promise and pitfalls of generative AI for education. In: 2023 IEEE Global Engineering Education Conference (EDUCON), Kuwait, Kuwait, 1-4 May 2023
- Qiu X (2020) Functions of oral monologic tasks: effects of topic familiarity on L2 speaking performance. *Lang Teach Res* 24(6):745–764
- Qiu X, Lo YY (2017) Content familiarity task repetition and Chinese EFL learners' engagement in second language use. *Lang Teach Res* 21(6):681–698
- Ramadhanti D, Yanda DP (2021) Students' metacognitive awareness and its impact on writing skill. *Int J Lang Educ* 5(3):193–206
- Ramakrishnan A, Zylich B, Ottmar E et al. (2023) Toward automated classroom observation: multimodal machine learning to estimate class positive climate and negative climate. *IEEE Trans Affect Comput* 14(1):664–679
- Ramnarain UD, Modiba M (2013) Critical friendship collaboration and trust as a basis for self-determined professional development: a case of science teaching. *Int J Sci Educ* 35(1):65–85
- Rankanen M, Leinikka M, Groth C et al. (2022) Physiological measurements and emotional experiences of drawing and clay forming. *Arts Psychother* 79:101899
- Rassaei E (2013) Corrective feedback, learners' perceptions, and second language development. *System* 41(2):472–483
- Rassaei E (2015) Journal writing as a means of enhancing EFL learners' awareness and effectiveness of recasts. *Linguist Educ* 32:118–130
- Rassaei E (2020) The separate and combined effects of recasts and textual enhancement as two focus on form techniques on L2 development. *System* 89:102193
- Révész A, Kourtali NE, Mazgutova D (2017) Effects of task complexity on L2 writing behaviors and linguistic complexity. *Lang Learn* 67(1):208–241
- Rietdijk S, van Weijen D, Janssen T et al. (2018) Teaching writing in primary education: classroom practice, time, teachers' beliefs and skills. *J Educ Psychol* 110(5):640
- Rissanen I, Kuusisto E, Hanhimäki E et al. (2018) The implications of teachers' implicit theories for moral education: a case study from Finland. *J Moral Educ* 47(1):63–77
- Rissanen I, Kuusisto E, Tuominen M et al. (2019) In search of a growth mindset pedagogy: a case study of one teacher's classroom practices in a Finnish elementary school. *Teach Teach Educ* 77:204–213
- Sanchez HS, Grimshaw T (2019) Stimulated recall. In: McKinley J, Rose H (eds.) *The Routledge handbook of research methods in applied linguistics*. Routledge, Abingdon, pp 312–323
- Sato R (2019) Fluctuations in an EFL teacher's willingness to communicate in an English-medium lesson: an observational case study in Japan. *Innov Lang Learn Teach* 13(2):105–117
- Schindler M, Lilienthal AJ (2019) Domain-specific interpretation of eye tracking data: towards a refined use of the eye-mind hypothesis for the field of geometry. *Educ Stud Math* 101(1):123–139
- Schindler M, Lilienthal AJ (2020) Students' creative process in mathematics: insights from eye-tracking-stimulated recall interview on students' work on multiple solution tasks. *Int J Sci Math Educ* 18(8):1565–1586
- Schneider D, Barth A, Getzmann S et al. (2017) On the neural mechanisms underlying the protective function of retroactive cuing against perceptual interference: evidence by event-related potentials of the EEG. *Biol Psychol* 124:47–56
- Shepherdson P, Oberauer K, Souza AS (2018) Working memory load and the retro-cue effect: a diffusion model account. *J Exp Psychol Hum Percept Perform* 44(2):286
- Shintani N (2016) The effects of computer-mediated synchronous and asynchronous direct corrective feedback on writing: a case study. *Comput Assist Lang Learn* 29(3):517–538
- Smagorinsky P (1998) Thinking and speech and protocol analysis. *Mind Cult Act* 5(3):157–177
- Smithers LG, Sawyer AC, Chittleborough CR et al. (2018) A systematic review and meta-analysis of effects of early life non-cognitive skills on academic, psychosocial, cognitive and health outcomes. *Nat Hum Behav* 2(11):867–880
- Souza AS, Oberauer K (2016) In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Atten Percept Psycho* 78:1839–1860
- Stolpe K, Björklund L (2013) Students' long-term memories from an ecology field excursion: retelling a narrative as an interplay between implicit and explicit memories. *Scand J Educ Res* 57(3):277–291
- Stough LM (2001) Using Stimulated Recall in Classroom Observation and Professional Development. Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA, 10-14 April 2001
- Sundberg B, Areljung S, Due K et al. (2018) Opportunities for and obstacles to science in preschools: views from a community perspective. *Int J Sci Educ* 40(17):2061–2077
- Tan ST, Tan CX, Tan SS (2021) Physical activity, sedentary behavior, and weight status of university students during the covid-19 lockdown: a cross-national comparative study. *Int J Environ Res Public Health* 18(13):7125
- Thararuedee P, Wette R (2020) Attending to learners' affective needs: teachers' knowledge and practices. *System* 95:102375
- Tiainen O, Korkeamäki RL, Dreher MJ (2018) Becoming reflective practitioners: a case study of three beginning pre-service teachers. *Scand J Educ Res* 62(4):586–600
- Ucan S, Webb M (2015) Social regulation of learning during collaborative inquiry learning in science: how does it emerge and what are its functions? *Int J Sci Educ* 37(15):2503–2532
- Van der Kleij F, Adie L, Cumming J (2017) Using video technology to enable student voice in assessment feedback. *Br J Educ Technol* 48(5):1092–1105
- Walan S, Enochsson AB (2019) The potential of using a combination of storytelling and drama, when teaching young children science. *Eur Early Child Educ Res J* 27(6):821–836
- Wang M, Yu H, Bell Z et al. (2022) Constructing an Edu-Metaverse ecosystem: a New and innovative framework. *IEEE Trans Learn Technol* 15(6):685–696
- Wass R, Moskal ACM (2017) What can Interpersonal Process Recall (IPR) offer academic development? *Int J Acad Dev* 22(4):293–306
- White C, Drenzo R, Bortolotto C (2016) The learner-context interface: emergent issues of affect and identity in technology-mediated language learning spaces. *System* 62:3–14
- Wijayasundara M (2020) Integration of ICT in teaching and learning in schools. *Int J Res* 1(10):198–209
- Wilby KJ, Govaerts MJ, Austin Z et al. (2017) Exploring the influence of cultural orientations on assessment of communication behaviours during patient-practitioner interactions. *BMC Med Educ* 17(1):1–10
- Wu SC (2019) Researching children's learning and play in a Chinese context: children's perspectives on their play and learning. *Eur Early Child Educ Res J* 27(4):551–565
- Yu S (2021) Giving genre-based peer feedback in academic writing: sources of knowledge and skills, difficulties and challenges. *Assess Eval High Educ* 46(1):36–53
- Yu S, Hu G (2017) Understanding university students' peer feedback practices in EFL writing: Insights from a case study. *Assess Writ* 33:25–35
- Zahid M, Khanam A (2019) Effect of reflective teaching practices on the performance of prospective teachers. *Turk Online J Educ Technol* 18(1):32–43
- Zhai X, Fang Q, Dong Y et al. (2018) The effects of biofeedback-based stimulated recall on self-regulated online learning: a gender and cognitive taxonomy perspective. *J Comput Assist Learn* 34(6):775–786
- Zhai X, Chu X, Meng N et al. (2022) The effect of multi-mode stimuli of feed-forward and eye tracking on metacognition—an exploratory study using digital dictionaries. *Educ Technol Soc* 25(1):213–227
- Zhai X, Xu J, Chen NS et al. (2022) The Syncretic Effect of Dual-Source Data on Affective Computing in Online Learning Contexts: A Perspective From

Convolutional Neural Network With Attention Mechanism. *J Educ Comput Res* 61(2):466–493

Acknowledgements

This research is funded by the National Science and Technology Major Project (Grant No. 2022ZD0115904), 2021 National Natural Science Foundation of China (Grant No. 62177042), and 2024 Zhejiang Provincial Natural Science Foundation of China (Grant No. Y24F020039).

Author contributions

All authors have contributed to material preparation and data analysis. Conceptualization, design and data collection: XS Zhai, and XY Chu. Methodology: All. Original draft: XS Zhai and XY Chu. Second draft: MJ Wang. Third draft: CC Tsai and JC Liang. Final round revision and quality check: JM Spector. All authors discussed the results and reviewed, edited, and approved the final version of the manuscript.

Competing interests

The author declares no competing interests.

Ethical approval

Ethical approval was not required as the study did not involve human participants.

Informed consent

This article does not contain any studies with human participants performed by any of the authors.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1057/s41599-024-02987-6>.

Correspondence and requests for materials should be addressed to Xiaoyan Chu or Minjuan Wang.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024