Received: 26 October 2022 | Revised: 28 June 2023 | Accepted: 10 July 2023

DOI: 10.1002/pits.23035

RESEARCH ARTICLE

WILEY

A multilevel framework for assessing self-regulated learning in school contexts: Innovations, challenges, and future directions

Timothy J. Cleary | **Michelle R. Russo**

Department of School Psychology, Rutgers, The State University of New Jersey, New Brunswick, New Jersey, USA

Correspondence
Timothy J. Cleary, Department of School Psychology, Rutgers, The State University of New Jersey, 152 Frelinghuysen Rd, Piscataway, NJ 08854, USA.
Email: timothy.cleary@rutgers.edu

Abstract

As the importance of student self-regulated learning (SRL) among school-aged populations continues to increase, there is an emerging need for school psychologists and other school professionals to develop the knowledge and skills to effectively assess this process. The current paper provides a conceptual and empirical overview of a multi-method SRL diagnostic assessment framework included as part of the Self-Regulation Empowerment Program (SREP). In addition to providing information about the operation and sequence of the SREP diagnostic assessment process, the current paper underscores the commonly used measures of SRL (e.g., self-report questionnaires, teacher rating scales, SRL microanalysis, think-aloud protocols) and provides evidence supporting their psychometric rigor and utility in school settings. This paper also serves as a roadmap that school psychologists can use to increase the frequency and quality of school-based SRL assessments. Future research directions regarding the SREP diagnostic assessment process are also emphasized.

KEYWORDS
academic, assessment, self-regulated learning

1 | INTRODUCTION

Self-regulated learning (SRL) can be defined as a goal-directed process through which individuals control or manage their behavior, thinking, or emotions in a purposeful and strategic way to optimize functioning (Greene, 2018). Most researchers and educators recognize SRL as an essential 21st-century skill given its relevance and applicability across academic, athletic, mental health, and physical health domains (Reddy et al., 2018; Schunk & Greene, 2018; Vohs & Baumeister, 2016). Specifically in academic contexts, SRL processes (e.g., goal setting, planning, strategy use, self-evaluation) have been linked with numerous academic and performance outcomes and have served as key components of empirically supported academic interventions (Dignath et al., 2008; Graham et al., 2005; Schunk & Greene, 2018). Despite increased interest in SRL, there is evidence that school psychologists and other school-based practitioners or educators do not routinely assess this process, nor do they possess the confidence or skills to develop and implement individualized intervention programs targeting student SRL (Cleary et al., 2010; Kramarski et al., 2013; Peters-Burton et al., 2022). A key objective of this paper is to provide an overview of an expanded and updated SRL assessment-to-intervention framework, Self-Regulation Empowerment Program (SREP), that practitioners can use as part of their professional assessment activities (Cleary & Zimmerman, 2004).

Assessment represents a fundamental and essential role of school psychologists (Dombrowski, 2020). Using comprehensive, multidimensional evaluations as part of a traditional refer-test-place framework, school psychologists seek to identify the underlying causes of student academic difficulties, determine student eligibility for special education services, and develop hypotheses regarding effective instructional programming (Erchul & Martens, 2010). With the emergence of the multitiered systems of support (MTSS) framework and other related initiatives (e.g., data-based decision-making, progress monitoring), school psychologists have increasingly turned to contextualized or situation-specific measures, such as curriculum-based measures (CBMs) and functional behavioral assessments (FBAs; Keller-Margulis et al., 2008; Wixson & Valencia, 2011). There has also been increased interest in assessing skills other than cognitive abilities, academic skills, and social–emotional skills, such as student motivation and SRL (Cleary et al., 2010; Osher et al., 2016).

A couple of decades ago, Cleary and Zimmerman (2004) developed SREP, an integrative assessment-to-intervention framework consisting of a four-level, multidimensional SRL assessment process (i.e., multiple sources, multiple methods, and multiple levels of SRL) and comprehensive SRL intervention program. The intervention component is designed to enhance students' awareness of their academic struggles while concurrently improving their strategic, metacognitive, and motivational skills to overcome these difficulties. Research shows promising results regarding the effects of the SREP intervention on student strategic, motivational, and academic functioning (Cleary et al., 2017). In terms of the SREP assessment process, the primary objective is to address a highly contextualized and situation-specific regulatory question: "To what extent does *this* student have knowledge of, select, and regulate the use of *these* specific study and self-regulation strategies to enhance his or her performance on *these* performance outcomes in *that* particular class?" (Cleary & Zimmerman, 2004, p. 541). The basic logic in addressing this question is that through gathering data about student SRL in the context of specific coursework and learning activities, psychologists and/or members of a child study team can generate more effective hypotheses about the regulatory determinants of students' academic struggles and customize the SREP intervention to optimally address their needs.

In the original SREP diagnostic assessment framework, Cleary and Zimmerman (2004) provided a general overview of the broad categories of SRL assessments used in the SREP diagnostic process. However, they did not provide specific information about the sequence of the assessment process, nor did they convey data regarding the utility and psychometric quality of the most common and relevant SRL assessment approaches. Over the past couple of decades, there has been substantial and rapid growth in the overall quality and innovativeness of SRL assessments, specifically in schools (Cleary et al., 2021; Schunk & Greene, 2018; Tise et al., 2019). Given that greater specificity in the SRL assessment process can enhance school-based practitioners' ability to link SRL assessment data to intervention planning, the current paper expands and updates the original SREP diagnostic

model in several ways. In addition to providing more detailed information about the operation and sequence of the SREP assessment process, the current paper provides evidence supporting the utility and psychometric rigor of specific tools emphasized within the SREP diagnostic framework (e.g., self-report questionnaires, teacher rating scales, SRL microanalysis, think-aloud protocols; Cleary et al., 2021; Tise et al., 2019). This paper also serves as a roadmap that school psychologists can use to increase the frequency and quality of school-based SRL assessments.

2 | OVERVIEW OF SRL

Self-regulation (SR) is a foundational concept underlying human functioning that researchers have studied across multiple fields and disciplines, including academic, behavioral, mental health, physical health, athletics, and music (Briesch & Chafouleas, 2009; McPherson et al., 2017; Schunk & Greene, 2018; Vohs & Baumeister, 2016). Many researchers use the term SR to describe individuals' overall capacity for self-control or their ability to resist impulses and temptations (e.g., management and control over health, weight, addictions, etc.; Greene, 2018; Vohs & Baumeister, 2016). However, SRL is the construct of interest in this paper given its association to human learning and performance within academic settings as well as its direct link to the SREP diagnostic assessment process.

Broadly speaking, SRL represents a contextualized process involving self-generated thoughts and actions that are planned and cyclically adapted to attain personal goals (Panadero, 2017; Zimmerman, 2000). SRL processes, which include goal setting, motivational beliefs, planning, monitoring, reflection, and strategic action, are malleable and influenced by contextual factors, such as teacher feedback, instructional quality, and classroom supports or intervention programs. These processes are naturally embedded and organized within an integrated *cyclical feedback loop*; that is, a process or mechanism through which individuals set goals and plans, use strategies to learn, and gather information and feedback to evaluate goal progress (Greene, 2018; Panadero, 2017). Thus, most SRL theorists recognize a natural temporal sequencing among regulatory processes (i.e., before, during, and after), and use well-defined terms to identify and describe these processes (e.g., monitoring, self-evaluation, or reflection; Greene, 2018).

The SREP diagnostic assessment process draws directly from Zimmerman's (2000) three-phase social-cognitive model of SRL. From this perspective, SRL is a goal-directed process that involves three interdependent and sequential phases: forethought, performance control, and self-reflection. *Forethought phase* processes set the stage for learning and include task analysis, goal setting, strategic planning as well as motivational beliefs (e.g., self-efficacy, goal orientation, and task values). When students display strong forethought, such as by establishing clear goals, developing effective strategic plans, and expressing a desire to perform well on academic activities, they are more likely to engage in those activities strategically and intentionally (Callan & Cleary, 2018, 2019).

During task engagement, or the *performance control phase*, self-regulated learners deploy self-control and self-observation tactics. Self-control processes, which include learning strategies (e.g., writing strategies, problem-solving heuristics, concepts maps) and regulatory strategies (e.g., attention control, organization, help-seeking, self-instruction), help students optimize their focus and ability to effectively manage their learning, thoughts, emotions, and actions. Self-observation is another performance control phase process that involves tracking or monitoring thoughts, emotions, and actions or performance, as well as the conditions or contextual factors surrounding performance (Zimmerman, 2000). The use of learning and regulatory strategies as one concurrently makes observations and gathers performance feedback sets the stage for the *self-reflection phase*. During self-reflection, students make judgments about their performance quality and/or goal progress (i.e., self-evaluation), identify potential reasons for their performance (i.e., attributions). The self-reflection process also emphasizes individuals' affective reactions (e.g., satisfaction) following performance and the types of conclusions they make about the necessity and utility of adapting or changing their strategic approaches before future task engagement (i.e., adaptive inferences). Collectively, these self-reflection phase processes influence students' subsequent forethought processes (i.e., strategic plans, self-evaluation) during future learning attempts. A core objective of the SREP

diagnostic assessment process is to gather information about the quality and sophistication with which students display regulatory and strategic skills relative to school-related activities.

3 | OVERVIEW OF SRL ASSESSMENT PRACTICES

To understand the SREP diagnostic assessment framework, it is important to first discuss a couple of important and overlapping issues. The SRL construct can be described at both a *macrolevel* and *microlevel* (Efklides, 2011; Greene, 2018; Panadero, 2017). At the macrolevel, SRL assessments focus on students' broad abilities, or their general or typical approach to regulating their lives. In contrast, microlevel SRL assessments examine students' thoughts or actions as they engage in a particular activity or situation. This difference between a macro- and microfocus is not trivial, as the two terms parallel two broad classes of measurement widely recognized in the SRL literature: (a) aptitude (i.e., macrolevel) and (b) event (i.e., microlevel; Callan & Cleary, 2018; Winne & Perry, 2000). *Aptitude measures* include an array of assessment approaches that generate broad information about individuals' regulatory behaviors or skills, typically in a retrospective fashion. The most common aptitude measures include questionnaires, rating scales, and certain types of interviews. When using questionnaires or rating scales, respondents use a Likert-type format (e.g., 1–5) to rate how frequently or well a set of items reflect their beliefs or behaviors. The item ratings are then aggregated (i.e., typically averaged) to create a composite score that is used for interpretation purposes.

In contrast, *event measures* generate information about students' regulatory skills and processes as they emerge or change during the completion of specific tasks and activities. Thus, event measures are context- and situation-specific assessment tools administered as students engage in a particular learning activity. Frequently used event measures include SRL microanalysis, direct observations, think-aloud protocols, and behavioral traces (Azevedo et al., 2010; Cleary et al., 2021; Greene et al., 2018). Collectively, event measures offer school-based practitioners the opportunity to generate fine-grained, real-time data about specific regulatory processes (e.g., goal setting, attributions) used during common academic activities, such as writing an essay or completing mathematics problems.

Although all event measures have a situational and task-specific focus, there are crucial differences among them. For example, some event measures rely on verbal responses and active engagement of the participant (e.g., think-aloud protocols, SRL microanalysis) while others are less obtrusive in nature and do not rely on self-reports (e.g., behavioral traces, direct observations; Schraw, 2010). Given that school-based practitioners tend to place particular emphasis on self-report methodologies within their assessment toolkit, the SREP diagnostic process places greater weight on SRL microanalysis and think-aloud protocols. *Think-aloud* techniques involve asking students to verbally express their natural thoughts as they perform a specific learning task but without providing an explanation or reasoning underlying their thoughts (Roth et al., 2016; Ward & Traweek, 1993). *SRL microanalysis* reflects an in-the-moment interview technique whereby an evaluator administers targeted questions about phase-specific SRL processes (e.g., goal setting in the forethought phase) as students learn or complete learning activities (Cleary et al., 2021). Although both types of event measures are used within the SREP diagnostic assessment process, SRL microanalysis is particularly relevant given its flexibility and capacity for targeting specific SRL processes that are directly linked to student referral concerns (e.g., focusing on forethought phase processes for students with poor impulse control and planning skills).

4 | SREP—DIAGNOSTIC ASSESSMENT PROCESS

4.1 | Multidimensional focus

The SREP diagnostic assessment is grounded in the general premise that a multidimensional assessment approach is optimal for addressing the different levels of SRL (i.e., macro and micro) as well as its various dimensions

(e.g., strategy, motivation, metacognition) and subprocesses (e.g., goal setting, attributions). Conceptually, administering *multiple methods* across *multiple sources* and *multiple contexts* is advantageous because it creates opportunities for triangulating data (i.e., identifying areas of convergence and divergence). Although we advocate using both aptitude and event SRL measures as part of the diagnostic process, we offer a couple of qualifications for using multidimensional SRL assessments. We do not consider one category of SRL measurement (i.e., aptitude or event measure) to be inherently superior or better in all situations (Callan & Cleary, 2018). The relevance of any assessment tool is determined by the types of inferences that evaluators wish to make about students' SRL skills. For example, aptitude measures are best suited to capture broad aspects of SRL, whereas event measures are structured to assess SRL processes as they emerge and shift during the completion of a particular task. Given this latter point, we also underscore that practitioners may opt to use only one category of SRL assessment depending on their specific objectives or goals for the assessment. For example, if a practitioner was only interested in examining student regulation as they write an essay for English Language Arts class, they may opt to only use event measures. These issues aside, we encourage practitioners to utilize a multidimensional assessment approach to more comprehensively and fully understand student SRL processes.

Researchers have generated data that support the use of a multimethod SRL assessment approach (i.e., use of both aptitude and event measures). For example, Cleary and Callan (2014) used multiple aptitude measures (i.e., self-report questionnaire and teacher rating scale) with a sample of middle school students. The authors found that both aptitude measures emerged as statistically significant and unique predictors of student mathematics classroom test scores, even after controlling for prior mathematics skills and motivation. Similar findings have been observed when using multiple event measures. Callan and Cleary (2019) showed that the quality of middle school students' metacognitive monitoring (i.e., as assessed with SRL microanalysis) and their strategy use (i.e., as assessed with behavioral traces) each uniquely predicted student performance on a mathematics problem-solving activity. In short, research has illustrated the additive value of using two or more SRL measures *within* a given assessment category (i.e., aptitude, event) in the prediction of student outcomes.

Importantly, there is also evidence that the combined use of aptitude and event measures can lead to robust predictions of student outcomes. Callan and Cleary (2018) included two aptitude measures (i.e., self-report, teacher ratings) and two event measures (i.e., SRL microanalysis, behavioral traces) to assess student SRL during a mathematics activity. Although all assessment approaches within each category did not emerge as significant predictors of achievement in the regression model, one aptitude (i.e., teacher ratings) and one event measure (i.e., SRL microanalysis) uniquely explained mathematics outcomes (Callan & Cleary, 2018). Thus, in this study, the combined use of an aptitude and event measure accounted for more variance in performance outcomes than either would have provided alone.

A related line of research has shown that the relations among SRL measures *within* a given assessment category are often stronger than the correlations of measures *between* categories. As an example, student self-report and teacher ratings of student SRL (i.e., both types of attitude measures) tend to be more highly related than aptitude and event measures (Callan & Cleary, 2018). Regarding the between-category relations, the evidence is complicated and somewhat mixed. In some situations, the relations between aptitude and event measures are small or negligible (DiBenedetto & Zimmerman, 2013; Winne & Jamieson-Noel, 2002), while at other times they reach statistical significance. For example, Cleary et al. (2015) showed that a microanalytic measure of strategy use correlated at a statistically significant level with one self-report measure of strategy use but not with other self-reports. Other research suggests that the nature of aptitude-event measure relations may depend on the granularity of the data used for the event measures (Cleary et al., 2015; Rovers et al., 2019). For example, Rovers et al. (2019) conducted a review of studies that examined the correlations between event measures of strategy use (e.g., behavioral traces and think-aloud data) and aptitude self-report measures. A key finding was that the correlation between self-report data and the event measures was meaningful only when the data for the event measures were coded at a more global level.

In short, both SRL aptitude and event measures can serve as unique and significant predictors of important student outcomes. They also overlap conceptually to some degree and there is strong evidence to support the validity of SRL inferences from scores generated from measures within both the aptitude and event categories (see next section; Cleary et al., 2021; Follmer & Sperling, 2016; Holland et al., 2018). Thus, it is reasonable to conclude that while event and aptitude measure appear to capture different levels or dimensions of the SRL construct, they collectively are able to provide greater coverage of SRL than the exclusive use or reliance on any single approach.

4.2 | Four level diagnostic assessment sequence

The SREP diagnostic assessment process consists of four levels that were originally labeled *Class* (Level 1), *Grading Criteria* (Level 2), *Strategy* (Level 3), and *Microanalytic* (Level 4; Cleary & Zimmerman, 2004). However, we updated these labels to enhance clarity and to more accurately align them with emerging trends and developments in the literature. The four-level diagnostic process is now labeled as *Challenge—Class* (Level 1), *Challenge—Task* (Level 2), *SRL—Aptitude* (Level 3), and *SRL—Event* (Level 4; see Table 1). As is reflected in these changes, the first two levels are designed to uncover the areas of difficulty or challenges that students experience in schools. These two levels do not target student SRL skills per se; rather, they generate information that sets the stage for the SRL assessments conducted at Levels 3 and 4. In other words, because SRL is conceptualized as a contextualized and situation-specific process (i.e., "To what extent does *this* student have knowledge of, select, and regulate the use of *these* specific study and self-regulation strategies to enhance his or her performance on *these* performance outcomes in *that* particular class?"), evaluators need to first identify the classes, content areas, or learning activities that are difficult for students before administering SRL assessments.

At Levels 3 and 4, evaluators target students' SRL skills relative to the previously identified challenging classes and/or activities (Levels 1 and 2). The descriptors for Levels 3 and 4 now include the *SRL* acronym to emphasize their collective focus on assessing student SRL. Further, the term strategy was replaced with *aptitude* (Level 3), and microanalytic with *event* (Level 4). These original terms (i.e., strategy and microanalytic) were appropriate but too narrow in scope. For example, while evaluators clearly need to assess students' use of learning or regulatory strategies at Level 3, they should also consider students' motivational beliefs (e.g., self-efficacy, mindset, etc.) or other regulatory processes. Thus, the aptitude label is more inclusive and parallels the terminology used in SRL assessment literature (Winne & Perry, 2000). For Level 4, the term event is also more inclusive and appropriate than microanalytic given that SRL microanalysis is only one of several prospective event measures that practitioners can use (e.g., think-aloud protocols, direct observations, behavioral traces). In short, the updated descriptors for Levels 3 and 4 (*SRL—Aptitude* and *SRL—Event)* are more effective because they clearly convey the focus on SRL assessments, are more consistent with terminology used in the literature, and are sufficiently broad to represent a wide range of assessment tools.

4.3 | Key features of challenge levels

The overarching goal for Levels 1 and 2 of the SREP diagnostic process is to accurately identify key areas of challenge that middle or high school students experience at the class (i.e., Level 1, overall performance within a class) and task levels (i.e., Level 2, performance within the class on assignments, tests, projects, etc.).

4.3.1 | Challenge—Class (Level 1)

The SREP diagnostic process begins by addressing a basic Level 1 (Challenge—Class) question: "In which classes is the student struggling to perform at an adequate level?" To examine this issue, evaluators can either

TABLE 1 Four levels of the SRL diagnostic assessment process.

Level of assessment	Assessment question(s)	Assessment tools
Challenge—Class level	*At the school level…* …in which class(es) does the student struggle? …is there convergence among sources regarding the classes identified as challenging?	Record review Teacher/parent/student interviews Informal questionnaire
Challenge—Task level	*Within an identified challenging class…* …on which tasks/activities/assignments does the student struggle the most? …is there convergence among sources regarding the challenging tasks/activities/assignments?	Record review Teacher/parent/student interviews Informal questionnaire
SRL—Aptitude level	*For a given class and/or each challenging task/activity/assignment…* …which *types* of learning or regulatory strategies does the student typically use? …how *frequently* does the student typically use the learning or regulatory strategies? …what is the student's level of motivation in those classes or when completing those tasks?	Self-report questionnaires Rating scales (e.g., teacher, parent) Situation-specific interviews
SRL—Event level	*As the student engages in the challenging task/activity/assignment…* …which strategies or tactics does the student use? …what is the nature and quality of the student's forethought, performance, or reflection processes?	SRL microanalysis Think-aloud protocols Behavioral traces Direct observations

Abbreviation: SRL, self-regulated learning.

review report card grades or other student records, engage in informal interviews, and/or use a brief questionnaire. An SREP Diagnostic Worksheet and corresponding three-point rubric (i.e., high, adequate, and low/struggling) is used for both Levesl 1 and 2 assessments. In the context of an informal interview, evaluators use this worksheet (see Table 2) to summarize the perspectives of each prospective respondent (i.e., teachers, parents, and/or students) about student performance. Respondents are also prompted to use a three-point rubric to enhance consistency in how the raters understand the meaning of diverse levels of student performance (see Table 3).

During these initial steps of the assessment process, evaluators may observe poor convergence among respondents (i.e., teachers, parents, and students) regarding their personal perspectives about student performance. Poor convergence could be the result of rater effects (e.g., parental responses may be influenced by social desirability), method effects (e.g., questionnaires and interviews may lead to discrepant findings), child effects (e.g., young children may not be able to accurately report behaviors), or a combination of these factors (Smith, 2007). Specifically regarding respondent perceptions about student performance, a lack of convergence may emerge due to differences in performance standards among the informants. For example, a teacher and parent may convey that a student earning a grade of C+ in mathematics class is underperforming (i.e., perceived minimum standard of B grade or higher) while the student feels completely satisfied with that grade (i.e., perceived minimum standard of C). When a lack of convergence among respondents emerges regarding student performance, evaluators can probe each respondent for additional information about the standards of success they use. Broadly speaking, making sense of discrepant assessment data from multiple informants is a complicated activity that often requires the use of a systematic process and clinical judgment (see Smith, 2007 and De Los Reyes et al., 2015, for an extended discussion of this topic).

TABLE 2 Challenge rubric for Levels 1 and 2 of the SREP diagnostic assessment process.

Low/struggling = report card grades or overall performance that are perceived to be below expected standards or lower than peers; follow-up assessment is needed or desired	Acceptable = report card grades or overall performance perceived as appropriate; some minor concerns may be expressed about this class	High = report card grades or overall performance is perceived to be at expected standards or is stronger than most other students in the same context; no concerns expressed

Abbreviation: SREP, Self-Regulation Empowerment Program.

4.3.2 | Challenge—Task (Level 2)

At the Challenge—Task level, evaluators focus on a more nuanced question: "In the classes identified as challenging, what are the key learning tasks, activities, or assignments that are most challenging for students?" Given that SRL represents a dynamic, task-specific process, evaluators should seek to identify learning activities (e.g., test/quiz preparation, homework completion, reading or written assignments, group projects) that are most heavily weighted or used by teachers in the target classes. As conveyed in the SREP diagnostic worksheet (see Table 2), each respondent (i.e., teacher, parent, or student) uses the same three-point rubric to identify their perceptions about the learning activities of greatest challenge for the target student.

4.4 | Empirical support for challenge levels

Although Levels 1 and 2 aspects of the diagnostic process have been used in clinical contexts and as part of the SREP intervention program (Cleary et al., 2017), there has not been any formal or systematic attempt to empirically evaluate the procedures used at these levels. This lack of research is not entirely surprising given that Levels 1 and 2 assessments are more informal and descriptive in nature. However, there are several fruitful lines of research that should be explored and considered (see Section 5).

4.5 | Key features of SRL levels

Over the past couple of decades, SRL assessment literature has expanded in several important ways. In this section, we provide an overview of commonly used SRL aptitude and event measures and describe the empirical foundation underlying these approaches.

4.5.1 | SRL—Aptitude level (Level 3)

At the SRL—Aptitude level, evaluators shift their attention to an SRL-focused question: "What types of learning, regulatory, or motivational processes and strategies do students generally use when completing work in the target classes?" Questionnaires, rating scales, and interview formats are used to address this question. Thus, SRL—Aptitude level data can help evaluators generate interpretations about students' general tendency to use regulatory behaviors and processes. However, because aptitude measures often vary in their level of specificity, they can be further classified into one of four categories: domain-general, domain-specific, context-specific, and task-specific.

Domain-general measures generate data about SRL processes without reference to any specific domain (e.g., academic, vocational). Such measures typically include standardized tests and some rating scales, such as the

TABLE 3 SREP Level 1 and 2 diagnostic worksheet.

Content area	Which of the levels of performance best represents the student?	What are the key assignments or activities that influence the overall grade?	Which of the assignments or activities are problematic for the student?	What are some of the reasons why you believe they are challenging?
Science	□ High □ Acceptable □ Low/struggling	□ Tests/exams/quizzes □ Homework completion and/or quality □ Individual writing assignment (e.g., essay, lab report, research paper, etc.) □ Group project □ Reading □ Other		
Social studies	□ High □ Acceptable □ Low/struggling	□ Tests/exams/quizzes □ Homework completion and/or quality □ Individual writing assignment (e.g., essay, lab report, research paper, etc.) □ Group project □ Reading □ Other		
Language arts	□ High □ Acceptable □ Low/struggling	□ Tests/exams/quizzes □ Homework completion and/or quality □ Individual writing assignment (e.g., essay, lab report, research paper, etc.) □ Group project □ Reading □ Other		
Mathematics	□ High □ Acceptable □ Low/struggling	□ Tests/exams/quizzes □ Homework completion and/or quality □ Individual writing assignment (e.g., essay, lab report, research paper, etc.) □ Group project □ Reading □ Other		
Other	□ High □ Acceptable □ Low/struggling	□ Tests/exams/quizzes □ Homework completion and/or quality □ Individual writing assignment (e.g., essay, lab report, research paper, etc.) □ Group project □ Reading □ Other		

Abbreviation: SREP, Self-Regulation Empowerment Program.

Developmental Neuropsychological Assessment—Second Edition (NEPSY-II) and the Behavior Rating Inventory of Executive Function—Second Edition (BRIEF-2). The NEPSY-II assesses neuropsychological development of children and adolescents across six general domains of cognitive functioning (e.g., executive function and attention, language, etc.; Korkman et al., 2007), while the BRIEF-2 is a questionnaire that assesses executive functions across several key areas (i.e., cognitive, behavioral, and emotional regulation; Gioia et al., 2000). Although domain-general measures provide valuable information about students' regulatory capabilities and can, to some degree, be linked to specific domains (e.g., teacher rating scales generate data about teacher perceptions of student regulation in school), neither the directions nor the test items specifically or consistently reference skills within a domain.

Domain-specific measures are administered with respect to a general domain (e.g., academics, sports, etc.) but do not target specific contexts or activities within that domain (e.g., science class). The Learning and Study Strategies Inventory—Third Edition (LASSI) is a 60-item self-report domain-specific measure that uses 10 subscales to assess learning strategies in high school or college students (e.g., "I try to find relationships between what I am learning and what I already know," "I have difficulty adapting my studying to different types of courses"; Weinstein et al., 2016). Students rate how strongly they relate to each item using a five-point Likert format, ranging from 1 *(not at all typical of me)* to 5 *(very much typical of me)*. While the LASSI was initially developed for college-aged populations in traditional academic settings, there are versions for high school students and online learners. The School Motivation and Learning Strategies Inventory (SMALSI) represents another domain-specific self-report questionnaire used to assess strategic and motivated behaviors of students in school-aged and postsecondary school settings (Chatham Stroud & Reynolds, 2006). The SMALSI uses a four-point Likert-style with frequency anchors (1 [*never*] to 4 [*almost always*]). Both the LASSI and SMALSI address SRL in school or academic domains but, to our knowledge, are not typically used with specific content areas or learning tasks in mind.

Although domain-general and domain-specific measures are quite useful for several purposes and situations, they do not generate data that directly address the regulatory question targeted in the SREP diagnostic assessment process. *Context-specific* and/or *task-specific* aptitude measures are more appropriate for addressing this question. Context-specific measures assess student SRL relative to a particular content area class or setting within an academic or school context. Several measures reported in the literature can be classified as context-specific, with two common ones being the Motivated Strategies for Learning Questionnaire (MSLQ) and the Self-Regulation Strategy Inventory (SRSI). The MSLQ is a self-report questionnaire that includes 15 scales targeting different SRL skills and processes (e.g., self-efficacy, control of learning beliefs, metacognitive SR, organization, effort regulation, and help-seeking) in college-aged students (Pintrich et al., 1991). Students use a seven-point Likert scale to describe how strongly they relate to each item, ranging from 1 *(not at all true of me)* to 7 *(very true of me)*. The MSLQ is widely considered one of the most popular and frequently used measures of SRL at the postsecondary level, although it has also been applied to other populations (Holland et al., 2018).

Of greater relevance to comprehensively assessing school-aged students' SRL processes is the SRSI. Although similar to the MSLQ in content and scope, the SRSI is unique in its focus on middle and high school populations and the inclusion of both adaptive and maladaptive subscales (Cleary, 2006). The SRSI is also particularly relevant to school-based practitioners because it includes a student, teacher, and parent version and focuses on the frequency with which students engage in regulatory behaviors (i.e., a five-point Likert scale ranging from 1 [*almost never*] to 5 [*almost always*]) rather than how typically an item represents a student (Chen et al., 2015; Cleary & Callan, 2014). Thus, inferences from SRSI scores can be made regarding how frequently students use specific strategies or processes in each content area or class rather than whether a strategy is like them or not or how much they agree or disagree with a statement.

Task-specific measures reflect the most granular and specific level of measurement within the aptitude category. Task-specific measures target students' regulatory processes or skills relative to specific situations or activities within an academic context. Although some context-specific measures reference learning activities in a general or ambiguous way (e.g., the SRSI references both studying and completion of coursework in the directions), task-specific measures are distinct in their explicit focus on a specific learning activity or situation in the directions

or items. The Self-Regulated Learning Interview Scale (SRLIS), developed by Zimmerman and Pons (1986), represents a task-specific aptitude measure. Using a structured interview format, the SRLIS examines students' strategic approaches to learning depicted across eight common academic tasks or situations in schools, such as preparing for a capstone final test in a particular class (e.g., English or history) or motivating oneself to focus on a homework assignment over a more preferred activity. Students' open-ended responses to these scenario prompts are coded into one of several strategy categories (e.g., help-seeking, time management, organization, etc.). Although the SRLIS and most SRL self-report questionnaires are similar in their assessment of regulatory beliefs, processes, or strategies that individuals typically display in school (Callan & Cleary, 2018), the information generated by the SRLIS is situation- and task-specific.

4.5.2 | Empirical support for SRL—aptitude level

There is extensive research examining the utility and psychometric rigor of aptitude measures. In this section, we provide a brief review of commonly used context-specific (i.e., MSLQ, SRSI) and task-specific (i.e., SRLIS) aptitude SRL measures reported in the literature.

In terms of the MSLQ, research is mixed, with some research supporting the measure's factor structure and reliability of scale scores (Pintrich et al., 1993; Roth et al., 2016). However, other research calls construct validity into question for certain components of the MSLQ (e.g., motivational component) and emphasizes that the MSLQ's predictive utility can vary greatly between subscales (Credé & Phillips, 2011; Hamilton & Akhter, 2009) Additionally, although some studies support the use of the MSLQ in primary and secondary school populations (Andreou & Metallidou, 2004; Ilker et al., 2014), the MSLQ was originally designed to examine SRL and motivational skills at the postsecondary level. Thus, most evidence for the reliability and validity for the MSLQ involves college populations (Holland et al., 2018). Taken together, these challenges suggest that although the MSLQ is a widely used measure and has several positive features, the results should be interpreted with caution and interpreted alongside other aptitude measures for assessments conducted at the middle or high school level.

The SRSI was developed with middle school and high school populations in mind. To our knowledge, it is the only context-specific aptitude SRL measure that includes a student self-report (SRSI-SR), teacher rating scale (SRSI-TRS), and parent rating scale (SRSI-PRS). The SRSI-SR has been studied extensively and has been used to assess student SRL across different content and academic skill areas (e.g., mathematics, science, history, reading) and in relation to student motivational and academic outcomes (e.g., self-efficacy, academic achievement, interest, office referrals, etc.; Cleary & Callan, 2014; Madjar et al., 2011). It has also been adapted for postsecondary populations (Follmer & Sperling, 2016; Thomas et al., 2020), translated into multiple languages, and applied to different learning environments (e.g., in-person vs. online; Khodarahmi & Zarrinabadi, 2015; Madjar et al., 2011; Zambrano et al., 2018). In terms of psychometric characteristics, there is evidence to support the reliability of scores and the validity of inferences regarding internal factor structure, and convergent and predictive validity (Cleary, 2006; Cleary et al., 2015; Tise et al., 2019). Although less attention has focused on the SRSI-PRS and SRSI-TRS, research has revealed similar levels of strong psychometric quality (Chen et al., 2015; Cleary & Callan, 2014).

The SRLIS (i.e., a type of structured interview) has also received attention in the empirical literature. This measure was originally developed for use in secondary school contexts, with initial research showing strong construct validity and support for its use in predicting high school students' academic course achievement (Zimmerman & Pons, 1986). Subsequent studies have provided evidence to support inferences regarding predictive and convergent validity (Kitsantas, 2002; Young & Ley, 2005). The SRLIS has also been used to examine SRL skills across school-aged populations (i.e., from the primary grades through the postsecondary years; Ablard & Lipschultz, 1998; Vandevelde et al., 2011), and has been adapted across cultures and translated into multiple languages (Metallidou, 2013; Nota et al., 2004; Pratontep & Chinwonno, 2008). In sum, the MSLQ, SRSI, and SRLIS are appropriate and psychometrically strong measures to use at the SRL—Aptitude level (Level 3) of the SREP

diagnostic assessment process. Collectively, they generate information that can be used to draw conclusions about students' general use of regulatory strategies within the context of specific content classes (e.g., SRSI, MSLQ) or academic situations (e.g., SRLIS).

4.6 | SRL–Event level (Level 4)

The SRL–Event level includes the most nuanced and situation-specific assessments of SRL. Collectively, these measures generate data about student SR as it occurs in real-time or *during* student engagement with specific academic activities, such as studying, writing an essay, or solving a mathematics problem. The most common event measures include SRL microanalysis, direct observations, behavioral traces, and think-aloud protocols. Research on event measures has increased over the past couple of decades, in part, because they are closely aligned with a process-oriented conceptualization of SRL (i.e., dynamic feedback loops) and can be readily incorporated into computer-based learning environments (CBLEs), a rapidly expanding and frequent pedagogical approach used in schools (Araka et al., 2020). Further, because event measures can reveal gaps or deficiencies in the processes used by students during actual learning activities, school-based practitioners can leverage this data to enhance intervention planning and development (Greene et al., 2018; Peters-Burton et al., 2022).

Although all event measures can be used to address the situation-specific regulatory question targeted in the SREP diagnostic process, we assert that *SRL microanalysis* is ideally suited for this purpose. In general, SRL microanalysis is a contextualized, task-specific interview used to examine students' regulatory processes across forethought (e.g., goal setting, planning), performance (e.g., strategy use, self-monitoring), and reflection (e.g., self-evaluation, attributions) phases as they engage in a specific learning activity. It is distinct from other event measures because it links the temporal aspects of the three-phase cycle of SRL (i.e., forethought, performance, and reflection) with the temporal dimensions of learning activities (i.e., before, during, and after). Thus, microanalysis protocols are developed so that forethought phase questions are administered *before* students begin an activity, performance phase questions are administered *during* task completion, and reflection phase questions are administered *after* learning or performance feedback is received (see Table 4 for phase-specific microanalytic questions). Given that students are often referred to a child study team due to difficulties linked to specific academic activities (e.g., reading, test performance, etc.), the use of SRL microanalysis procedures can enhance an evaluator's understanding of students' strategic and regulatory approaches before, during, and after engagement in those activities (Callan & Cleary, 2018).

Think-aloud protocols represent another viable event measure. This assessment approach has a long history that predates its use as an SRL assessment tool (Ericsson & Simon, 1980). Think-alouds prompt students to verbalize their thoughts during task performance without providing an explanation or reasoning about their thinking (Greene et al., 2018). The think-aloud process begins by creating a one-on-one environment between an evaluator and student that is quiet and free from distractions (Greene et al., 2018). Evaluators position themselves in a manner that dissuades interactions with students during the primary task (e.g., out of the student's view). As part of the think-aloud training procedure, evaluators typically provide students the opportunity to practice thinking aloud, usually in a domain different from the primary task. For example, Heirweg et al. (2020) allowed students to practice thinking aloud while folding an origami swan before beginning a think-aloud protocol for the target activity. Another key aspect of the think-aloud procedure involves using prompts to ensure students continuously report their thoughts throughout the entire activity. These prompts should be neutral and simple in nature, (e.g., "keep talking") so as not to bias or influence students' natural thinking processes. Like SRL microanalysis data, the information generated from think-alouds are interpreted through coding and qualitative analysis. For example, verbalizations such as, "I will first read the text, then underline keywords and try to memorize the text" may be coded as part of the SRL planning process, whereas "I think I have studied the text thoroughly" may be coded as self-evaluation (Heirweg et al., 2020). In short, the primary goal in using think-aloud protocols as an SRL assessment is to elicit

TABLE 4 Example microanalytic questions across cyclical SRL phases.

SRL phase	SRL process	Example question
Forethought	Goal-setting	"Do you have a goal in mind as you study for this exam?" "What are you trying to accomplish when complete your homework?"
	Planning	"Do you have a plan for how best to remember the information as you study?" "What do you need to do to successfully complete these mathematics problems?"
	Self-efficacy	"How confident are you that you can get an A in your social studies class?" "How sure are you that you can effective make this oral presentation?"
Performance	Strategy use	"Tell me all of the things you did to solve these physics problems."
	Self-observation or metacognitive monitoring	"What is the primary thing you are thinking about as you solve the problems?" "Are you keeping track of anything as you write the essay?"
Reflection	Self-evaluation	"What did you use to judge how well you wrote this essay?" "How well did you perform on this exam?"
	Attributions	"Why do you think you got a 62 on this math test?" "Why do you think you struggled with writing the persuasive essay?"
	Adaptive inferences	"What do you need to do to improve your performance on the next science lab?" "Is there anything you would do differently when preparing for your next mathematics test?"
	Reflective affect	"How satisfied are you with your essay grade?" "What are you feeling when you look at this grade?"

Note: Forethought questions are administered before beginning a task, performance questions during the task, and reflection questions after task completion or provision of feedback.

Abbreviation: SRL, self-regulated learning.

information about students' real-time thoughts during a particular learning task and to interpret these thoughts relative to SRL processes.

A third category of event measures include *behavioral traces*: data about student actions as they engage in specific learning activity. Behavioral traces are typically recorded as a data log within a CBLE (Araka et al., 2020; Schraw, 2010) but can also be gathered from paper and pencil products (e.g., highlights in a textbook). Common examples of trace data include the number of clicks a student makes on webpage or the number of questions they ask a teacher using the chat feature of a software program. To enhance the interpretability of behavioral traces, technology is intentionally designed to reveal aspects of SRL processes, such as planning, help-seeking, or reflecting (Bernacki, 2018). Thus, if a communication hub is created on a software application (i.e., a function that enables teachers and students to correspond and share information), behavioral traces that occur within that hub (e.g., number of times students ask questions, the nature of feedback provided by teachers) can often be interpreted as help-seeking behaviors. Behavioral traces are unique relative to many other event measures because they directly examine student behaviors rather than their verbal reports or perceptions about their behaviors. Despite this advantage, behavioral traces represent a high inference assessment approach in that interpretations are made

about underlying SRL processes based on overt actions. Although behavioral traces can be used as part of the SREP diagnostic assessment approach, this type of data is often not easily accessible by school-based practitioners as part of a school-based evaluation. Thus, behavioral traces may not represent the most practical assessment approach at this time.

4.6.1 | Empirical support for common SRL—Event level measures

In this section, we focus on research utilizing SRL microanalysis or think-aloud protocols given their direct relevance to examining student regulation during common school-related asks and activities (unlike gathering trace behaviors through web-based learning). The use of SRL microanalysis has proliferated over the past couple of decades. Cleary et al. (2021) conducted a systematic review of the SRL microanalysis literature to examine trends in its use and applicability across multiple domains (e.g., academic, athletic) and tasks within each of those domains (e.g., writing, reading, studying in academic domains). Broadly speaking, this review provided substantial data to support the relevance, use, and applicability of SRL microanalysis in school settings. In fact, SRL microanalysis has been used for several school-related purposes, such as examining the effects of academic and SRL interventions in middle and high school settings (Cleary et al., 2017; Kolovelonis et al., 2020; Zimmerman & Kitsantas, 1999), describing the SRL characteristics of specific student populations, differentiating important subgroups (e.g., high vs low achieving), and examining its link with other constructs, such as motivation and achievement (Artino et al., 2014; Callan & Cleary, 2018; DiBenedetto & Zimmerman, 2010). Across these studies, evidence exists to support claims about content validity, predictive and concurrent validity, convergent validity, and consequential validity of scores generated by SRL microanalysis. When considering that SREP is an assessment-to-intervention program, it is noteworthy that SRL microanalysis has also been used as a formative assessment and intervention tool with various groups, such as high school teachers as part of professional development (PD) experiences (Peters-Burton et al., 2022), middle school and high school students during test reflection activities (Cleary et al., 2017), and college-level musicians as they self-directed musical instrument practice sessions (McPherson et al., 2017; Osborne et al., 2020).

Cleary et al.'s (2021) systematic review also showed that SRL microanalysis has been used extensively across common academic (e.g., reading comprehension, writing, math problem solving, studying, developing science lessons) and nonacademic tasks (e.g., basketball shooting, venipuncture, practicing a musical instrument). Thus, microanalysis is a highly flexible and adaptable approach that can be linked to different learning contexts and numerous activities within those contexts—a key objective of the SREP diagnostic process. Given the inherent flexibility and adaptability of SRL microanalysis as well as the general premise of Level 4 assessments to capture real-time, task-specific data about student SRL, we believe that SRL microanalysis is one of the most important Level 4 assessment tools that evaluators can use in a school setting.

The flexibility and adaptability of SRL microanalysis are notable strengths, but these features can also present challenges for school-based evaluators. Because microanalysis questions are developed to directly correspond with a target academic activity in mind, evaluators need to have intimate knowledge of these learning activities and then customize the questions to fit those activities. When this knowledge is lacking, evaluators are encouraged to consult with teachers who possess content expertise and intimate knowledge about the target activity. Given that SRL microanalysis generates qualitative data about SRL processes, evaluators also need to devote time to coding student responses into meaningful SRL categories. Importantly, Cleary et al. (2012) offered specific guidelines for constructing and using SRL microanalysis protocols including: (1) identifying a specific task with a clear beginning, middle, and end, (2) identifying the specific SRL subprocesses that one wishes to assess (e.g., goal-setting, planning, attributions, etc.), (3) refining or customizing questions from existing SRL protocols, (4) aligning the phase-specific SRL processes targeted by the phase-specific questions (i.e., forethought, performance, and reflection) to the temporal dimensions of the task (i.e., before, during, and after), (5) administering the interview, and (6) coding the responses (see Cleary et al., 2012, for expanded discussion).

Like SRL microanalysis, think-aloud protocols have been used to assess student SRL across various academic areas and activities. Think-aloud protocols have been used with diverse samples across grade levels (e.g., primary and secondary students, undergraduate students, medical students), languages (e.g., English, Flemish, German), and learning domains (e.g., CBLEs; Azevedo, 2005; Fox, 2009; Heirweg et al., 2020; Lajoie et al., 2021; Lim et al., 2021; Vandevelde et al., 2017). There is extensive evidence to support claims about its convergent and criterion-related validity, as reflected by its strong correlations with other SRL event measures (e.g., behavioral traces) and achievement outcomes across reading, science, mathematics, and history content areas (Azevedo et al., 2004; Fox, 2009; Greene et al., 2018; Muis, 2008; Rovers et al., 2019).

Like any assessment approach, however, there are a couple of potential limitations when using think-aloud protocols as part of the SREP diagnostic process. First, there is debate about whether this assessment procedure leads to reactive effects (i.e., assessment procedures that lead to changes in respondents' behaviors). Although reactivity is a possibility, allowing the student and evaluator to engage in practice and training before completing the protocol is an important part of the think-aloud procedure that can mitigate such risks. For example, there is compelling evidence showing that there are no reactive effects when evaluators follow recommended procedures for implementing a think-aloud protocol (i.e., not prompting students to engage in reasoning or elaboration; Greene et al., 2018). Second, unlike SRL microanalysis, think-aloud protocols do not directly target specific processes within the three-phase model of SRL (i.e., planning and self-evaluation). Thus, while think-aloud protocols can reveal information about student SRL, the nature of information can vary greatly across students and may or may not reflect specific SRL processes. However, this unguided feature of think-aloud protocols also represents a strength relative to SRL microanalysis. That is, because think-aloud protocols are not constrained by specific assessment questions or prompts, as is the case with SRL microanalysis, respondents have greater freedom to express all their thoughts and reactions that naturally occur during the performance of an activity.

In sum, although SRL microanalysis and think-aloud protocols are not without limitations, there is a strong empirical literature supporting their use in schools. They both offer evaluators the opportunity to generate nuanced and situation-specific information about student SRL that is not possible with SRL—Aptitude level measures.

5 | CONCLUSION: CHALLENGES, SOLUTIONS, AND FUTURE DIRECTIONS

5.1 | Challenges

Over the past couple of decades, SRL assessment practices have been refined and improved in many respects. In fact, there is empirical evidence showing that SRL aptitude (e.g., questionnaires) and event measures (e.g., microanalysis, think-aloud protocols) can be used to generate unique yet complementary information about students' SRL functioning. Whereas aptitude measures provide data about the frequency or tendency of students to use regulatory processes in a general sense, event measures generate more fine-grained, task-specific, and process-oriented SRL information. Despite these important developments, the use of SRL tools as part of school-based assessments is not commonplace (Cleary et al., 2010). There are several potential explanations for this research-to-practice gap, including conceptual ambiguity about SRL, inadequate training and knowledge of SRL assessments, poor knowledge of SRL assessment methods, lack of fit between SRL assessments and school psychologists' assessment requirements, and insufficient access to SRL event measure protocols and procedures.

In terms of construct ambiguity, there are several constructs referenced in the literature related to the SRL process: SR, self-control, executive functions, and self-management (Greene, 2018; Panadero, 2017). Although these terms are similar in their focus on how individuals control or manage aspects of their lives, they are grounded in distinct theoretical and research paradigms; as a result, they typically emphasize different dimensions or aspects of regulation (Greene, 2018). For example, while *executive functions* represent broad cognitive abilities that have

been linked to brain functions that are regulatory in nature (e.g., cognitive flexibility, selective attention, impulse control; Blair, 2016), *self-management* has been used extensively in the behavior modification literature to represent how individuals engage in behavioral self-control (e.g., self-monitoring, self-evaluating, or using self-reinforcement (Briesch & Chafouleas, 2009). If a school psychologist wanted to assess students' executive functions, they would typically use a standardized test or broad rating scale. If self-management skills were the target, such as staying on task or managing the speed at which students complete class assignments, the school psychologist would likely use a different approach, such as direct observations. Thus, school psychologists' knowledge or preference for a particular regulatory-related construct will likely have a strong influence on their approach to assessing and making interpretations about student regulation.

Another reason for the research-to-practice gap is that most school psychologists and educators do not typically receive sufficient training or practice opportunities in SRL assessments as part of their professional training programs (Cleary, 2009; Cleary et al., 2010). Although preservice school psychologists and teachers may be exposed to assessment approaches that address some aspects of human regulation (e.g., executive functions, self-recording behaviors), to our knowledge, most teacher education or school psychology training programs do not emphasize comprehensive approaches to examining SRL dimensions and processes. When knowledge about SRL processes or assessment approaches is not well-developed, it is difficult for school-based practitioners to obtain the information needed to effectively develop SRL intervention plans. The lack of clear guidelines for integrating multiple SRL assessment tools as part of a broader psychoeducational evaluation is another likely contributing factor to the research-to-practice gap in SRL assessments. In other words, even if school psychologists possess sufficient awareness and knowledge of SRL assessment tools, they may not be skilled in effectively using these assessments and corresponding data as part of their routine assessment batteries. Further, although some survey research shows that school psychologists perceive SR to be a valuable and important skill to assess in schools (Cleary, 2009; Cleary et al., 2010), it is likely that not all school-based practitioners hold such beliefs; thus, some professionals may not seek to conduct SRL assessments. Finally, school psychologists and teachers may struggle to administer and analyze SRL assessment data for logistical reasons, such as their extensive professional responsibilities and demands on their time.

The SREP diagnostic assessment process provides clarity about the use of multiple SRL assessments, but there are a couple of issues that still need to be addressed. At the SRL–Aptitude level (i.e., Level 3), several context-specific SRL questionnaires can be used by school psychologists, with the SRSI representing a multisource system (i.e., student, teacher, and parent versions). However, developing national norms for any context-specific measure will likely prove to be quite challenging given that separate norms are needed for multiple content areas and learning contexts (e.g., nature of activities and assignments in mathematics classes is distinct from those in social studies classes). This norming issue is a particular problem for school psychologists given that they typically rely on norms to interpret scores from questionnaires. Without norms or another clear approach to interpretation, school psychologists may disregard or not utilize many SRL aptitude measures reported in the literature.

In terms of SRL—Event level measures (i.e., Level 4), we previously noted that SRL microanalysis is ideally suited to address the nuanced and situation-specific aspects of SRL. However, because microanalysis protocols need to be customized to the specific learning tasks, school psychologists will need to devote atleast some time to either developing new microanalytic interviews that align with their assessment focus or to modifying existing interviews to meet their purposes. This allocation of time can be challenging for school psychologists who are already stretched thin by extensive caseloads and other professional responsibilities.

5.2 | Solutions

To address these ongoing challenges regarding school-based SRL assessments, we recommend a three-pronged approach: (1) emphasize SRL assessment activities during preservice training, (2) provide SRL PD experiences to

current school-based practitioners, and (3) provide school psychologists and educators with greater access to SRL assessments, supports, and materials. Underlying all three initiatives is needed to help teachers and school psychologists recognize the critical importance of student regulatory thinking and action.

Students in all preservice training programs (e.g., school psychology, school counseling, teacher education) are required to develop a broad array of knowledge and skills relevant to their respective professions. Although adding new content or curriculum to an already jam-packed degree program is not an easy endeavor, the best way to increase the future and sustained use of SRL assessments in schools is to provide explicit training in these areas. Specifically for school psychology programs, it is recommended that experiences related to the integrated use of SRL aptitude (e.g., self-report questionnaires, teacher or parent rating scales, retrospective interviews) and event measures (e.g., SRL microanalysis, think-aloud protocols) be included, perhaps in courses for academic assessments and/or interventions, social–emotional assessment approaches, and/or students with learning or behavioral challenges.

On a different level, we recommend that researchers and/or SRL experts provide school psychologists PD experiences or other supports regarding SRL assessments. Over the past few decades, there has been an upsurge of SRL PD programs designed to enhance the knowledge and application of SRL principles and practices among school-based personnel (Darling-Hammond et al., 2017; Kramarski et al., 2013; Peters-Burton et al., 2022). More broadly, a recent report was published regarding the key features of PD initiatives that produce the most desirable effects on learning (Darling-Hammond et al., 2017). Generally speaking, when PD programs include features, such as active learning, collaboration, modeling of effective practices, and extensive opportunities for feedback and reflection, participants will typically display enhanced SRL knowledge, self-efficacy for teaching SRL, and skills in applying SRL principles to practice situations (Allshouse, 2016; Peters-Burton et al., 2022). Further, while these SRL PD experiences often vary from single-day workshops to multiday or multiweek formats, best practice guidelines suggest developing PD programs that extend over multiple sessions and include formative feedback. Interestingly, some researchers have begun using SRL assessments, such as SRL microanalysis, as instructional and formative assessment tools during PD programming. These assessments have been used to help teachers become more self-aware, goal-directed, and strategic in their teaching (Peters-Burton et al., 2022). The use of SRL microanalysis, other SRL assessments (e.g., think-aloud protocols), and alternative instructional approaches (e.g., videos of class teaching and/or student learning) offers much potential for PD programs involving school psychologists and other school personnel.

Finally, due to the natural disconnect that often occurs between research and practice, researchers need to make a more concerted and organized effort to give school psychologists and other school-based professionals access to relevant SRL assessment materials, resources, and guidelines. As an example, we previously noted that SRL microanalysis can be time intensive in terms of administration and coding procedures. To optimally facilitate and encourage the use of this innovative and novel assessment practice, school psychologists need access to existing protocols and relevant coding schemes. One viable option is the creation of an open-access, web-based library that houses the increasing number of existing microanalytic protocols (see Cleary et al., 2021). Researchers using SRL microanalysis can populate this web-based library with example protocols targeting different academic activities or situations (e.g., studying, writing an essay, reading textbooks). School-based practitioners can then either use the microanalytic protocol as part of an evaluation or as an exemplar for adapting or customizing new protocols to match the target academic activity of interest (Cleary et al., 2012).

Relatedly, there is a need for researchers to help school psychologists and local school districts interpret scores generated from SRL measures, particularly for the *context-specific SRL questionnaires* used at the SRL—Aptitude level (i.e., Level 3). Because it is quite challenging to develop national norms for context-specific assessment tools (i.e., due to the idiosyncratic nature of different learning activities and contexts), there is a need for researchers to assist local school districts in properly developing local norms for the various aptitude measures. Developing local norms can enable school psychologists to make interpretations about the quality or frequency with which students use SRL strategies and processes relative to their same age peers in the same schools. Alternatively, one can develop and use broad cutoff scores to guide interpretations about students' regulatory skills and processes and to provide feedback about how to improve these skills. This suggested use of cutoff scores mirrors the assessment and

feedback approach used by Bryer et al. (2023) within their computer-based Diagnostic Assessment and Achievement of College Skills (DAACS) system. In general, students complete a self-report questionnaire about their SRL strategies and skills. Using cutoff scores, these data are then transformed into profiles of strengths and areas in need of improvement that both students and their college advisors can access. Students also receive actionable feedback and suggestions for directly improving their SRL skills as part of this system.

5.3 | Future research directions

To enhance the use of the SREP diagnostic assessment process in schools, additional research is needed to address several emergent issues across each of the four levels. The *Challenge—Class* and *Challenge—Task* levels (Levels 1 and 2) represent an informal and descriptive dimension to the SREP diagnostic process. Although we discussed these levels and provided example worksheets and rubrics in this paper, research is needed to examine the feasibility and utility of these processes from the perspectives of school-based practitioners. Further, because the SREP diagnostic process emphasizes the triangulation of data from multiple sources, future research needs to quantitatively and/or qualitatively examine levels of convergence among parents, teachers, and students regarding expectations for student success. The level of agreement among assessment sources is particularly important given that prior research reveals a low level of correspondence among respondents when rating internal and externalizing student issues (De Los Reyes et al., 2015).

At the SRL—Aptitude level (i.e., Level 3), we again emphasize the importance of researchers assisting local school districts with developing norms for different context-specific SRL questionnaires. Further, because there are increased expectations among middle school and high school students for studying and work completion at home (e.g., research papers, writing science investigation reports, studying for broad exams), there needs to be a greater focus on research examining parent perceptions and attitudes about students' regulatory skills.

The SRLIS is another aptitude measure that school-based practitioners can use as part of the SREP diagnostic assessment process. Despite its many strengths, the SRLIS presents only a single question for each of the eight academic situations (e.g., completing a test, writing an essay, motivating oneself). When conducting an evaluation targeting students' use of SRL strategies for challenging academic activities, it is important for interviews to generate more detailed information. Several semistructured interviews (Level 3, SRL—Aptitude) have been developed and used in applied contexts to address SRL skills relative to test preparation, homework completion, reading, and writing assignments, but they have not yet to be rigorously evaluated (Cleary, 2022). These SRL—Aptitude level interviews are important given their potential for generating comprehensive information about students' strategic and regulatory approaches in the context of specific academic situations. A fruitful line of research involves generating evidence to support the validity of inferences from the interview and to examine their utility and acceptability among school-based professionals.

Regarding the SRL—Event level (i.e., Level 4), research has devoted substantial attention to SRL microanalysis, behavioral traces, and think-aloud protocols. The emergence of these measures has occurred, in part, due to the realization that the dynamic, fluid aspects of SRL captured by event measures are especially relevant to intervention planning and development. For example, if a student struggles to write high-quality essays in their English Language Arts class, an evaluator can use SRL microanalysis or a think-aloud protocol to gather data about how the student initiates, manages, and adapts their thinking and behaviors while writing. Gaining clarity about the processes used by students during learning can be quite helpful to school-based practitioners for developing valid hypotheses about the most effective approaches to assist them (Goh & Bambara, 2012). Specifically regarding SRL microanalysis, we reviewed data to support its strong potential as a Level 4 assessment tool. However, research on SRL microanalysis is still emerging; thus, researchers need to expand and build on the existing microanalytic literature base to identify ways to optimize and enhance its applicability to schools and across various academic activities.

From a holistic perspective regarding the SREP diagnostic process, there is a need to examine whether including multiple-aptitude measures (e.g., teacher ratings, student self-reports, student interviews) and multiple event measures (e.g., think-aloud protocols, microanalysis) facilitates the diagnostic process and/or improves the prediction of student outcomes. Very few studies have explored this issue. Additionally, there is evidence that teacher beliefs regarding SRL (e.g., growth vs. fixed mindset) can predict their own use of cognitive and metacognitive strategies in the classroom, which may impact how effectively they promote strategies during classroom instruction (Vosniadou et al., 2021). Given the potential importance of teacher or even parent beliefs and practices of SRL, researchers can expand beyond the student-focused SREP diagnostic assessment process by evaluating adult perceptions of SRL or other contextual processes linked to student behaviors in the classroom. Finally, there is the need to directly examine the feasibility, utility, and effectiveness of the entire four-level SREP diagnostic process within a school-based setting. Although we reviewed research examining selected aspects of this process (i.e., aptitude and event measures), to our knowledge no studies have examined the implementation of the entire four-level process. In addition to using this four-level process to assess student SRL at a single time point, it may also be fruitful for researchers to evaluate the impacts of this assessment process over time alongside corresponding team-based decisions and recommendations. The value and importance of the SREP diagnostic process will ultimately be determined by its cost-effectiveness and utility for providing clinically relevant information about students' underlying regulatory challenges and the generation of ideas to optimize intervention programming for those students.

ACKNOWLEDGMENTS

We are grateful to Susan Forman, Adam Lekwa, and Rui Zhang for their helpful comments and suggestions on earlier drafts of this manuscript.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

ORCID

Timothy J. Cleary http://orcid.org/0000-0002-3222-2391

REFERENCES

Ablard, K. E., & Lipschultz, R. E. (1998). Self-regulated learning in high-achieving students: Relations to advanced reasoning, achievement goals, and gender. *Journal of Educational Psychology*, *90*(1), 94–101. https://doi.org/10.1037/0022-0663.90.1.94

Allshouse, A. D. (2016). *Professional development in self-regulated learning: Effects of a workshop on teacher knowledge, skills, and self-efficacy, and the development of a coaching framework* (Publication No. 10297495) [Doctoral dissertation, Rutgers: The State University of New Jersey]. ProQuest Dissertations and Theses Global.

Andreou, E., & Metallidou, P. (2004). The relationship of academic and social cognition to behaviour in bullying situations among Greek primary school children. *Educational Psychology*, *24*(1), 27–41. https://doi.org/10.1080/0144341032000146421

Araka, E., Maina, E., Gitonga, R., & Oboko, R. (2020). Research trends in measurement and intervention tools for self-regulated learning for e-learning environments—Systematic review (2008–2018). *Research and Practice in Technology Enhanced Learning*, *15*(1), 6. https://doi.org/10.1186/s41039-020-00129-5

Artino Jr., A. R., Cleary, T. J., Dong, T., Hemmer, P. A., & Durning, S. J. (2014). Exploring clinical reasoning in novices: A self-regulated learning microanalytic assessment approach. *Medical Education*, *48*, 280–291. https://doi.org/10.1111/medu.12303

Azevedo, R. (2005). Computer environments as metacognitive tools for enhancing learning. *Educational Psychologist*, *40*, 193–197. https://doi.org/10.1207/s15326985ep4004_1

Azevedo, R., Guthrie, J. T., & Seibert, D. (2004). The role of self-regulated learning in fostering students' conceptual understanding of complex systems with hypermedia. *Journal of Educational Computing Research*, *30*(1), 87–111. https://doi.org/10.2190/DVWX-GM1T-6THQ-5WC7

Azevedo, R., Moos, D. C., Johnson, A. M., & Chauncey, A. D. (2010). Measuring cognitive and metacognitive regulatory processes during hypermedia learning: Issues and challenges. *Educational Psychologist*, *45*(4), 210–223. https://doi.org/10.1080/00461520.2010.515934

Bernacki, M. L. (2018). Examining the cyclical, loosely sequenced, and contingent features of self-regulated learning: Trace data and their analysis. In D. H. Schunk & J. A. Greene (Eds.), *Educational psychology handbook series. Handbook of self-regulation of learning and performance* (pp. 370–387). Routledge/Taylor & Francis Group.

Blair C. (2016). Developmental science and executive function. *Current Directions in Psychological Science*, *25*(1), 3–7. https://doi.org/10.1177/0963721415622634

Briesch, A. M., & Chafouleas, S. M. (2009). Review and analysis of literature on self-management interventions to promote appropriate classroom behaviors (1988–2008). *School Psychology Quarterly*, *24*(2), 106–118. https://doi.org/10.1037/a0016159

Bryer, J. M., Andrade, H. L., Cleary, T. J., Lui, A. M., Franklin, D. W., & Akhmedjanova, D. (2023). *The efficacy and predictive power of the Diagnostic Assessment and Achievement of College Success (DAACS) on academic success indicators*. City University of New York. Manuscript submitted for publication.

Callan, G. L., & Cleary, T. J. (2018). Multidimensional assessment of self-regulated learning with middle school math students. *School Psychology Quarterly*, *33*(1), 103–111. https://doi.org/10.1037/spq0000198

Callan, G. L., & Cleary, T. J. (2019). Examining cyclical phase relations and predictive influences of self-regulated learning processes on mathematics task performance. *Metacognition and Learning*, *14*, 43–63. https://doi.org/10.1007/s11409-019-09191-x

Chatham Stroud, K., & Reynolds, C. R. (2006). *School Motivation and Learning Strategies Inventory manual*. Western Psychological Services.

Chen, P. P., Cleary, T. J., & Lui, A. M. (2015). Examining parents' ratings of middle-school students' academic self-regulation using principal axis factoring analysis. *School Psychology Quarterly: The Official Journal of the Division of School Psychology, American Psychological Association*, *30*(3), 385–397. https://doi.org/10.1037/spq0000098

Cleary, T. J. (2006). The development and validation of the Self-Regulation Strategy Inventory—Self-Report. *Journal of School Psychology*, *44*(4), 307–322. https://doi.org/10.1016/j.jsp.2006.05.002

Cleary, T. J. (2009). School-based motivation and self-regulation assessments: An examination of school psychologist beliefs and practices. *Journal of Applied School Psychology*, *25*(1), 71–94.

Cleary, T. J. (2022). *The Self-Regulation Empowerment Program (SREP) Task Specific Regulatory Interviews* [Unpublished manual]. Department of School Psychology, Rutgers, the State University of New Jersey.

Cleary, T. J., & Callan, G. L. (2014). Student self-regulated learning in an urban high school: Predictive validity and relations between teacher ratings and student self-reports. *Journal of Psychoeducational Assessment*, *32*(4), 295–305. https://doi.org/10.1177/0734282913507653

Cleary, T. J., Callan, G. L., Malatesta, J., & Adams, T. (2015). Examining the level of convergence among self-regulated learning microanalytic processes, achievement, and a self-report questionnaire. *Journal of Psychoeducational Assessment*, *33*(5), 439–450. https://doi.org/10.1177/0734282915594739

Cleary, T. J., Callan, G. L., & Zimmerman, B. J. (2012). Assessing self-regulation as a cyclical, context-specific phenomenon: Overview and analysis of SRL microanalytic protocols. *Education Research International*, *2012*, 1–19. https://doi.org/10.1155/2012/428639

Cleary, T. J., Gubi, A., & Prescott, M. V. (2010). Motivation and self-regulation assessments: Professional practices and needs of school psychologists. *Psychology in the Schools*, *47*(10), 985–1002. https://doi.org/10.1002/pits.20519

Cleary, T. J., Slemp, J., Reddy, L. A., Alperin, A., Lui, A., Austin, A., & Cedar, T. (2021). Characteristics and uses of SRL microanalysis across diverse contexts, tasks, and populations: A systematic review. *School Psychology Review*, *52*(2), 1–21. https://doi.org/10.1080/2372966X.2020.1862627

Cleary, T. J., Velardi, B., & Schnaidman, B. (2017). Effects of the Self-Regulation Empowerment Program (SREP) on middle school students' strategic skills, self-efficacy, and mathematics achievement. *Journal of School Psychology*, *64*, 28–42. https://doi.org/10.1016/j.jsp.2017.04.004

Cleary, T. J., & Zimmerman, B. J. (2004). Self-Regulation Empowerment Program: A school-based program to enhance self-regulated and self-motivated cycles of student learning. *Psychology in the Schools*, *41*, 537–550. https://doi.org/10.1002/pits.10177

Credé, M., & Phillips, L. A. (2011). A meta-analytic review of the Motivated Strategies for Learning Questionnaire. *Learning and Individual Differences*, *21*(4), 337–346. https://doi.org/10.1016/j.lindif.2011.03.002

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute.

DiBenedetto, M. K., & Zimmerman, B. J. (2010). Differences in self-regulatory processes among students studying science: A microanalytic investigation. *International Journal of Educational and Psychological Assessment*, *5*, 2–24.

DiBenedetto, M. K., & Zimmerman, B. J. (2013). Construct and predictive validity of microanalytic measures of students' self-regulation of science learning. *Learning and Individual Differences*, *26*, 30–41. https://doi.org/10.1016/j.lindif.2013.04.004

Dignath, C., Buettner, G., & Langfeldt, H.-P. (2008). How can primary school students learn self-regulated learning strategies most effectively?: A meta-analysis on self-regulation training programmes. *Educational Research Review*, *3*(2), 101–129. https://doi.org/10.1016/j.edurev.2008.02.003

Dombrowski, S. C. (Ed.). (2020). *Psychoeducational assessment and report writing* (2nd ed.). Springer International Publishing. https://doi.org/10.1007/978-3-030-44641-3

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist*, *46*(1), 6–25. https://doi.org/10.1080/00461520.2011.538645

Erchul, W. P., & Martens, B. K. (2010). *School consultation: Conceptual and empirical bases of practice* (3rd ed.). Springer. https://doi.org/10.1007/978-1-4419-5747-4_4

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*, 215–251. https://doi.org/10.1037/0033-295X.87.3.215

Follmer, D. J., & Sperling, R. A. (2016). The mediating role of metacognition in the relationship between executive function and self-regulated learning. *British Journal of Educational Psychology*, *86*(4), 559–575. https://doi.org/10.1111/bjep.12123

Fox, E. (2009). The role of reader characteristics in processing and learning from informational text. *Review of Educational Research*, *79*(1), 197–261. https://doi.org/10.3102/0034654308324654

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Test review Behavior Rating Inventory of Executive Function. *Child Neuropsychology*, *6*(3), 235–238. https://doi.org/10.1076/chin.6.3.235.3152

Goh, A. E., & Bambara, L. M. (2012). Individualized positive behavior support in school settings: A meta-analysis. *Remedial and Special Education*, *33*(5), 271–286. https://doi.org/10.1177/0741932510383990

Graham, S., Harris, K. R., & Mason, L. (2005). Improving the writing performance, knowledge, and self-efficacy of struggling young writers: The effects of self-regulated strategy development. *Contemporary Educational Psychology*, *30*(2), 207–241. https://doi.org/10.1016/j.cedpsych.2004.08.001

Greene, J. A. (2018). *Self-regulation in education*. Routledge. https://doi.org/10.4324/9781315537450

Greene, J. A., Deekens, V. M., Copeland, D. Z., & Yu, S. (2018). Capturing and modeling self-regulated learning using think-aloud protocols. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 323–337). Routledge/Taylor & Francis Group.

Hamilton, R. J., & Akhter, S. (2009). Construct validity of the Motivated Strategies for Learning Questionnaire. *Psychological Reports*, *104*(3), 711–722. https://doi.org/10.2466/PR0.104.3.711-722

Heirweg, S., De Smul, M., Merchie, E., Devos, G., & Van Keer, H. (2020). Mine the process: Investigating the cyclical nature of upper primary school students' self-regulated learning. *Instructional Science*, *48*(4), 337–369. https://doi.org/10.1007/s11251-020-09519-0

Holland, D. F., Kraha, A., Zientek, L. R., Nimon, K., Fulmore, J. A., Johnson, U. Y., Ponce, H. F., Aguilar, M. G., & Henson, R. K. (2018). Reliability generalization of the Motivated Strategies for Learning Questionnaire: A meta-analytic view of reliability estimates. *SAGE Open*, *8*(3), 2158244018802334. https://doi.org/10.1177/2158244018802334

Ilker, G. E., Arslan, Y., & Demirhan, G. (2014). A validity and reliability study of the Motivated Strategies for Learning Questionnaire. *Educational Sciences: Theory & Practice*, *14*(3), 829–833.

Keller-Margulis, M. A., Shapiro, E. S., & Hintze, J. M. (2008). Long-term diagnostic accuracy of curriculum-based measures in reading and mathematics. *School Psychology Review*, *37*(3), 374–390. https://doi.org/10.1080/02796015.2008.12087884

Khodarahmi, E., & Zarrinabadi, N. (2015). Self-regulation and academic optimism in a sample of Iranian language learners: Variations across achievement group and gender. *Current Psychology*, *35*(4), 700–710. https://doi.org/10.1007/s12144-015-9340-z

Kitsantas, A. (2002). Test preparation and performance: A self-regulatory analysis. *The Journal of Experimental Education*, *70*(2), 101–113. https://doi.org/10.1080/00220970209599501

Kolovelonis, A., Goudas, M., & Samara, E. (2020). The effects of a self-regulation learning teaching unit on students' performance calibration, goal attainment, and attributions in physical education. *The Journal of Experimental Education*, *90*(1), 1–18. https://doi.org/10.1080/00220973.2020.1724852

Korkman, M., Kirk, U., & Kemp, S. (2007). *NEPSY–Second Edition (NEPSY-II)*. Harcourt Assessment.

Kramarski, B., Desoete, A., Bannert, M., Narciss, S., & Perry, N. (2013). New perspectives on integrating self-regulated learning at school. *Education Research International*, *2013*, 1–4. https://doi.org/10.1155/2013/498214

Lajoie, S. P., Zheng, J., Li, S., Jarrell, A., & Gube, M. (2021). Examining the interplay of affect and self regulation in the context of clinical reasoning. *Learning and Instruction*, *72*, 101219. https://doi.org/10.1016/j.learninstruc.2019.101219

Lim, L., Bannert, M., van der Graaf, J., Molenaar, I., Fan, Y., Kilgour, J., Moore, J., & Gašević, D. (2021). Temporal assessment of self-regulated learning by mining students' think-aloud protocols. *Frontiers in Psychology*, *12*, 749749. https://doi.org/10.3389/fpsyg.2021.749749

De Los Reyes A., Augenstein T. M., Wang M., Thomas S. A., Drabick D. A. G., Burgers D. E., Rabinowitz J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, *141*(4), 858–900. https://doi.org/10.1037/a0038498

Madjar N., Kaplan A., Weinstock M. (2011). Clarifying mastery-avoidance goals in high school: Distinguishing between intrapersonal and task-based standards of competence. *Contemporary Educational Psychology*, *36*(4), 268–279. https://doi.org/10.1016/j.cedpsych.2011.03.003

McPherson, G. E., Osborne, M. S., Evans, P., & Miksza, P. (2017). Applying self-regulated learning microanalysis to study musicians' practice. *Psychology of Music*, *47*(1), 1–15. https://doi.org/10.1177/0305735617731614

Metallidou, P. (2013). Epistemological beliefs as predictors of self-regulated learning strategies in middle school students. *School Psychology International*, *34*(3), 283–298. https://doi.org/10.1177/0143034312455857

Muis, K. R. (2008). Epistemic profiles and self-regulated learning: Examining relations in the context of mathematics problem solving. *Contemporary Educational Psychology*, *33*(2), 177–208. https://doi.org/10.1016/j.cedpsych.2006.10.012

Nota, L., Soresi, S., & Zimmerman, B. J. (2004). Self-regulation and academic achievement and resilience: A longitudinal study. *International Journal of Educational Research*, *41*(3), 198–215. https://doi.org/10.1016/j.ijer.2005.07.001

Osborne, M., McPherson, G. E., Miksza, P., & Evans, P. (2020). Using a microanalysis intervention to examine shifts in musician's self-regulated learning. *Psychology of Music*, *49*(4), 1–17. https://doi.org/10.1177/0305735620915265

Osher, D., Kidron, Y., Brackett, M., Dymnicki, A., Jones, S., & Weissberg, R. P. (2016). Advancing the science and practice of social and emotional learning: Looking back and moving forward. *Review of Research in Education*, *40*(1), 644–681. https://doi.org/10.3102/0091732x16673595

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*(422), 1–28. https://doi.org/10.3389/fpsyg.2017.00422

Peters-Burton, E. E., Goffena, J., & Stehle, S. M. (2022). Utility of a self-regulated learning microanalysis for assessing learning during professional development. *The Journal of Experimental Education*, *90*, 523–549. https://doi.org/10.1080/00220973.2020.1799314

Pintrich, P. R., Smith, D. A. F., García, T., & McKeachie, W. J. (1991). *A manual for the use of the Motivated Strategies Questionnaire (MSLQ)*. University of Michigan, National Center for Research to Improve Postsecondary Teaching and Learning. https://doi.org/10.13140/RG.2.1.2547.6968

Pintrich, P. R., Smith, D. A. F., Garcia, T., & Mckeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement*, *53*(3), 801–813. https://doi.org/10.1177/0013164493053003024

Pratontep, C., & Chinwonno, A. (2008). Self-regulated learning by Thai university students in an EFL extensive reading program. *MANUSYA*, *11*(2), 104–124. https://doi.org/10.1163/26659077-01102006

Reddy, L. A., Cleary, T. J., Alperin, A., & Verdesco, A. (2018). A critical review of self-regulated learning interventions for children with attention-deficit hyperactivity disorder. *Psychology in the Schools*, *55*(6), 609–628. https://doi.org/10.1002/pits.22142

Roth, A., Ogrin, S., & Schmitz, B. (2016). Assessing self-regulated learning in higher education: A systematic literature review of self-report instruments. *Educational Assessment, Evaluation, and Accountability*, *28*, 225–250. https://doi.org/10.1007/s11092-015-9229-2

Rovers, S. F. E., Clarebout, G., Savelberg, H. H. C. M., de Bruin, A. B. H., & van Merriënboer, J. J. G. (2019). Granularity matters: Comparing different ways of measuring self-regulated learning. *Metacognition and Learning*, *14*(1), 1–19. https://doi.org/10.1007/s11409-019-09188-6

Schraw, G. (2010). Measuring self-regulation in computer-based learning environments. *Educational psychologist*, *45*(4), 258–266. https://doi.org/10.1080/00461520.2010.515936

Schunk, D. H. & Greene, J. A. (Eds.). (2018). *Handbook of self-regulation of learning and performance* (2nd ed.). Routledge.

Smith, S. R. (2007). Making sense of multiple informants in child and adolescent psychopathology: A guide for clinicians. *Journal of Psychoeducational Assessment*, *25*(2), 139–149. https://doi.org/10.1177/0734282906296233

Thomas, A. O., Voils, S. A., & Childs-Kean, L. (2020). Modification of the Self-Regulated Strategy Inventory-Self Report for blended and team-based learning. *Currents in Pharmacy Teaching and Learning*, *12*(9), 1093–1100. https://doi.org/10.1016/j.cptl.2020.04.021

Tise, J., Follmer, D. J., & Sperling, R. A. (2019). A review of the Self-Regulation Strategy Inventory—Self-Report (SRSI-SR). *Psychology*, *10*(03), 305–319. https://doi.org/10.4236/psych.2019.103022

Vandevelde, S., Van Keer, H., & Merchie, E. (2017). The challenge of promoting self-regulated learning among primary school children with a low socioeconomic and immigrant background. *The Journal of Educational Research*, *110*(2), 113–139. https://doi.org/10.1080/00220671.2014.999363

Vandevelde, S., Van Keer, H., & De Wever, B. (2011). Exploring the impact of student tutoring on at-risk fifth and sixth graders' self-regulated learning. *Learning and Individual Differences*, *21*(4), 419–425. https://doi.org/10.1016/j.lindif.2011.01.006

Vohs, K. D., & Baumeister, R. F. (2016). *Handbook of self-regulation: Research, theory, and applications* (3rd ed.). Guilford Press.

Vosniadou, S., Darmawan, I., Lawson, M. J., Van Deur, P., Jeffries, D., & Wyra, M. (2021). Beliefs about the self-regulation of learning predict cognitive and metacognitive strategies and academic performance in pre-service teachers. *Metacognition and Learning*, *16*(3), 523–554. https://doi.org/10.1007/s11409-020-09258-0

Ward, L., & Traweek, D. (1993). Application of a metacognitive strategy to assessment, intervention, and consultation: A think-aloud technique. *Journal of School Psychology*, *31*, 469–485. https://doi.org/10.1016/0022-4405(93)90032-E

Weinstein, C. E., Palmer, D. R., & Acee, T. W. (2016). *LASSI user's manual: For those administering the Learning and Study Strategies Inventory* (3rd ed.). H & H Pub.

Winne, P. H., & Jamieson-Noel, D. (2003). Self-regulating studying by objectives for learning: Students' reports compared to a model. *Contemporary Educational Psychology*, *28*, 259–276. https://doi.org/10.1016/S0361-476X(02)00006-1

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 531–566). Academic Press.

Wixson, K. K., & Valencia, S. W. (2011). Assessment in RTI: What teachers and specialists need to know. *The Reading Teacher*, *64*(6), 466–469. https://doi.org/10.1598/RT.64.6.13

Young, D., & Ley, K. (2005). Developmental college student self-regulation: Results from two measures. *Journal of College Reading and Learning*, *36*(1), 60–80. https://doi.org/10.1080/10790195.2005.10850180

Zambrano, C., Rojas, D., Díaz, A., & Salcedo, P. (2018). Psychometric properties of the Inventory of self-regulation strategies in Chilean pedagogy students. *Formación Universitaria*, *11*(5), 85–92. https://doi.org/10.4067/S0718-50062018000500085

Zimmerman, B. J., & Kitsantas, A. (1999). Acquiring writing revision skill: Shifting from process to outcome self-regulatory goals. *Journal of Educational Psychology*, *91*(2), 241–250. https://doi.org/10.1037/0022-0663.91.2.241

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

Zimmerman, B. J., & Pons, M. M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, *23*(4), 614–628. https://doi.org/10.3102/00028312023004614

How to cite this article: Cleary, T. J., & Russo, M. R. (2024). A multilevel framework for assessing self-regulated learning in school contexts: Innovations, challenges, and future directions. *Psychology in the Schools*, *61*, 80–102. https://doi.org/10.1002/pits.23035