Investigating Relationships between Self-Regulated Learning Processes and Formal Classroom Science Assessment

Maansa Bajaj Prakash

A thesis submitted in fulfilment of the requirements to the degree of

Doctor of Philosophy in Education

University of Canterbury, 2023

Acknowledgement

All our dreams come true if we have the courage to pursue them.

-Walt Disney

When I left India to come to New Zealand to attend a PhD program, I came with the hope of my dream finally coming true. Today, I realise this dream was only made possible because of the special people who have taught me to be courageous every step of the way.

First, I would like to express my profound appreciation for my supervisory team, Valerie Sotardi and Misty Sato. Thank you for the guidance and encouragement you have given me through the years. I am grateful for all the times you have challenged my thinking and reasoning while at the same time teaching me to believe in myself and to stand by my decisions. I'm also thankful to the school leaders, teachers, and students who believed in this thesis and shared their experiences with me.

There are several people at the University of Canterbury to who I would like to extend my gratitude for their support and advice. My 4th floor PhD mates, Kate and Ciara, you guys get what this is like, so thank you for sharing the ups and downs of this incredible journey. To Sriparna, thank you for keeping me grounded in my Indian roots and teaching me how to grow into my person. William and Mengping, engaging with you during your final stage helped me persevere through mine; thank you for lighting the path. I am incredibly grateful to you for those basketball mornings and weekend shenanigans. In those times, I found deep friendships that I am sure will last a lifetime.

I am grateful to Ishaan, Harshitha, and Pranav for having best friends like you. Despite being miles away and the difference in time zones, you have been with me every step of the way. Ishaan, thank you for the Sunday calls filled with laughter and deep conversations. Harshitha, it is astonishing that you were able to care for and listen to me every day from halfway across the world. And, to Pranav, thank you for your patience, the guitar tunes that helped reduce my anxiety, and the unconditional love and belief you have in me. You made being away from home more manageable

Finally, my family. Shans and Chach, you sparked the PhD dream, which gave me a glimpse of a world I wanted to be a part of. Shibho, my first best friend, thank you for virtual yoga sessions and being the comedian during our family group calls. And to the parents, Deepa and Prakash, thank you for believing in my dreams and helping make them come true. I hope I make you proud.

Abstract

Self-regulated learning (SRL) processes contribute to the short- and long-term academic success of young people. In particular, students' SRL processes before and after a task predict learning and performance. The research on SRL suggests that environmental conditions such as classroom assessments can help students develop and implement effective strategies for learning before and after the task. Consequently, identifying how classroom assessments can be designed and administered that account for these SRL processes has gained increasing attention. Classroom assessments may be characterized as formal or informal based on their purpose (e.g., formative and summative) and format (e.g., quiz, observations, tests, etc) across disciplines. The current study focuses on formal assessments which typically take the form of an exam or test and tend to serve a summative function. There are compelling reasons to use such summative assessments for developing SRL, especially in India, where such assessments are more common. In particular, evidence suggests that teachers build tasks and make decisions about classroom assessments that often inform students about how to learn and perform. Yet, few studies have explored these aspects concerning students' SRL in the forethought and self-reflection phases. Therefore, in this thesis, I investigate an underexplored type of classroom assessment-formal classroom assessments—and their relationship to students' processes in the planning and reflection stages. This thesis offers the first substantial work to examine SRL for formal classroom assessments in India.

I used a mixed methods design to explore classroom science assessment and SRL processes across two studies. To begin, I developed a microanalytic interview protocol to

gather rich data on SRL processes in the forethought as well as self-reflection phases described in Zimmerman's SRL model (2009). The sample comprised 229 high school students from India. Results indicated that students' goal-setting and self-evaluation strategies were largely focused on performance, and study strategies reflected surface learning approaches (e.g., rehearsal). On average, students felt confident about their abilities to learn and perform on the assessment. They were also moderately interested in the subject. Even though students reported confidence and interest in learning science, they were more likely to set performance goals than mastery goals. This approach to goal-setting can weaken other SRL processes which include their motivation, monitoring and regulating capacities, and ways in which they reflect on their learning. Self-regulated students who focus on acquiring content knowledge are more likely to optimize their learning for success than those learners who focus on achieving a grade or score.

The second part of my research was focused on the relationships between reported SRL processes and the assessment task. Given that SRL processes are determined by the ongoing interaction between the learner and task, it is necessary to understand the characteristics of the task presented to them. Researchers have identified task conditions that support and promote SRL among students in the classroom, but were not necessarily assessment tasks (Perry et al., 2006). Therefore, I integrated insights from previous research and developed an instrument that measured the design features of the teachers' assessment task. Findings indicated that the assessment design is associated with how students think and act in learning and performance settings. For instance, students were less likely to feel efficacious or interested if the number of questions that required higher-order thinking skills were more than the number of questions focused on lower-order thinking skills. The results also indicated that students were more likely to focus on the journey of learning (process goals) than the destination (outcome goals) when the task covered a range of topics and questions and demanded higher-order thinking skills.

The third and final part of my research aimed to investigate teachers' decisions concerning the assessment task and the impact of these decisions on students' SRL. This series of questions taps into the foundation of students' assessments: what did teachers think about when designing assessments and in what ways did SRL differ based on these reported intentions? Within student group differences indicated that students who reported adaptive strategies and higher motivational beliefs belonged to the classroom in which the teacher reported a learning-focused orientation toward assessments. Students' goal-setting and strategy selection differed between teachers' stated intentions regarding task design. Chi-square (X^2) tests for independence indicated teachers' intentions regarding design and evaluation styles were associated with how students chose to attribute their failure. Overall, the results suggest that teachers' reported assessment decisions can contribute to students' approaches before and after a formal assessment task.

In sum, my findings revealed that students' SRL processes appeared to be less adaptive for a classroom assessment task, which could mean that they lack the flexibility in analysing the situation and identifying necessary strategies for learning and performance success. Although this exploratory research does not explain the causal effects of teachers' design decisions on students' SRL, it highlights that multiple factors are at play. Consistent with a social cognitive framework (Zimmerman, 2013), it is likely that the assessment context creates an environment in which students think about and (attempt to) learn. How we think about and approach assessment design, therefore, matters.

This research has several novel elements. First, I developed an instrument to measure task features for a formal assessment which offers a new way to understand the design of a task and how it relates to SRL processes. Researchers and educators could use this tool to analyse and design assessments that help promote SRL within students. Second, I provide validity for the SRL microanalysis protocol (Cleary, 2011)—a relatively new methodological approach for academic tasks in a classroom context. Third, this research provides insights into an under-researched assessment context: India. Given that most SRL research is conducted in developed countries, this thesis provides significant implications for improving learning and performance for millions of students in the Indian context. The findings provide substantial evidence on the types of tasks designed, students' approaches to learning, and how teachers make decisions concerning formal classroom assessments.

In this thesis, I argue for the use of formal assessments as a promising event for promoting and sustaining SRL forethought and self-reflection processes. In particular, the results and findings have implications for practice and policy in the Indian context. Based on the current research, I present an initial checklist that teachers could realistically use to facilitate assessment design decisions with an SRL-focused lens. I contend that the structured nature of a formal classroom allows for a systematic process to evaluate the task design and integrate practices that promote planning, strategy selection, and self-reflection. I propose a framework that integrates SRL processes into formal classroom assessment decisions. More research is needed to identify how and when SRL-promoting methods can be introduced into the formal assessment process. Future research could also shed more insight on assessment task design features across academic subjects to help distinguish appropriate practices for developing SRL. Through intentional assessment design, teachers and researchers in India can raise students' scientific knowledge, competencies, and attitudes to become successful lifelong learners.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 My position in the research

My past experiences shaped my understanding of what it meant to learn. Even though I was able to "succeed" at school, I felt I lacked the capacity to learn. In high school, I chose to study the "hard" subjects (i.e., physics, chemistry, math, and biology). And my school decided that I was competent to pursue those disciplines because I fared a decent average in my secondary school national examinations. Within the first few months, I struggled with physics and found it challenging to cope with the subject. I joined after-school tuitions to help me perform on examinations, but they did not help me grasp the subject matter. Even though the laws of physics and fundamental reactions of chemistry fascinated me, I often felt incompetent, anxious, and nervous in the classroom and before examinations. These feelings of stress and anxiety were amplified by the pressure to perform and achieve high scores on exams, which are India's dominant form of assessments. Unfortunately, these feelings around science assessments are still prevalent among many students.

As a primary school teacher at an International Baccalaureate (IB) School in India, I learned how to encourage autonomy for learning among young children using effective pedagogies, but the most striking revelation during my teaching was the power of assessments. As a student, my perceptions of assessments were limited to performance. However, as a teacher, at a non-traditional school in India, I was fortunate to be a part of professional development modules that improved my understanding of the relationships among teaching, learning, and assessment. Through these professional development workshops, I gained a fresh and uplifting perspective on assessments, especially formative assessments. I was inspired by the potential of assessments while, at the same time, heartbroken about the millions of children who do not have this experience because of the traditional pedagogies used within schools in India. Therefore, in considering moving to doctoral research, I wanted to understand the experiences of students in a typical mainstream school affiliated with national or state educational boards. In particular, I wanted to explore formative assessment practices associated with decreased anxiety levels.

In the beginning stages of the doctoral program, I surveyed the literature on anxiety and assessments and stumbled upon self-regulated learning (SRL). As I delved into the vast research on the topic, I felt a renewed sense of hope for millions of students in India. Even though SRL dates back to the 1970s and research continues to reinforce its significance, there have been few studies in India on this topic. In the Indian context, studies are limited to associations among anxiety, high-stakes examinations, and parental expectations. The lack of research on SRL in the Indian context motivated me to design a research project that focused on understanding how students in India demonstrate SRL and potential ways in which they can be supported to become strong, autonomous learners. During my review of the SRL literature, I understood that certain conditions need to be met for students to become effective independent learners. And it was this understanding that led me to focus my research aims and objectives on SRL and formal classroom assessments in the Indian education context.

1.2 Research Problem

For decades, educational psychology research has focused on understanding how students gain mastery of their learning and identifying distinct attitudes and behaviours that result in academic achievement. The literature demonstrates that students who know *how* to learn (i.e., use subject knowledge and self-awareness in optimal ways) achieve success. This approach is referred to as self-regulated learning (SRL). SRL is an intentional process during which students manage, regulate, and evaluate their thought, motivation, and action to optimize their learning and attain academic success (Pintrich, 2002; Schunk, 2005; Zimmerman, 2013). Such intentional and proactive learners develop and exercise agency in the learning process, have increased self-efficacy, and gain lifelong learning competencies (Perry, 2015; Pintrich, 1996; Schunk & Pajares, 2005; Zimmerman & Schunk, 2011). Research examining SRL has also been linked to scientific literacy and conceptual understanding (Sinatra et al., 2012). Consequently, scholars advocate that self-regulated learning skills are necessary for learners to accommodate and adapt to scientific tasks such as problem-solving and critical inquiry (Andrzejewski et al., 2016; Sinatra & Taasoobshirazi, 2018). As such, SRL can help young people become intelligent consumers and producers of scientific understanding.

International and national achievement surveys paint a dire picture of students' scientific capabilities. A recent report by the Programme for International Student Achievement (PISA) report (2018) provided insights into students' performance across member countries of the Organization for Economic Co-operations and Development (OECD). The report stated that although over 78% of the student sample were able to explain and demonstrate an understanding of scientific concepts, only 6% were proficient in using scientific data to offer novel solutions through logical reasoning and critical thinking. In India, a recent National Achievement Survey by the National Council for Education Research and Training (NCERT, 2017) revealed similar findings to the PISA report. Over 50% of the sampled population of students were able to identify and describe scientific concepts, but only 30% of students were able to demonstrate higher-order thinking capacities. India's response to this growing concern is recognised in the recent National Education Policy (2020). This policy seeks to reform pedagogy and curriculum to promote scientific knowledge and skills among young learners. Given the evidence that SRL is pertinent to academic success and that students struggle to achieve success on assessments, the current thesis aimed to investigate students' SRL processes for a science assessment and how they relate to aspects of the assessment. The findings from this thesis have implications for developing countries (such as India) that are revising their policies.

A vast body of literature links SRL skills and processes with scientific competencies such as evaluating evidence, manipulating data, and arriving at unique and novel solutions (Sinatra & Taasoobshirazi, 2018). Although studies have associated SRL with science-related outcomes, the enactment of effective SRL processes is dependent upon environmental and individual characteristics (Hadwin et al., 2001; Pintrich, 1999; Zimmerman, 1989). Despite the best intentions of students and teachers to improve learning processes and achieve academic success, there are considerable challenges in adopting and implementing successful SRL processes. Many students have little understanding of how to approach learning in adaptive ways, so ensuring classroom processes effectively and explicitly support the development of SRL is critical (Omrod, 2011; Paris & Paris, 2001). Consequently, copious research has been conducted on identifying the classroom characteristics and practices that increase the likelihood of students developing appropriate SRL processes (Perry et al., 2006; Perry et al., 2020). Because classroom assessments provide information for teachers and students, they have begun receiving significant attention from SRL researchers (Greene, 2020).

In the recent decade, scholars have explicitly advocated for using classroom assessments to improve students' SRL (Panadero et al., 2018). Assessments, when done well, entail a vast set of processes, including establishing a purpose, assigning a task, setting criteria for student performance, providing feedback, and appraising student performance. Classroom assessments, broadly defined, involve gathering information from various sources to inform teaching and learning (Kane & Wools, 2020). These assessments differ in their forms, such as quizzes, tests, assignments, and projects, and can also serve multiple purposes (e.g., diagnosing gaps and problems in learning, motivating students, and monitoring learning progress; Kane and Wools, 2020). As such, classroom assessments may be informal or formal to fulfil their formative and summative functions. For instance, informal formative assessments may be characterised as unstructured interactions and conversations between students and teachers in the classroom (Bell & Cowie, 2002). Formal formative assessments, on the other hand, are planned and structured activities, and students are informed of them before the assessment (Chen & Bonner, 2020). Ultimately, the purpose of formative assessments is to gather information to inform future learning and teaching.

Black and Wiliam (1998; 2018) also argue that shifting the use of classroom assessments purely from an evaluative function to a more formative practice can help students improve their learning. As such, using data from assessments helps guide teaching and learning in the classroom (Black & Wiliam, 1998; Black & Wiliam, 2018; Black et al., 2011; Boekarts & Corno, 2005). Such a perspective on assessments provides students and teachers with opportunities to enact SRL processes effectively. For example, teachers can share performance criteria with students, and doing so will support students' SRL processes, such as goal-setting, monitoring, and self-reflection (Andrade & Brookhart, 2020). Recognising these overlapping processes has encouraged researchers to identify specific ways in which formative practices of classroom assessment promote SRL. This shift is opportune as many Asian-Pacific countries are moving toward low-stakes assessments that enhance students' learning processes (Harris & Brown, 2009; Koh & Luke, 2009).

There is extensive literature committed to understanding assessment-related interactions and SRL. Researchers found that sharing assessment criteria (e.g., rubrics) and providing quality feedback helps students develop SRL processes in the planning and reflection stages (Andrade & Valtcheva, 2009; Panadero & Alonso-Tapia, 2014). There is also evidence that assessments can be motivating for students, enhancing their self-efficacy, promoting a growth mindset, and help them become successful learners (Brookhart, 1999; Granberg et al., 2021). Furthermore, professional development for pre-service and in-service teachers that integrate self-assessment and feedback practices have been successful in improving students' SRL (Brookhart & Chen, 2015; Brown & Harris, 2014; Panadero et al., 2017). The scholarship on classroom assessments and SRL has grown substantially, and empirical evidence points to the value of assessments in improving students' SRL processes.

Much of the SRL and assessment literature is focused on formative classroom assessments, and formal summative assessments have largely been neglected except for Chen and Bonner's assessment framework (2020). Formal summative classroom assessments are used for evaluating and grading student learning at the middle or end of a school term (Chen & Bonner, 2020). These assessments may be characterized as cumulative assessments that aim to document how much students have learned and judge their performance against some standards (Gardner, 2010). Formal summative assessments differ from formative assessments primarily in their function and timing of administration. In terms of functionality, summative assessments are typically used to report on student performance and are high-stakes because of the grading attached to them (Gardner, 2010). As far as the timing of administration is concerned, these assessments are less frequent and occur at the end of instruction, such as in the middle of a term or the end of a term. Finally, these assessments commonly take the form of exams and term papers (NCERT, 2020). Table 1 highlights the distinguishing features of the different types of classroom assessment.

Table 1

Characteristics of the Different Types of Classroom Assessment

	Informal	Formal
Formative	Spontaneous, unstructured interactions with students, not graded	Planned, structured activities and students are informed, not graded
Summative	Planned, low-stakes, non-standardized, graded (e.g., weekly tests, projects)	Planned, structured, standardized, high-stakes, and graded (e.g., exams)

Understandably, informal formative assessments are more viable for developing SRL skills, while formal assessments are ideal for students to implement SRL skills to achieve success (Chen & Bonner, 2020; Clark, 2012). In India, informal, classroom-based formative assessments are conducted through Continuous and Comprehensive Evaluation (CCE). This system was introduced in 2009 by the Government of India to encourage teachers to use assessments for formative purposes (Berry et al., 2020). Research on the efficacy of this programme shows that this system failed to achieve its objectives due to limited professional development and learning opportunities provided to teachers (Berry et al., 2020; Yagnamurthy, 2017). On the other hand, formal summative assessments are compulsory, planned, and regularly occurring assessments in classrooms across India. Formal assessments in India tend to take the form of examinations or tests in the middle and end of an academic term. These aspects make formal summative assessments prime avenues for integrating SRL practices to support students to become strong and independent learners in India. For the remainder of the thesis, formal summative assessments refer to the exams.

Unpacking classroom assessment processes and linking them with SRL provides meaningful avenues for improving and supporting teachers' assessment practices and students' SRL processes. A critical constituent of the assessment process is the task presented to students. Early research highlights the central role of academic tasks in the classroom (Davis & McKnight, 1976; Doyle & Carter, 1984; King, 1980). Since formal assessment tasks are used to measure and grade student performance students spend considerable time focusing on how best to accomplish the task (Crooks, 1988). Research findings suggest that the design of an academic task influences students' interactions with learning and performance in the classroom (Lodewyk et al., 2006; Perry et al., 2004). For example, when tasks required independent work and engagement in higher-order thinking skills, students were more likely to seek guidance from teachers (Doyle & Carter, 1984; Lodewyk et al., 2006). Because the nature of the task was ambiguous, students struggled to accomplish the task without help-seeking strategies. However, there is evidence linking these tasks to improved self-regulated learning processes because students extend autonomy over their learning (Lodewyk, et al., 2006; Perry et al., 2006). Designing formal classroom assessment tasks that facilitate the development of SRL is, therefore, essential.

Furthermore, literature alludes to the preference of some assessment practices over others, particularly those that provide quality feedback, foster a growth mindset, and facilitate reflection (Andrade & Brookhart, 2020; Hattie & Timperley, 2007; Panadero et al., 2018). This means that teachers must make crucial decisions about their classroom assessment, from identifying its purpose to assigning the task to evaluating student performance. These decisions significantly impact instructional practices in the classroom and subsequently contribute to students' understanding of the assessment process and improved academic achievement (Bearman et al., 2016; McMillan, 2003; Stiggins, 2001). It is also established that teachers provide students with directions and instructions on how to approach and complete the task (Doyle & Carter, 1984). Students learn how to approach learning, including self-regulatory processes, through such engagement. These formal classroom assessment decisions are essential components that contribute to student learning because, in addition to the content, it expresses implicit messages of expectations and goals, leading to internalising learning and performance standards (Brookhart, 2004).

Therefore, understanding teachers' design decisions regarding assessment tasks is essential, as they can affect students' learning processes and outcomes.

Although limited, the literature on learning and assessment in India indicates that teachers and students experience significant challenges. For instance, studies indicate that students face pressure to perform on assessments because outcomes determine their opportunities and matter to their family's social status (Deb et al., 2015; Siahi & Maiyo, 2015). Such emphasis on external rewards and pressures likely results in over-testing in the classroom and students typically use rote memorization to achieve performance outcomes (Berry et al., 2020; Julius & Evans, 2015; NCERT, 2005). Furthermore, educational policies also advocate using classroom assessments for formative purposes; however, they provide teachers with little guidance on how to implement these practices (NCERT, 2005; 2020). These challenges are evident in the discourse on learning and assessment. For instance, Crooks (1988) found that teacher-designed tasks are less focused on higher-order thinking capacities, thus limiting how students approach their learning. There is also evidence to indicate that students choose ineffective regulatory strategies (e.g., crammed study sessions) because they are unaware of appropriate strategies, or the task is poorly designed (Omrod, 2011). And finally, national and international achievement tests continue to highlight concerns regarding student performance in scientific literacy and competencies (PISA, 2018; NAS, 2017). The breadth of research on assessments in Western educational contexts illustrates how fundamental assessment practices are to learning and the importance of context for understanding how assessment practices impact students' SRL.

With this in mind, there is scarce literature on teachers' classroom assessment decisions and task design in compulsory educational settings, particularly in the Indian context. Indeed, researchers have strived to understand and identify potential pathways that promote SRL and vouch for classroom assessment practices as useful to achieving this end. The research linking SRL to formative classroom-based assessments provides reasonable grounds to expand our understanding of the relationship SRL can have with formal classroom assessments. Formal classroom assessments are critical to students and teachers in India and widely across contexts. Therefore, by understanding the relationships between assessment task design and teachers' decisions and students' existing SRL approaches, it may be possible to design assessment practices that intentionally promote stronger SRL processes. In response to these areas of inquiry, the primary objective of this thesis is to understand and analyse students' SRL for a science classroom assessment concerning how tasks were designed and why teachers designed them that way. The following research questions guided this project:

1. What SRL processes do students demonstrate in the forethought and self-reflection phases of Zimmerman's model for a formal science classroom assessment?
2. In what ways do the design features of the assessment task relate to students' SRL processes in the forethought and self-reflection phases of Zimmerman's model?
3. What are teachers' intentions regarding decisions of a formal classroom science assessment?
4. In what ways do teachers' assessment intentions relate to students' SRL processes in the forethought and self-reflection phases of Zimmerman's Model?

1.3 Research Context

India's educational context is complex, comprising private and government (public) schools. Public schools are expected to follow the guidelines in the National Curriculum Framework (NCF, 2005) and National Education Policy (NEP, 1986; 2020). These documents provide a framework for curricular objectives at the different stages of schooling and assessment expectations. Public schools are affiliated with either the state board or the Central Board of Secondary Education (CBSE). Private schools have more freedom in choosing their affiliation, which includes the Indian Certificate of Secondary Education (ICSE), the Indian equivalent of the Cambridge curriculum. Private schools can also affiliate themselves with internationally recognized organizations, which include the Cambridge International General Certificate of Secondary Education (IGCSE) and the International Baccalaureate (IB).

For this thesis, I have narrowed the research context to the three national educational boards: state board, CBSE, and ICSE. Schools affiliated with these boards might vary in the depth of the content they are expected to teach and how subjects are grouped, but they follow the NCF and NEP guidelines on the topics to be covered and the assessment formats. In India, formal classroom assessments (e.g., term exams) for science subjects depend on how the subjects are grouped, and informal classroom assessments (e.g., weekly tests, unit tests) for science subjects depend on how they are taught. For example, in state board schools, science is grouped into two: physical and biological sciences for formal assessments such as exams but not for informal assessments. In these situations, there will be two formal assessments: one for physical science and another for biological sciences.

While informal classroom assessments maybe vary based on how the subjects are taught. For instance, there may be one teacher for physics, chemistry, and biology, and teachers may choose how and when they conduct informal classroom assessments (NCERT, 2005). Formal classroom assessments are typically higher in stakes because they are graded and serve a summative function. Meanwhile, informal classroom assessments can serve a formative or summative function (and are lower in stakes). This thesis focuses on formal classroom assessments, in particular examinations, conducted during the middle of the term or at the end of the term. In the next paragraph, I explain how exams are similar and different among state board, CBSE, and ICSE schools.

As previously mentioned, state board schools have grouped physics and chemistry together as physical science, and botany and zoology, and CBSE schools assess science as one subject. And finally, ICSE schools have a distinct assessment for physics, chemistry, and biology. In most schools, two mid-terms occur during the term and one end-of-term exam. All of these assessments are graded, and the task (i.e., question paper) is designed, conducted, evaluated, and graded by the subject teacher and approved by the head of the department or senior teachers (NCERT, 2020). These science assessments are scheduled in the academic calendar for the year along with the other subjects (e.g., math, languages, social sciences). For mid-term assessments, students sit for exams between a week to 10 days, while for end-of-term exams, it is over a more extended period (between two – three weeks).

Since this research project was conducted during the Covid-19 pandemic, I will briefly explain the impact of the pandemic on the examination process. All formal assessments were conducted online via video conferencing platforms (e.g., Zoom, and Google Classrooms). Some schools had only one exam at the end of the year, while others tried to meet the schedule set out in the calendar. Depending on the school and socio-economic conditions, students had access to a computer, laptop, or smart phone. Teachers adapted to teaching online and conducted assessments as closely as possible to the pre-pandemic scenarios. For their formal assessments, students were presented with a question paper shared via Google Classrooms or other electronic communication (e.g., e-mail). They were given a limited timeframe to respond to a set of questions. Students returned their answer scripts via similar electronic communication. Teachers evaluated and graded students' performance and then reported back to students and parents.

1.4 Overview of the Thesis

This thesis comprises two distinct studies. As a result, the structure of the thesis is designed to develop a comprehensive narrative of each of the investigations conducted. Given that self-regulated learning is the core focus of the project, I provide an overview of the available literature in Chapter 2. I briefly position my argument for assessments because I cover them in further detail in the literature review for each of the two studies. I have endeavoured to reduce overlap among Chapters; however, the core ideas foundational to the thesis are repeated. With respect to methods, I have explained my methodological decisions in Chapter 2, which include my philosophical foundations as a researcher and how SRL is situated in the context of measurement. The details regarding how I conducted the research are included in the methods sections of the relevant studies.

Chapters 3 and 4 are separate studies conducted on SRL for this thesis. In Chapter 3, I begin with a thorough description of students' planning and reflection processes for the science classroom assessment and differences in these processes based on educational boards. I also provide an in-depth examination of the assessment task structure and its relationship to SRL. Chapter 4 is an investigation into teachers' assessment decisions and students' SRL processes. Similar to the literature review, each study has its own discussion section which reflects on the findings in that study. In Chapter 5, I provide a comprehensive discussion of SRL in the context of classroom assessments. I draw on the conclusions of the two studies to describe a working model that captures the key assessment decisions to be made by teachers and how they can positively contribute to the development of independent learners who use strong SRL processes. I conclude the Chapter with implications for future educational practice and research.

Chapter 2: Literature Review and Methodological Framework

In the previous Chapter, I introduced the topic for this multi-study thesis, and in this Chapter, I provide a background for this thesis project. I begin by describing SRL and the social cognitive framework used in this thesis. I proceed to draw attention to the areas of inquiry in the context of SRL for formal classroom assessments and argue for using contextualised measurement tools to investigate relationships between the two. In this Chapter, I also explain the research objectives accompanied by methodological decisions.

2.1 Self-regulated Learning

Self-regulated learning (SRL) is a dynamic and complex construct. SRL describes how individuals use their agency to plan, strategise, and sustain their actions to reach their desired learning goals (Pintrich, 2000; Zimmerman, 1989). The multifaceted nature of SRL allows researchers to examine this process from various perspectives (e.g., cognitive, metacognitive, motivational, and emotional). This diversity has resulted in different models of SRL being proposed over the years (Boekarts, 1996; Efklides, 2011; Jarvela & Hadwin, 2013; Pintrich, 2000; Winne & Hadwin, 1998; Zimmerman & Moylan, 2009). While the models differ on how SRL occurs, there is an agreement on the core elements that characterise the phenomenon. For example, all models describe SRL as purposeful and deliberate, which means SRL is a goal-directed process. Additionally, theorists agree that students self-regulate before, during, and after completing a learning activity or task. And finally, researchers acknowledge that contextual, individual, and biological factors impact SRL processes.

Although all models are validated and substantiated with empirical evidence, this dissertation is grounded in Zimmerman's SRL model (2008). As an early researcher of SRL, I found Zimmerman's model accessible while still capturing the complex processes within the SRL phenomenon. Additionally, a large body of literature across academic and non-academic disciplines have contributed to the validity and reliability of this SRL model (Panadero, 2017). And finally, the social cognitive framework from which this model was developed acknowledges the environmental factors that influence SRL (Zimmerman, 1989). According to the social cognitive perspective, SRL is a process wherein individuals use their agency to manage and regulate personal and environmental resources to achieve a learning goal (Schunk, 1989; Zimmerman, 1989). Social cognitive theory emphasises that individuals use their cognitive skills (e.g., goal setting) and summon motivation and volition (e.g., self-efficacy) to enact actions (e.g., stay on task) in intentional and strategic ways to optimise their learning or modify their processes when challenged (Bandura, 1986; Pintrich, 2000; Zimmerman, 1989). From this view, Zimmerman (2001, 2009) proposed a cyclical model to describe the cognitive processes and motivational beliefs that individuals activate for learning. According to this model, SRL comprises three distinct but interrelated phases: forethought, performance, and self-reflection. Before unpacking this model, I explain social cognitive theory in relation to SRL and, in the process, justify my reason for choosing this conceptualisation of SRL.

2.1.1 Social Cognitive Theory and SRL

In the 1970s, social cognitive theory challenged the then-dominant behavioural and information-processing theories of learning. According to social cognitivism, individuals learn by observing and modelling others' behaviours and can learn vicariously through others' behavioural consequences. In contrast, behaviourists postulate that learning results from trial and error and the response to the resulting consequences of the behaviour. Bandura (1977) theorised human beings have an agentic capacity, which implies they can set goals and meet them strategically. Human beings regulate, monitor, and evaluate their thoughts, feelings, and actions to achieve predetermined goals. Also, human functioning is embedded in social structures, which means that individuals interact with each other in complex social systems. These interactions create opportunities, sustain and impose constraints for regulation, and strengthen the social cognitive perspective of learning as interactive and reciprocal between human beings and the social structure. In this way, this theory challenged the assumptions of behaviourism and cognitivism that learning was unidirectional. Instead, social cognitive theorists proposed a unifying view of learning that encompasses reciprocal interactions among personal agency, social systems, and behaviour (Bandura, 1991). This core assumption of social cognitive theory distinguishes it from other approaches. It makes it the most suitable for this thesis project because classroom assessment practices are a social contract between teachers and students.

Bandura (1988) describes this reciprocal relationship among persons, environment, and behaviour using a triadic causal determinism model. According to this model, behaviour, cognition (personal factors), and the environment interact bi-directionally, and these reciprocal interactions determine an individual's behaviour. Although each factor can influence one another, it is not necessary that these influences are simultaneous or that the strength of the influence is equal among the three factors (Bandura, 1986, 1989). For example, a student may answer questions (behaviour) when a teacher uses open-ended prompts (environment) to facilitate student responses, thus increasing the behaviour. Or a student who feels confident (personal) in their abilities to grasp abstract subjects (e.g., physics) might work harder (behaviour), and in turn, their success in the subject increases their confidence. The inclusion of environmental influences differentiates the social cognitive theory from other learning frameworks. By recognising the impact of the environment on an individual's learning and behaviour, the social cognitive theory emphasises that individuals exhibit their competencies (e.g., skills, beliefs) within a specific context rather than as a generalised approach. Therefore, the social structure, task type, or even task characteristics (e.g., difficulty) often determine the degree to which one applies their competencies or whether to choose to use them at all (Bandura, 1986; Usher & Schunk, 2018).

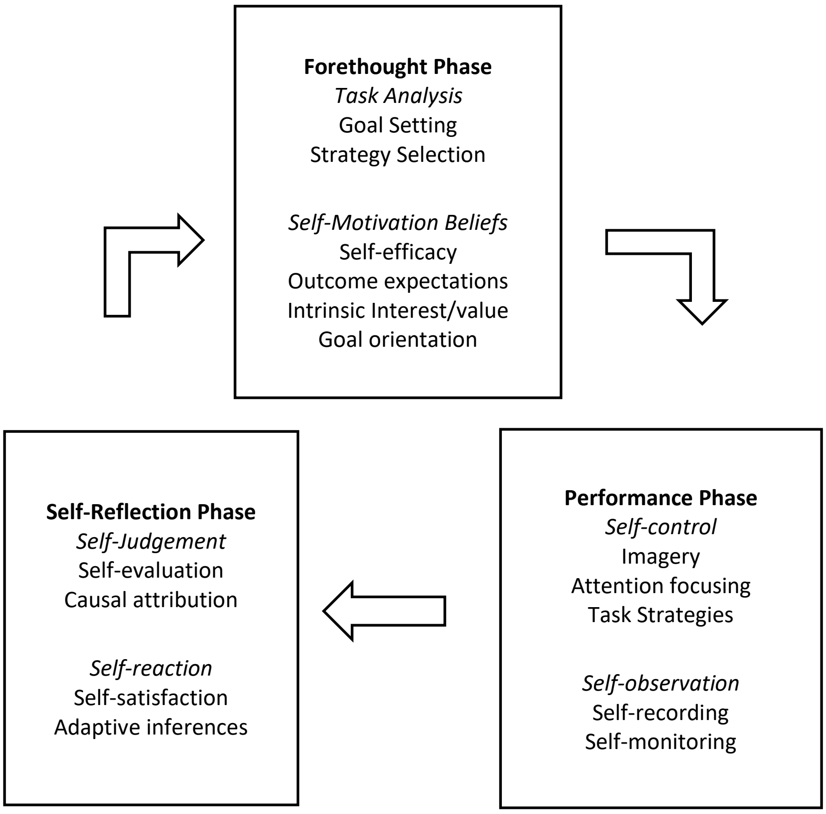
Social cognitive theory also differed from behavioural theories by assuming that individuals exercise personal agency, which means that human behaviour is goal-directed. Bandura (1977) proposed that humans can choose to restructure their environment and employ individual strengths to meet their desired goals. This process is referred to as self-regulation. According to Bandura (1986), self-regulation occurs due to self-observation, self-judgement, and self-reactions. Individuals observe their own behaviours over time and reiterate goals and strategies to meet specific self-set standards. These personal standards are also the basis on which individuals compare their behaviours and outcomes. Finally, they engage in self-reaction processes that function as feedback regarding their performance and self-judgements. Depending on the performance or outcomes, individuals feel satisfied and continue to reiterate their behaviour or experience dissatisfaction which may result in defensive behaviours such as self-handicapping. These aspects of social cognitive theory and self-regulation influenced Zimmerman's model (1989; 2009), which is the foundational framework of this thesis.

2.1.2 Overview of Zimmerman's Model (2009)

Zimmerman (2009) adapted Bandura's self-regulation model to account for motivational and cognitive constructs in the learning context. In his model, SRL comprises three cyclical and reciprocal phases: forethought, performance, and self-reflection (see Figure 1). In each stage, self-regulated learners activate cognitive processes and summon motivational beliefs. In the forethought phase, self-regulation occurs through task analysis and self-motivating beliefs. Ideally, effective self-regulated learners deliberate on the task and accordingly set goals and select strategies, which are informed by self-beliefs (e.g., self-efficacy), achievement values, and outcome expectations (Zimmerman, 2009). In the performance phase of the SRL cycle, students invoke volition to sustain their strategic actions in their goal pursuit. Zimmerman (2005) describes two volitional control processes characterising this phase: self-observation and self-control. This process allows learners to engage in the third phase, self-reflection, to make necessary changes to optimise performance.

Figure 1

Zimmerman's Self-Regulated Learning Model (2009)



The third phase in the cyclical SRL model is self-reflection comprising two closely related processes: self-judgment and self-reaction. After completing a task or a learning session, students appraise their performance according to a set of standards (e.g., normative, mastery, or social comparison) and attribute their performance to internal or external factors (Zimmerman, 2005). The self-judgment processes are crucial because they influence how learners react to their own performance. Self-reaction is a powerful indicator of how students will approach a similar task in the future (Bandura, 1981; Zimmerman & Moylan, 2009). Self-reaction, according to Zimmerman (1989), manifests in two steps. First, learners perceive a sense of satisfaction or dissatisfaction; second, they proceed to make inferences which may be adaptive (e.g., modify their strategy) or defensive (e.g., procrastination, avoidance). Research suggests that learners' perceived satisfaction is generally found to be influenced by the value of the task (Greene, 2018). A learning cycle is understood to be complete once the processes in the self-reflection phase inform those in the forethought phase.

Research examining SRL from this perspective reiterates its context-specific nature, suggesting that students' learning approaches vary across academic tasks (Hadwin et al., 2001). For example, a student is likely to self-regulate differently for a math task compared to a science task. These approaches are different because of the nature and demands of different disciplines (math vs science) and students' personal dispositions (e.g., motivational beliefs) concerning the task (Usher & Schunk, 2018). A student who feels efficacious in science is likely to be more motivated to regulate their learning in successful ways than a student who does not feel confident about the subject. In addition to students' responses to the task, social cognitive theory emphasises the reciprocal nature of learning. This would mean that environmental factors (e.g., teachers' decisions, task structure) are likely to inform students' SRL processes (Bandura, 1986; Lodewyk et al., 2006; Perry et al., 2004; Zimmerman, 1989). Therefore, it is equally important to understand the context within which students are situated to investigate students' SRL processes.

2.1.3 SRL in the Science Learning Context

In the case of science learning, Sinatra and Taasoobshirazi (2018) elucidate the implicit connections between SRL processes and scientific competencies. For instance, SRL entails using cognitive functions such as memory and sense-making, which aids in conceptual knowledge acquisition. This relationship can be seen in physics, wherein students who deeply understand the subject matter have enhanced strategy use

(Taasoobshirazi & Carr, 2009). Moreover, having a solid grasp of the subject is necessary for students to engage in complex thinking processes such as argumentation, inquiry, and reasoning (Anderman et al., 2012; Asterhan & Schwarz, 2016; Schraw et al., 2006). Evidence suggests the inverse is also true: students with a sophisticated repertoire of strategies are more likely to obtain scientific knowledge (Kuhn, 2006).

In addition to strategy use, scientific thinking requires goal-setting strategies and self-reflection skills to modify and improve learning (Hoffman & Spatariu, 2009). Much of the literature on SRL and science focuses on the metacognitive component, which includes strategy use and monitoring for problem-solving. In particular, scholars have focused on understanding the what, how, why, and when of strategy use when solving problems (Sinatra & Taasoobshirazi, 2018). Findings suggest that stronger problem-solving skills are associated with metacognitive processes, and students with weaker problem-solving skills struggle to explain their approaches to learning (Elfkides, 2001).

Indeed, substantial evidence indicates that SRL is necessary for success in scientific competencies. But the development of effective SRL skills is, in part, dependent upon tasks presented to students and teachers' instructions (Lodewyk et al., 2005; Perry et al., 2020). In India, students experience fewer opportunities to problem-solve or engage in complex scientific tasks because formal classroom assessments such as exams receive far more attention (NCERT, 2005). Instead, teachers focus on completing the syllabus and helping students acquire low-level learning objectives that are likely to be tested (Berry et al., 2020). The significance of performance outcomes is pervasive in Indian society and is, therefore, unsurprising that teachers and students focus on learning and instruction that helps meet the test requirements (Geist, 2010). Pellegrino (2013) argues for external and internal tests to match the competencies required for science. The author suggests that assessment design needs to consider how best to develop assessments that measure these skills and for resources to support instruction. Given the impact of assessment design on learning approaches, several aspects of this relationship need to be understood further to develop practices that support SRL.

2.2 Linking SRL with the Assessment Context

Classroom assessments are opportunities for students and teachers to gather and use information to improve learning and performance. Teachers may use the information to provide students with feedback, report student progress to parents and school management, and also gain valuable insights for instruction (Andrade & Brookhart, 2020; Peggy & Chen, 2021). For students, classroom assessments might be an opportunity to *show* teachers and parents, or check for themselves, how much they have learned (Cowie, 2005). Students could also use classroom assessments to check their understanding and strategies and identify areas for improvement. In other words, classroom assessments are prime opportunities for students to develop and demonstrate effective SRL processes such as goal-setting, strategy selection and implementation, and self-reflection. This recognition of overlapping processes between SRL and classroom assessments has led to theoretical frameworks that link the two, which outline how classroom assessment practices help improve students' SRL processes (Andrade & Brookhart, 2020; Bonner & Chen, 2019; Clark, 2012; Panadero et al., 2018; Wiliam, 2007).

Contemporary frameworks stress the role of classroom assessments in developing successful SRL processes among students by combining the principles of behaviour, cognition, metacognition, and motivation. For instance, Andrade and Brookhart (2020) use a social cognitive model of SRL to establish a relationship to classroom assessments. They propose a collaboration between students and teachers throughout the assessment process. For example, at the beginning of a unit or lesson, teachers help students set goals using success criteria (i.e., learning objectives) that need to be achieved. During and after completing a learning activity, teachers can ask students to reflect and evaluate their progress and develop a course of action for the next learning phase. The authors advocate for teachers to provide timely constructive feedback to help students monitor and regulate their cognition, motivation, and behaviour. Research suggests that content-specific feedback which supports students' self-assessment strategies improves academic achievement and enhances students' volition (Chamberlin et al., 2018; Moos, 2011; Panadero et al., 2018). Indeed, classroom assessments grounded in practices that seek to improve student learning are powerful opportunities to develop SRL.

Chen and Bonner (2020) conceptualise classroom assessments as a wider variety of activities to account for the theoretical and practical limitations of existing frameworks. For example, they recognise that even though formative uses of assessments are preferred, teachers also use classroom assessments for summative purposes such as evaluating and reporting student achievement. Evidence indicates the negative impact of grading and evaluation on feelings of anxiety and avoidance (Chamberlin et al., 2018). However, empowering students on how to enact learning strategies that strengthen motivation can mitigate the stress around assessment (Kramarski et al., 2010). Ample evidence in Western countries provides evidence that SRL interventions help improve achievement gaps (Cleary & Platten, 2013; Panadero et al., 2017). Yet, explicit teaching of SRL strategies for formal assessments in Eastern countries such as India remains unexplored.

Research on SRL indicates that students benefit from explicit instruction, without which they continue to use faulty strategies (Greene, 2021; Perry et al., 2021). Although teachers endorse learning views consistent with SRL, they rarely integrate strategies into their instruction and less for formal assessments (Dignath-van Ewijk & van der Werf, 2012). Chen and Bonner (2020) suggest incorporating instruction with opportunities for students to develop SRL strategies through intentional assessment design to account for this challenge. For instance, prior knowledge assessment could serve forethought processes such as goal-setting, consistent monitoring and reflection of learning, and strategy use during classroom activities and tests.

So far, theoretical frameworks neatly tie SRL processes with informal, formative classroom assessment practices. However, little attention is paid to the design of assessment tasks, particularly formal assessments, such as those that occur at the end of an instructional unit or school term and their relation to SRL (Chen & Bonner, 2020). This limited focus on formal assessments may be due to its performative nature, which perhaps reduces its potential for the development of SRL. However, it is for these assessments that students often need to demonstrate a sophisticated approach to learning and engage in effective SRL skills (Chen & Bonner, 2020; DiBenedetto & Schunk, 2021). Indeed, students should enact successful strategies, but evidence suggests that the task needs to warrant the use of effective SRL practices (Perry & Dowler, 2004). Teachers typically develop and administer formal classroom assessments, and early studies highlight that such tasks primarily focus on recall and retrieval of subject matter rather than re-constructing understanding (Crooks, 1988; Doyle, 1982). These assessments are less likely to trigger or encourage students to activate self-regulated learning processes because students might adopt a surface approach to learning (remembering and understanding) the subject matter rather than deeper approach (analysing and reasoning) of the concepts (Hattie & Donoghue, 2016).

Studies suggest that students prepare for assessments based on task criteria (Perez-Garcia, 2020). For example, undergraduate students use surface-level strategies such as rehearsal and comprehension when studying for multiple-choice quizzes. In comparison, deeper learning strategies are preferred by students for essays that focus on organisation and elaboration (Garcia-Perez et al., 2020). In school, formal assessments are perceived as performance tasks that encourage students to use SRL strategies that focus on outcomes (Pintrich, 2000). Some scholars examined SRL for tasks distinguished as ill-structured or well-structured and found that goals and motivation differed based on the structure (Lodewyk & Winne, 2005). Accordingly, it could be reasoned that students must align their strategies to the task, but the task must also evoke desirable SRL processes.

The impact of task design on SRL is critical, especially when considering scientific competencies and skills. Tasks that focus on declarative content knowledge with little attention given to higher-order processes skills do not support deep learning approaches, including self-regulated learning processes (Crooks, 1988; Hattie & Donoghue, 2016). These tasks limit students' scientific competencies, such as solving equations, hypothesising, and evaluating evidence (Chi, 2006; Kuhn, 2006; Lombardi et al., 2018).

Despite this evidence, exams dominate formal assessment tasks in India and primarily focus on evaluating students' content knowledge (NCERT, 2020). This narrow focus encourages students, teachers, and parents in India to view assessments as high stakes instead of a more diagnostic view (Alhuwalia & Singh, 2022). Such a performance-oriented perspective reduces the likelihood of using assessment data to acquire mastery of the subject matter. Nonetheless, teachers and students could re-direct their attention from performance to mastery through careful investigation of the assessment process. Given the established evidence regarding task design and SRL and the ubiquitous use of formal assessments in India, it would be useful to investigate how the task design of a formal classroom assessment in science design relates to SRL processes.

Furthermore, formal assessment tasks are of immense significance to students and teachers in the Indian context because performance outcomes determine future opportunities (Deb et al., 2015). These assessments measure state or national standards and cover multiple objectives across topics to grade students' performance. Teachers inform students about upcoming examinations and advise them on what will be tested. Indeed, a vast body of literature supports formative assessment practices that promote SRL, such as types of feedback and strategy instruction (Andrade & Brookhart, 2020; Hattie & Timperley, 2007). Moreover, it is well-accepted that formal assessments, which often serve a summative function, could also be used for formative purposes (Black & Wiliam, 2018;

Broadbent et al., 2018). This rich literature base provides sufficient evidence to consider teachers' practices for formal assessments and how they might support SRL.

Research on teachers' classroom assessment literacy indicates that intentional design decisions significantly impact learning outcomes (Mellati & Khademi, 2018). Understanding how teachers make these design decisions for formal assessments is significant because teachers have autonomy in the classroom and their decisions impact students' learning and performance outcomes (McMillan, 2002). In India, teachers administer formal assessments at the middle and end of a term. These exams are scheduled into the academic calendar for the year. Therefore, teachers make crucial decisions regarding formal assessments. These decisions include setting an intention, designing the task, conducting the assessment, interpreting, and evaluating student performance, and feeding back data for future learning and instruction (Bearman et al., 2016).

Even though national policies in India make recommendations on enacting these steps, there is little research examining the teachers' formal assessment decisions in the classroom. According to a study by Brown et al. (2015), teachers in India believe that by teaching for exams, they are improving student learning. Although curricular frameworks push for a formative view of assessments and implement practices that aid in diagnosing learning, few teachers implement assessments this way. The national and state governments launched an initiative called the Continuous and Comprehensive Evaluation (CCE) system to support and encourage formative assessment practices. Still, schools have interpreted this suggestion as frequent, periodic testing through different 'activities' (e.g., oral tests, projects). Although these activities may be perceived as opportunities for developing SRL, it is not necessarily true that all students view these tests as opportunities. Additionally, policies that focus on frequent testing are likely to encourage teachers and students to focus more on performance and may limit their ' time and instruction to support SRL (NCERT, 2020). Such views contradict those supporting the development and enactment of effective SRL processes (Berry et al., 2020). Ultimately, in India, performance outcomes of formal assessments or exams supersede the mastery view of learning because they determine learning achievement even though such perspectives strongly influence students' self-concept as a learner and instructional strategies for teachers (Brookhart, 2001).

In sum, SRL processes are essential for young people's short- and long-term academic success. Extant literature indicates the value of the SRL sub-processes for science learning, mainly because scientific tasks require effortful deliberation, problem-solving, and critical thinking. More importantly, how students approach tasks and reflect on their performance is, in part, determined by environmental conditions. Therefore, when investigating students' SRL processes, exploring the context within which they are enacted is useful. Therefore, the current thesis investigates students’ SRL processes in the forethought and self-reflection phases outlined in Zimmerman’s model (2009). These two phases were chosen for practical reasons. First, these sub-processes were measurable for the task identified. The subprocesses in the performance phase were more challenging to measure because they were related to metacognitive monitoring and self-control that students would employ while studying for the exam at home. The second reason was to identify ways in which these phases can be integrated into design and development of formal classroom assessment practices that encourage students to become strong, independent learners.

In India, schools and families emphasise assessment outcomes, and much of the classroom learning and instruction serves to help students succeed in their examinations. Despite examinations receiving overwhelming attention, students' scientific understanding and competencies are declining (NAS, 2017). Moreover, students in India continue to employ surface-learning habits such as rote memorisation. Western literature on SRL suggests that classroom assessments are promising for the development of SRL and are prime opportunities for students to reflect upon and improve their learning strategies. In particular, research suggests that task characteristics and specific teacher assessment practices effectively promote and sustain SRL.

From this analysis of the literature, it is clear that more research is needed to understand students' self-regulated learning processes for formal classroom assessments, particularly in India. Despite policy recommendations and changes in curricular frameworks, students continue to struggle with science learning achievement. Therefore, a deeper exploration of the assessment context in India is warranted. This thesis attempts to build on these areas in the literature and offer insights into the Indian context, including investigating assessment tasks and examining teachers' assessment decisions to help empower students' SRL processes to achieve success.

2.3 Methodology

In the previous sections, I have described relational processes between SRL and classroom assessments. I have argued that a more contextualised understanding of classroom assessments, particularly formal ones, is needed to further our understanding of how SRL can be promoted among students. In the recent decade, multiple approaches to measuring SRL have been developed that pay attention to the context-specific nature of SRL. Additionally, formal classroom assessment aspects such as task design and teacher decisions are relatively under-researched and require the development of instruments for analyses. Therefore, I had to decide how to best conduct this research and guide my inquiry. I chose a mixed-methods approach guided by a pragmatic worldview, which I explain in detail in this section.

I begin by outlining a pragmatic worldview that has influenced my interpretations of SRL literature and my methodological decision-making for this thesis. After this, I discuss the research methods that have dominated SRL research and the use of microanalytic protocols that will advance the field, mainly because it enables SRL practicality due to its context-specific nature. I follow this with an explanation of the ethical considerations undertaken, with particular attention given to how I approached the collection and use of data in an ethical way. Finally, I describe how I integrated my research questions within the microanalytic protocol and the methodological decisions for the classroom assessment factors.

2.3.1 Philosophical Paradigm

There are several paradigms in which educational research is grounded. Essentially, a paradigm is a worldview that constitutes a set of beliefs that informs the researcher about what counts as knowledge (Cohen et al., 2018; Kuhn, 1962). A paradigm guides the research process with shared beliefs about what knowledge is and how one comes to know it (Cohen et al., 2018). Although there are several paradigms, this thesis is guided by the assumptions rooted in pragmatism, particularly Deweyan Pragmatism (Morgan, 2017), which rests upon the belief that knowledge results from the relationship between reflective actions and their consequences.

Pragmatism

Pragmatism refers to a process of knowing that arises from action and consequences. Deweyan Pragmatism comprises three core notions. First, behaviours and actions are inseparable from one's context and situation, so there can never be one universal truth. Since knowledge is the relationship between actions and consequences, what we learn are possibilities rather than certainties (Biesta, 2015). From this view, pragmatists stress the concept of warranted beliefs. In other words, as we repeat actions, we learn the outcomes of acting in one way, and these repeated experiences produce warranted beliefs. The second notion follows closely with the previous one: because knowledge comes from the consequence of actions, it evolves. Since experience is rooted in situations and contexts, no two experiences can be the same (Biesta, 2007; Morgan, 2017). This variation of experiences comes from the belief that actions taken in a situation are conditional, and one can only act from what you know about the likely outcomes of that action. Therefore, the meaning attached to an outcome changes over time as the consequences of those actions change. And finally, the third belief is that all actions depend on socially shared beliefs about worldviews. Simply put, pragmatists view the world with a transactional lens and suggest that we can only know the world in the way we interact with it, and because we need to interact with each other, we tend to share beliefs and construct an “intersubjective” world (Biesta, 2015, p.19).

Since pragmatists view knowledge as the *relationship* between actions and consequences, it liberates one from seeing the world through an objective or subjective lens. And instead, allows the researcher to understand knowledge as the consequence of an action that can change as actions change (Biesta, 2018).

Pragmatism in the current research

This doctoral thesis aimed to understand students' self-regulated learning processes and how they relate to classroom assessments. Essentially, the purpose reflects a Dewyen perspective, which is understanding or gaining knowledge about a relationship that has consequences for future actions. I subscribe to a pragmatic paradigm because it allows me to combine the epistemologies of post-positivism and interpretivism to address the research questions. On the one hand, self-regulated learning (SRL) is theory-laden and observed as a universal phenomenon dependent on the context: a post-positivist notion of knowledge. On the other hand, classroom assessments are subjective events that are heavily influenced by one's social context and interpretations of the process and are, thus, best examined from an interpretivist lens.

Furthermore, a pragmatic approach to research follows the process of inquiry that begins with identifying a problem or question that needs to be addressed and choosing appropriate methods to achieve that goal. This approach links to the notion described above of actions and consequences. It was crucial for me to combine the different strengths of qualitative and quantitative methods to answer the research question. Upon understanding the theoretical foundations of SRL and classroom contexts, I was intent on ensuring that the context was included to help establish an understanding of the relationship between the two. As a result, a mixed-methods design that integrated quantitative and qualitative methods to drive the inquiry process was most suitable for this thesis (Morgan, 2018).

Finally, a pragmatist approach entails selecting methods for collecting and analysing data based on a thorough reflection of actions and potential consequences. Simply put, the decision-making process is needed to ensure a fit between the options available and the likely outcomes of each choice (Morgan, 2018). I use qualitative and quantitative data collection methods in this doctoral thesis but rely on quantitative methods for data analyses. This is explained in detail under the methods section in the studies described in Chapters 3 and 4.

2.3.2 Methodology

The current study aimed to understand how students self-regulate their learning and determine if a relationship exists between SRL and an assessment task. A mixed-methods research (MMR) design was the most suitable choice for the current thesis project because it enabled me to answer the research questions by integrating aspects of qualitative and quantitative approaches. MMR designs are differentiated from other mono-methods by the timing of the data types. Determining when which type of data is most useful is a critical decision to be made by the researcher. More specifically, I chose to use a conversion mixed methods design which is characterised by transforming one type of data into another (Teddlie & Tashakkori, 2009).

2.3.3 Methods

In this section, I briefly describe the methods of data collection and analyses used in this dissertation. I begin by describing methods for collecting and measuring SRL and the instrument chosen for this study. Following this discussion, I explain the methodological decisions for measuring classroom assessment factors: the task design and teachers' assessment decisions. I conclude this section with a justification of why the methodological choices in this study are most suitable for the identified research objectives.

SRL Methods

SRL is theorised as a set of processes (e.g., planning, monitoring, regulating, and evaluating) that individuals use across tasks to achieve learning goals. Although common, these processes vary across contexts and situations and are characterised as essential skills for academic success (Hadwin et al., 2002). As a result, researchers have developed several measurement tools ranging from self-report questionnaires and event measures to data mining using learning analytics to gather as much insight into these processes to support student outcomes and success (Greene et al., 2018). Although self-report questionnaires have been the dominant approach to collecting data on students' SRL processes, they are primarily retrospective accounts. Consequently, they do not provide evidence of these processes as they closely relate to a particular task. In contrast, event measures such as think-aloud protocols and microanalytic interviews offer researchers the potential to gain granular details about these processes in real time (Greene et al., 2018). Since the research aim of the current thesis was to examine students' SRL for a specific classroom assessment, an event measure protocol was identified as the most suitable instrument.

Using an event measure protocol enables the researcher to access context-specific processes students use as they approach, engage with, and reflect on their behaviours and thoughts on particular tasks. Some examples of event measures are think-aloud protocols, behaviour traces, direct observations, and microanalytic interviews. Think aloud protocols (TAPs) generally allow for richer data on students' SRL processes as they engage in a task (Greene et al., 2018). For example, for a math problem, participants are encouraged to 'think aloud' or verbalize their thoughts as they engage in the task. The researcher has the opportunity to collect overt (e.g., verbal statements) and covert data (e.g., nonverbal data, observations). Similar to this methodology is the microanalytic protocol which is a structured interview that uses specific questions for collecting data on SRL processes for a specific task before, during, and after completion (Cleary & Callan, 2018).

The SRL microanalysis protocol differs from TAPs by its structuring and administration of the interview. TAPs allow participants to freely express themselves with a general instruction and modelling of thinking aloud (Greene et al., 2018). For example, "tell me what you're doing" and supports open questions with prompts such as "keep talking." While microanalytic protocols comprise questions that are targeted to specific SRL processes (Cleary & Callan, 2018). For instance, if the researcher selected goal-setting as an SRL process, they would ask the question, "do you have a goal in mind as you prepare to read this textbook?" Additionally, microanalytic protocols provide scope for using both open-ended questions and Likert-type rating questions. Finally, TAPs are not grounded in any specific SRL framework, but the microanalytic protocol is based on Zimmerman's model (2011). While all measures have strengths and limitations, I decided to employ microanalytic interviews. In this study, the microanalytic interview protocol is a structured interview comprising questions targeting processes in Zimmerman's SRL model. Cleary (2011) outlines the protocol's characteristics and describes the procedure for implementing this protocol which is described in the subsequent paragraphs.

SRL Microanalysis Protocol

This protocol is a structured interview developed to measure the sub-processes in the three phases of Zimmerman's cyclical model. Essential characteristics of this protocol are its individualised administration, content and structure of the questions, and its potential to measure multiple phase-specific processes (Cleary, 2011). The protocol is also distinguished by its temporal dimension, which means the protocol is linked to the SRL phases. For example, questions targeting forethought measures are administered before the individual engages with the task (see Figure 2). And finally, the protocol has straightforward coding and scoring procedures for the microanalytic questions. Cleary (2011) reorganises these characteristics into five steps that help plan and implement this protocol, which I describe below.

Step 1. Select a well-defined task. The first step in this protocol involved identifying the task or activity for which SRL would be measured. A task is defined as an activity with a clear beginning, middle, and end. By choosing a task with clear temporal dimensions the researcher can develop questions embedded in Zimmerman's model. Previous literature in SRL microanalysis has used a range of targeted tasks such as volleyball serve practice (Kitsantas & Zimmerman, 2002), science learning engagement (DiBenedetto & Zimmerman, 2013), creative problem-solving (Callan et al., 2021), music practice (Osborne et al., 2021), reading comprehension (Follmer & Sperling, 3018) and basketball free-throw practice (Cleary, Zimmerman, & Keating, 2006).

Step 2. Identify target SRL processes. This step in the microanalysis protocol consists of selecting phase-specific self-regulated learning sub-processes and motivational beliefs from the cyclical model. The cyclical model consists of three phases that form a feedback loop, forethought, performance, and self-reflection. Furthermore, each phase has different self-regulatory cognitive, metacognitive, and motivational processes (see figure 1).

Although it is recommended that all three phases are measured, they are not always assessed, and researchers only choose to examine specific phases (Cleary, Platten & Nelson, 2008; Kitsantas & Zimmerman, 2002). Selecting only specific phases from the SRL may be due to the nature of the task or authors' interests in investigating specific sub-processes within the model (Cleary, et al., 2012). The phases and their self-regulatory sub-processes are described in the methods section of the study in Chapter 3.

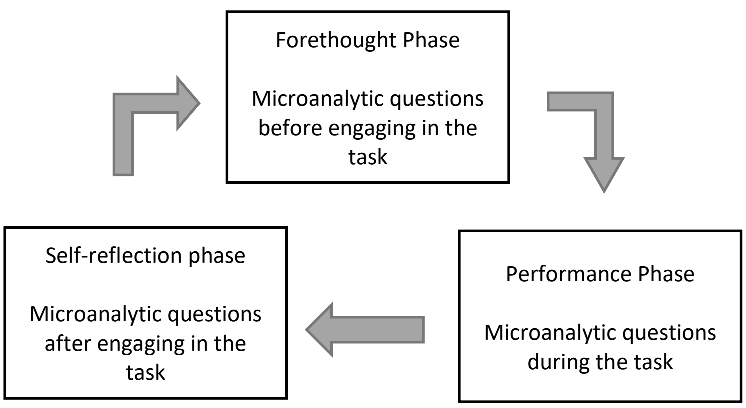
Step 3. Development of microanalytic questions. Once the task is selected, and the self-regulatory sub-processes to be measured are identified, the researcher either customizes pre-existing microanalytic questions to the task or develops new task-specific questions. It is recommended that all microanalytic questions are brief, short, and directly linked to the task and self-regulatory sub-process (Cleary et al., 2012). For example, previous research (Kitsantas & Zimmerman, 2002) for the targeted process of goal-setting, asked the question: "Do you have a goal when practising for free throws?" This process will be applied to the other targeted self-regulatory processes. Furthermore, the questions customized or developed may be close-ended or open-ended questions (Cleary, 2011).

Generally, responses to close-ended questions are recorded on a Likert scale or forced-choice structure, whereas open-ended questions are free responses. Given the nature of responses for the open-ended questions, researchers devised a contextualized coding and scoring schemes which is explained in the final step (Cleary et al., 2012)

Step 4. Link cyclical phase processes to task dimensions. The SRL microanalysis protocol is characterized as a tool that measures self-regulated learning in real-time. Therefore, it is imperative that questions targeting the self-regulatory phase are linked to the temporal dimensions of the task (see figure 2). For example, researchers targeting the self-regulatory sub-process goal-setting will have an interview with the participants prior to engaging in the task. Similarly, a study by DiBenedetto and Zimmerman (2013) examined students' performance phase, using questions focused on strategy use by interviewing participants during their engagement in the task. Questions exploring participants' self-evaluation practices will be administered on completing the task (Cleary et al., 2012).

Figure 2

Temporal Sequencing of Questions for SRL Microanalysis Protocol



Step 5. Coding and scoring microanalytic responses. The microanalysis protocol will consist of qualitative and quantitative questions for the identified task. Measurement of the quantitative questions is based on the participants' response on a Likert scale. Due to the use of single-item questionnaires, reliability of measures is not available. However, high alpha coefficients (.89 - .95) are reported by studies using multi-item self-efficacy scales (Cleary, 2011; Cleary & Zimmerman, 2001). The open-ended qualitative questions are independently coded into distinct categories by two or more coders. Similar to the development of the questions, the coding categories also stem from empirical and conceptual research. For instance, Cleary and Zimmerman (2001) referred to the literature that highlights features of goal-setting to guide the development of categories (e.g., process versus outcome; general versus specific). Previous studies (Cleary & Zimmerman, 2001; Cleary et al., 2006; Kitsantas & Zimmerman, 2002) have calculated the kappa coefficient and demonstrated a high inter-rater agreement (.81 - .98). Details describing the coding and scoring scheme was modified for this thesis and is described in the methods section of the study described in the next chapter.

Reliability and Validity of SRL Microanalysis Protocol

Emerging evidence suggests strong reliability and validity properties of an SRL microanalytic protocol. Concerning reliability, kappa coefficient and percent agreement are key metrics used to measure inter-rater agreement (Cohen et al., 2018). A review of past studies demonstrates strong interrater agreement due to the highly structured and explicit coding formats and manuals (DiBenedetto & Zimmerman, 2010; Cleary, Zimmerman & Keating, 2006). Self-efficacy constructs are numerical, and alpha coefficients will be used to examine reliability and previous studies reflected high internal consistency (alpha = 0.95; Cleary, Callan, & Zimmerman, 2012).

The validity of microanalytic protocols have been examined in terms of construct, predictive and differential validity. There have been few studies in which microanalytic measures reliably distinguish SRL processes between high performing and poor performing students (Cleary & Zimmerman, 2001; DiBenedetto & Zimmerman, 2010; Kitsantas & Zimmerman, 2002). Furthermore, microanalytic measures in such studies provide the opportunity to differentiate the nature and type of self-regulatory processes based on participants' performance levels (Cleary, Callan, & Zimmerman, 2012). There is also evidence from studies suggesting SRL microanalysis is a reliable predictor of task performance. Specifically, findings showed processes within task analysis and self-evaluation to predict performance (Cleary et al., 2012; Cleary et al., 2018; Kitsantas & Zimmerman, 2002). The reliability and validity of the modified protocol in the current study is described in the methods section of Chapter 3.

2.3.4 Methodological Decisions for Classroom Assessment Factors

This thesis was guided by an overarching question regarding the relationships between students' SRL processes and classroom assessments. The microanalysis protocol provided a fine-grain contextual measurement of SRL processes and was, therefore, a central measurement tool in this thesis. Moreover, it was essential that the data obtained from this protocol could be examined in relation to classroom assessment factors, which are task design and teachers' assessment decisions. It was essential to make decisions regarding the methods for classroom assessment factors to investigate the research aims of this thesis. I explain the methodological choices for these two distinct studies below.

Study 1. Examining the Relationship Among SRL Processes, Classroom Assessment Design, and Student Performance.

The data collected were the assessment tasks which were essentially text documents. Since there are no existing measures to measure task design, I had to develop an instrument to measure these assessment tasks. There were several modifications made to this instrument. The first attempt was a rating scale to capture the frequency of the different task features. For example, to determine the cognitive complexity of the assessment, one item was 'The assessment allows for deep cognitive processing' that was accompanied by a frequency scale from 1 (never) to 4 (always). However, this approach did not sufficiently describe the task feature in a way that represented the complexity.

To account for the description and level of the task features, I decided to develop an analytical framework similar to a rubric. This framework was designed to describe the level to which the features were evident in the assessment task. I relied on SRL literature to determine the characteristics and then developed descriptive levels for each element. I used a frequency distribution count to establish the extent to which each task feature was present in the assessment. For example, the aspect of cognitive complexity was measured using the structured observable learning outcomes (SOLO) taxonomy proposed by Biggs (1983). The descriptive levels reflected the various levels of the SOLO taxonomy (unistructural to extended abstract). Each level of the taxonomy increases in cognitive complexity regarding the learning objectives. A unistructural item focuses on a single unit of knowledge, e.g., a definition. In comparison, extended abstract items require learners to generalise and transfer their integrated knowledge to novel and unfamiliar situations. After analysing every question on the assessment task, I tallied the number of questions for each level of the taxonomy. This framework was developed to suit the current study's assessment tasks. Adopting a pragmatic approach to developing the instrument allowed me to quantify the assessment task for data analysis. More details on the instrument are provided in Chapter 3 as part of the methods section of the study and in Appendix B.

Study 2. Impact of Teachers' Classroom Assessment Decisions on Students' SRL Processes.

The second research aim for this thesis was to examine differences in students' SRL processes based on teachers' assessment decisions. I decided to use a semi-structured interview to gain insights into teachers' assessment decisions. Interviews are commonly employed data collection methods that recognise human beings as knowledge generators. Essentially interviews allow the researcher and participant to exchange views on a topic, thus making it a social, interpersonal encounter (Cohen et al., 2011; Kavle, 1996).

In contrast to surveys, interviews are tools that help generate in-depth information and offer exploratory and explanatory insights (Hochschild, 2009). On the one hand, interviews provide rich data about how and why people represent ideas, demonstrate behaviours, and make connections between various events. On the other hand, they also pose limitations to the researcher, such as being subject to researcher bias, social desirability, and being time-consuming.

Although interviews pose challenges and limitations, they were the most suitable tool for the objective of the second study, which was to gain insights into teachers' decisions when designing, conducting, and evaluating classroom assessments. By using a semi-structured interview, I could select specific questions and design prompts and probes for elaboration (Cohen et al., 2018). Since an assessment is a process, it is relatively straightforward with common steps to its design and implementation. Typically, teachers would establish a purpose, identify learning objectives, design the task, conduct the assessment, evaluate student work, and provide feedback (Andrade, 2018; Bearman, 2016). From this perspective, the questions for the interview were designed as open-ended questions and aimed to elicit specific responses regarding these different aspects of assessment design. Open-ended questions present participants with a frame of reference but also restrict their responses to a specific context (Cohen et al., 2018; Kerlinger, 1970).

In addition to open-ended questions, semi-structured interviews consider prompts and probes (Morrison, 1993). A researcher may use prompts for several reasons. For instance, prompts can help the researcher clarify any response they may not have understood or misunderstood. Additionally, a prompt may also aid the researcher in gaining further insight into the response by using an example, rephrasing the response, or repeating the question using different phrases to clarify the participants' response (Denscombe, 2014). While prompts help the researcher clarify participant responses, probes aid in extending, exemplifying, or elaborating their responses. As such, probes help the researcher gain richer insights into responses, dig deeper, and understand participants' thought processes with greater detail and comprehensiveness (Patton, 1980; Priede, 2014; Wellington, 2015). For the current study, I used anticipated probes that were 'pre-scripted to follow up on an initial question' (Beatty & Willis, 2007) and expansive probes that elicit additional details, such as seeking examples (Priedre et al., 2014). I designed a semi-structured interview protocol using these guidelines to explore teachers' assessment decisions (see Appendix C).

The data analysis for this study included a transformation from qualitative into quantitative data because a core objective of the study was to examine differences in students' SRL processes based on teachers' assessment decisions and intentions. This transformation was achieved using content analysis, wherein I identified top-level codes based on the interview questions and then further organised the codes into specific categories. For example, one of the questions required teachers to explain their approach to choosing the items for the assessment. In this case, the top-level code identified was design style and teacher responses were categorised into three distinct groups based on the approach they used to design items for the assessments such as innovative, challenging, or by-the-book. The first category (innovative) included teachers who intended the questions to be unfamiliar and novel, the second category (challenging) comprised teachers who designed their questions to be difficult but interesting for students and the final category (by-the-book) included teachers who selected pre-designed questions found in textbooks or other prescribed sources (e.g., question banks). I followed this systematic approach for all responses and developed five top-level codes related to assessment decisions: purpose, design style, strategy instruction, evaluation style, and feedback strategies (Appendix E). The coded teacher data were then matched to students' SRL data and entered into SPSS for further analyses, which can be found in the study described in Chapter 4.

2.3.5 Ethical Considerations and Management of the data

Ethical approval was received from the Human Research Ethics Committee at the University of Canterbury (See Appendix E). The original ethics approval was obtained for in-person data collection. I requested an amendment due to the covid-19 pandemic, allowing me to collect data online through video conferencing platforms. Before recruiting students, I introduced myself to school principals and directors. I explained my research and obtained their permission to speak with teachers and students. I entered virtual classrooms to introduce myself and my study and assured them that participation is voluntary, and should they participate, their responses would remain confidential. Since the study recruited participants under 18, I needed parental consent and students' assent. I requested all participants and their parents to sign, scan, and email class teachers with attached consent and assent forms. At the end of the recruitment, four schools agreed to participate in the study. Each school had an average of 30 students, which required me to seek assistance in administering the interview protocol.

An incoming graduate student was contacted to support me with the data collection. The research assistant signed a confidential agreement to ensure confidentiality was maintained and respected. After signing the document, we practised the microanalysis protocol several times to ensure we used the same prompts.

2.3.6 Summary

Self-regulated learning (SRL) processes are critical for acquiring scientific knowledge and competencies. The development and enactment of SRL rely on personal dispositions and environmental conditions such as task characteristics and teachers' instructions. There is a long history of scholars unpacking various classroom interactions that inhibit or promote successful student learning processes. Recently, classroom assessments have gained attention as they provide a systematic structure in which SRL processes may be integrated. Findings from some studies indicate that formative classroom assessment practices, such as feedback, have a positive impact on students' SRL. Despite the advocacy and evidence in favour of formative practices, teachers in India struggle to implement these practices due to the overwhelming focus on performance outcomes derived from formal assessments. The scarce educational research in India and the immense value of SRL for science learning helped design a project that could have practical and theoretical implications.

In designing this thesis, I had to make several methodological decisions to form a coherent whole. This included deliberating how SRL is measured in contemporary research and how novel methods can offer deeper insights into the context-specific nature of SRL. Additionally, I had to consider how best to measure classroom assessment factors (i.e., assessment design and teachers' decisions), and importantly my own beliefs about research epistemology. Through this process, I was able to design and conduct two research studies using mixed methods for data collection and analysis to make a significant and original contribution to the SRL literature with implications for practice. The rest of this thesis is dedicated to the two studies which aim to answer critical research questions on the relationship between SRL and formal classroom assessment factors in the Indian context with focus and intent.

Chapter 3: Examining Relationships among Self-regulated Learning, Student Performance, and Assessment Design Features

Abstract

Academic outcomes such as reading comprehension, mathematical problem-solving, and deep conceptual understanding have often been linked to successful self-regulated learning (SRL). Students who demonstrate SRL have higher self-efficacy, better academic performance, and are lifelong learners. Despite these outcomes, not all students demonstrate effective SRL strategies. Contemporary literature suggests that classroom assessment practices contribute to the development and enactment of SRL; however, few studies have examined a formal classroom assessment task with regard to how it supports SRL. This research uses a mixed-methods approach to understand students' SRL processes for a formal science assessment and examine relationships among SRL processes, academic performance, and design features of the assessment task. The sample comprises 229 high school students from Southern India. A microanalytic protocol was used to measure students' SRL and an analytical framework was developed to measure assessment task features. The results indicate meaningful relations between task features such as cognitive complexity and types of questions and SRL processes. Correlational analyses revealed negative associations between task features and students' motivational beliefs. Furthermore, assessment features such as cognitive complexity and types of questions predict students' SRL processes such as goal setting, self-efficacy, and strategy selection. The findings and methodologies from this research have implications for educational research and practice. For instance, teachers can use instructional cues from task features to support students' planning processes. The methodologies from this study can be used to improve our understanding of SRL across diverse assessments.

Key words: Self-regulated learning, classroom assessment, task design

3.1 Introduction

The importance of students assuming personal responsibility and control for the acquisition of knowledge and skill is paramount today. Decades of research have helped identify how students exercise agency and become masters of learning, resulting in an umbrella of processes referred to as self-regulated learning. Self-regulated learning (SRL) is when students plan, monitor, regulate, and reflect on their actions and strategies to attain their learning goals (Pintrich & Zusho, 2002; Zimmerman, 1989). SRL encompasses critical elements of learning including cognition, metacognition, and motivation, which have been linked to academic outcomes such as mathematical problem-solving, conceptual understanding in science, and reading comprehension (Cleary & Kitsantas, 2017). Therefore, improving students' SRL is of significance for their short and long-term academic success.

Classroom assessments are regularly occurring situations that require students to engage in SRL processes such as goal setting, deploying strategies, and reflecting on their performance. Recent literature has focused on bridging classroom assessment practices with self-regulated learning processes by using feedback strategies and self-assessment tools (Andrade & Brookhart, 2016; Panadero et al., 2017; Panadero et al., 2018). More importantly, self-regulated learning is strengthened when students plan, adapt, and reflect on their approaches based on their interactions with the task. Despite this evidence, scholars have paid relatively little attention to SRL processes in relation to formal classroom assessment tasks.

From a social cognitive perspective, self-regulated learning is contextualized, and depends largely upon personal and environmental factors (Bandura, 2001; Pintrich, 2000;

Zimmerman, 1989). This would mean that SRL is determined by individual characteristics and how these characteristics interact with an external factor, such as an assessment task. Contemporary socio-cognitive models of SRL indicate that these interactions occur before (e.g., goal setting), during (e.g., strategy use), and after (e.g., cognitive appraisal) a learning task (Pintrich, 2000; Zimmerman & Moylan, 2009). Much of the literature on SRL and assessment tasks are limited to students' study strategies for assessments tasks, which are the processes enacted during learning (Garcia-Perez, 2021; Smith & Miller, 2005; Rovers et al., 2018), or measure students' SRL using self-report questionnaires for various tasks. SRL encompasses a wide range of processes that also occur before (e.g., goal setting and strategy selection) and after the learning activity (e.g., cognitive appraisal) which feed the next learning cycle (Zimmerman, 2013). In addition to understanding SRL processes in these two phases, I argue that it is also important to examine how they are related to the task structure. Research has established links between task structure and SRL process (Lodewyk & Winne, 2005), but this study seeks to refine the level of detail of the task structure to provide a nuanced view of a formal assessment task structure. Therefore, by focusing on classroom assessment tasks from a granular perspective, the current study adds to the understanding of the assessment task structure in relation to SRL. Findings from the study have implications for educational practice and research on SRL processes that occur before and after the task in relation to assessment task structure.

According to national and international assessments, academic achievement in science is a concern (NCERT, 2017; OECD, 2019). For example, a recent survey in India indicated that a meagre 15% of the sampled student population were able to respond to items that required logical reasoning, problem-solving, and critical thinking (NCERT, 2017).

Researchers advocate for improving students' self-regulated learning skills to improve scientific proficiency. Moreover, self-regulated learning and assessment design are under-researched areas in the Indian context, where much of the literature in India is focused on pressure from assessments rather than understanding how students engage with assessments (Bodas et al., 2008). Therefore, the current study focuses on high school students' self-regulated learning processes for science assessment tasks in India. Findings from our study have implications for theory and future research, particularly for the design of SRL interventions and evaluation of assessment structures within and across academic disciplines.

3.2 Literature Review

3.2.1 Self-Regulated Learning

Self-regulated learning (SRL) is a systematic approach to learning whereby individuals organize and optimize their cognition, motivation, and behaviour to attain academic goals (Winne & Perry, 2000; Zimmerman, 2000). Although there are several theoretical models describing SRL, researchers generally agree upon specific core characteristics (Boekarts, 2011; Efklides, 2011; Pintrich, 2000; Winne & Hadwin, 1998; Zimmerman, 2009). For example, most models illustrate SRL as cyclical, and comprises phases in which cognitive, metacognitive, and motivational processes are situated. Theorists also agree that SRL is influenced by a myriad of biological, contextual, and developmental factors (Boekarts & Corno, 2005).

Several decades of scholarship have established links between SRL and academic outcomes in science. For example, Greene et al. (2012) found that students who use effective SRL approaches gain deep conceptual understanding of the content. In contrast, students tend to acquire content knowledge at the surface level when they do not regulate their learning effectively. Sinatra and Taasoobshirazi (2018) argue that for students to successfully address scientific tasks such as problem-solving and logical reasoning, they require a deep conceptual understanding of the subject matter. For example, in physics and chemistry students are required to manipulate either side of an equation to identify unknown quantities (Chi, 2006). Students equipped with sufficient knowledge will analyse the task and select a range of strategies such as inductive, deductive, analogical, and abductive reasoning to address scientific problems (Sternberg & William, 2002). Elfkides (2011) found that self-regulated learners develop better metacognitive awareness and control, and as a result engage in better planning and regulation during problem-solving tasks.

So far, literature in SRL and science has been limited to investigating relationships between students' metacognitive and cognitive factors and science achievement. Scientific tasks demand considerable effort and deliberate practice which require regulating one's motivation to stay focused and engaged when practice becomes frustrating or boring (Ericsson, 2006; Ericsson et al. 1993; Sinatra et al. 2015; Zimmerman & Campillo, 2003). Over two decades ago, Zimmerman (1995) encouraged researchers to broaden the scope of SRL and view it as a complex and interactive process encompassing motivational, social, and behavioural elements. Since then, empirical evidence has emerged to support the significant role contextual factors and personal motivations play in implementing cognitive and metacognitive processes (Hoffman & Spatariu, 2008; Schunk & Usher, 2013).

The promise of SRL for academic success has led to 55copious research in promoting and developing self-regulatory processes among students. For example, Lombardi et al. (2013) recommend tailoring learning environments that challenge students to consider alternate theories when learning scientific concepts as this process encourages SRL. Other researchers have found that when teachers instruct students to use specific strategies to learn science, students improve their self-regulated learning processes (Kuhn, 2009; Olakanmi & Gumbo, 2017; Zepeda et al., 2015). More recently, researchers have found strong associations between SRL and classroom assessment practices (Andrade & Brookhart, 2016; Panadero et al., 2016).

3.2.2 Classroom Assessment and SRL

Classroom assessments have long since been recognized as tools to improve student learning. Black and Wiliam (1998) challenged the dominant evaluative function of assessments and encouraged educators to view classroom assessments as formative opportunities for teachers to improve their instruction and help students identify learning gaps by providing qualitative feedback. Black et al. (2004) reported extensive evidence supporting feedback quality. In particular, authors contend that students who receive comments in their feedback and without grades outperformed students who received only grades. Hattie and Timperley (2007) extended feedback practices to strengthen the self-regulatory processes among students by using three succinct questions: “Where am I going?", "How am I going?", and "Where to next?" These conceptualizations of classroom assessments reflect their inherent potential to promote self-regulated learning.

Logically, much of the recent scholarship has focused on integrating SRL processes into classroom assessment practices. Allal (2019) and Andrade and Brookhart (2019) have developed frameworks advocating for classroom assessments as co-regulating processes to improve students' SRL. The framework positions teachers as co-agents of SRL, thus helping students set goals, choose strategies, and use feedback to help monitor and evaluate learning progress. Panadero et al. (2018) provide a systematic review of the empirical literature on classroom assessment practices such as feedback and self-assessment and their relation to SRL. Importantly, the literature offers further evidence for the significance of contextual factors. Teachers' feedback practices and self-assessment tools for assessment support the development of SRL. Given the value of contextual factors, the current study aims to deepen this understanding by focusing on the structure of an assessment task presented to students.

Classroom Assessment Task Structure

Sanford (1985) describes an academic task as an assignment on which students interact with content and are held accountable for delivering a tangible outcome. As such, a task is characterized as having a product/output, using resources to produce the output, using a set of cognitive operations to produce the output, and is accounted for in the form of a grade (e.g., minor, major, extra credit; Sanford, 1985). There are generally two types of tasks—well-structured tasks (WST) and ill-structured tasks (IST)—that are distinguished based on the degree to which they meet a set of features on a continuum. Content-related characteristics of WST include employing factual data in student responses (e.g., correct and incorrect responses) and defining specific evaluation criteria (e.g., rubric, answer key; Frederickson, 1984; Spiro et al., 1988). In contrast, ill-structured task (IST) features are open-ended problems for which there are fewer sub-goals mentioned in the description of the assignment (e.g., develop an output for a funding program to allocate resources for cancer; Lodewyk et al., 2009). Such tasks typically require higher-order thinking skills such as synthesis and integration of learning. Students engage in these cognitive processes by drawing information from various sources, apply and reiterate their knowledge and skills. There is less attention to right and wrong answers, as there may be multiple solutions or perspectives (Roth, 1994).

Early studies found significant associations between task structure and student learning before and after completing an academic task (Chin & Chia, 2005; Lodewyk & Winne, 2005; Shin et al., 2003). Generally speaking, open-ended, complex, challenging, and relevant tasks with opportunities to self-reflect are more likely to promote SRL (Perry et al., 2004). Importantly, these differences depend on how individuals perceive, interpret, and execute the task. These interactions between SRL and the task structure are better understood when it is viewed through the lens of a model describing how SRL occurs. The current study adopts Zimmerman's (2009) cyclical model of SRL grounded in a social cognitive perspective because the learning processes identified in the model account for the interactions between the person (student) and their environment (task; Badura, 1986; Zimmerman, 1989).

According to Zimmerman, SRL comprises three phases that encompass a cyclical feedback loop: (1) forethought, (2) performance, and (3) self-reflection. In the forethought phase, learners engage in task analysis sub-processes, which includes setting goals and selecting strategies. Furthermore, students' task analysis is informed by motivational beliefs such as self-efficacy, interest, and task value. In the subsequent performance phase, individuals utilize metacognitive monitoring for self-control and self-observation. For example, a student may use strategies such as rewards to keep their action focused on the task and self-explanation to check their understanding. And in the self-reflection phase, learners appraise their performance using judgement criteria based on mastery (e.g., was I successful in learning?) or outcome (e.g., how many mistakes did I make?). In this phase, students also tend to have affective responses to their performance. Positive reactions are more likely to inspire subsequent learning cycles on similar tasks, while negative reactions may lead to defensive behaviours such as learned helplessness. Ultimately, all these processes are interactive and reciprocal, which would enable self-regulated learners to become more adaptive in their approach to learning. The current study focuses on the interactions between the task structure the sub-processes found in the forethought and self-reflection phases of Zimmerman's model (see Table 2). The study aims to gain a deeper understanding of students' approaches to and reflections on their learning and performance for a formal assessment task in science.

Table 2

Targeted SRL Phases, Processes, and Sub-Processes Based on Zimmerman's Model of SRL (2009)

SRL Phase	Sub-Processes	Sub-Processes
Forethought	Task analysis	Goal-setting, strategy selection
	Motivational beliefs	Self-efficacy, interest, perceived importance
Self-reflection	Self-evaluation	Self-judgement, causal attributions
	Self-reaction	Perceived satisfaction, adaptive inferences

Forethought Phase and Assessment Task Structure

This phase comprises two processes: task analysis and summoning motivational sources. In the task analysis process, effective self-regulated learners set goals and select strategies that will help achieve the goal (Zimmerman, 2009). For an assessment task, students set different goals which include process goals, mastery goals, outcome goals, or a combination of goals. Classroom assessments offer an ideal opportunity for students to set learning goals that describe outcomes, skills, and knowledge (Andrade & Brookhart, 2019). However, students do not tend to set learning goals for assessments, but instead, their goals are influenced by outcome expectations which could be a form of external regulation by teachers or parents (Butler & Schnellert, 2015; Butler et al., 2016); thus, limiting their autonomy (Ryan & Deci, 2000). For instance, in India, students feel pressured to perform on assessments from a very young age which is likely to result in students setting goals that are focused on performance (Bodas et al., 2008; Singhal & Misra, 1994).

Furthermore, the task structure could facilitate goal-setting processes. Complex tasks, such as ISTs are more likely to encourage students to set sub-goals that tend to be process-focused. In contrast, less flexible tasks such as WSTs suggest the sub-goals for students, such as completing the syllabus for the test (Lodewyk & Winne, 2006). In general, research finds that self-regulated processes tend to be more effective when goals are aligned with the task and outcome expectations and provide standards against which one can monitor and evaluate progress (Pintrich, 2000; Zimmerman, 2000; 2008). Moreover, to succeed in scientific tasks, students need to set goals that help develop an awareness of subject knowledge and related skills to inform their strategy selection.

Concerning strategic selection, assessment literature has examined learning strategies and their relation to task interpretation, assessment type, and performance outcomes. There is a general agreement that one of the most salient contextual variables to influence students' approaches to learning is the assessment method (Biggs, 1998; Garcia-Perez, 2021; Segers et al., 2006). Research indicates that students shift from surface to deep learning approaches to meet the expectations of the assessment (Hattie & Donoghue, 2016; Segers et al., 2006). For example, when the assessment is a multiple-choice quiz, students are more likely to use surface learning approaches, which include rehearsal and comprehension strategies and not much critical thinking (Crede & Phillips, 2011; Garcia-Perez et al., 2021; Rovers et al., 2018). In contrast, some scholars have found that students rarely modify their learning strategies to suit the assessment expectations and are less likely to change their learning approaches because they remain unaware of effective learning strategies (Blasiman et al., 2017; Karpicke et al., 2009; Ormrod, 2011).

Research findings indicate that complex tasks increase students' use of knowledge to integrate and synthesise information (Jonassen, 1997). However, there is also evidence that suggests students' self-regulatory processes are not always productive (Lodewyk et al., 2009; Winne, 1995; Winne & Perry, 2000). Researchers have also identified that many of the tasks designed and assigned by teachers do not support the development of effective SRL processes (Doyle, 1984). In such cases, using meaningful strategies may be counterproductive or impossible. For example, when tasks require students to recall knowledge verbatim, students are less inclined to use strategies that require deeper conceptual learning and elaboration (Thomas & Rohwer, 1986; Turner, 1995; Van Meter et al., 1994). Also, if students need to learn expansive material for a test, then they tend to engage with the content in a general and superficial manner rather than gain an in-depth understanding of the concept (Thomas, 1993).

In India, national policies recommend assessing higher-order thinking skills, however, assessments continue to focus on measuring surface-level knowledge of the subject matter (Kapur, 2008; NCERT, 2005; MHRD, 2020). These assessment practices reflect the findings from studies that indicate that teachers tend to teach to the test and struggle to develop test items that measure deep conceptual understanding (Ananthakrishnan, 2019; Crooks, 1988; Jensen et al., 2014). Studies highlight that an overwhelming majority of students in India and more widely employ rote memorization techniques or rehearsal strategies to prepare for their assessments (Biggs, 1998; Crooks, 1988; Garcia-Perez et al., 2021; Gupta & Mehtani, 2017; Segers et al., 2006). But using such strategies while effective for short-term success is unlikely to translate into long-term deep conceptual understanding required for scientific proficiencies (Boud & Falchikov, 2006; Hattie & Donoghue, 2016).

In the forethought phase, motivational beliefs play an important role as they influence students' task analysis sub-processes. For example, Andrade and Brookhart (2016) highlight aspects such as task value, interest, importance, and perceptions of ability contribute to students' willingness to learn and their academic performance. Usher and Schunk (2018) highlight studies in which students' motivational beliefs, such as self-efficacy, goal orientations, and intrinsic interest, are strongly associated with task characteristics. In particular, self-efficacy and task value are directly associated with students' interpretations of the task, which in turn affects performance, future accomplishments, and other self-regulatory processes (Pajares, 2008; Durik et al., 2006; Zimmerman & Cleary, 2009). There is an external value placed on formal assessment tasks, which implies that performance outcomes are important (either to them or to those around them). Studies have found mixed results between self-efficacy and task structure. For example, most students tended to feel more efficacious for highly structured tasks; and poor self-regulated students perceived lower competencies for complex tasks that are less structured (Lodewyk & Winne, 2005). Perry et al. (2004) corroborates the impact of complex tasks on students' intrinsic interest which is associated with motivation. Together, these sub-processes in the forethought phase of SRL impact students' engagement with the task, determine knowledge acquired, and influence other SRL processes.

Self-Reflection Phase and Assessment Task Structure

Upon completing the learning task, students appraise their performance in the self-reflection phase. Teachers and students can gain useful information regarding student learning through performance on assessments. Therefore, the interpretation of their performance which would include how they learned, and the product of their learning is crucial to self-reflection sub-processes (Andrade & Brookhart, 2019; Black & Wiliam, 2018; Panadero et al., 2018; Stiggins & Chappuis, 2005). Different types of tasks provide students and teachers with useful information for feedback on learning and performance. For instance, structured tasks which have clear expectations with easily identifiable solutions or correct answers inform students on specific learning outcomes (Crooks, 1988; Lavonen & Laaksonen, 2009). In contrast, tasks that are complex and have multiple approaches to arriving at solutions provide students and teachers with mastery and process-based feedback. For example, when attempting to solve a problem with more than one solution, students learn to apply their knowledge in different ways and as a result are likely to reflect on the efficacy of their process in relation to the knowledge used (Lodewyk & Winne, 2005).

Although both tasks provide opportunities for critical reflection, the use of information to improve learning depends on the teachers' evaluation criteria and students' self-judgement standards (Brookhart & Chen, 2014; Zimmerman, 2005). Ideally, mastery-based judgement criteria are more useful for subsequent forethought processes because they provide clarity for goals and strategies (Perry et al., 2006; Zimmerman, 2005). However, there is a general tendency for students to use normative and social comparison as criteria for self-judgement because classroom assessments tend to reflect performance evaluation (Black et al., 2004; Brookhart, 2013). Such self-judgement criteria are likely to limit students' approaches to scientific tasks. For example, problem-solving skills rely on abilities to judge the efficacy of the solution using reasoning skills that stem from sound conceptual understanding (Sinatra & Taasoobshirazi, 2018).

When students engage in self-evaluation processes, they also make attributions of their performance. Effective self-regulated learners tend to attribute their success or failure to the strategies they used in the performance stage. For example, a student may realise they used an ineffective learning technique for reasoning, and as a result they change their strategy for the next learning cycle. In contrast, students who attribute their success or failure to ability, or difficulty are more inclined to feel discouraged (Zimmerman & Moylan, 2005). For example, a student might conclude that they performed poorly because the test was difficult. Even though these judgements are made by the individuals, research examining students' attributions suggests that the design of the task may be an influencing factor on causal attributions. For example, tasks that are vague and ambiguous with poor instructional clarity may result in students feeling confused and perceiving the task as complex and insurmountable (Malmberg et al., 2014). Moreover, these self-judgement processes, (i.e., self-evaluation and causal attributions), are closely associated with motivational constructs in the form of self-reactions which are critical to subsequent learning cycles.

Students' reactions to their performance tend to determine adaptive or defensive behaviours for future learning tasks (Bandura, 1997). Feelings of satisfaction are positively associated with encouraging behaviour such as adaptive strategy use, or a try harder attitude. In contrast, feelings of dissatisfaction of performance may lead individuals toward self-handicapping behaviours (e.g., procrastination; Bandura, 1989; Schunk & Usher, 2018; Zimmerman, 1995). Although it is desirable that learners adapt their actions and affect to facilitate the continuance of the cycle, the task structure could play a role in students' actions. For example, Malmberg et al. (2014) found positive associations between ISTs and students' help-seeking behaviours, which are adaptive strategies when teachers are perceived as supportive. If there is a lack of adequate and meaningful support, then ISTs may also lead to self-handicapping strategies. Thus, teachers and educators need to balance various characteristics of task structure to support SRL carefully.

3.2.3 The Current Study

This study builds on Zimmerman's (2009) cyclical model of SRL to explore potential links between the structure of assessments and its relationship to students' SRL strategies in the forethought and self-reflection phases of the model. The research seeks to contribute to the existing literature in multiple ways. First, this study focuses on an underrepresented population and offers a unique perspective on SRL from an assessment standpoint in India. Second, the instrumentation in the research draws on recent approaches to capture SRL in a specific context, i.e., formal classroom assessment. Third, the research uses a strong empirical and theoretical foundation to link SRL to the structure of assessment tasks. There are no hypotheses for the current study because of the exploratory nature of the project. The study explores an under-researched context within SRL and assessments and therefore seeks to understand relationships more thoroughly. The following research questions will guide the study:

1. Which SRL strategies reflected in Zimmerman's cyclical model (2009) do students demonstrate for a science summative assessment task?
2. To what extent does SRL predict students' academic performance on science summative assessment tasks?
3. In what ways do design features of an assessment task relate to students' SRL strategies reflected in Zimmerman's cyclical model (2009)?

3.3 Methodology and Methods

3.3.1 Study Population and Setting

The sample comprised 229 students (108 boys and 121 girls) from grades 8 and 9 across four schools that participated in the study. Student age ranged from 13 – 15 years (*M* = 13.58 *SD* = .64).

All students belong to a school affiliated with the state or central government of India. This affiliation determined the school's syllabus, teaching, and assessment practices. Although all schools follow the National Curriculum Framework (NCERT, 2005), which outlines the learning objectives, the depth of the content that is taught differs across schools. Furthermore, the assessment format is the same across the schools; but they may differ in structure. All assessments are tests which comprise a set of questions that range from one-word responses to whole constructing paragraphs. The assessments differ in how the questions are weighed and organized.

Each school serves 1000+ students from kindergarten to Grade 12 and the primary medium of instruction is English. As science is a mandatory subject, all students attended science-related classes which included physics, chemistry, and biology. Based on the school's affiliation to the government (central or state), science subjects are grouped in different ways. Schools affiliated to the Central Board of Secondary Education (CBSE) combine physics, chemistry, and biology into one assessment; State Board schools have two assessments for science, one for physics and Chemistry combined, and one for Biology; and schools affiliated to the Indian Certificate of Secondary Education (ICSE) have three distinct assessments for each science discipline (physics, chemistry, and biology). Due to the Covid-19 pandemic, all classes and assessments were held virtually.

Table 3

Distribution of Sample Student Population based on Assessments

School Board	Subject	Assessment	Number of Students
State board	Physics	Assessment A	62 (Year 8)
ICSE	Physics	Assessment B	65 (Year 8)
		Assessment C	52 (Year 9)
CBSE	Physics	Assessment D	15 (Year 8)
	Biology	Assessment E	14 (Year 9)
	Science	Assessment F	21 (Year 8)
Total		9	229

Students were scheduled to perform a classroom summative assessment in science designed by the teacher. Each assessment included in this study required students to write their answers on foolscap sheets while being supervised by teachers online. All students were expected to submit their answer scripts either by scanning the images and emailing them or posting the scripts to their respective science teachers. Teachers shared their respective summative assessments with me after students completed the assessment.

3.3.2 Methodology and Methods

The study employed a mixed-methods design. More specifically, I chose to use a conversion mixed methods design which is characterised by transforming one type of data into another (Teddlie & Tashakkori, 2009). This methodology is best suited to address the research questions outlined above because it allows me to collect data that are qualitative and quantitative in nature. Furthermore, this methodology also supports data transformation which is crucial to answering the research questions. I integrated qualitative and quantitative data at the interpretation and reporting level through data transformation (Fetters et al., 2013).

3.3.3 Interview Procedure

Once schools consented to participate in the study, I contacted teachers or guidance counsellors in each school to meet the students and introduce the research project. I explained the study, participation procedure, and use of data. Upon receiving assent from students and consent from primary caregivers, individual interviews were scheduled toward the end of the academic year (November–April) and were administered using secure video links. The initial procedure was to collect data in-person at the school; however, due to the global pandemic, I used online video conferencing platforms (e.g., Zoom/Google meet) to conduct the interviews.

3.3.4 Instruments

In this study, data included: (1) SRL microanalysis protocol; (2) student performance outcomes, and (3) assessment tasks as designed by each school.

SRL Microanalysis Protocol

In this study, I used the SRL microanalysis protocol, a structured interview grounded in Zimmerman's model to measure SRL processes (Cleary, 2011). I followed the protocol guidelines to develop and select questions and associated coding and scoring strategies. The subsequent paragraphs describe the modifications made to the protocol (explained in Chapter 2) for the current study.

Step one: Selecting a task. The current study focused on measuring students' SRL for a formal classroom assessment (i.e., an exam). This task is developed, administered, and evaluated by respective science teachers. The task is considered a measurable task because it has a beginning (time before the assessment), during (performing on the assessment), and an after (time after the assessment), as reflected by Cleary (2011).

Step 2: Identifying target SRL processes. The SRL sub-processes targeted in the current project include the sub-processes in the forethought and self-reflection phases of Zimmerman's model (2009). The forethought phase comprises two sub-processes: goal-setting and strategy selection, and motivational beliefs, which include: self-efficacy, intrinsic motivation, and perceived importance of the task. The self-reflection phase consists of two self-judgement sub-processes: self-evaluation and causal attributions, and self-reactions which are perceived satisfaction and adaptive inferences.

Step 3: Developing microanalytic questions. The questions for this current study were developed using the protocol guidelines. As such, the microanalytic questions for the sub-processes in the forethought and self-reflection phases are either based on previous studies or grounded in conceptual or empirical research. The development of questions for each sub-process in detail after the description of Step 5.

Step 4: Linking the microanalytic questions to the temporal dimension. In this step, it was important that the microanalytic questions were linked to the temporal dimension of the task. As such, the questions in the forethought phase were administered before students performed on the assessment and the self-reflection microanalytic questions were asked after students completed the assessment and received their performance outcomes.

Step 5: Coding and Scoring Scheme. According to Cleary (2011), the coding and scoring scheme for each open-ended question needed to be grounded in empirical research or based on existing studies. In the current study, where possible, I used codes based on existing studies and developed codes with clear guidelines when no pre-existing codes were available. In the current study, I opted to measure students' responses to open-ended questions using an ordinal scale for specific reasons. First, ordinal scaling allows the researcher to classify qualitative data into an order that enables the ranking of variables without a degree of difference (Cohen et al., 2018). In essence, ordinal data can be placed in a definitive order such as strongly agree, agree a little, disagree a little, strongly disagree. However, it is not possible to indicate that the difference between strongly agree and agree a little is the same as disagree a little and strongly disagree (Cohen et al., 2018). Second, a 3-point ordinal scale was most appropriate because establishing the degree of difference between the points on larger 5-point scale is not possible due to the qualitative nature of the data. Given distinct SRL processes have been identified as superior to others, the scoring procedure employs a 3-point ordinal scaling rather than categorical for the open-ended questions was appropriate (Callan, 2014; DiBenedetto & Zimmerman, 2013; Follmer & Sperling, 2019).

The subsequent paragraphs are organised based on the temporal feature outlined in Step 4. I describe, in detail, steps 3 (i.e., the microanalytic question) and 5 (i.e., coding and scoring scheme) for all the sub-processes selected in the forethought and self-reflection phases.

Forethought Phase

The first round of interviews was conducted a week before the assessments were scheduled. During this interview, we asked students to respond to microanalytic questions in the forethought phase. The duration of each interview was between five to seven minutes for each student. The sub-processes targeted in this phase were: goal-setting, strategy selection, self-efficacy, interest, and perceived importance.

Goal-setting. In the current research, I sought to capture the nature of students' goals according to each participant's responses to the following microanalytic question *"Do you have a goal in mind as you begin preparing for this assessment? Please explain."* This question was developed keeping in mind similar research (e.g., Callan, 2014; Cleary & Zimmerman, 2001) and age-appropriate language. Participants freely shared their comments without interference or bias from the interviewer; thus, data came from students' authentic and candid dialogue, and were then coded and aligned to be consistent with the extant literature. The current study adopted a coding scheme relating to students' goals similar to the work done by Callan (2014) and Zimmerman & Kitsantas (1996). Existing research (e.g., Callan, 2014) has distinguished goals based on two features: (a) the nature of students' goals (e.g., process and performance goals), and (b) students' goal specificity.

The nature of students' goals. According to research, the nature of students' goal can be conceptualised in terms of process and/or performance foci. A process goal is that which focuses on series of actions or steps that help the student learn; outcome goals are characterized as those that focus on the performance such as the grade or score (Zimmerman & Kitsantas, 1997). Process goals set by the individual effectively sustain and regulate action, cognition, and motivation for successful performance (Dweck, 1986; Schunk, 1989, 2001; Locke & Latham, 2002). Outcome goals are less favoured because the feedback is delayed, and they students remain unsure about the steps or actions needed for successful learning (Zimmerman & Kitsantas, 1997).

Students' goal specificity. In additional to the nature of students' goals, I sought to capture the specificity of their goals according to their responses to the interview question. Student responses were coded as specific or general based on the degree to which the goals were detailed, explicit, and precise (see Table 4).

I used these two goal-setting aspects to develop a coding scheme: (1) process-specific, (2) process-general, (3) outcome-specific, (4) outcome-general, (5) non-task goal, (6) no goal (see Table 4). It is worth noting that those students who reported a combination of goals were rated the highest goal described. For example, if the goal included process and outcome, the response would be coded at the process level (i.e., 3; Callan, 2014). Furthermore, students' responses could be non-task related goals, or have no indication of goals at all. Therefore, categories for these possibilities were also made available.

Table 4

Coding and Scoring Scheme for goal-setting

Nature and specifity of goal	Definition	Example
Process-specific (3)	Statements that focus on the process of studying for the assessment and also identify the use of specific study strategies, tactics, or procedure as the primary focus of the studying sessions	• I'll probably solve the problems first and then go to the theory • I'll make a timetable to study • I'll mark the important points
Process-general (3)	Statements indicating a focus on a process in general but does not identify any particular procedures.	• I'll finish the portions • I want to study well • I will understand the concepts • I'll work on the numericals/problems
Outcome-specific (2)	Statements that identify a clear and measurable outcome as the focus of the studying session	• I want get to an A/green card/excellence • I want to get at least 65+; 75+; 80+;etc.. on the exam
Outcome-general (2)	Statement identifies as an outcome that is unclear, not quantifiable, or not directly measurable as the focus of the study session.	• I want to do the best I can • I want to get the highest I can • I want a good grade • I want to get good marks
Non-task (1)	Statements that indicate a goal that is so incongruent with the current assessment task that the goal reflects inadequate understanding of the task.	• To become a doctor • To get into college
No goal (1)	Statement indicates that the student doesn't have a goal for the assessment.	• No • I don't really have a goal • I don't know • Not really

Given that research has helped identify optimal goals and studies that have validated this approach, I opted to adopt an ordinal scale to measure students' goal-setting. The scale ranged from process goals (high; 3), Outcome goals (medium; 2), and other or no goals (low; 1).

Strategy selection. In the current study, I used a question similar to the one used by DiBenedetto & Zimmerman (2013): *"Do you have any plans on how you will prepare for this assessment? Please explain."* For the coding and scoring scheme, I used a two-part rubric grounded in literature (Follmer and Sperling, 2019). The first part of the rubric included the number of strategies or actions employed by the student to help their learning or preparation for the assessment (Mayer, 1996; DiBenedetto & Zimmerman, 2013). The scoring for this aspect was the total number of strategies reported by students.

The second part of the rubric involved evaluating the strategies grounded in the learning frameworks proposed by Hattie and Donoghue (2016) and Mayer's Selection-Organization-Integration model (SOI Model; Follmer & Sperling, 2019). These frameworks suggest that learning strategies can be distinguished based on the level of information processing. At the lower end of the continuum is rehearsal or rote memorization strategies followed by comprehension and organizational strategies and ends with integration strategies that help transfer of learning (Hattie & Donoghue, 2016). Therefore, students' learning strategies can be ranked from lower-order processes to higher-order processes using an ordinal rating scale. The scale was ranked as: (1) rehearsal, (2) comprehension, (3) organization, and (4) integration). Students who reported multiple strategies of varying processing levels were rated with the highest level of processing mentioned (see table 5).

Table 5

Coding and Scoring Scheme for Strategy Selection – Processing Level

Strategy type based on processing level	Definition	Examples
Rehearsal or memorization (1)	Statements that reflect the use of strategies that focus on committing information to memory.	• Flashcards • Mnemonics • Writing practice of answers
Comprehension (2)	These statements reflect those strategies that describe learning the content that focuses on understanding the subject matter.	• Summarizing and paraphrasing • Reading for comprehension
Organization (3)	A statement describing organization or consolidation of ideas or checking of conceptual understanding	• Concept maps • Mind maps • Self-explanation
Integration (4)	Statements that describe approaches to learning that deepen their understanding of the concept such as elaboration	• relating content to prior knowledge • relating content to real-life situations

Students' responses were further grouped into two larger strategy groups: surface and learning strategies. According to Hattie and Donoghue (2016), rehearsal and comprehension strategies belong to the surface learning approach category because they focus on acquisition of content knowledge; while organization and integration strategies reflect deeper learning strategies as they go beyond content acquisition and facilitate transfer of learning.

Motivational Beliefs. Based on recommendations in the literature (Callan & Cleary, 2018; Cleary, 2011; Cleary et al., 2012), the questions for motivational beliefs were developed utilizing close-ended questions in the Likert-scale formats. I included two questions targeting self-efficacy: the first question targeted efficacy for preparation and the second was to measure efficacy for performance (DiBenedetto & Zimmerman, 2013). The questions were worded as: *"On a scale from 1 (not confident at all) to 5 (very confident), how confident do you feel about preparing for the assessment?"* and *"On a scale from 1 (not* *confident at all) to 5 (very confident), how confident do you feel about performing or doing the assessment?"* The questions for perceived interest and importance were developed similarly and were finalized as: "*On a scale from 1 (not interested at all) to 5 (very interested), how interested are you in this subject?"* and *on a scale from 1 (not important at all) to 5 (very important), how important is this assessment for you?"* Since these questions utilized the Likert-scale, the scoring reflected students' reported responses.

Self-Reflection Phase

The second round of interviews was conducted after students completed the assessment and received their evaluations from their teachers. For this round, we asked students to respond to microanalytic questions in the self-reflection phase. The duration of this interview for each student was between five to seven minutes. The sub-processes measured in the self-reflection phase include self-evaluation, causal attributions, perceived satisfaction, and adaptive inferences.

Self-evaluation. The self-evaluation sub-process had two questions. The first question was a Likert-scale question that targeted students' perceptions regarding how well they prepared for the assessment (e.g., DiBenedetto & Zimmerman, 2010) and was worded as, *"Now that you have received your results, how well do you think you prepared for the assessment?"* Given that this question utilized the Likert-scale, the scoring reflected students' reported responses.

The second question self-evaluation measured students' criteria for self-judgement (e.g., Cleary et al., 2006; Cleary & Sanders, 2011), and the question was, *"What did you use* *to judge how well you have performed on the assessment?"* Students were presented with choices to reflect the potential criteria for a formal assessment that were based existing literature. For instance, a study by Cleary, Zimmerman & Keating (2006) asked students to choose a response from a multiple-choice response item which included options such as performance of others, number of attempts, use of correct plan/technique, other factors and do not know. The current study presented students with the following choices: mastery-based, prior performance, normative comparison, social comparison, and no-self-evaluative strategy (Zimmerman & Moylan, 2009).

Table 6

Coding and Scoring Scheme for Self-Evaluation Criteria

Self-evaluation Criteria	Definition	Presented Choice
Mastery-based (3)	A statement that describes mastery-focused standards, such as skill, knowledge, or competency	I did well because was able to answer the questions well
Prior Performance (3)	A statement that describes evaluating one's performance based on how they performed on previous assessments/exams/tests	I did well because I did better than last time
Normative Comparison (2)	A statement that describes evaluating one's performance only the score, regardless of their past performance, or the performance of their peers.	I did well because of the grade/score I received
Social Comparison (1)	A statement that describes evaluating one's performance to their peer or classmate.	I did well because I did better than my friends
No evaluative strategy (1)	Statements that describe **no** evaluation criteria or being unaware of a strategy they use.	I don't know what I use to tell how well I have done, or I don't care how well I have done

The rating scheme for this microanalytic item reflected the extent to which the evaluation strategy reflected strong SRL processes. According to research, experts are more likely to evaluate their performance on mastery-based criteria or use their past performance to indicate how they have done (Zimmerman & Moylan, 2005). Novice learners are more likely to use their performance such as scores or successful attempts as strategies for self-evaluation, while poor regulators compare their performance to that of others or do not engage in self-evaluation at all (Zimmerman & Moylan, 2009). Therefore, an ordinal scale was used and students who reported strong evaluation strategies were rated 3, while those who used their performance scores were rated 2, and poor self-evaluation strategies such as social comparison or no evaluation strategy were rated at one.

Causal Attributions. Regarding causal attributions, students were asked two questions that targeted their reasons for success and/or failure (DiBenedetto & Zimmerman, 2013). The questions were: *"Why do you think you were able to perform successfully on some items in your assessment?"* and *"Why do you think you were not able to perform as successfully on some other items in your assessment?"* Students' responses were recorded verbatim, and codes were developed to reflect existing frameworks on causal attribution.

According to SRL research, effective and strong self-regulated learners attribute their success and/or failure of their performance on internal factors that are controllable such as effort and strategy use (DiBenedetto & Zimmerman, 2013). In the current study, students' responses to both microanalytic questions (i.e., for success and failure) were coded using Weiner attribution theory (1972) framework. The codes were established keeping in mind the three main aspects of the attribution theory: locus, stability, and control.

Locus. The locus refers to the location of the causal factor: internal (to the individual) or external (others; environment).

Controllability. This construct refers to the volitional control that the individual can exert on the causal factor. In other words, controllability indicates whether the causal factor was subject to the actor's control (controllable) or beyond it (uncontrollable).

Stability. This dimension refers to whether the causal factor remains stable over time or changes (i.e., variable).

Based on the above three dimensions, eight codes were established for classification. The description of each code is described in table 7 below. There is one additional code for when students do not report a causal attribution. The score for the ordinal-scale is the numerical value assigned beside each category.

Table 7

Coding and Scoring Scheme for Causal Attributions (Success and Failure)

Casual Attributions	Definition	Examples
Internal, stable, and controllable (3)	Statements that describe causes that are stable over time, ascribed to the individual and are controllable by the individual (e.g., long-term effort)	• I've been studying every day for the exam • I've been working hard from the beginning
Internal, unstable, controllable (3)	Statements that describe the cause to be attributed to the individual are controllable and but unstable over time (e.g., situational/temporary effort, skills/knowledge)	• I used the right strategies to learn • I made silly mistakes when writing the exam • I put in a lot of effort before the exam
Internal, stable, uncontrollable (1)	Statements that ascribe the cause to the individual, are stable over time, however uncontrollable. E.g., ability, aptitude	• I'm very good at numericals • Physics is not my thing
Internal, uncontrollable, and unstable (1)	Statements that ascribe cause to the individual, are uncontrollable and unstable over time. E.g., health on the day of the exam, mood	• I wasn't feeling it this morning • I wasn't feeling well this morning

External, Controllable, and Stable (2)	Statements that ascribe cause to sources outside the individual, are controllable, and stable over time. E.g., teacher bias, favouritism	• This teacher always marks low
External, controllable, unstable (2)	Statements that ascribe cause to sources outside the individual, are controllable, but unstable over time. E.g., help from friends/ teacher	• Sir explained this topic well • My friend helped me study
External, uncontrollable, stable (1)	Statements that ascribe cause to sources outside the individual, are uncontrollable, but stable over time. E.g., ease/difficulty of course or exam.	• The exam was very difficult • The questions were easy
External, uncontrollable, unstable (1)	Statements that ascribe cause to sources outside the individual, are uncontrollable, and unstable over time. E.g., luck, chance.	• I was lucky that all the questions I studied for came in the exam
No attributions (1)	Statements that describe no reason for success and failure on the task items. Only code if no other category fits.	• I don't know

The total number of responses for each factor was small, and therefore codes were grouped into an ordinal scale based on the degree to which they supported strong SRL processes. Students' responses that were indicative of internal, controllable factors (e.g., effort, strategy use) were rated at 3; responses that were internal, uncontrollable (e.g., ability) and external but controllable (e.g., help-seeking) were rated at 2, and those responses that were internal uncontrollable (e.g., mood/health) or external uncontrollable (e.g., luck) were rated at one on the scale.

Perceived Satisfaction. An important SRL sub-process is students' perceived satisfaction because their reactions to their performance influence future feelings for similar tasks or subject matter (Zimmerman, 2013). To find out students' perceived satisfaction, I asked the question, *"Are you satisfied with your performance?"* Previous studies have used a 100-point scale (e.g., DiBenedetto & Zimmerman, 2010). However, this scale was too large and was therefore modified to an open-ended question for the current study. Student responses were coded into an ordinal scale which included 3 for yes satisfied, 2 for somewhat satisfied, and 1 for not satisfied. Responses such as “kind of” were coded as somewhat satisfied.

Adaptive inferences. Closely related to perceived satisfaction is the type of inference students make based on their performance and outcome. In the current research, I sought to determine the degree to which participants showed an ability to critique the effectiveness of their learning strategies and in what ways they might have changed their learning strategies at the self-reflection stage (i.e., after completing their assessed tasks and seeing their grade). Participants were asked: *“If you were given another chance to do this assessment again, would you do anything differently?”* This question was modified from the one used in the study by DiBenedetto and Zimmerman (2010) to reflect the targeted task in this study.

The current study adopted a coding scheme different from previous studies to represent the range of student responses captured. I created an ordinal scale designed to estimate the extent to which each participant’s adaptive thinking and strategy improvement were described as precise, explicit, and detailed. Codes, therefore, ranged from specific inferences (high) general inferences (medium), and no inferences (low). Responses that reflected clear actions for future learning and performance were rated at 3 (e.g., I’ll practice diagrams), those responses that were vague and general (e.g., I’ll work harder) were rated at 2, and those responses that were not indicative of any inferences were rated at one (e.g., I won’t change/do anything different). A potential problem with this coding is that students who report no adaptive inferences may not want to change their approach because, in their view, they performed well. DiBenedetto and Zimmerman (2010) code this aspect as ability, which means that students may believe this is the best they can do and found that students with this reaction did not differ in their performance outcome. However, there were significant associations between specific strategies or general statements such as effort. Moreover, an inference related to ability is less adaptive than strategy use or effort (DiBenedetto & Zimmerman, 2010). Therefore, the coding and rating established for the study is appropriate. The table below describes the coding scheme and the ordinal rating for adaptive inferences.

Table 8

Coding and Scoring Scheme for Adaptive Inferences

Adaptive Inference	Definition	Example	Rating
Specific Inference	Statements that describe a particular change to their study plan or strategies.	• I'll work on my word problems • I'll do more practice tests	3
General Inference	These statements reflect general adaptations	• I'll study harder • I'll put in more effort	2
No inference	Statements that indicate no inferences	• I don't know • I won't change anything	1

Reliability and Validity

The current study used similar measures to establish reliability and validity as previous studies (e.g., Callan, 2014; DiBenedetto & Zimmerman, 2010; Kitsantas & Zimmerman, 2012). The SRL microanalytic protocol was piloted with 12 students in India for to test the instrument (Cohen et al., 2018). Only one question was modified based on the feedback. For example, the original question: "Do you have any plans on how you are going to prepare for this assessment? Please explain" was modified to include strategies. Additionally, if the original responses were generic such as *reading or studying,* two prompts were included to gain more granularity: *"how would you go about doing that?" and "what would that look like?"* The coding scheme for attribution and adaptive inferences were also modified to meet possible responses. The original coding scheme for attributions and inferences were based on the study by Kitsantas & Zimmerman (2010), which were: effort, ability, other, don't know. In the current study, student responses were more varied, and these codes did reflect those responses. For example, students reported attributions such as "I got help from my friend" which would be categorized into other. However, the category did not describe the action authentically. Thus, the coding scheme was modified to use attribution theory (Weiner, 1972).

Similarly, inferences made by students did not fit the established codes and were changed to capture students' responses as authentically as possible. Student responses included "I will work harder next time" which could be coded as effort, while others said, "I'll work harder on my word problems" which could be coded as effort as well. Inferences are critical to carry forward the learning cycle because they inform the following forethought stage (Zimmerman, 2013). As such, making specific inferences is likely to inform goal-setting and strategy selection scores. Therefore, the coding scheme for the current study was modified to reflect the nature of inferences.

The student interviews were coded by two independent coders who were also PhD students enrolled in the School of Educational and Leadership at the University of

Canterbury. I set up a meeting to explain the coding scheme and practiced with 10 random samples. After the initial meeting, I randomly selected 40 student responses and copied them on an excel sheet that was shared with the two independent coders. After they completed their coding, discrepancies in the codes were discussed until there was 100% agreement with all the responses (e.g., Callan, 2014). The discussion helped explain any cultural references that could have been interpreted in different ways. For example, in India, students often use the phrase ‘mug up’ to refer to rote memorization. This phrase was not familiar to the one of the coders, and a discussion on cultural references helped reach an agreement.

Analytical Framework for Summative Assessment Design Features

Developing an analytical framework helped me conceptualize the variety of features of the assessment tasks. By breaking down the task into its features, I was able to examine relations between task features and self-regulated learning sub-processes in the forethought and self-reflection phases. Based on the literature on task design and an analysis of the expectations for self-regulated learning, I identified five dimensions of task design (Lodewyk et al., 2006; Perry et al., 2004). Table 9 outlines those dimensions. Each of the five dimensions was developed into a five-point score scale. These descriptions were used to analyse the qualities of each assessment task. The assessment tasks were collected from participating schools. For a sample analysis of an assessment task, please refer to Appendix B.

Table 9

Analytical Framework for Summative Assessment Design Features

Criterion	1	2	3	4	5
Distribution of SOLO	Overall, the items on this assessment are focused on the unistructural and multistructural levels	There are more unistructural and multistructural items than relational and extended abstract	There is a moderate mixture of all four levels of SOLO	There are more relational and extended abstract than unistructural and multistructural items	The majority of items in the assessment are focused on the relational and extended abstract levels
Familiarity of Questions	Overall, the relational items are familiar to the students	There are more familiar relational items than unfamiliar relational items	There is a moderate mix of familiar and unfamiliar relational items	There are more unfamiliar relational items than familiar relational items	Overall, all the relational items are unfamiliar to the students.
Types of Questions	The items in the assessment are all close-ended requiring one word or phrase	There are more close-ended items than open-ended items	There is a moderate mixture of close and open-ended questions	There are more open-ended items than close-ended items	The items in the assessment are all open-ended requiring students to construct a response
Distribution of sub-topics	The items in the assessment focus on very few sub-topics from each unit	The items on this assessment unit focus on a few sub-topics from the unit	The items on this assessment focus on some sub-topics from each unit	The items on this assessment focus on most sub-topics from each unit.	The items in this assessment have a good distribution of all sub-topics from each unit
Relevance to rea-life/ application of knowledge to real-life situations	Students are never provided with opportunities to apply their learning to real-life situations	Students are rarely provided with opportunities to apply their learning to real-life situations	Students are occasionally provided with opportunities to apply their learning to real-life situations	Students are often provided opportunities to apply their learning to real-life situations	Students are frequently provided with opportunities to apply their learning to real-life situations

Dimensions of a Summative Assessment Task. The first criterion in the rubric is the distribution of learning outcomes based on the SOLO taxonomy (Biggs & Collins, 1989). This taxonomy was used to help examine the degree to which the task facilitated surface versus deep learning. The SOLO taxonomy classifies learning outcomes into four levels: unistructural, multistructural, relational, and extended abstract. Biggs and Collins use the depth of cognitive processing to distinguish these levels. At the unistructural level, the learning outcome focuses on a single element of a concept, such as a definition of a concept (e.g., metals and non-metals). The multistructural level refers to several aspects of a concept, for example properties of metals. At relational level, multiple aspects of the concept are understood in relation to each other (e.g., compare and contrast metals from non-metals). And the final level, extended abstract, refers to the transfer of knowledge, such that the student uses knowledge to extend understanding beyond what is presented on the topic. For example, reflecting on the use of metals and non-metals in the development of electric appliances would require students to apply their understanding of the concept. In this case, each question on the assessment was analysed, and the distribution was tallied; based on this distribution a score was assigned for this criterion.

The second criterion in the rubric is the familiarity of questions which examined how many items in the assessment were known to the students. This criterion, while not used in previous studies, is critical to help understand how students may engage with the content. For example, if a question is classified as relational (according to SOLO), the purpose of the question is for the student to integrate two elements of a concept that deepen conceptual grasp. But if students have the content that describes the interrelationship in the question, they might not need to fully understand the relationship, but rather familiarise themselves with the content through memorization techniques such as rehearsal. Early studies in educational research (e.g., Crooks, 1988) suggest that those students who need to only reproduce content from textbooks are more likely to engage in surface learning approaches. As previously mentioned, surface learning approaches are less likely to promote effective self-regulated learning processes. Based on the understanding from these findings, I incorporated a criterion that captured the number of questions that were from the textbook in the rubric. To determine whether the items were from the textbook, I examined the prescribed textbooks for students. If the answer to the question in the assessment was present in the content, the assessment item was classified as textbook-based. If there was no indication of the answer in the textbook, the item was categorized as non-textbook based. Upon completing this analysis, the number of items that were textbook-based and non-textbook based were tallied and assigned a score on the rubric based on the overall distribution of the items. It is important to mention that the textbook-based criterion of items was examined only for those items that were categorised as relational or extended abstract of the SOLO taxonomy. It would be unreasonable to expect students to recall unistructural and multistructural items that are not known to them; however, since the purpose of relational and extended abstract is to measure deeper conceptual understanding, it is reasonable to examine whether they were sourced from the textbook of these items.

The third criterion was developed based on prior research suggesting associations between the nature of the question and students' self-regulated learning. For example, task items that are moderately ill-structured require students to use effective SRL processes; this is because the task gives students limited guidance on what is expected in their response (Lodewyk et al., 2006). In the current study, items on the assessment tasks were analysed based on their nature, but were modified to suit the constraints (e.g., school policy) of the assessment. The basic premise of using ill-structured tasks is to allow students to respond such that they demonstrate their abilities to integrate, evaluate, and describe knowledge. As the task measures learning outcomes related to physical and natural science domains, students are likely to be assessed on subject–related skills which included identifying, describing, and explaining scientific phenomena as well as represent the working of science using diagrams or, in the case of physics, establish relationships between various concepts using mathematical derivations (Crooks & Collins, 1986; Doyle, 1983). Given this understanding, the nature of the questions was investigated based on students' expected response, which included close-ended (one sentence/phrase/multiple choice), open-ended (construct a response), or other (e.g., diagrammatic or computational). A criterion describing the distribution of the different questions was incorporated into the rubric.

The fourth and fifth criteria were included as a result of findings from a study by Perry et al. (2004), which highlights the association between assessing multiple learning outcomes, including real-life questions and students' SRL engagement making two additional criteria to the analytical framework. To identify the learning outcomes, I used the questions presented in the task and the content from the prescribed textbook to determine the distribution of the content. Based on this distribution, a score was assigned based on the total number of sub-topics in the textbook and the number assessed in the assessment. To examine real-life questions, I checked the content of the question. In particular, the question would require students to use a scientific concept to explain its application in real life. For example, in a physics exam, one question asked students to justify why cinema screens were made rough using the concept of reflection/refraction of light. In all, I identified five aspects of the assessment and computed an overall score for each assessment. Following this, each of these five criteria were used to examine how they might relate to students' SRL.

Reliability and Validity. The analytical framework was coded by one other independent coder who has worked extensively with teachers in India and delivered professional learning development workshops on assessment and learning. I shared the analytical framework and shared one assessment with the independent coder. Inter-rater agreement was high (86%). Any discrepancies were resolved in a discussion. Based on the discussion, a description for the levels of SOLO taxonomy and nature of questions were made explicit and examples were included to facilitate the coding (See Appendix B).

Student Performance

I used students' performance outcomes as measures of student performance. The scores were teacher evaluations of student work. Science teachers had their unique practices of student evaluation and provided me with the raw scores. Teachers assessed student work based on how accurately student answers represented the expectations of the question and assigned them scores. Students' teacher-assigned scores were converted to percentages for further analyses.

3.3.5 Analysis

I analysed SRL measures and assessment tasks independently using the established coding strategies and rating scales. Descriptive and inferential statistical tests were conducted to address the research questions. Descriptive tests were conducted to gain an overview of the data. Pearson's product moment coefficient was used to measure correlations among SRL processes, student performance, and assessment design features (Pallant, 2012). Preliminary analyses were performed to ensure no violation of the assumptions of normality, linearity, and homoscedasticity. Additionally, regression analyses were used to check for predictive patterns in the data (Pallant, 2012). The next section reports the findings of these tests. Post-hoc logistic regressions were also performed to strengthen the established coding and rating scheme.

3.4 Results

3.4.1 RQ 1: Which SRL Strategies Reflected in Zimmerman's Cyclical Model do Students Demonstrate for a Science Summative Assessment Task?

Analyses began with a bird's eye view of SRL processes in the forethought and self-reflection phases (Table 10). In the forethought phase, I elaborate on goal-setting, strategy selection, and motivational beliefs (self-efficacy, intrinsic interest, and perceived value). And in the self-reflection phase, I report on students' self-evaluation, causal attribution, and self-reactions (perceived satisfaction and adaptive inferences).

Table 10

Descriptive Statistics for Forethought and Self-Reflection Sub-Processes (n=229)

SRL Phase	Sub-processes	Minimum	Maximum	*M*	*SD*
Forethought Task Analysis	Goal-setting	1	3	2.04	.48
	Strategy Selection (breadth)	1	5	1.72	.654
	Strategy Selection (depth)	1	4	1.78	.931
Forethought	Self-efficacy for preparation	1	5	3.63	.93
	Self-efficacy for performance	1	5	3.65	.99
Motivational beliefs	Perceived Interest	1	5	3.74	1.33
	Perceived importance	1	5	4.36	.80
Self-reflection	Self-evaluation	1	5	3.61	.98
	Judgement criteria	1	3	2.27	.64
Self-judgement	Causal attribution for success	1	3	2.22	.95
	Causal attribution for failure	1	3	1.98	.98

Self-reflection	Perceived satisfaction	1	3	2.08	.895
Self-reaction	Adaptive inferences	1	4	2.12	1.09

Most students had outcome goals (general and specific) for the assessment task and reported a moderate breadth of strategies (i.e., between two and three strategies), which included rehearsal and comprehension techniques. With respect to students' motivational beliefs, on average, students reported feeling competent about preparing and performing on the assessments. Students' interest in and value of the assessment were higher than the scale's midpoint (Table 10). Moreover, it can be seen from Table 10 that students' goal-setting and strategy selection scores (breadth and depth) had lesser variation when compared to their motivation.

With respect to self-reflection measures, students reported that they felt they had adequately prepared for the assessment after receiving their performance outcomes (M = 3.61 and SD = .98). Most students (49.3%) used teacher-assigned evaluations to judge their performance and felt a sense of satisfaction. I also asked students to report on their attributions for success and failure. Most students attributed their success and failures to situational and temporary aspects, e.g., "I practiced a lot for the exam," or "I made some careless mistakes." A majority of students reported adaptive inferences that were either specific (e.g., "I'll practice the word problems more") or general (e.g., "I'll work harder").

3.4.2 RQ 2: To What Extent Does SRL Predict Students' Academic Performance on Science Summative Assessment Tasks?

Forethought Measures

Pearson coefficient was conducted to check for associations between sub-processes of task analysis and academic performance. No correlation was found between goal setting and academic performance, but a small positive correlation between strategy selection (breadth and depth) and academic performance (Table 6).

The difference in mean scores between students who used surface-level and deep-level strategies was also examined. An independent-samples *t*-test was conducted to compare the performance scores for surface and deep learning strategies. There was a significant difference in scores for students who used surface level strategies (M = 71.55, SD = 17.92) and students who used deep level strategies, M = 82.02, SD = 15.36; t (227) = -2.53, p = .01 (two-tailed). The magnitude of the differences in the means (mean difference = -10.47, 95% CI: -18.64 to -2.29) was very small (eta squared = .03).

Student performance on the assessment positively correlated with motivational beliefs in the forethought phase. Notably, students who reported having higher interest and perceived self-efficacy for preparation performed better. As can be seen from Table 11 correlations between most forethought sub-processes and performance were statistically significant except for goal-setting scores. A regression analysis was performed to check if the relationship was non-linear: no cubic or quadratic association was found between goal-setting and performance outcomes. The relationship between goal-setting and performance could be non-significant because there was lesser variability in students' goals for this specific assessment task.

Table 11

Pearson's Correlation Coefficient for Forethought Measures and Performance Outcome (n = 229)

	Self-efficacy preparation	Self-efficacy performance	Interest	Perceived importance	Goal-setting Score	Strategy Selection, Breadth	Strategic Selection Depth
Performance outcome	.394**	.251**	.446**	.186**	.050	.161*	.163*

*Correlation is significant at the .01 level (2 tailed)
** Correlation is significant at the 0.01 level (2-tailed).

Self-Reflection Measures

Self-reflection measures comprised students' self-evaluation and self-reaction reports. Self-evaluation measures included perceived preparedness, self-judgement criteria, and causal attributions. Students' perceived preparedness positively correlated with academic performance, with high levels of self-evaluation associated with high levels of performance (Table 12). No statistically significant associations were found among students' self-judgment scores, causal attribution scores, and academic performance**.**

Table 12

Pearson's Correlation Coefficient for Self-Reflection Measures and Performance Outcome (n = 229)

	Self-evaluation	Self-judgement	Causal Attribution - Success	Causal Attribution - Failure	Inferences
Performance outcome	.559**	-0.006	0.078	-0.106	-.132*

** Correlation is significant at the 0.01 level (2-tailed).

Self-reaction processes were perceived satisfaction and adaptive inferences. There was a medium, positive correlation between students' perceived satisfaction scores and academic performance, with higher satisfaction scores associated with higher academic performance. There was a small, positive correlation between students' inference scores and academic performance, with higher inference scores associated with higher academic performance.

Standard multiple regression was used to assess the ability of the SRL measures to predict levels of student performance. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity, and homoscedasticity. The total variance explained by the model as a whole was 45.3%, F (12228) = 15.35, $p < .001$. Students' perceived self-evaluation scores ($\beta = .36$, $p < .001$), perceived interest ($\beta = .29$, $p < .001$), perceived satisfaction ($\beta = .19$, $p < .001$), and goal-setting ($\beta = .14$, $p < .01$) were found to predict student performance. In the model, self-efficacy showed no significant relation to student performance (see Table 13).

Table 13

Standard Multiple Regressions for SRL Sub-processes Predicting Performance Outcome (n = 229)

	Variable	*B*	*SE*	*B*	R^2	*t*	95% CI		*Sig.*
							Lower	Higher	
1					.429				
	Goal-setting	5.06	1.96	.14		2.57	1.20	8.91	.01
	Strategy Selection (depth)	1.40	1.46	.05		.96	-1.48	4.27	.34
	Strategy Selection (breadth)	-.10	1.03	-.01		-.10	-2.13	1.92	.92
	Self-efficacy for preparation	.90	1.24	.05		.72	-1.55	3.34	.47

Self-efficacy for performance	-.71	1.11	.047	-.642	-1.555	3.344	.52
Interest	3.92	.85	-.039	4.591	-2.891	1.470	<.001
Perceived importance	-.33	1.25	.292	-.261	2.238	5.606	.80
Self-evaluation	6.57	1.20	-.015	5.475	-2.783	2.133	<.001
Self-judgement scores	-1.38	1.44	-.05	-.96	-4.21	1.451	.93
Causal Attribution for Success	.14	.98	.01	.14	-1.80	2.07	.92
Causal Attribution for Failure	-.39	.94	-.02	-.41	-2.25	1.47	.94
Perceived Satisfaction	3.89	1.13	.19	3.43	1.65	6.12	<.001
Adaptive Inferences	1.77	.874	.11	2.03	.05	3.49	.04

These results indicate that forethought and self-reflection measures positively correlate with academic performance. Group differences based on strategy selection were also evident, with students employing deeper learning approaches performing on average better than surface-level learning approaches. Overall, SRL motivational beliefs were found to be stronger predictors of academic performance than the task analysis or self-judgment sub-processes for the assessment task in the current study.

3.4.3. RQ 3: In what ways do design features of an assessment task relate to students' SRL strategies reflected in Zimmerman's cyclical model (2009)?

Descriptive Statistics for Assessment Design

From Table 14, it can be seen that the assessment design features tended to fall close to minimum value for SOLO distribution, which indicates that the assessments measured more multistructural items that required description and application of knowledge. For the criterion referring to familiarity, the descriptive data indicate that the assessment items were moderate with some unfamiliar items. On average, the assessments scored above the mid-point on aspects related to types of questions and distribution of sub-topics. This descriptive data suggests that the assessment had a good mixture of question types (e.g., one-word questions, descriptive, computational questions), and covered a large portion of the unit of learning. The aspect criterion of real-life questions fell below the mid-point, suggesting that few items on the assessment related to real-life contexts or situations.

Table 14

Descriptive Statistics for Assessment Design Features (n = 9)

Design features	Min.	Max.	Mean	SD
Distribution of SOLO	2	4	2.89	.741
Familiarity of Items	1	3	2.54	.617
Types of Question	2	4	3.01	1.002
Distribution of sub-topic outcomes	2	4	3.60	.500
Relevance to rea-life/ application of knowledge to real-life situations	2	3	2.37	.483

A power analysis was conducted to check for the limited sample number required for multiple regressions at 90% (Cohen, 1977). From Table 15, it can be seen that the sample size has power for conducting multiple regressions.

Table 15

Power Analysis of Sample Size for Standard Regressions

			Predictors		Test Assumptions		
	N	Actual Power[b]	Total	Test	Power	Effect Size[c]	Sig.
Type III F-test[a]	79	.904	5	5	.9	.236	.05

a. Intercept term is not included.
b. Predictors are assumed to be random.
c. Cohen's f-squared.

The section below describes investigations between assessment design features and various sub-processes in the SRL phases.

Assessment Design and Forethought Measures

Relationships between the different criteria and task analysis sub-processes (goal-setting and strategy selection) and motivational beliefs (self-efficacy, interest, and perceived importance) of the forethought phase were examined. Notably, goal-setting scores were positively correlated with assessment design features. For example, process goals were associated with design features such as types of questions (i.e., more variety in questions as one-words, open-ended questions). A small negative correlation was found among SOLO levels, types of questions, and strategy selection (breadth).

On average, motivational beliefs were associated with all design features. Self-efficacy for preparation and performance negatively correlated with most design features. Markedly, lesser interest in science was associated with higher score features such as SOLO distribution, textbook-based items, types of questions, and distribution of sub-topic outcomes (i.e., number of questions that covered several sub-topics was high). Concerning perceived task importance, medium, negative correlations were found between overall assessment score and reported importance, with higher assessment scores associated with lower levels of task importance. Table 16 highlights the correlations among assessment design features and forethought measures.

Table 16

Correlations Among Assessment Design Features and Forethought Measures (n = 229)

	Distribution of SOLO	Textbook based Items	Types of Question	Distribution of sub-topic outcomes	Relevance to real-life/ application of knowledge to real-life situations
Self-efficacy preparation	-.328**	-.327**	-.301**	-.298**	.173**

Self-efficacy performance	-.341**	-.297**	-.366**	-.330**	.280**
Interest	-.519**	-.421**	-.577**	-.465**	.508**
Perceived importance	-.349**	-.308**	-.294**	-.274**	.180**
Goal Setting Score	.135*	.172**	.233**	.163*	-.144*
Strategy selection (depth)	-.014	-.078	.212	.161*	-.040
Strategy selection (breadth)	-.142*	-.103	-.178**	-.066	.231**

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Assessment Design and Self-Reflection Measures

I also examined relationships between the assessment design features and self-reflection sub-processes, which included components of self-judgment (self-evaluation and causal attributions) and self-reaction (perceived satisfaction, adaptive inferences). Overall, there was a medium, negative correlation between students' perceived levels of preparation and the assessment design features. From Table 17, it can be seen that there were medium, negative correlations between students' perceived self-evaluation and most assessment design features. But there was a medium, positive correlation with the number of real-life questions, with higher levels of perceived self-evaluation associated with a higher score of real-life questions. Overall, no correlations were found among students' self-judgment criteria, causal attributions, and most assessment design features. Table 17 highlights the correlations among assessment design features and self-reflection measures.

Table 17

Correlations among Assessment Design Features and Self-Reflection Measures (n = 229)

	Distribution of SOLO	Textbook-based Items	Types of Question	Distribution of sub-topic outcomes	Relevance to real-life/ application of knowledge to real-life situations	Sum

Self-evaluation	-.333**	-.243**	-.356**	-.251**	.369**	-.300**
Causal Attribution-Success	-0.01	-0.072	-0.092	-0.097	0.012	-0.08
Causal Attribution-Failure	0.058	-0.01	0.031	-0.002	-.165*	-0.005
Inferences	.214**	.193**	.258**	.200**	-.227**	.226**

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

On average, a small, positive correlation was found between students' adaptive inference scores and assessment design features, with higher assessment design scores associated with higher adaptive inference scores. The relationship between students' perceived satisfaction and assessment design features negatively correlated with each other. Due to the contradictory nature of this result, I conducted a partial correlation to explore the relationship between perceived satisfaction and assessment design features, while controlling for performance outcome measures. Results indicated no significant association between assessment design features and students' perceived satisfaction.

The current study was interested in examining factors that predict sub-processes in the forethought and self-reflection phases in Zimmerman's model of SRL. A series of multiple regression analyses were performed for which assessment design features and performance scores were independent variables and SRL sub-processes were the dependent variables. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity, and homoscedasticity.

Results indicated that some assessment design features predicted motivational beliefs in forethought and self-reflection measures. Among the motivational beliefs in the forethought phase, the total variance explained by the model for interest was 41.5%, *F* (5,

228) = 8.77, *p* < .001. One assessment design feature, types of questions, was found to contribute more to interest recording a higher beta value (*β* = -.39, *p* < .005) than another assessment design feature, distribution of SOLO (*β* = -.21, *p* < .05) and performance outcome measures (*β* = .28, *p* < .001). The total variance in cognitive processes such as goal-setting and strategy selection was explained to a small extent by the regression models. For instance, the design feature, distribution of sub-topics recorded a higher beta value for goal-setting (*β* = .45, *p* < .005), questions relevant to real-life recording a higher beta value for breadth of strategy selection (*β* = .24, *p* < .05), and distribution of sub-topics recorded a higher beta value for depth of strategy selection (*β* = .35, *p* < .005).A summary of the multiple regressions is described in Tables 18-24.

Table 18

Summary of Multiple Regression Analysis for Task Features and Performance Predicting Self-Efficacy for Learning Processes (n=229)

	Variable	*B*	*SE*	*Β*	R^2	*t*	*95% CI*		*Sig.*
							Lower	*Upper*	
1					.218				
	Distribution of SOLO	-.34	0.13	-.27		-2.55	-.594	-.076	.01
	Types of Question	-.22	0.13	-.23		-1.65	-.479	.043	.10
	Distribution of sub-topics	.07	0.23	0.04		0.32	-.379	.523	.75
	Relevance to real-life	-.53	0.19	-.27		-2.68	-.915	-.140	.01
	Performance	.02	0.00	0.35		5.58	-.012	.025	.00

Table 19

Summary of Multiple Regression Analysis for Task Features and Performance Predicting Self-Efficacy for Performance (n=229)

Variable	*B*	*SE*	*Β*	R^2	*t*	*95% CI*		*Sig.*
						Lower	*Upper*	

1				.146				
Distribution of SOLO	-.19	.15	-.15		-1.34	-.486	.092	.18
Types of Question	-.25	.15	-.25		-1.68	-.540	.043	.09
Distribution of sub-topics	-.01	.26	-.01		-.05	-.516	.489	.96
Relevance to real-life	-.13	.22	-.06		-.59	-.562	.302	.56
Performance	.01	.00	.15		2.34	.001	.016	.02

Table 20

Summary of Multiple Regression Analysis for Task Features and Performance Predicting Interest (n=229)

Variable	*B*	*SE*	*β*	R^2	*t*	*95% CI*		*Sig.*
						Lower	*Upper*	
1				.415				
Distribution of SOLO	-.38	.16	-.21		-2.36	-.703	-.062	.02
Types of Question	-.52	.16	-.34		-3.10	-.843	-.201	.00
Distribution of sub-topics	.31	.28	-.11		1.08	-.251	.862	.28
Relevance to real-life	.09	.24	-.03		.41	-.380	.577	.67
Performance	.02	.00	.28		5.25	.013	.029	.00

Table 21

Summary of Multiple Regression Analysis for Task Features and Performance Predicting Perceived Importance (n=229)

Variable	*B*	*SE*	*β*	R^2	*t*	*95% CI*		*Sig.*
						Lower	*Upper*	
1				.133				
Distribution of SOLO	-.41	.12	-.39		-3.48	-.648	-.349	.00
Types of Question	-.22	.12	-.28		-1.87	-.459	-.294	.06
Distribution of sub-topics	.25	.21	.15		1.18	-.163	.652	.24
Relevance to real-life	-.37	.18	-.24		-2.23	-.746	-.046	.03
Performance	.01	.00	.11		1.69	-.001	.011	.09

Table 22

Summary of Multiple Regression Analysis for Task Features and Performance Predicting Goal-Setting (n=229)

	Variable	*B*	*SE*	*β*	R^2	*t*	*CI Intervals* Lower	*CI Intervals* Upper	*Sig.*
1					.057				
	Distribution of SOLO	-.03	.08	-.04		-.36	-.175	.121	.72
	Types of Question	-.22	.08	.45		-2.89	.069	.367	.004
	Distribution of sub-topics	-.09	.13	-.10		-.72	-.350	.163	.47
	Relevance to real-life	.09	.11	-.09		.76	-.135	.306	.45
	Performance	.00	.00	.13		1.89	.000	.007	.06

Table 23

Summary of Multiple Regression Analysis for Task Features and Performance Predicting Strategy Selection - Breadth (n=229)

	Variable	*B*	*SE*	*B*	R^2	*t*	95% CI Lower	95% CI Higher	*Sig.*
1					.044				
	Distribution of SOLO	.001	.131	.001		.008	-.256	.259	.993
	Types of Questions	-.077	.156	-.083		-.495	-.384	.230	.621
	Distribution of sub-topics	.137	.178	.091		.772	-.213	.487	.441
	Relevance to real-life	.333	.233	.173		1.43	-.125	.792	.154
	Performance	.005	.004	.099		1.44	-.002	.012	.151

Table 24

Summary of Multiple Regression Analysis for Task Features and Performance Predicting Strategy Selection – Depth (n = 229)

	Variable	*B*	*SE*	*B*	R^2	*t*	95% CI Lower	95% CI Higher	*Sig.*
1					.075				
	Distribution of SOLO	-.221	.090	-.250		-2.44	-.399	-.043	.015
	Types of Questions	-.047	.108	-.072		-.437	-.259	.165	.663
	Relevance to real-life	-.246	.161	-.182		-1.532	-.563	.070	.127
	Distribution of sub-topics	.371	.123	.350		3.024	.129	.613	.003
	Performance	.008	.002	.210		3.102	.003	.013	.002

Among the self-reflection measures, the cognitive sub-processes of self-evaluation and causal attribution was examined using multiple regressions. The total variance explained by the model as a whole for self-evaluation was 36.8%, F (5, 228) = 25.95, $p <$ .001. Performance was found to be the only contributor to the self-evaluation measures recording a high beta value (β = .48, p < .001). And the total variance explained by the model as a whole for self-judgement scores was 36.8%, F (5, 228) = 1.99, p > .08. The assessment design feature, types of questions, was the only contributor to the self-judgment measures recording a high beta value (β = .36, p < .05). For causal attributions, there were no significant associations.

With respect to self-reaction sub-processes, perceived satisfaction and adaptive inferences, academic performance (β = .41, p < .001) was found to predict students' perceived satisfaction scores, while there were no significant associations with adaptive inferences. The Tables 25 and 26 summarise the results for self-reflection processes. Post-hoc power analyses for multiple regressions were conducted, for which r squared values were entered as the effect size (Pallant, 2018). Results indicated that the minimum needed was $N = 79$. Given that the sample size for the current study was 229, the multiple regressions have the necessary power.

Table 25

Summary of Multiple Regression Analysis for Task Features and Performance Predicting Self-Evaluation Measures (n=229)

Variable		*B*	*SE*	*β*	R^2	*t*	*95% CI Intervals*		*Sig.*
							Lower	*Upper*	
1					.354				
	Distribution of SOLO	-.23	.13	-.17		-1.83	-.479	.018	.07
	Types of Question	-.16	.13	-.16		-1.28	-.411	.088	.20
	Distribution of sub-topics	.35	.22	.18		1.59	-.081	.781	.11
	Relevance to real-life	.14	.19	.07		.76	-.227	.514	.45
	Performance	.03	.00	.48		8.53	.020	.033	.00

Table 26

Summary of Multiple Regression Analysis for Task Features and Performance Predicting Self-Judgement Measures (n=229)

Variable		*B*	*SE*	*β*	R^2	*t*	*CI Intervals*		*Sig.*
							Lower	*Upper*	
1					.021				
	Distribution of SOLO	-.17	.10	-.21		-1.82	-.387	.016	.07
	Types of Question	.23	.10	.36		2.24	.027	432	.03
	Distribution of sub-topics	-.00	.18	.00		-.00	-.350	.349	.99
	Relevance to real-life	.08	.15	.06		.52	-.222	.380	.61
	Performance	.00	.00	.03		.44	-.004	.006	.66

Post-hoc Analyses

Direct logistics regressions were performed to assess the impact of a set of variables (design features and performance) on the odds that respondents would report stronger SRL strategies (Pallant, 2018). These post-hoc analyses were conducted to confirm that the ordinal nature of scaling did not affect the interpretation of the results. The results were found to generally add support for the ordinal scaling used.

Logistic regressions for forethought measures

Direct logistic regressions were performed to assess the impact of a set of variables on the odds that respondents would report that they set higher level goals (i.e., process goals). The model contained three independent variables (distribution of SOLO, distribution of sub-topics, and relevance to real-life). The full model containing all the predictors was statistically significant, χ^2 (3, *N* = 229) = 11.682, $p < .05$, indicating that the model was able to distinguish between participants who reported process goals versus reported poor goal-setting strategies (e.g., other, non-task related goals). The model as a whole correctly classified 86% of cases. As shown in Table 27, only one independent variable made a unique statistically significant contribution to the model (i.e., distribution of sub-topics or breadth of content), recording an odds ratio of 3.73. This indicated that the odds are 3.73 times greater than participants who had a broader content coverage would report process goals than those participants who did not have as much content coverage in their assessment. This finding is consistent with the ordinal-based analyses performed earlier.

Table 27

Logistic Regression Predicting Likelihood of Reporting Process goals

	B	*SE*	Wald	*df*	*P*	Odds Ratio	95% CI for Odds Ratio	
							Lower	Upper
Distribution of SOLO	-.519	.433	1.43	1	.231	.56	.255	1.39
Distribution of Sub-Topics	1.315	.636	4.28	1	.039	3.73	1.072	12.96
Number of Questions Relevant to real-life	-.837	.571	2.14	1	.143	.14	.14	1.32

Direct logistic regressions were performed to assess the impact of a set of variables on the odds that respondents would report deep learning strategies. The model contained three independent variables (distribution of SOLO, distribution of sub-topics, types of questions, and relevance to real-life). The full model containing all the predictors was statistically significant, chi square (4, N = 229) = 12.033, $p < .01$, indicating that the model as able to distinguish between participants who reported deep learning strategies versus reported surface learning strategies. The model as a whole correctly classified 91.3% of cases. As shown in Table 28, only one independent variable made a unique statistically significant contribution to the model (i.e., distribution of sub-topics or breadth of content), recording an odds ratio of 16.27. This indicated that the odds are 16.27 times greater than participants who had a broader content coverage would report deep learning strategies than those participants who did not have as much content coverage in their assessment. This result was also found to be consistent with the ordinal-based analyses performed earlier.

Table 28

Logistic Regression Predicting Likelihood of Reporting Deep Learning Strategies (n = 229)

	B	*SE*	Wald	*df*	*P*	Odds Ratio	95% CI for Odds Ratio	
							Lower	Upper
Distribution of SOLO	-.756	.696	1.19	1	.278	.47	.120	1.838
Distribution of Sub-Topics	2.789	1.012	7.60	1	.006	16.27	2.239	118.21
Types of questions	-.715	.669	1.14	1	.285	.49	.132	1.814
Number of Questions Relevant to real-life	-.248	1.197	.04	1	.836	.78	.08	8.155

Logistic regressions for self-reflection measures

Direct logistic regressions were performed to assess the impact of a set of variables on the odds that respondents would report that they use their prior performance or mastery-based criteria to judge their performance. The model contained three independent variables (distribution of SOLO, types of questions, and relevance to real-life). The full model containing all the predictors was statistically significant, chi square (3, *N* = 229) = 8.876, *p* < .05, indicating that the model as able to distinguish between participants who reported higher self-evaluation strategies versus those who reported poor self-evaluation strategies (e.g., normative comparisons). The model as a whole correctly classified 61.1% of cases. As shown in Table 29, only one assessment design feature made a unique statistically significant contribution to the model (i.e., types of questions), recording an odds ratio of 2.028. This indicated that odds are 2.028 times greater participants would report higher self-evaluation strategies when the assessment covered several types of questions than those participants who had fewer diverse questions on their assessment. This result is consistent with the analyses performed using the ordinal scale.

Table 29

Logistic Regression Predicting Likelihood of Reporting Higher-level Self-Evaluation Strategies (n = 229)

	B	*SE*	Wald	*df*	*P*	Odds Ratio	95% CI for Odds Ratio Lower	Upper
Distribution of SOLO	-.415	.291	2.028	1	.154	.660	.373	1.169
Relevance to real-life	.365	.499	.535	1	.464	1.441	.542	3.830
Types of questions	.707	.278	6.456	1	.011	2.028	1.176	3.500

Direct logistic regressions were performed to assess the impact of a set of variables on the odds that respondents would report that they satisfied with their performance. The model contained three independent variables (distribution of SOLO, types of questions, and relevance to real-life) and performance outcomes. The full model containing all the predictors was statistically significant, chi square (4, *N* = 229) = 50.027, *p* < .001, indicating that the model as able to distinguish between participants who reported being satisfied with their performance versus those who reported not being satisfied with their performance. The model as a whole correctly classified 61.1% of cases. As shown in Table 30, only performance outcome made a unique statistically significant contribution to the model, recording an odds ratio of 1.061. This indicated that odds are 1.061 times greater participants would report being satisfied when their performance outcomes were higher than those participants who performed poorly. This result is consistent with the analyses performed using an ordinal scale.

Table 30

Logistic Regression Predicting Likelihood of Reporting as being Satisfied with their Performance (n = 229)

	B	*SE*	Wald	*df*	*P*	Odds Ratio	95% CI for Odds Ratio Lower	Upper
Distribution of SOLO	.079	.315	.063	1	.802	1.082	.584	2.005
Relevance to real-life	-.519	.532	.950	1	.330	.595	.210	1.690
Types of questions	-.430	.290	2.199	1	.138	.651	.369	1.148
Performance outcomes	.058	.011	29.968	1	<.001	1.060	1.038	1.082

Direct logistic regressions were performed to assess the impact of a set of variables on the odds that respondents would report that they attribute their failures to controllable factors (i.e., effort, strategies). The model contained three independent variables (distribution of SOLO, types of questions, and relevance to real-life). The full model containing all the predictors was statistically significant, chi square (4, N = 229) = 12.053, $p <$ .01, indicating that the model as able to distinguish between participants who reported controllable factors as attributions versus reported uncontrollable attributions (e.g., task difficulty). The model as a whole correctly classified 54.8% of cases. As shown in Table 31, two of the three independent variables made a unique statistically significant contribution to the model (i.e., types of questions and questions relevant to real-life), recording an odds ratio of .208 and .561.

More importantly, the *B* values for these predictor variables are negative which suggest that an increase in the independent variable (i.e., assessment design features) will result in a decreased probability of the case recording a score of 1 in the dependent variable (indicating controllable attributions, in this case). Since these values are less than one, it can be inverted, to 4.81 and 1.78 by dividing the values by 1. This indicated that the odds are 4.81 times greater that participants who had less questions that are relevant to real-life situations would attribute their failures to controllable factors than those participants who had more questions that were relevant to real-life situations in their assessment. This finding was not evident in the analyses performed with the ordinal scale, and thus adds additional insights into the relationships between students' attributions and assessment design features.

Table 31

Logistic Regression Predicting Likelihood of Reporting Controllable Attributions (n = 229)

	B	*SE*	Wald	*df*	*P*	Odds Ratio	95% CI for Odds Ratio Lower	Upper
Distribution of SOLO	.099	.696	.119	1	.730	1.10	.629	1.94
Relevance to real-life	1-.568	.498	9.93	1	.002	.209	.079	.553
Types of questions	-.579	1.900	7.34	1	.007	.561	.327	.960

Direct logistics regressions were also conducted to assess the impact of assessment design features and performance outcomes on adaptive inferences. There were no statistically significant models for causal attributions for success and adaptive inferences. This was consistent with previous analyses performed using an ordinal scale.

In summary, descriptive statistics indicate that participants in this study have a moderate approach to self-regulated learning, which means that some SRL processes are adaptive and stronger than others. One of the objectives for the current study was to investigate the relationship between SRL and academic performance. Findings suggest SRL sub-processes such as self-efficacy, interest, and strategy selection relate to performance, but interest, self-evaluation, and perceived satisfaction emerged as key predictors of performance. Another objective of the study was to examine the relation between assessment design features and SRL. Results suggest that assessment design features relate to SRL in both positive and negative ways. For instance, self-efficacy and interest scores were lesser among students when the cognitive complexity of the assessment increased. Furthermore, design features such as cognitive complexity and types of questions predict cognitive processes (e.g., goal setting and strategy selection) and motivational beliefs (e.g., self-efficacy and interest). Logistic regressions also indicated similar results to multiple regressions, with distribution of sub-topics emerging as a predictor that increases the likelihood of students reporting stronger SRL processes in the forethought phase (e.g., goal-setting, deep learning strategies) and types of questions emerged as a predictor that increases the likelihood of students reporting stronger SRL processes in the self-reflection phase (e.g., causal attributions and self-evaluation strategies).

3.5 Discussion

In India, student achievement in science-related outcomes is declining (NAS, 2017). Research suggests that SRL processes promote academic success; however, not all students employ effective SRL strategies and attitudes due to personal or environmental factors. One critical factor that contributes to student learning is classroom assessment processes. Therefore, the current study aimed to understand students' SRL processes and whether these processes relate to performance and assessment design. Overall, SRL was associated with both student performance and assessment design in meaningful ways.

The picture that emerges from the analyses above is one of SRL being moderately adopted as students prepare for science assessments. I established a link between SRL approaches and academic performance. Most notable associations with academic performance were the breadth of learning strategies and motivational beliefs. While it is beneficial for students to have a large repertoire of strategies, they predominantly reported using surface learning strategies. Although surface learning strategies are integral to the learning process, they are inadequate for complex problem-solving, critical thinking, and logical reasoning for science assessments (Hattie & Donoghue, 2016; Sinatra & Taasoobshirazi, 2018). If students are likely to resort to rehearsal and comprehension learning strategies, it may limit their ability on science achievement tests that expect more complex reasoning. Future research can focus on gaining a deeper understanding of not just how students approach learning tasks, but also the reasoning for their approach. For instance, researchers can design questions that vary in complexity and investigate environmental factors (question design) and personal factors (awareness and purpose of strategies) that contribute to students' learning approach. Furthermore, this link, between performance and strategy use, provides scope for improving student learning and performance. Educators can teach students deep learning strategies such as elaboration and organizational strategies that are pertinent to scientific thinking (Sinatra & Taasoobshirazi, 2018).

Students' motivational beliefs in the forethought phase which include perceived competencies in learning and performance, interest, and perceived value, positively correlated with academic performance. Indeed, these results are consistent with scholarly work on motivational beliefs and their relation to achievement and SRL (Schunk & Zimmerman, 2007; Zimmerman, 2002). Among all motivational beliefs analysed in the current study, intrinsic interest was associated with better strategy use, self-efficacy, perceived value, and academic performance. This finding is of significance to both educators and researchers. For instance, educators can focus on increasing students' interest in the subject using a teaching approach such as project-based learning which emphasizes the meaningfulness of the science outcomes by fostering a deep engagement with the content. Future research can focus on understanding how interest is linked to other motivational constructs in the SRL model for science achievement. A broader perspective of these interactions can provide useful directions for interventions that could improve students' scientific thinking and achievement.

A few additional points should be made about the observations between the self-reflection phase of SRL and academic performance. Most students reported using their performance outcomes as the standard for self-judgement. This finding is not surprising, especially because most students reported having outcome goals, and performance outcomes are inextricably linked to summative assessments (Andrade & Brookhart, 2016). Furthermore, this self-judgement standard may also explain the link between students' self-evaluation measure and academic performance; students who perceived they were well-prepared also secured higher outcomes on the assessments. However, normative reference for self-judgement limits students in using more adaptive SRL approaches, which are crucial for academic performance and deep conceptual learning (Zimmerman, 2013). For instance, using the grade as a criterion for self-evaluation may not inform students about their content knowledge. For scientific thinking, students need to use factual knowledge to hypothesize, reason, and problem-solve (Sinatra & Taasoobshirazi, 2018). Teachers can support students in using mastery-based criteria as they engage in self-assessments (Panadero et al., 2018). Additionally, policymakers and other stakeholders could de-emphasize the focus on performance outcomes by implementing better assessment practices. For instance, in India, it might be useful to consider performance assessments and the use of rubrics to improve self-evaluation strategies.

The study also established a link between students' affect reactions and performance outcomes. As expected, students with higher performance outcomes were more likely to be satisfied, and less likely to report adaptive inferences. In the current study, students' use of performance standards as goals or judgment criteria could explain this association. However, deriving a sense of satisfaction from performance outcomes may result in self-handicapping such as learned helplessness and task avoidance (Zimmerman, 2013). To support students with successful SRL approaches, it might be useful for educators to empower students with a growth mind-set. For instance, using formative assessment practices such as self-assessment tools to teach students how to reflect on their performance to identify areas for improvement for subsequent learning cycles could be beneficial (Panadero et al., 2017).

A core objective of the current study was to understand how assessment design features related to SRL approaches. In the forethought phase, features such as cognitive complexity, content breadth, and types of questions predicted students' cognitive processes and motivational beliefs. Students in this study were more likely to set process-based goals (e.g., step-by-step plan to finish the syllabus) when the questions on the assessment required students to relate and extend their learning, covered a range of sub-topics within a single unit, and had a distribution of questions (e.g., open-ended, multiple-choice, and computational). This evidence indicates that students' interpretation of the task requirements relates to their goal-setting strategies. Indeed, process goals indicate an adaptive SRL approach. However, it might be useful to broaden students' goal-setting strategies suitable for science, which combine process and mastery goals. For instance, encouraging students to use success criteria related to the topic (e.g., I can explain the different types of forces with examples) and identifying specific strategies that would help them achieve those learning goals (e.g., using a mind map) would help improve SRL. Furthermore, the interaction between task and SRL approaches confirms the dynamic nature of SRL (Lodewyk et al., 2006). Future research can focus on using this features-based approach to examine diverse assessment tasks and how these assessment task features relate to SRL.

The current study also found that strategy selection is inversely related to types of questions and cognitive complexity. This finding is important because strategy selection positively correlated with performance, implying that students might require further support in implementing successful SRL strategies that are task specific. Further, the distribution of sub-topics predicted strategy depth among students in this study; when the assessment covered a broad range of topics, students were more likely to use deeper processing strategies such as organization and integration. These findings confirm previous studies that indicate learning strategies vary based on task characteristics (Garcia-Perez et al., 2021). Nonetheless, this relationship between strategy selection and task design requires further clarity because several other factors may play a role. For example, students' strategy planning might be less related to the task and more about their awareness of strategies. Equipping students with a broad range of strategies for various tasks can foster the development of problem-solving and reasoning skills (Sinatra & Taasoobshirazi, 2018). Therefore, future research can focus on gaining an in-depth understanding of students' interpretations of the task, their strategy choices, and interactions with the task through multiple measures such as interview protocols and rating scales.

The link between motivational beliefs and assessment design was inversely related, and some features predicted students' perceived self-efficacy and interest. This relationship is of significance because self-efficacy and interest were strong factors in predicting performance. This finding, in tandem with existing literature in SRL, is critical for researchers. For instance, the results provide further evidence for the social cognitive theory of SRL; task characteristics interact with students' self-beliefs (e.g., perceived competency). This finding provides reasonable support for examining differences in motivational beliefs on diverse assessment tasks, students' interpretations of the assessment task and how those interpretations relate to their SRL approaches. These results have implications for educators. Teachers can consider improving students' competencies for challenging assessment tasks by teaching them effective learning strategies, which would improve their overall SRL approach.

The current study also established links between self-reflection processes and assessment design features. In this study, lower self-evaluation measures were associated with higher scores on assessment design features. This finding could be explained when coupled with students' self-efficacy measures and performance scores. Post-hoc analyses revealed that students were more likely to report attributions that were controllable when the assessment had fewer questions related to real-life situations or were less varied in nature. When this finding is viewed with students' reported learning strategies, it may be inferred that students report uncontrollable attributions when questions in the assessment require deeper conceptual understanding or transfer of learning to a wider variety of questions (Hattie & Donoghue, 2016).

Furthermore, students' adaptive inferences are also related to design features. Students who performed on assessments that required deeper processing also appeared to report higher adaptive scores. There are several possible reasons why students' reactions are higher when the assessment is more challenging. Perhaps students did not perform well on more challenging assessments, which may have prompted their response to adapt their approach. Nonetheless, this is critical for educators because an adaptive approach facilitates lifelong learning competencies. Such skills could benefit complex problem-solving, which requires adapting to uncertain outcomes. Therefore, teachers and educators could support students to make adaptive inferences from their performance by scaffolding students with self-reflective questions or self-assessment strategies focused on the assessment design (Perry, 2006). For example, some students may have performed poorly on items that require reasoning skills. A teacher or tutor could guide the student to understand why they may have performed poorly and how they can adapt their approach to achieve success in the future.

3.5.1 Limitations

There are limitations to this study. For instance, the criteria for assessment design were derived from literature to suit a traditional formal assessment task in the Indian education context. Therefore, the measure requires further testing in a broader context before a definitive case can be made whether it accurately captures the assessment design. Furthermore, the study examined summative assessments that measured student learning for varying science topics at different points in the academic year, which could lead to certain inconsistencies. For example, the amount and type of content varied across schools which might have interfered with the task design and students' SRL processes; in some schools, there was only one unit, while in others there were two or three chapters. Additionally, time restrictions such as how much time schools provided students with to prepare for the assessment may also have varied. It is also possible that differences in grades/marks for the assessment could somehow interfere with the assessment design and students' SRL because some schools had assessments that totalled to 50 marks while others had it for 80 marks either due to the assessment policy in the school or the pandemic.

It is also possible that students' perceptions of the assessment and interpretations of these factors (e.g., grades) could impact their SRL (Greene et al., 2012). Another limitation of the study was the small sample size for each assessment which meant that the data could not be nested for more rigorous analyses. But instead, the data had to be viewed collectively, and analyses were performed on the data set as a whole.

An underlying factor that may have hindered the researcher's ability to thoroughly capture the SRL processes in Zimmerman's model was the language setting. The students belonged to schools from suburban to urban settings where English is a second language. These factors could have influenced how students responded to microanalytic questions and subsequently the coding and scoring of their verbal responses. There is also the possibility that administration of the protocol may not have captured students' SRL sub-processes in real-time but only what they could report in anticipation of the assessment. The microanalytic protocol also uses leading questions to elicit responses from students, which might have limited their responses when compared to a think-aloud protocol that allows for free expression. As a result, the quality of students' SRL processes that were measured may have been compromised. Additionally, the limited sample size and exploratory nature of the study reduce the generalizability of the findings.

Finally, I would like to acknowledge that data for this study were collected through interviews conducted online due to the global pandemic. Students were dealing with new and uncertain circumstances around learning and assessment. For instance, the digital divide was brought to the forefront for many students. While some students were learning from the comfort of their homes on a computer screen, others had to share a mobile phone with limited internet among siblings. Moreover, students were expected to adhere to similar protocols to a regular school day but from their homes, which was challenging for many. Students wrote their summative assessments online and in front of the camera despite these external, uncontrollable, and unusual circumstances. Undoubtedly, this strange and complex situation impacted students' responses to many of the microanalytic questions and potentially their performance on the assessment.

3.5.2 Conclusion

The current study reiterates that self-regulated learning is a context-specific, task-focused, and dynamic process. Students' motivational beliefs are related to their performance, assessment design, and other SRL processes. This reciprocity between SRL and task features indicates a direction for future research. Further, findings from the current study reaffirm contemporary literature highlighting the significance of situational interest and its relation to self-regulated learning, making it a potential avenue for more research.

I also provide evidence to substantiate interactions between assessment design features and cognitive and motivational sub-processes within Zimmerman's SRL phases. Goal-setting strategies were associated with features such as the depth of SOLO levels, questions diversity, and sub-topics coverage. These features could provide students with cues on how to set adaptive, process-based goals for assessments, increasing students' metacognitive skills. Teaching students a range of learning strategies would benefit learning and performance as students' knowledge of strategies may be limited to rehearsal and comprehension tactics. Increased forethought-process competencies would benefit deeper engagement with science concepts, problem-solving and reasoning skills, and overall academic performance.

In terms of measurement, there are several important directions for future research. Indeed, additional studies need to be conducted to validate and improve the scoring rubric used for assessment design features. Additional criteria on the format and implementation of assessments could shed insight on how task variation interacts with SRL. Using the microanalysis protocol to examine students' SRL for various assessments over time would also be useful to examine differences in SRL. A more nuanced approach to measuring SRL, such as the microanalysis protocol, could also inform teachers and students on areas for improvement in the learning process, both in classroom settings and individual study sessions.

Chapter 4: Examining the Impact of Teachers' Classroom Assessment Decisions on Students' Self-Regulated Learning Processes

Abstract

Classroom assessments are foundational to developing and sustaining students' self-regulated learning (SRL) processes. In India, classroom assessment is dominated by examinations. However, few studies have focused on design decisions for formal classroom assessments and how they relate to SRL. Therefore, the current study used a mixed-methods research design to investigate teachers' assessment decisions and their impact on students' SRL processes. The sample comprises nine high school science teachers and 229 students from their respective classes. Interview protocols were used to collect data from participants. Results indicate that students' SRL processes differ in relation to teachers' assessment decisions. Findings suggest that students reported using deep learning strategies and also felt more confident when teachers' decisions were centred on learning instead of performance. Additionally, students also felt confident and identified adaptive strategies for future learning when the assessment design focused on familiar questions. Insights from the study provide guidelines for teachers to consider when designing assessments that promote SRL within students. The discussion concludes with a checklist for classroom assessment decisions for practical implications.

Key words: self-regulated learning, classroom assessments, assessment design decisions, learning and teaching

4.1 Introduction

Although self-regulated learning (SRL) has been linked to academic success for decades, it has never been more important for young people than today. The world is changing rapidly, and students need adequate skills and strong self-beliefs to help them keep up with these changing strides. SRL refers to how students manage, regulate, and reflect upon their thoughts, feelings, and actions to achieve their desired goal (Winne & Perry, 2000; Zimmerman, 1990). A vast literature base indicates that SRL predicts academic outcomes, including scientific knowledge, mathematical problem-solving, and literacy among students (Callan, 2014; Cleary et al., 2017; Rozencwajg, 2003; Taasoobshirazi & Carr, 2009). Furthermore, research suggests that such processes can be enhanced in the classroom through direct instruction and intentional environmental design (e.g., Cleary et al., 2017; Dignath & Buttner, 2008; Perry et al., 2004), for example, encouraging students to set task goals and reflect on their performance using self-assessments. Such practices are inherent to classroom assessment practice, thus encompassing processes parallel to SRL, and are promising avenues to promote and sustain students' independent learning strategies and attitudes.

Classroom assessments are conducted for various purposes but are primarily developed, administered, evaluated, and interpreted by teachers. Such a comprehensive and demanding process requires sound decisions about the task, evaluation practices, feedback, and instructional strategies (Bearman et al., 2016). Research suggests that each aspect of classroom assessment has the potential to develop SRL, and contemporary frameworks have been developed to integrate SRL processes into classroom assessment practices (Andrade & Brookhart, 2020; Chen & Bonner, 2020; Panadero et al., 2018). Self-assessment practices and feedback are associated with improved SRL processes such as goal-setting and self-evaluation (Panadero & Romero, 2014). In previous frameworks (Andrade & Brookhart, 2020; Chen & Bonner, 2020), authors illustrate how teachers can use classroom assessment processes to help support students' SRL processes. Such frameworks guide how teachers can use assessment practices to improve students' SRL processes.

Furthermore, teachers agree formative assessment practices develop SRL even though they believe that such practices might not always be feasible due to school expectations. (McMillan, 2001). For instance, students' performance scores are used for ranking schools, teacher evaluation, and for student admissions. This situation is especially the case in India, where performance outcomes on science assessments are of utmost importance: grades are a matter of social status. Parents have high expectations of performance scores because it also determines future academic opportunities, especially for medical and engineering professions (Dwivedi, 2012; Subramani & Venkatachalam, 2019). These high stakes associated with assessments encourage teachers to teach to the test and evaluate only specific curricular objectives (Harlen, 2005). This emphasis on performance is likely to play a role in the decision-making process for classroom assessments in India.

Understanding teachers' decisions related to formal classroom assessments is important to develop design strategies for assessments that promote SRL. Therefore, the objectives of this study are twofold: (1) to explore decisions made by science teachers in India and examine how they relate to students' SRL, and (2) to recommend a framework for formal classroom assessments that support teachers' decision-making to include practices that promote SRL. My intention for this study is that by better understanding teachers' decisions for classroom assessments and their relationship to SRL processes, it would be possible to develop practices for teachers and educators to improve science outcomes in India.

4.2 Literature Review

4.2.1 Self-Regulated Learning

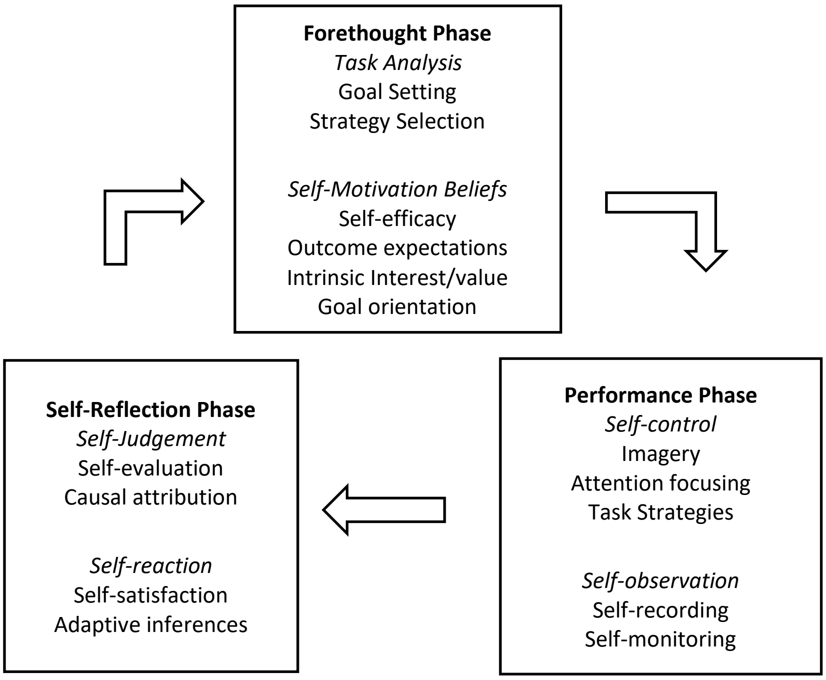
SRL is the active participation of learning whereby students cognitively, motivationally, and metacognitively manage and regulate their learning (Pintrich, 1999; Zimmerman, 1989). Multiple theories describe SRL as goal-directed and occurring in phases that include cognitive processes and motivational beliefs. For instance, Zimmerman (1989) posited that SRL is a process that arises from the interactions among cognitive and metacognitive strategies, self-beliefs, and volition. This framework is grounded in the social cognitive theory of learning, which assumes that learning results from reciprocal interactions between the individual, behaviour, and the environment (Bandura, 1989; Zimmerman, 1989). From this point of view, individuals do not act independent of the context in which they are present (self-determination), nor do their actions result from external conditioning (e.g., behaviourist). Instead, the social cognitive view posits that internal and external factors determine individuals' behaviour and thought (Bandura, 1989). Based on the principles of social cognitive learning, Zimmerman proposed a cyclical model to illustrate cognitive processes and motivational beliefs and their interactions that guide SRL.

Overview of Zimmerman's Model

This model comprises three phases: forethought, performance, and self-reflection (see Figure 1 presented here for ease of interpretation). In the forethought phase, students analyse the task to set goals and choose appropriate strategies. This analysis is guided by students' self-beliefs related to learning, such as perceived competency, the value of the task, and intrinsic interest. In the performance phase, students employ metacognitive processes (e.g., monitoring) and volitional strategies (e.g., self-control) to regulate and sustain their actions. Finally, in the self-reflection phase, effective self-regulated learners evaluate their performance, including their learning and regulatory strategies based on the degree to which they could meet their goal(s) using different criteria (e.g., mastery, performance). During this phase, students also make inferences to carry forward information for the next learning cycle. Ideally, self-regulated learners will view any mistake as an opportunity to improve their learning and derive satisfaction from their performance.

Figure 1

Zimmerman's Self-Regulated Learning Model (2009)



Empirical findings provide evidence for Zimmerman's SRL model and indicate that strategic, deliberate, and motivated students are more likely to experience gains in conceptual understanding and academic achievement. Meta-analyses found that students trained in SRL strategies outperformed their peers who were part of the control group (Dignath & Buttner, 2008; Graham & Harris, 2003; McDonald & Boud, 2003). Such findings are encouraging and suggest that the environment can be tailored to help improve students' learning skills. Recently, scholars have been paying considerable attention to the role of classroom assessments in developing and supporting SRL (Andrade & Brookhart, 2020; Chen & Bonner, 2020; Panadero et al., 2018).

4.2.2 Classroom Assessment Decisions

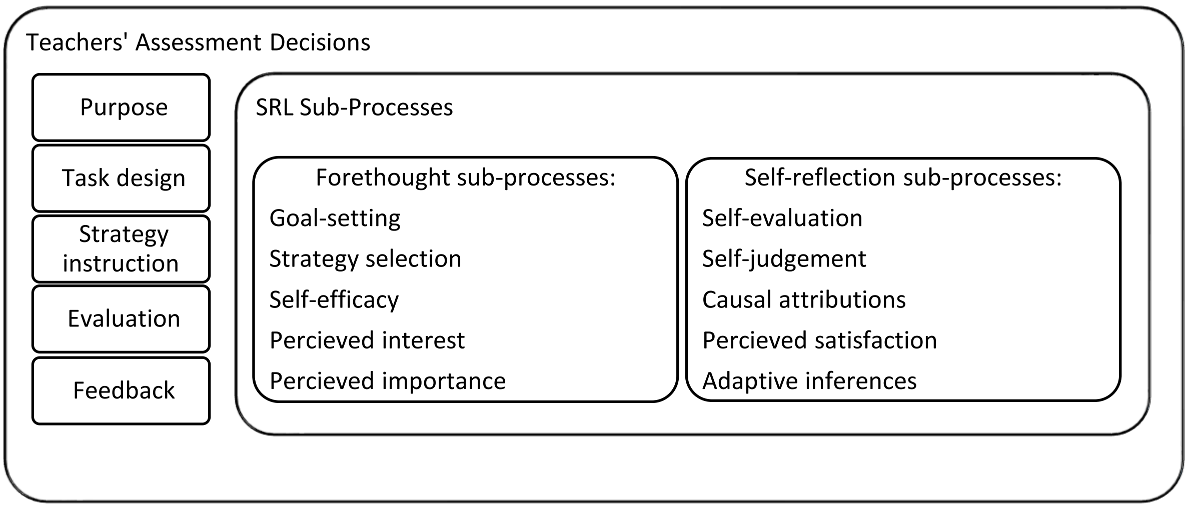
Classroom assessments are regularly occurring academic tasks that teachers use for various purposes. According to Black and Wiliam (2018), classroom assessments are “those assessments where the main decisions about what gets assessed, how the students will be assessed, and the scoring of the students’ responses, is undertaken by those who are responsible for teaching the same students” (p. 554). Studies suggest this decision-making process is complex as teachers strive to balance their choices between internal beliefs with external expectations (e.g., school’s assessment policy; Boud & Malloy, 2013; McMillan, 2003). Teachers’ decision-making rationale is ambiguous, often reflecting a ‘hodgepodge’ of influences such as their prior experience and using strategies they believe increase student engagement (McMillan, 2003). Therefore, identifying how teachers make such decisions will help us understand this process more clearly and establish strategies to support SRL development.

As previously mentioned, creating a classroom assessment involves identifying its purpose, designing the task, teaching students learning strategies for preparation, setting evaluation criteria, and determining feedback strategies. Research suggests that each element could impact student learning and achievement. Much of this evidence is isolated and focused on formative assessment practices such as self-assessments and feedback (Harris et al., 2018; Hattie & Timperley, 2007; Panadero et al., 2017). Due to the significant promise of these formative practices for learning, most theoretical models integrate SRL into classroom assessments with this perspective (Andrade & Brookhart, 2020; Chen & Bonner, 2020). Chen and Bonner (2020) highlight that these models fail to account for the wider range of classroom assessments. In particular, there is insufficient attention paid to formal classroom assessments, especially those which are evaluated for a grade or score and are thus high stakes for teacher and students.

I argue for the need to view formal assessments as an additional classroom assessment practice that can support SRL development in the forethought and self-reflection phases described in Zimmerman's model. In this study, I focus on the forethought and self-reflection sub-processes because they complement the assessment process (Figure 3). For instance, identifying an assessment's purpose, instructing students on strategy use, and giving feedback can directly impact sub-processes in the forethought phase (e.g., goal-setting) and self-reflection phase (e.g., self-judgement). In the following paragraphs, I provide an account of the research on the various assessment decisions and their relationship to SRL.

Figure 3

Teachers' Assessment Decisions and SRL Sub-Processes Investigated in the Study



Assessment Purpose

There is a general agreement that assessment purposes reflect one of three themes: for learning, as learning, or of learning (Gardner, 2011; Stiggins, 2002). When data (e.g., evidence of incorrect/inaccurate responses) are used to inform teaching and improve learning outcomes, it is typically *for* learning and takes a formative purpose. Assessment *as* learning may also be linked to a formative function because it refers to the processes (e.g., self-assessment) used to help students improve their learning strategies and engage in monitoring and evaluation (Hayward, 2015). In contrast, when assessment information is used to establish how much students know (e.g., to assigning a grade or score), it reflects an assessment *of* learning and serves a summative function (Black & Wiliam, 2018). Although such distinctions in purpose appear straightforward in theory, these purposes can, and often do, co-exist (Black et al., 2011). Therefore, teachers need to discern student data on assessments to evaluate and support their learning and performance.

Contemporary frameworks of classroom assessments and SRL make a strong argument for using formative practices to develop SRL processes. Some scholars suggest a formative lens can encourage students to set learning goals for themselves (Andrade & Brookhart, 2020; Heritage, 2010; Moss et al., 2013). Doing so can help improve calibration and metacognitive strategies. Students become better at judging their learning and gaining a mastery orientation (Brookhart & Durkin, 2003). According to Zimmerman's model (2009), goal-setting occurs in the forethought phase, where students analyse the task and summon motivation beliefs. Research highlights the importance of this phase because it defines a student's initial task engagement which consequently has an impact on their academic success (Pajares, 2008). For example, helping students identify their strengths and weaknesses for a topic can help them set goals that challenge them adequately. Such objectives are associated with increased self-beliefs such as higher self-efficacy, interest, and mastery orientation (Butler & Schnellert, 2015). Establishing a purpose that encourages a mastery orientation can help students view assessments as an opportunity to improve their learning processes

This approach justifies the use of assessments as serving not only teachers, but also students. Helping teachers specify the purposes of an assessment could also offer implications to support students in setting goals that they find meaningful thus supporting their motivation. This distinction is less evident in the Indian context where the education system is dominated by an evaluation system. There has been considerable effort to reform an exam-oriented assessment practice by distinguishing assessment for summative and formative purposes in the national policies (National Council for Educational Research, 2005). As a result, state-level and national-level education boards developed an assessment system for teachers with recommendations in the national curriculum framework. However, these created systems that reflected a test-based system which confused teachers, thus leading to implementation challenges (Ratnam & Tharu, 2018). The ambiguity in prescribed assessment procedures, constraints from schools, and lack of clarity on formative and summative purposes will likely contribute to teachers' perceptions of assessment in India. Given that identifying a formative purpose relates to goal setting and achievement motivation, it is worthwhile to consider how the intentions of formal assessments impact students' SRL.

Assessment Task Design

Teachers generally develop assessment tasks, and their design contributes to how students approach their learning. Researchers have highlighted concerns with teacher-developed assessments and assignments, stating that teachers struggle to design tasks that require higher-order cognitive processes and tend to administer tasks requiring factual and procedural knowledge (Crooks, 1988; Fleming & Chambers, 1983; Haertel, 1986). Fleming and Chambers (1983) revealed that out of 8,800 test questions designed by teachers, nearly 80% were at the knowledge level (recall and retrieval). Crooks (1988) suggested that such assignments are less challenging to evaluate when compared to those that seek high-order processes such as problem-solving. In India, formal assessments, such as those conducted for mid-terms and at the end of a term, follow a typical examination structure. Although India's educational policies (NCERT, 2005; 2020) encourage teachers to use assessments that discourage rote memorisation and enhance scientific skills such as inquiry and reasoning, examinations that tend toward knowledge recall continue to dominate the assessment context. Such tasks, especially those that do not require more than recall, retrieval, and application are less likely to promote SRL (Crooks, 1988; Lodewyk & Winne, 2009).

Perry et al. (2007) developed a guiding framework to mentor student teachers in designing and implementing literacy tasks that support SRL. Similarly, Case and Gunstone (2002) modified assessment tasks to support metacognition and a positive relationship between task design and students' SRL. These findings corroborate with Lodewyk et al. (2009). They found that tasks focusing on mastery of learning using open-ended and challenging tasks are associated with strategic approaches that employ higher-order cognitive processes such as problem-solving and critical thinking. Children develop metacognition when given challenging, relevant tasks that address multiple learning goals (Perry et al., 2004). Such task features improve intrinsic motivation, cognitive, and metacognitive strategies, thus enhancing students' self-efficacy (Neuville et al., 2007; Perry et al., 2004; Pintrich & Degroot, 1990).

Although challenging tasks support motivation and deep learning strategies, researchers caution that students feel less motivated to perform if tasks are perceived as too difficult. Lodewyk and Winne (2005) focused on the structure of tasks and students' SRL. Their research findings suggest that well-structured tasks are more likely to enhance student performance but are less favourable in promoting SRL. However, this depended on students' characteristics. For instance, students with low SRL benefitted from well-structured tasks, whereas high-performing students experienced boredom (Lodewyk & Winne, 2005). It might not be possible to tailor a formal assessment design to suit the needs of different children; however, it might be worth understanding how teachers approach the creation of a somewhat standardised assessment given the implications of design for students' SRL processes.

Assessment Strategy Instruction

Once familiar with the task design, teachers tend to share different strategies for learning and performance with students. Research by Perry and Drummond (2002) identified specific interactions between teachers and students that promote SRL. For instance, when teachers provide targeted instruction and strategies to learn, students are more likely to engage in SRL-related processes such as strategy use. In another study by Perry et al. (2008), students adopted effective SRL approaches when teachers prompted them to choose and implement strategies independently instead of those classrooms where teachers provided instructions to complete procedural tasks (e.g., worksheets). It is noteworthy that teachers' strategy instruction depended on their awareness of learning strategies and the alignment of the strategies to the task design. Such competencies would help teachers guide students in selecting a strategy and how to use it. Moreover, exchanges between teachers and students are significant because students pay attention to how teachers interact with them in the classroom, particularly to the instruction given about learning (Dignath and Veenman 2021; Greene, 2021).

Teaching explicit SRL strategies is underscored in research, yet it appears challenging to implement. Current SRL frameworks provide a compelling case for classroom assessments as an effective process through which teachers can elevate and enhance students' SRL. For example, Chen and Bonner (2020) suggest that informal assessments are opportunities to teach students how to track their learning progress by employing various cognitive (e.g., goal-setting, reflection) and metacognitive strategies (e.g., monitoring, control). While these strategies are crucial to SRL development, there is less understanding of how teachers instruct students to learn or prepare for an assessment.

Strategy selection is a core sub-process in the task analysis process which requires careful deliberation to reach the desired learning or performance goal. Furthermore, learning strategies help students judge their learning and evaluate their understanding, making it critical for them to gain a sophisticated repertoire of strategies (Dignath &

Buttner, 2018; Tuckman & Kennedy, 2011; Veenman, 2011). Formal assessments typically require students to cover large amounts of content and respond to the task within a constrained time limit. Such an assessment task will undoubtedly require strategies suited to its design (Sebesta & Speth, 2017). Therefore, students must be taught how to approach this task as effectively as possible (Dignath & Buttner, 2018; Tuckman & Kennedy, 2011). With this in mind, this research project seeks to uncover teachers' typical instructional practices associated with preparing for and completing a formal assessment. It could help establish effective practices and techniques that can help teachers improve their students' SRL for formal assessments.

Assessment Evaluation Criteria and Grading Practices

Another critical aspect of classroom assessments is evaluating and grading student work. Some studies show that grading practices differ across subjects, and sometimes teachers consider non-academic evidence while assessing student work (Brookhart, 2016). For example, Pilcher (1994) highlighted that mathematics teachers rely less on traditional assessments because they want to understand how students arrive at an answer and assign a grade or mark, thus evaluating students' cognitive abilities. In contrast, other subject teachers account for classroom participation, motivation, and effort while assessing student work. This combination of cognitive and non-cognitive factors contributes to the inconsistency and ambiguity in the criteria used by teachers for grading.

In part due to the complex education system, grading practices in India have been unexplored. Private schools can be affiliated with state or national boards in addition to international educational boards. In this paper, I focus on schools that subscribe to state or national education boards. Typically, schools have autonomy regarding their assessments. It is worth noting that these school hold state-and nation-wide exams in years 10 and 12 (Brown et al., 2014). Due to the high stakes associated with these standardised exams, students are pressured by parents to score high marks. As a result, the country is dominated by an exam-oriented system that determines evaluation criteria and grading practices. This assessment format focuses on the response to a question that typically has a right or wrong answer. For national-level and state-level examinations, teachers use marking schemes developed by the respective education boards. These marking schemes tend to have the expected answer, including key words and total mark allotment for each question. This is the only set of assessment design guidelines available for teachers, and many of them tend to use the same approach in their classrooms for assessments they create. As a result, students aim meet those desired expectations outlined by standardized examinations and teachers are constrained by the same expectations in their evaluation.

The impact of grading practices on student learning is documented in two distinct categories. On the one hand, measurement scholars highlight concerns about the reliability of teachers' grading in terms of its accuracy and validity of student achievement (Brookhart, 2016; Pellegrino, 2016). On the other hand, researchers emphasise that grades reduce intrinsic motivation because they are extrinsic rewards used for control and management purposes (Brown et al., 2012). Shephard et al. (2018) reiterate that motivating students with points is detrimental to learning outcomes and argue for a coherent link between formative assessments, grading practices, and large-scale assessments. They posit that large-scale, standardised assessment should be designed to drive learning and motivation, and formative assessment practices should be the bedrock for this objective. There have been numerous arguments for implementing standards-based evaluation practices, which assist in evaluating student work according to grade-level standards with clear indicators of performance (e.g., basic, developing, proficient, and advanced; Brookhart, 2011; Guskey, 2009). A compelling reason to use standards-based grading is its potential to offer high-quality information to students as feedback and for parents as information on their child's achievement (Brookhart et al., 2016).

The current study builds on the above perspective: teachers' evaluation practices have implications for learning and motivation, both of which are core elements of SRL. Some researchers suggest that using rubrics and sharing performance criteria with students is associated with more robust SRL processes of planning and self-reflection (Andrade & Du, 2005; Fraile et al., 2017; Reynolds-Keefer, 2010). Rubrics also promote self-assessment practices, which inform students about their achievement and contribute to SRL development. These self-assessment practices promote a mastery goal orientation, self-reflection, and ultimately, autonomy and responsibility of their learning (Andrade et al., 2008). Criterion-referenced evaluation based on a mastery approach is most likely to contribute to a mastery approach of goal orientation (Andrade & Brookhart, 2008). Such practices enhance students' metacognition, deepen learning, and help develop appropriate SRL practices (Panadero & Romero, 2014; Perry et al., 2002). With a clear and structured approach to evaluation, students can receive corrective, meaningful feedback. Evaluation benefits student learning and performance, making it necessary for SRL development.

Assessment Feedback Strategies

In the context of learning, feedback may be characterised as information about a student's performance or understanding that the teacher transmits with a purpose in mind (such as formative or summative). Hattie and Timperley (2007) articulate the effect and value of different types of feedback on learning. Meta-analyses found feedback strategies reflective of instruction and how to perform the task were far more effective than those related to praise, rewards, and punishments (Kluger & DeNisi, 1996). As such, Hattie and Timperley (2007) suggest a feedback model focused on four levels: task, process, self-regulation, and self, and each contributes to learning and performance to varying degrees. For instance, descriptive feedback with a corrective course of action, coupled with reflective questions, are far more effective than feedback limited to the evaluation (e.g., grade) of performance (Brookhart & Chen, 2015; Brown et al., 2012; Butler & Winne, 1995; Davis & Neitzel, 2011). This type of feedback is also preferred by teachers as they tend to provide feedback at the task or process level (Harris et al.,2014; Irving et al., 2011).

Elsewhere, Guo (2017) proposed a framework encompassing five different types of feedback suited to a Chinese educational context. The author suggested that feedback may take the form of verification (evaluating student work as correct or incorrect), directive (informing students of the correct answer), scaffolding (using prompts and cues), praise, and criticism. Teachers' use of these various types of feedback tends to vary based on the assessment purpose and evaluation practices. For instance, teachers are likely to use scaffolding feedback if the focus is mastery-oriented. When the assessment objective is to improve test performance, teachers may use verification or directive feedback strategies.

Studies suggest that effective feedback helps students to identify their learning gaps, choose appropriate methods for future tasks and monitor their performance (Andrade & Brookhart, 2016; Black & Wiliam, 1998; Sadler, 1989; Schunk & Ertmer, 2000).

Identifying the type of feedback based on students' needs has powerful implications for SRL (Hattie et al., 1996). For instance, if the student has trouble discerning the task's demands, it would help the student set goals related to the task, referring to forethought and self-reflection processes in Zimmerman's model (2009). Additionally, the nature of feedback given to students may also determine their judgement of attributions and self-reactions. For instance, statements such as "you are smart" instil a fixed mindset. In contrast, feedback statements focused on process (strategies), effort, and persistence are more likely to contribute to a growth mindset (Andrade & Brookhart, 2016; Dweck, 1986). SRL literature reiterates the importance of causal attributions which influences one's motivational behaviours (Zimmerman, 2013). More importantly, when feedback is solely focused on evaluative judgement, it can have a negative impact on SRL, resulting in handicapping strategies (e.g., task avoidance) and reduced self-efficacy (Kluger & DeNisi, 1996). It is also possible that students are non-reactive to feedback and are reluctant to adaptive modifications to their learning to protect their well-being or avoiding further hard work and doing just enough instead of their best (Harris et al., 2018; Meyers et al., 2009). When opportunities for self-reflection are inherent in feedback practices, students' SRL processes are enhanced, making this a crucial decision by teachers, and making it worthwhile to examine for formal assessments.

4.2.3 The Current Study

Given the general agreement that assessment practices can deepen student learning and develop SRL, the current study examines teachers' decision-making for a formal classroom assessment in science. This study builds on social cognitive theory (Zimmerman, 1989) and classroom assessment frameworks (Allal, 2020; Andrade & Brookhart, 2020; Chen & Bonner, 2020) to explore potential links between teachers' decisions and students' SRL. The research seeks to contribute to the gap in literature from the Indian cultural perspective dominated by performance-focused assessments. The study adds to the existing SRL literature on an under-researched topic and population by providing insights into teachers' assessment design decisions for formal assessments. Finally, the research seeks to link specific teacher practices and decisions to particular processes in the forethought and self-reflection phases of Zimmerman's SRL model. The following research questions guided this study:

1. What decisions (purpose, task design style, strategy instruction, evaluation style, and feedback strategy) do teachers make regarding formal classroom assessment design?
2. Which SRL strategies reflected in Zimmerman's cyclical model do students demonstrate for a science summative assessment task?
3. In what ways do teachers' assessment decisions impact students' SRL sub-processes in forethought and self-reflection phases described in Zimmerman's model?

4.3 Methodology and Methods

4.3.1 Participants and Setting

The sample comprised nine teachers and 229 students from their respective classrooms (108 boys and 121 girls) from four schools in Southern India affiliated with India's state or central government. Each school serves more than 1,000 students from kindergarten to Grade 12, and the primary medium of instruction is English. As science is a mandatory subject in middle school, all students attended science-related classes, including physics, chemistry, and biology.

Teacher Sample

Several schools in Southern India were invited to participate in the study. Science teachers from schools who consented to participate were contacted for recruitment. Teachers were informed about the purpose of the study and expectations for participation. Only those teachers who volunteered to participate were included in this study. Teachers who participated in the study were all science teachers with varying years of experience (M = 15 years), with 30 years as the most experienced teacher and six years as the least experienced teacher. Among the teachers, two belonged to a state-board school, three were from schools affiliated with the Central Board of Secondary Education (CBSE), and the remaining four teachers worked in a school affiliated with the Indian Certificate for Secondary Education (ICSE).

Table 32

Distribution of Sample Population Including Teachers and Students

School board	Subject	No. of teachers	Students
State board	Physics	Teacher A	15
		Teacher B	47
ICSE	Physics	Teacher C	38
		Teacher D	21
		Teacher E	36
		Teacher F	22
CBSE	Physics	Teacher G	14
	Biology	Teacher H	15
	Biology	Teacher I	21
Total		9	229

Student Sample

Only those students whose teachers consented to participate were invited to participate in the study. In total, I recruited 332 students, but the final student sample comprised 229 students. Students belonged to grades 8 and 9 from the four different schools that agreed to participate in the study. Student age ranged from 13 to 15 years (*M* = 13.58 *SD* = .64).

Assessment Setting

Teachers conducted formal classroom assessments for students as part of the academic calendar. The current study focused on a formal examination in a science class (physics, chemistry, biology) that was scheduled either in the middle or end of the school term.

There were several key commonalities among these assessments making it appropriate for them to be analysed together. For example, every assessment included in this study was a pencil-paper test conducted online with teacher supervision. Although the science subjects were grouped differently, the structure of the assessment remained the same. This meant that all assessments were tests which included the same range of questions (e.g., one-word/objective questions and paragraph questions). This form of supervised exam would be the case for in-person examinations had the pandemic not forced the assessments to be administered online. All students were expected to submit their answer scripts by scanning the images and emailing them or posting the answer scripts to their respective schools.

4.3.2 Methodology

The current study employed a mixed-method design to address the research questions identified. The data collection process involves interviews with teachers and students. By using a transformation mixed-methods design, I was able to convert one type of data (qualitative) into another (quantitative; Teddlie & Tashakkori, 2009). I integrated qualitative and quantitative data at the interpretation and reporting level through data transformation (Fetters et al., 2013).

4.3.3 Data Collection

In this study, there were three different types of data from the two populations. (1) SRL measures were obtained from microanalytic protocol with students, (2) student performance outcomes were gathered from their respective science teachers, and (3) teachers' assessment decisions were investigated using semi-structured interviews.

SRL Microanalysis Protocol

The information below is the same protocol described in Chapter 3. This study used students' SRL from the previous study to investigate differences based on teachers' reported assessment intentions. As such the design, coding and scoring of the microanalytic responses are identical to the study described in Chapter 3. Please refer to method and methodology section in Chapter 3 for a detailed explanation of the SRL microanalysis protocol.

Semi-Structured Interview with Teachers

I developed a semi-structured interview that comprised questions and prompts related to teacher demographics, professional experience, and assessment decisions. The questions were developed using the assessment decisions framework outlined by Bearman et al., (2016) Examples of questions included, "Could you please walk me through the assessment process at your school?" And "What was the purpose of the assessment?" I also included prompts and probes to clarity responses, elicit richer and more comprehensive data (Cohen et al., 2018; Wellington, 2015). Refer to Appendix C for all the questions in the semi-structured teacher interview. I consulted teachers regarding their availability to schedule the interview.

I scheduled interviews after teachers completed the assessments on which students performed. Each interview lasted between 35 and 40 minutes. I used online video-conferencing platforms to interview teachers, and all interviews were audio-recorded with the participant’s permission and transcribed for analyses.

4.3.4 Analysis Process

Since this study employed a mixed-methods design, qualitative and quantitative analyses were involved. In particular, I used data transformation methodology to analyse the impact of assessment decisions on students' SRL processes (Teddlie & Tashakkori, 2009). The first step involved measuring students' SRL, for which I followed the microanalytic protocol and coded qualitative data into pre-determined categories grounded in literature (refer to Chapter 3 and Appendix A). Regarding teacher interviews, I used the interview questions to establish top-level codes (e.g., purpose, design, evaluation, and feedback practices). Following this, I used a content analysis approach to determine sub-categories (Cohen et al., 2018).

In the current study, the codes were based on the data and not decided in advance. After the interviews were transcribed, a careful reading of the transcript was done several times. After re-reading the transcriptions, I created five separate sheets and pulled out the relevant statements from each interview for the top-level codes mentioned previously. For instance, for the question, 'What is the purpose of this assessment?,' I collated the responses from all the nine teachers into one document. After completing this step for all the top-level codes, I used analytical coding strategies to break the data into smaller units for further analysis. This coding procedure enabled me to interpret the words beyond their descriptive nature (Cohen et al., 2018).

For example, the question regarding assessment purpose as learning-focused, performance-focused, or combination of purposes. The two purposes were distinguished based on whether the statements and phrases were beneficial for students' learning or if performance outcomes were the focus. Phrases such as 'to know much,' 'evaluate' or 'training' was coded as performance-focused purpose, while a phrase such as 'students' to maintain connectivity with the subject matter' was coded as learning-focused. I provide a detailed description of these sub-categories in the results section of this study and the coding scheme is described in Appendix D. These sub-categories were dummy coded to serve as categorical variables for statistical analyses.

Validity and Reliability

After the interviews were transcribed, the transcripts were shared with the participating teachers to validate that their responses were captured authentically (Cohen et al., 2018). Once the codes and categories were established, I shared samples of teacher responses and coding guide with a peer at the University of Canterbury. They have 15 years of teaching experience in science at an international school and were enrolled in a doctoral program in educational studies. The coding framework was discussed, and the teachers' responses the coding was performed by the independent coder (Cohen et al., 2018).

Descriptive statistics of students' SRL processes were conducted, and teachers' responses were examined to gain a bird's eye view of the data. Since the objective of this study was to investigate differences among students who were grouped according to teachers' intentions for assessments, I completed a one-way between-group analysis of variance (ANOVA; Cohen et al., 2018). to examine differences in students' SRL processes and academic performance based on teachers' assessment decisions. Additionally, I used the chi-square (X^2) test for independence to investigate associations between categorical variables derived from the teacher interview and SRL processes (Pallant, 2007).

4.5 Results

4.5.1 Assessment Context

In this section, I describe the context based on teachers' descriptions of the assessment. As I analysed the data, it became evident that teachers made assessment decisions within a specific context. Therefore, before describing the results regarding teachers' assessment decisions, I elaborate on the assessment context that was derived from the data.

Teachers explained the intention and format of their recent classroom assessment which is the same one that students completed. Based on their responses it appeared that teachers' decisions were, in-part, determined by their school's expectations. Teachers reported that they conducted the assessment because it was part of the school's academic calendar and counted toward students' final grades. This might indicate that the assessment is high-stakes for students and teachers. Teachers also stated that they used an examination format because this was typically expected for formal assessments. This format meant that students were presented with a question paper and were expected to respond to those questions within a specified timeframe. Teachers also mentioned that the question paper was designed according to national or state expectations. One CBSE physics teacher explained by saying, "before question paper is being prepared, they will give us the blueprint… the department head will give us the blueprint. We will have to follow the blueprint. It will be marks-based."

Overall, teachers explained similar aims for the assessment questions. The first aim was to measure conceptual understanding and subject-specific skills. For instance, biology teachers reported that they were purposeful in their decisions to include diagram-based questions, while physics teachers explicitly stated that they incorporated computation and application-based questions; one teacher from a state board school pointed out, "Even as a physics teacher, I will try to pose some mathematical questions to students." From the teachers' descriptions on the assessment process, it appeared that the second aim was for questions to focus on higher-order thinking processes such as reasoning, logical thinking, and problem-solving and cover every part of the topic. Another biology CBSE teacher said, "We'll just frame the questions in such a way that it should have logical thinking and higher order thinking questions, and generalised questions, descriptive questions and multiple-choice questions."

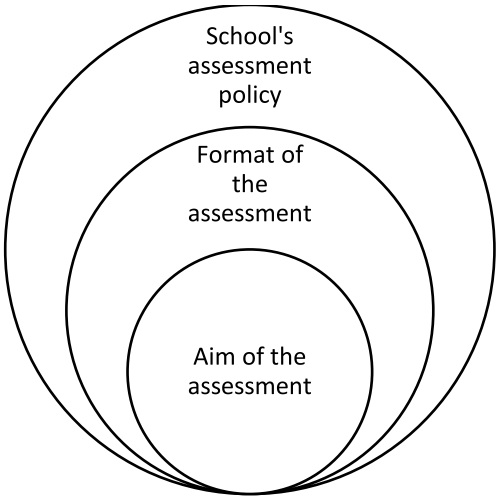
Additionally, teachers mentioned that they were forced to make some changes this year due to the Covid-19 pandemic. Since schools were not accommodating in-person learning, teachers reported that they conducted assessments online via video conferencing platforms (e.g., Google Classrooms and Zoom). They explained that online tests were challenging because it allowed students to engage in malpractices such as copying off the internet, contacting their friends, or referring to the textbook. For example, one ICSE physics teacher said, "Now when it is online, we realised over the year that children are getting better and better at copying. So [we] had to fix a test where they just couldn't take it from the net." This teacher’s perspective indicates that they were concerned students were not learning the subject, but instead were finding ways to score marks.

Teachers also reported taking various measures to mitigate the challenge of students' copying behaviours. For example, they said their schools chose to conduct assessments similar to in-person examinations. Teachers explained that they instructed students to have their webcam switched on such that supervision was achievable, write their answers rather than type, and upload their answer scripts on the Google Classroom site. Teachers reported that the assessments went as smoothly as possible, and they were content with the process and outcomes. Finally, teachers also reported paying attention to students' responses, such as sentence structure and vocabulary, to judge the authenticity of their answers. It appears that teachers wanted to evaluate students' knowledge and competencies as accurately as possible, but they do not appear to completely trust students' academic integrity.

Based on teachers' descriptions of the assessment process in their schools, some phrases were coded into context and policy. These were common across all teachers and can be described as an embedded diagram (Figure 4). The largest circle is the school's assessment policy which dictates when the assessment occurs and expectations from it. The middle inner circle reflects the format of the assessment (i.e., exams) and in the inner most circle sits the aim of the exam.

Figure 4

Interpretation of Inter-Related Elements in the Assessment Context



The data suggested that the aim of the assessment may be influenced by the school’s expectation and the preferred assessment format as described one teacher:

When we are given the schedule of the timings, we are going to have the periodic test timetable […] Before question paper is being prepared, they will give us the blueprint, the department head will give us the blueprint. We will have to follow the blueprint. It will be marks -based. We have different standards are given. Seven standards are given by CCE. 7 standards need not be introduced in one section. All the standards have to be represented, there is percentage for each standard, that has to be followed. For example, writing a poem on eye. Generally, we do not, when we were preparing, why English -language skills - will come in this. But now we have to make them write a poem on eye because human eye and colourful world is related it is all related to defects in the vision [..]. It is part of academic standard A6 (aesthetic sense). And questioning skills, one question should be asked using questioning skills. For example, we have mirrors: concave, convex mirrors where you see the utensils around the house, prepare some questions that make these utensils be like/behave like mirrors.

This uni-directional influence on the assessment purpose is critical because it could constrain teachers in their decision-making process. Moreover, teachers' descriptions of the assessment aims indicate only a summative function which may limit the function of assessment *for* learning and as learning, which are vital for the support and development of SRL. Perhaps an indication of assessment for learning would be for the school's expectation to report on student progress in terms of learning objectives or success criteria, in addition to performance outcomes.

4.5.2 RQ 1: Teachers' Assessment Decisions Regarding Formal Classroom Assessments

This section highlights the main categories that emerged from teacher interviews regarding purpose, design, evaluation strategies, and feedback practices. Regarding the assessment's purpose, teachers' responses generally fell into three categories: performance, learning, and a combination of both. Most teachers mentioned performance-related intentions, which meant they wanted to assess how much students have learned, and only one teacher set the objective as being focused on maintaining students' engagement with the subject matter. Sample quotes and teacher distribution are provided in Table 33.

Table 33

Distribution and Sample Codes of Teachers' Responses for Assessment Decisions (Purpose) (n = 9)

Codes for assessment decisions and number of teachers	Code description	Sample quote
Performance (3)	Evaluate how much students have learned Evaluate what they have understood Future exam training	"In written examination only, we will find out in which part the children are lagging. So, we want to evaluate the students, in which part, theory part or numerical part, they are lagging."
Learning (1)	Maintain connectivity with the subject matter Opportunity to go through the content before moving on	"[we] have to access students with assessments otherwise their connectivity will be lost."
Combination (5)	A combination of the goals mentioned above	"This assessment is only to [...] Know what they learned […] and they should also know what they are learning."

Teachers' descriptions of their task design decisions and strategy instruction reflected three distinct approaches. Some teachers were innovative with their questions in that they did not rely on the prescribed textbook or resource materials. Some other teachers took a more balanced approach considering students of all abilities. And only one teacher used the questions from the textbook. Teachers' instructions for assessment preparation to students included surface learning, deep learning strategies, and a combination of both. All teachers except one reported sharing performance strategies that would help students improve their outcomes. Table 34 details the distribution of design strategies and sample quotes.

Table 34

Distribution and Sample Codes of Teachers' Responses for Assessment Decisions (Design)(n=9)

Codes for assessment decisions and number of teachers	Code Description	Sample quote
Innovative (5)	The teacher who chooses questions that are from the textbook, or creates questions based on examples from the textbook. This teacher also uses standard resources to select questions, such as old question papers and question banks.	"I refer [to] many textbooks and then I choose the questions. Sometimes it might be even my own question but based on the concept."
Challenge, balanced (3)	This teacher creates challenging questions that increase in cognitive depth and complexity. Also, this teacher considers the abilities of all students and therefore strives to achieve a balance between straightforward questions (e.g., define, explain, describe) and higher-order thinking questions (e.g., application, reasoning, problem-solving).	"We need to go for some creative questions, we need to go for some simple questions because all the students will be there, no? They should not be disappointed by seeing the assessment."
By the book (1)	The teacher who develops their own questions that meet the requirements of the test format and academic standards set forth by the educational board.	"Book back questions surely. We give only book back questions."

Teachers also explained the different strategies they advised students to use to prepare for the assessment. Teachers' strategy instruction ranged from surface learning approaches through deep learning approaches to a combination of both. For example, some teachers instructed their students to read the textbook and underline key points and some advised students to write answers verbatim. Other teachers advised students to use deeper learning strategies such as elaboration which included finding examples from their surroundings that relate to the subject matter. A few teachers explained that they gave students a combination of the two approaches. Table 35 describes the frequency distribution of teachers' strategy instruction along with codes and sample quotes. Although teachers provided guidance to students to help them learn and prepare for their assessments, the approach seemed ad-hoc with little instruction on *how* to adopt and adapt strategies based on the learning goal or objective.

Table 35

Distribution and Sample Codes of Teachers' Responses for Assessment Decisions (Strategy Instruction) (n=9)

Codes for assessment decisions and number of teachers	Descriptions for codes	Sample quote
Surface learning approaches (3)	Surface only (Rehearsal and comprehension)	"Reading is the best way of getting the knowledge" Year 8, Physics Teacher, state board school
Deep learning approaches (2)	Deep only (Organization & Elaboration)	"They relate it to their day-to-day life" Year 8, Physics Teacher, CBSE school
Combination of learning approaches (3)	Combination of Surface learning & Deep learning strategies	"Remembering the concepts with side heading and key points is most important" Year 8, Physics teacher, ICSE school

Teachers' evaluation styles commonly included checking for key terms, however, they differed based on how they awarded student responses with marks. Some teachers were rigid and only scored students' responses if they fully met their expectations. Other teachers took a more subjective approach based on their experience and perceived satisfaction with students' responses. The remaining teachers used an answer key to objectively measure all responses and provided students with self-assessment opportunities (see Table 36).

Table 36

Distribution and Sample Codes of Teachers' Responses for Assessment Decisions (Evaluation Style) (n=9)

Codes for assessment decisions and number of teachers	Code description	Sample quote
Rigid (3)	This marker evaluates student work with strict expectations and no leniency.	"During correction, I used to show them I will cut the marks. I will tell them, in the class I am very lenient, but for exams I am very strict in correction" Year 8, Physics, state board school
Subjective (3)	This marker evaluates student work based on their past experience and grades student work based on personal expectations for each question.	"We will not say you should write the answer which is there in the textbook. They can write in their own way, they can go for a pictorial representation, they can go mind-mapping, something they can show, their creativity, if we are satisfied with their answer and if that answer covers the required content, then it is okay." Year 8, Physics teacher, state board school
Objective (3)	This marker develops an answer key to evaluate student work. The answer key derived from past experience and based on standards followed in national standardized exams is more objective than the subjective marker because the key is explicit and is used against each student paper.	"Answer key is only the best method because the students, they can evaluate themselves also. Because if we are sending the answer key means, they will do their own correction." Year 9, physics teacher, CBSE school

An aspect close to evaluation practices is feedback strategies. Most teachers focused their feedback on performance. Such feedback informed students where they lost their marks and how they could improve their performance scores. A few teachers' feedback reflected a focus on learning, whereby they discussed answers and clarified students' doubts. Only one teacher mentioned that they did not provide feedback because of the school's policy not to provide feedback for the last exam. See Table 37 for sample quotes from teachers and frequency distribution of feedback strategies.

Table 37

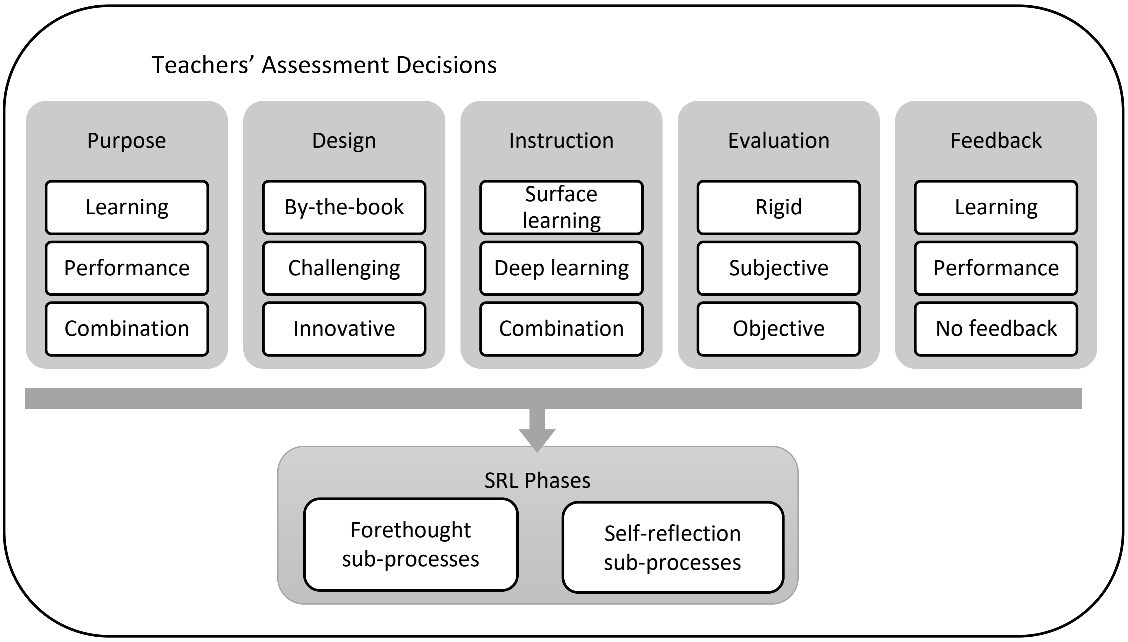
Distribution and Sample Codes of Teachers' Responses for Assessment Decisions (Feedback Strategy) (n=9)

Codes for assessment decisions and number of teachers	Code description	Sample quote
Performance (5)	The teacher focuses on where and why marks were lost. .	"[we] tell them where they go wrong. So, by that way they understand what the answer should be." Year 8, Physics teacher, ICSE school
Learning (3)	The teacher focuses on provide + clarify answers	"[we] give the answer scripts to them and they go through their paper, and then they come with the questions, like why was the mark cut, or what should they do so they can improve?" Year 8. Biology teacher, CBSE school
No feedback (1)	No indication of feedback	"For this exam, students were not given any feedback, because it is the end of the year assessment" Year 9, Physics teacher, ICSE school

These top-level codes that emerged from analysis provides a deeper understanding of the broad assessment design decisions teachers made for their formal science assessment. As such, this analysis allowed me to expand on teachers' assessment decisions that are identified in Figure 5 to include these insights and thus explore relationships to SRL sub-processes in the forethought and self-reflection phases. Furthermore, these categories for the different assessment decisions were used for quantitative analyses, the results of which are explained later.

Figure 5

Summary of Teachers' Assessment Decisions Derived from Content Analysis (n = 9)



4.5.3 RQ 2: Students' SRL Processes Based on Zimmerman's SRL Model (2009)

Descriptive analyses suggest that students demonstrated moderate SRL processes in the forethought and self-reflection measures. In the forethought phase, most students set outcome goals (general and specific) and used a moderate breadth of strategies which primarily included memorisation and comprehension techniques. In reference to motivational beliefs in the forethought phase, students, on average, reported feeling efficacious for the preparation of and performance on the assessment. Students' reported interest in and perceived value of the task were higher than the scale's midpoint.

In the self-reflection phase, students responded to questions on self-judgement and self-reaction processes. On average, students believed they were well-prepared for the assessment and primarily used performance outcomes to evaluate their performance. With respect to causal attributions, students chose more robust strategies for success when compared to failure. In terms of their reactions, most students were satisfied with their performance. However, their responses reflected moderate inferences, which were generic and vague. Table 38 highlights the descriptive statistics for students' SRL sub-processes in the forethought and self-reflection phases.

Table 38

Descriptive Statistics for Forethought and Self-Reflection Sub-Processes (n=229)

SRL Phase	Sub-processes	Minimum	Maximum	*M*	*SD*
Forethought Task Analysis	Goal-setting	1	3	2.04	.48
	Strategy Selection (breadth)	1	5	1.72	.654
	Strategy Selection (depth)	1	4	1.78	.931
Forethought Motivational beliefs	Self-efficacy for preparation	1	5	3.63	.93
	Self-efficacy for performance	1	5	3.65	.99
	Perceived Interest	1	5	3.74	1.33
	Perceived importance	1	5	4.36	.80
Self-reflection Self-judgement	Self-evaluation	1	5	3.61	.98
	Judgement criteria	1	3	2.27	.64
	Causal attribution for success	1	3	2.22	.95
	Causal attribution for failure	1	3	1.98	.98
Self-reflection Self-reaction	Perceived satisfaction	1	3	2.08	.895
	Adaptive inferences	1	4	2.12	1.09

4.5.4 RQ 3: Teachers' Assessment Decisions and Students' SRL Processes

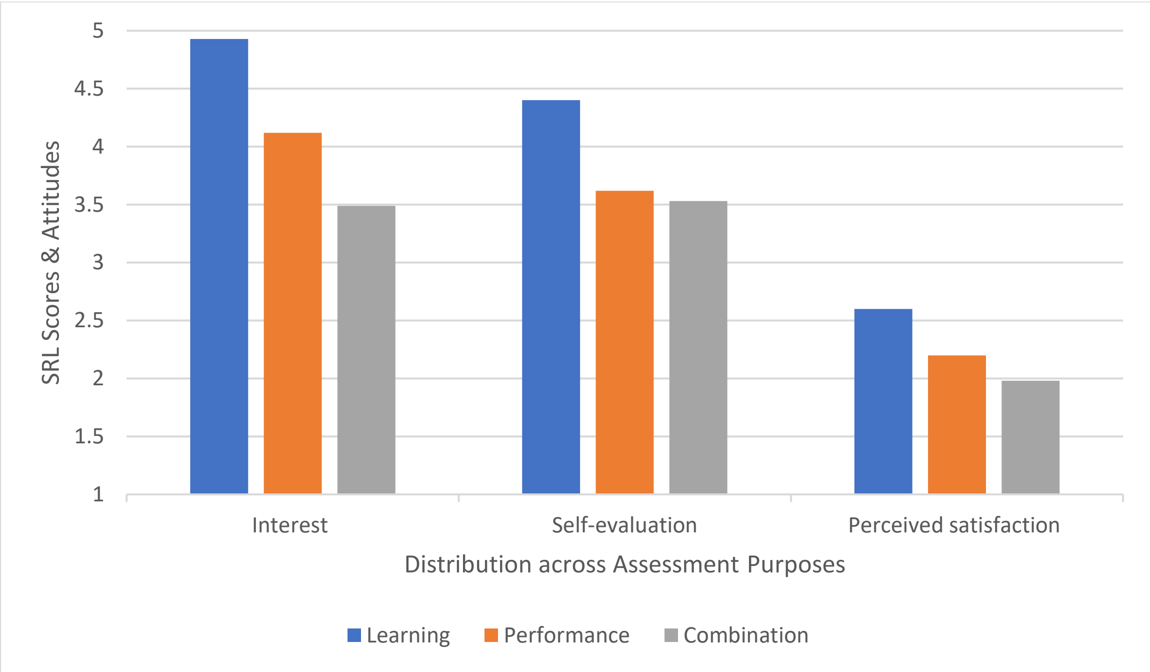
This section presents the ANOVA and chi-square (X^2) tests for independence that examined differences and associations between teachers' assessment decisions and students' SRL processes, respectively. I organise this section based on the questions from the interview regarding assessment decisions and explain the differences in SRL processes among students.

Assessment Purpose

Students were divided into three groups according to the purpose set by their teacher (Group 1: learning; Group 2: performance; Group 3: combination). There were statistically significant differences for strategy breadth interest scores in the forethought phase; and self-evaluation and perceived satisfaction scores in self-reflection phases as shown in figure 9. Table 36 describes post-hoc comparisons using Tukey HSD test and shows that mean scores for Group 1 (learning) were statistically different from Group 3 (combination). The maximum scores for perceived satisfaction were 3 and minimum scores were 1; and the maximum scores for interest, and self-evaluation scores were 5 and minimum scores were 1. The maximum score for strategy selection breadth was 7 and minimum score was 1; and the maximum score for strategy selection depth was 4 and minimum score was 1.

Figure 6

Means Plot for SRL Scores, Attitudes, and Assessment Purpose (Learning, Performance, Combination) (n = 229)



From Figure 6 it can be seen that mean scores are higher for strategy selection, interest, self-evaluation, and perceived satisfaction for learning-focused teachers. Students who felt more interested and satisfied with their performance were among that one teacher who perceive a learning-focused intention for assessments. It can also be seen from Figure 6 that students who selected deep elaborative learning strategies were among the same group of teachers (i.e., learning-focused). Table 39 provides the means, standard deviations, and one-way analyses of variance in SRL measures for assessment purpose.

A chi-square (X^2) test for independence indicated a significant association between assessment purpose and strategy selection, $\chi^2(4, n=229) = 23.19$, $p < .001$, $\phi = .32$ (Cramer's V = .23). This suggests that teachers' assessment purposes have a small to moderate effect on students' strategy selection (breadth). These results imply that *why* teachers assess students is linked to *how many* strategies students report when approaching tasks. There were no significant associations between teachers' assessment purpose and students' goal-setting, self-judgement strategies, causal attributions, or adaptive inferences. It is possible goal-setting and self-judgement sub-processes were not impacted by teachers' assessment purpose because formal assessments in India are evaluated and assigned scores.

Table 39

Means, Standard Deviations, and One-Way Analyses of Variance in Self-Regulated Learning Measures in the Forethought and Self-Reflection Phases for Purpose (n = 229)

SRL phase	Measure	Learning		Performance		Combination		*F*(2,228)	η^2
		M	*SD*	*M*	*SD*	*M*	*SD*		
Forethought	Goal-setting	2.07	.46	1.95	.44	2.08	.50	1.514	.01
	Strategy selection (breadth)	2.40	1.24	1.88	.751	1.68	.937	4.712*	.04
	Strategy selection (depth)	1.53	.516	1.74	.784	1.74	.613	.686	.01
Task analysis and motivational beliefs	Self-efficacy for preparation	4.13	1.06	3.72	.95	3.55	.90	3.076	.03
	Self-efficacy for performance	4.07	.88	3.81	1.05	3.54	.97	2.980	.03
	Interest	4.93	.26	4.12	1.14	3.49	1.37	12.266**	.09
	Perceived importance	4.60	.51	4.38	.88	4.33	.79	.819	.01
Self-Reflection	Self-evaluation	4.40	.83	3.62	.97	3.53	.97	5.633*	.05
	Self-judgement criteria	2.33	.48	2.22	.70	2.32	.64	.490	.00
Self-judgement and self-reaction	Causal attribution for success	2.47	.92	2.32	.94	2.16	.96	1.176	.01
	Causal attribution for failure	1.73	.96	2.21	.99	1.93	.98	2.235	.02
	Perceived satisfaction	2.60	.73	2.21	.86	1.98	.89	4.188*	.04
	Adaptive inferences	2.47	1.35	2.10	1.07	2.09	1.07	.821	.01

*p < .01

Assessment Design

Students were divided into three groups based on the teachers' design approaches derived from the content analysis (Group 1: by-the-book; Group 2: innovative; Group 3: challenging but balanced). There were statistically significant differences at the $p < 0.05$ level in most of the forethought sub-processes including goal-setting scores, strategy selection, self-efficacy for performance, and interest scores. Maximum scores and minimum scores for task analysis sub-processes (i.e., goal-setting) was 3 points and 1 point, respectively. For the motivational beliefs (self-efficacy and interest), maximum and minimum points were 5 and 1, respectively. The maximum score for strategy selection breadth was 7 and minimum score was 1; and the maximum score for strategy selection depth was 4 and minimum score was 1. Post-hoc comparisons using Tukey HSD test indicated differences across design groups as can be seen in Table 40.

From Figures 7 it can be seen that an innovative approach had higher goal-setting scores when compared to a challenging and balanced approach. The Figure also denotes higher mean scores for strategy selection in the balanced and challenging teacher designs when compared to using a book-based design approach. This finding indicates that students selected elaborate deep learning strategies when teachers designed the assessment with questions that had a good mix of questions (e.g., familiar and challenging). Mean scores for self-efficacy for performance were lower for an innovative approach when compared to the other two approaches. Interest scores were also lower for the innovative design approach compared to a more challenging, yet balanced design style (Figure 8). These results imply that students were less motivated among those teachers who chose questions that were unfamiliar to students.

Figure 7

Mean Plot for SRL Tasks Analysis Scores and Teachers' Design Styles (By-the-book, Innovative, Challenging) (n = 229)

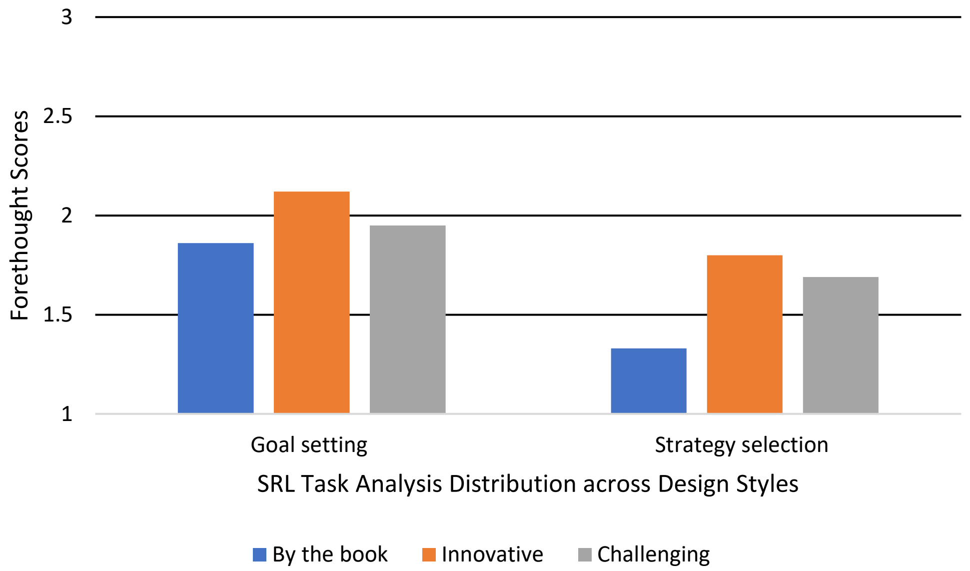
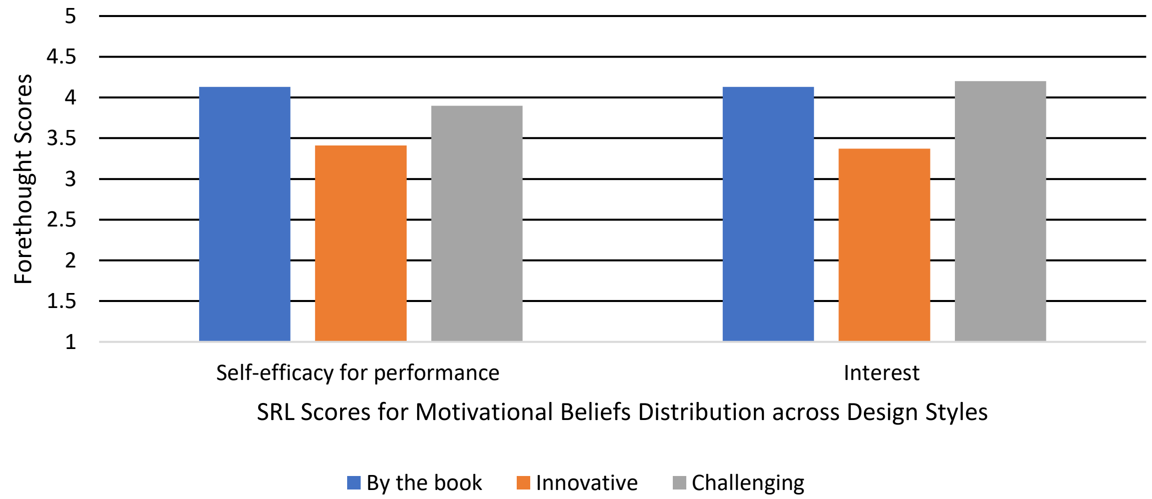


Figure 8

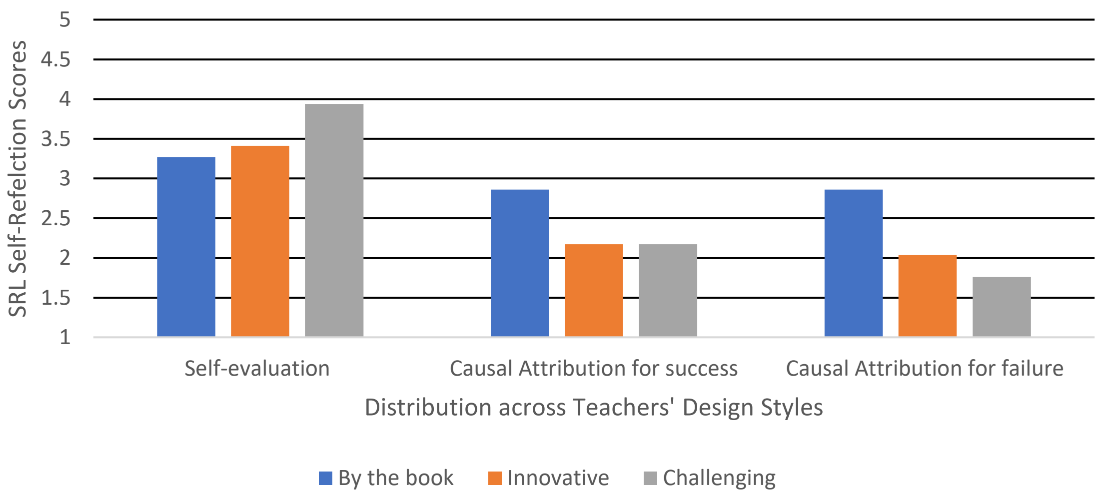
Mean Plot for SRL Motivational Belief Scores and Teachers' Design Styles (By-the-book, Innovative, Challenging) (n = 229)



There were also statistically significant differences at the $p < 0.05$ in the self-reflection phase which include subprocesses of self-evaluation, causal attribution for success and failures. Post-hoc comparisons using Tukey HSD test indicated differences in mean scores between the groups as described in Table 40. From Figure 9, it can be seen that self-evaluation scores were higher for the challenging, balanced design approach in comparison to the other two. And causal attribution scores for success and failure were higher for the book-based design in comparison to the other approaches. This result indicates that students who attributed their success and failure to factors within their control were among those teachers who chose questions from familiar sources. The minimum and maximum scores for causal attribution scores were 1 and 3, respectively. The minimum and maximum scores for the self-evaluation scale were 1 and 5, respectively.

Figure 9

Means Plot for SRL Self-Reflection Scores and Teachers' Design Styles (By-the-book, Innovative, Challenging) (n = 229)



Concerning the forethought measures, only goal-setting strategies were significantly associated with teachers' design style. A chi-square (X^2) test for independence indicated that teachers' assessment design style has a small to moderate effect on students' goal-setting strategies, χ^2 (2, *n*=229) = 11.41, $p < .05$, $\phi = .22$ (Cramer's V = .16).

With respect to self-reflection measures, causal attribution for failure and adaptive inferences were significantly associated with teachers' design style. Chi-square (X^2) tests for independence indicated that teachers' design styles had a small to moderate effect on causal attribution for failure, χ^2 (2, *n*=229) = 17.84, $p < .005$, $\phi = .28$ (Cramer's V = .19), and adaptive inferences, χ^2 (2, *n*=229) = 12.60, $p \leq .05$, $\phi = .24$ (Cramer's V = .17). These results indicate that students attributing their failures to controllable internal attributions and making adaptive inferences is related to assessment design. Therefore, how teachers designed their assessment was related to how students reflected on the success and failures of their performance.

Table 40

Means, Standard Deviations, and One-Way Analyses of Variance in Self-Regulated Learning Measures in the Forethought and Self-Reflection Phases for Teachers' Design Style (*n=229*)

SRL Phase	Measure	"By the book"		Innovative		Challenge, balanced		$F(2,228)$	η^2
		M	*SD*	*M*	*SD*	*M*	*SD*		
Forethought	Goal-setting	1.86	.35	2.12	.52	1.95	.42	4.515*	.04
	Strategy selection (breadth)	1.33	.62	1.76	.97	1.88	.931	2.56	.02
	Strategy selection (depth)	1.33	.49	1.80	.62	1.69	.70	3.758*	.03
Task analysis and motivational beliefs	Self-efficacy for preparation	3.87	.83	3.54	.96	3.72	.09	1.418	.01
	Self-efficacy for performance	4.13	.99	3.41	1.00	3.90	.91	8.789*	.07
	Interest	4.13	1.30	3.37	1.36	4.20	1.13	11.890*	.10
	Perceived importance	4.47	.74	4.32	.78	4.39	.85	.362	.00
Self-Reflection	Self-evaluation	3.27	1.22	3.41	.09	3.94	.85	9.304*	.08
	Self-judgement criteria	2.00	.53	2.33	.65	2.29	.64	1.816	.02
	Causal attribution for Success	2.86	.51	2.18	.96	2.17	.96	3.743*	.03
Self-judgment and self-reaction	Causal attribution for Failure	2.86	.52	2.04	.98	1.76	.95	9.049*	.07
	Perceived satisfaction	2.20	.86	1.95	.88	2.23	.89	2.809	.02
	Adaptive inferences	2.06	1.03	2.04	1.04	2.23	1.17	.854	.01

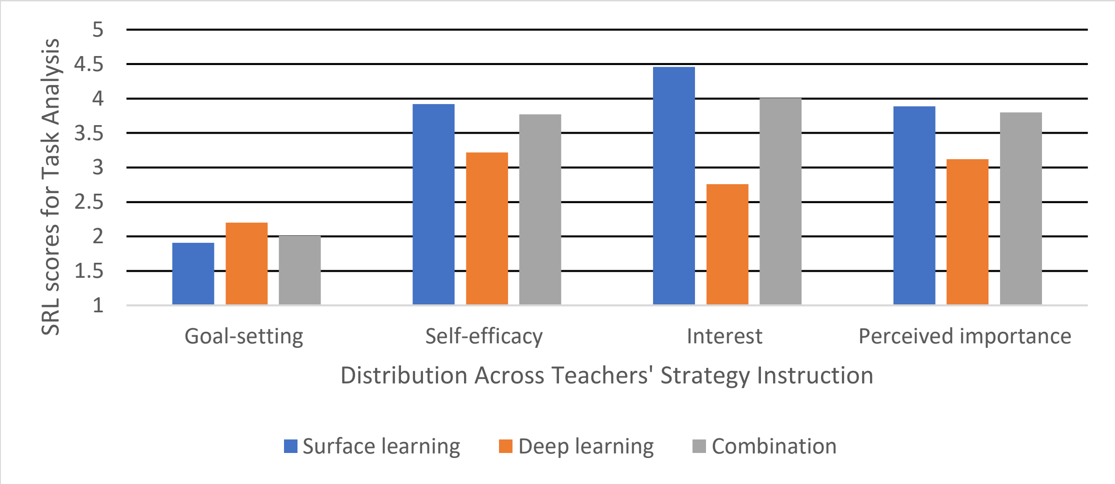
*$p < .05$

Assessment Instruction Strategies

Students were divided into three groups according to the feedback strategies set by their teacher (Group 1: surface learning; Group 2: deep learning; Group 3: combination). There were statistically significant differences at the $p < 0.05$ level in scores of forethought measures. Post-hoc comparison using Tukey HSD test indicated differences across groups (see Table 30). From Figure 10, it can be seen that mean scores for self-efficacy, interest, strategy selection (breadth), and perceived importance were higher for teachers who reported surface learning strategies compared to deep learning strategies. However, goal-setting scores were higher for teachers who instructed students about deep learning strategies in comparison to surface learning strategies (see Table 41). The minimum and maximum scores for goal-setting was 1 and 3, respectively. The minimum and maximum scores for motivational beliefs (self-efficacy, interest, perceived importance) scale were 1 and 5, respectively. The maximum score for strategy selection breadth was 7 and minimum score was 1; and the maximum score for strategy selection depth was 4 and minimum score was 1.

Figure 10

Mean Scores for SRL Task Analysis Scores and Attitudes based on Teachers' Strategy Instruction (Surface learning, Deep learning, Combination) (n = 229)



There were statistically significant differences at $p < 0.05$ level for two sub-processes in the self-reflection phase: self-evaluation and perceived satisfaction. The minimum and maximum scores for self-evaluation scores were 1 and 5, respectively. The minimum and maximum scores for perceived satisfaction scale was 1 and 3, respectively. Post-hoc comparisons using Tukey HSD tests indicated differences in mean scores for self-evaluation and perceived satisfaction scores were lowest for deep learning strategy instruction (see Table 41). This finding indicates that students who believed they were less prepared and felt less satisfied were among teachers who reported advising students to use deep learning strategies.

Figure 11

Mean Scores for SRL Self-Reflection Scores based on Teachers' Strategy Instruction (Surface Learning, Deep Learning, Combination) (n = 229)

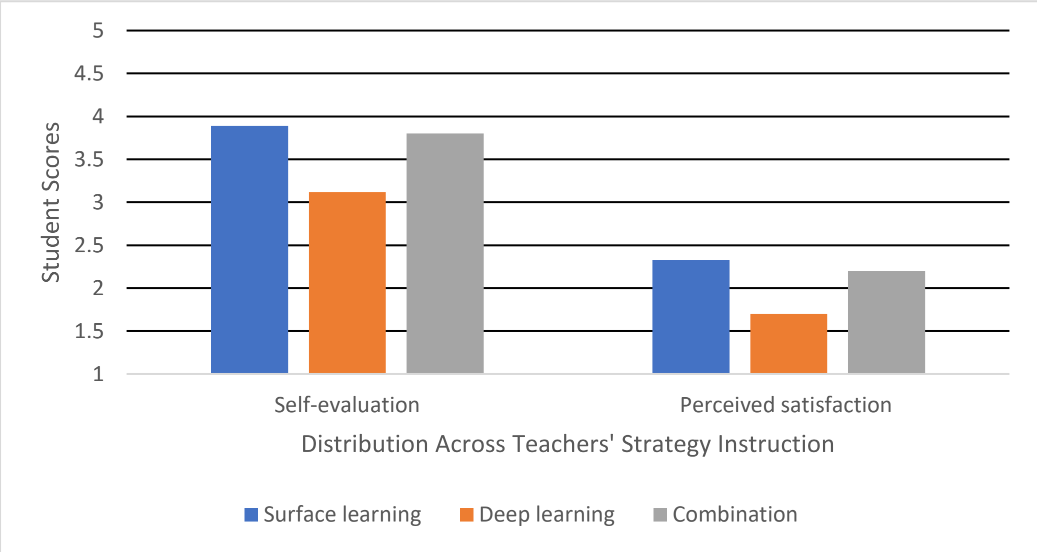


Table 41

Means, Standard Deviations, and One-Way Analyses of Variance in Self-Regulated Learning Measures in the Forethought and Self-Reflection phases for Cognitive Strategy Instruction (n = 229)

SRL phase	Measure	Surface learning		Deep learning		Combination		$F(2,228)$	η^2
		M	*SD*	*M*	*SD*	*M*	*SD*		
Forethought Task analysis and motivational beliefs	Goal-setting	1.91	.46	2.20	.55	2.01	.41	7.105*	.06
	Strategy selection (breadth)	2.00	.94	1.53	.86	1.82	.94	4.792*	.04
	Strategy selection (depth)	1.66	.71	1.81	.63	1.70	.63	1.007	.00
	Self-efficacy for preparation	3.92	1.03	3.22	.91	3.77	.75	12.565*	.10
	Self-efficacy for performance	4.02	.93	3.14	.94	3.80	.93	17.475*	.13
	Interest	4.46	.92	2.78	1.25	4.01	1.20	41.041*	.27
	Perceived importance	4.60	.75	4.15	.81	4.36	.80	5.708*	.05
Self-Reflection Self-judgment and self-reaction	Self-evaluation	3.89	1.09	3.12	.92	3.80	.80	15.146*	.12
	Self-judgement criteria	2.15	.64	.31	.68	2.38	.61	2.540	.02
	Causal Attribution for Success	2.43	.90	2.14	.97	2.13	.96	2.185	.02
	Causal Attribution for Failure	2.12	.99	2.06	.98	1.82	.96	2.151	.02
	Perceived satisfaction	2.20	.86	1.95	.88	2.23	.89	2.809	.02
	Adaptive inferences	2.06	1.03	2.04	1.04	2.23	1.09	.854	.01

*$p < .05$

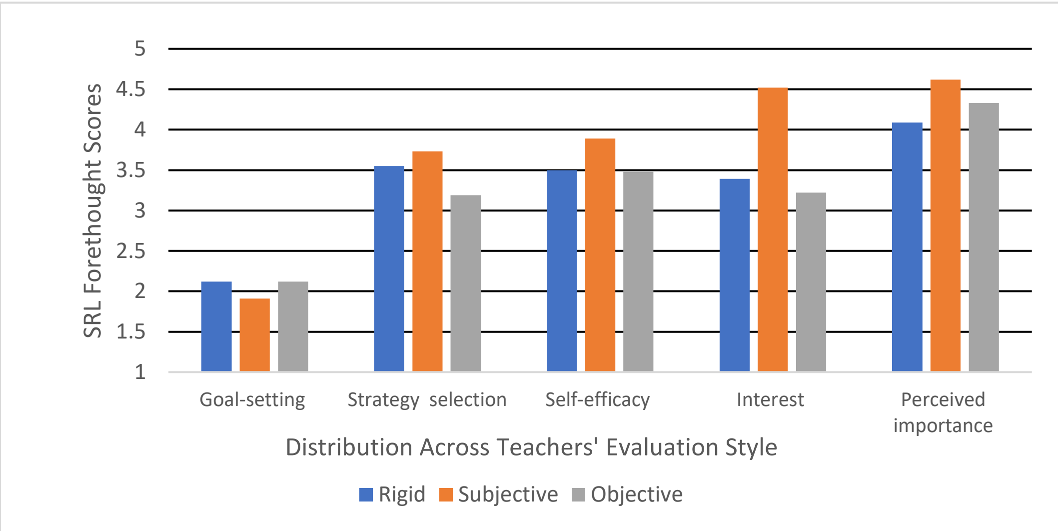
A chi-square (X^2) test for Independence indicated a significant association between teachers' strategy instruction and students' goal setting, $\chi^2(2, n=229) = 18.33$, $p < .005$, $\phi = .28$ (Cramer's V = .20) suggesting that teachers' strategy instruction had a small to medium effect on the type of goals students reported. This finding implies that how teachers advise students to learn and prepare for their assessments relates to students focusing on their learning on the journey of the learning (process goals). There was no significant association between cognitive strategy instruction and other SRL processes, including strategy selection, judgement criteria, causal attributions, adaptive inferences, or perceived satisfaction.

Assessment Evaluation Strategies

Students were divided into three groups according to the evaluation strategies set by their teacher (Group 1: Rigid; Group 2: Subjective; Group 3: Objective). There were statistically significant differences at the p <0. 05 level in most of the forethought sub-processes including goal-setting scores, strategy selection, self-efficacy for preparation and performance, interest, and perceived importance scores based on how teachers evaluated student performance. Post-hoc comparison using Tukey HSD test indicated that mean scores differed significantly across groups as reported in Table 42. For instance, Figure 12 indicates that mean scores for goal-setting were low for subjective evaluators compared to rigid and objective evaluation styles. Other SRL sub-processes including strategy selection (breadth), self-efficacy, perceived importance, and interest scores were higher for subjective markers when compared to objective or rigid evaluations. In particular, interest had a high effect size (0.19). The minimum and maximum scores for goal-setting scores was 1 and 3, respectively. The minimum and maximum scores for motivational beliefs (self-efficacy, interest, perceived importance) scale were 1 and 5, respectively. The maximum score for strategy selection breadth was 7 and minimum score was 1; and the maximum score for strategy selection depth was 4 and minimum score was 1.

Figure 12

Means Plot for SRL Forethought Scores and Teachers' Evaluation Style (Rigid, Subjective, Objective) (n = 229)



For the self-reflection phase, there was a statistically significant difference at the $p < 0.05$ for self-evaluation, causal attribution, and adaptive inference scores. Post-hoc comparisons using Tukey HSD test indicated differences in mean scores among the evaluation groups (see Table 42). From Figure 13 it can be seen that self-evaluation mean scores were low among teachers whose marking style was coded as objective when compared to those teachers whose marking styles were coded as rigid or subjective evaluation styles and were highest for the subjective evaluation style.

Figure 13

Means Plot for SRL Self-Reflection Scores and Teachers' Evaluation Style (Rigid, Subjective, Objective) (n = 229)

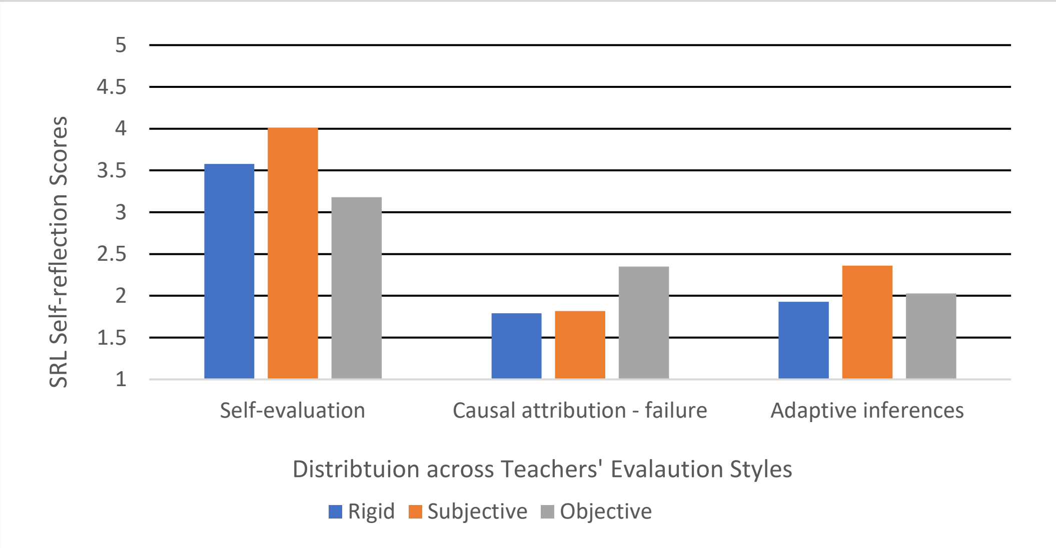


Figure 13 denotes that mean scores for adaptive inferences were higher for subjective evaluations, but mean scores for causal attribution for failure were low subjective evaluation styles when compared to the other evaluation styles. The minimum and maximum scores for self-evaluation scores were 1 and 5, respectively. The minimum and maximum scores for perceived satisfaction scale was 1 and 3, respectively. Table 42 provides the means, standard deviations, and one-way analyses of variance in SRL measures for assessment evaluation strategies.

Concerning forethought measures, chi-square (X^2) tests for independence indicated a significant association for goal-setting strategies and teachers' evaluation styles, χ^2 (2, n=229) = 10.21, p < .05, ϕ = .21 (Cramer's V = .15). This result indicates that how teachers chose to evaluate student performance had a small to moderate effect on the types of goals students set. For self-reflection measures, there were significant associations between teachers' evaluation styles and causal attribution for failure, χ^2 (2, n=229) = 16.73, p < .005,

ϕ = .27 (Cramer's V = .19), adaptive inferences, χ^2(2, n=229) = 13.98, p < .01, ϕ = .25 (Cramer's V = .16), and perceived satisfaction, χ^2(2, n=229) = 11.95, p < .05, ϕ = .23 (Cramer's V = .16). These results suggest that teachers' evaluation styles related to how students attributed their failures, adaptive choices, and their perceived satisfaction.

Table 42

Means, Standard Deviations, and One-Way Analyses of Variance in Self-Regulated Learning Measures in the Forethought and Self-Reflection Phases for Teachers' Evaluation Styles (n = 229)

SRL Phase	Measure	Rigid		Subjective		Objective		*F*(2,228)	η^2
		M	*SD*	*M*	*SD*	*M*	*SD*		
Forethought Task analysis and Motivational Beliefs	Goal-setting	2.12	.45	1.91	.45	2.12	.53	4.692*	.04
	Strategy selection (breath)	1.78	.98	1.98	.89	1.55	.88	4.191*	.03
	Strategy selection (depth)	1.77	.65	1.76	.69	1.64	.61	.830	.00
	Self-efficacy for preparation	3.50	.95	3.89	.94	3.48	.85	5.010*	.04
	Self-efficacy for performance	3.47	.98	3.99	.88	3.44	1.04	7.987*	.06
	Interest	3.39	1.42	4.52	.78	3.22	1.33	27.411*	.19
	Perceived importance	4.09	.89	4.62	.71	4.33	.71	9.083*	.07
Self-reflection Self-judgement and Self-reaction	Self-evaluation	3.58	.99	4.01	.93	3.18	.83	15.719*	.12
	Self-judgement criteria	2.32	.59	2.21	.66	2.35	.67	.955	.01
	Causal Attribution for Success	2.28	.94	2.18	.95	2.20	.97	.233	.00
	Causal Attribution for Failure	1.79	.97	1.82	.96	2.35	.91	8.025*	.07
	Perceived satisfaction	2.16	.87	2.18	.93	1.87	.84	2.780	.02
	Adaptive inferences	1.93	1.03	2.36	1.20	2.03	.97	3.508*	.03

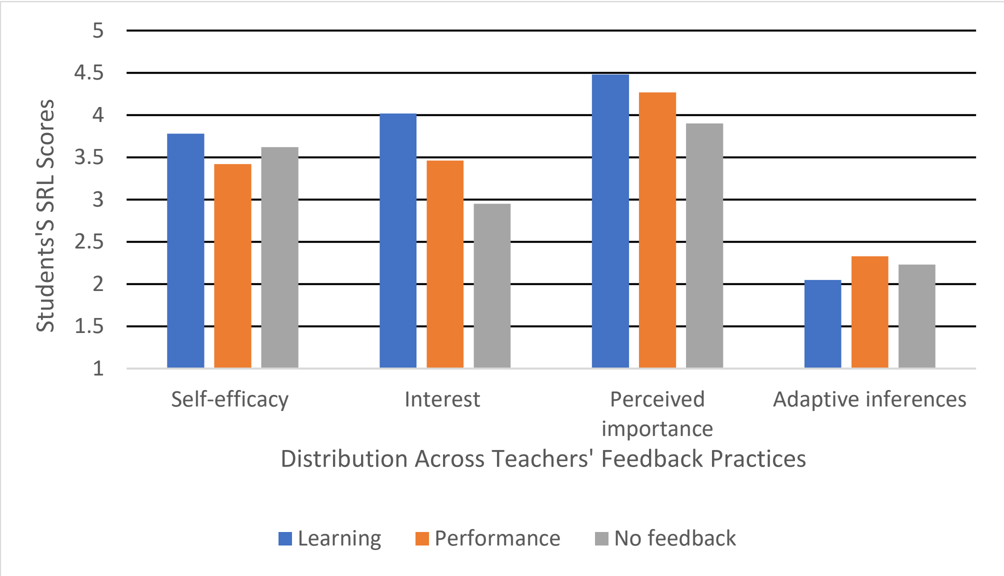
*$p < .05$

Assessment Feedback Practices

Students were divided into three groups according to the feedback strategies set by their teacher (Group 1: performance; Group 2: learning; Group 3: no feedback). There were statistically significant differences at the p <0. 05 level in most of the forethought sub-processes including self-efficacy for performance, interest, and perceived importance scores and only one sub-process in the self-reflection phase (i.e., adaptive inferences). Post-hoc comparison using Tukey HSD test, described in Table 43, indicated that mean scores of the forethought subprocesses differed across groups. Figure 14 indicates that mean scores for self-efficacy for performance, interest, and perceived importance were higher for feedback that focused on learning in comparison to performance or no feedback. With respect to strategy selection (depth), there were statistically significant differences at the $p < .001$ level. Post-hoc comparison using Tukey HSD test, indicated that the mean scores for strategy selection (depth) was higher for feedback focused on performance when compared to learning-focused feedback or no feedback.

Figure 14

Means Plot for SRL Scores and Attitudes for Teachers' Feedback Strategies (Learning, Performance, No Feedback) (n = 229)



Additionally, there were differences in mean scores for adaptive inferences with lower scores for learning-focused feedback when compared to performance-specific feedback. This finding indicated that students who chose specific strategies that they would adapt were among teachers who focused on performance. In Figure 14, the maximum and minimum scores for self-efficacy, interest and perceived importance was 5 and 1, respectively. The maximum scores for strategy selection breadth and depth were 5 and 4 respectively; and the minimum score for strategy selection breadth and depth were 1. And for adaptive inferences, the maximum score was 3 points, and the minimum score was 1. Table 43 provides the means, standard deviations, and one-way analyses of variance in SRL measures for assessment evaluation strategies.

The chi-square (X^2) test for independence indicated no significant association between feedback strategies and goal setting or strategy selection scores. There was also no significant association between feedback strategies, judgment criteria, causal attributions, adaptive inferences, or perceived satisfaction.

Table 43

Means, Standard Deviations, and One-Way Analyses of Variance in Self-Regulated Learning Measures in the Forethought and Self-Reflection Phases for Teachers' Feedback Strategies (n = 229)

SRL Phase	Measure	Performance		Learning		No feedback		$F(2,228)$	η^2
		M	*SD*	*M*	*SD*	*M*	*SD*		
Forethought Task analysis and motivational beliefs	Goal-setting	2.08	.49	2.02	.50	2.05	.38	.348	.00
	Strategy selection (breadth)	1.81	.86	1.79	.97	1.57	.99	.574	.01
	Strategy selection (depth)	1.95	.64	1.63	.64	1.52	.60	6.827	.06
	Self-efficacy for preparation	3.62	.93	3.67	.97	3.43	.59	.620	.01
	Self-efficacy for performance	3.42	1.00	3.78	.99	3.62	.86	3.130*	.03
	Interest	3.46	1.32	4.02	1.27	2.95	1.24	8.870*	.07
	Perceived importance	4.27	.88	4.48	.70	3.90	.94	5.495*	.05
Self-reflection Self-judgement and Self-reaction	Self-evaluation	3.42	.96	3.72	1.03	3.57	.59	2.219	.02
	Self-judgement criteria	2.29	.67	2.26	.64	2.52	.51	1.496	.01
	Causal Attribution for Success	2.18	.96	2.24	.95	2.19	.98	.098	.00
	Causal Attribution for Failure	1.97	1.00	2.02	.97	1.76	.99	.682	.01
	Perceived satisfaction	2.00	.91	2.05	.89	2.47	.74	2.418	.02
	Adaptive inferences	2.33	.74	2.05	.83	2.23	.70	3.033*	.03

*$p < .05$

4.5.5 Summary of Results

Overall, from teachers' reports it appeared that assessment decisions were guided and constrained by school policy and national standards. Teachers' responses to the questions in the interviews regarding assessment decisions were coded into categories. Based on their statements, assessment design aspects including purpose, task, evaluation, feedback, and strategy instruction varied among them. Further quantitative analyses indicated that that students' SRL processes and academic performance differed based on the categories. Students' strategy selection (breadth), interest, self-evaluation scores, and perceived satisfaction were statistically different when comparing students whose teachers stated their assessment purpose as focused on learning outcomes to those teachers who stated that the focus was on performance or a combination of both. It was also seen that each teacher's reported design style contributed to differences in students' SRL in the forethought and self-reflection phases. For example, strategy selection scores (depth) were higher among teachers who reported an innovative style than those who reported using the textbook. Students SRL scores differed based on the type of strategies teachers stated. For example, students mean scores for goal-setting, interest, self-efficacy for learning and performance differed between teachers whose advice was coded as surface learning instruction from those teachers' whose advice was coded as deep learning instruction.

Furthermore, students' motivational scores differed among teachers who reported that they evaluated students based on their own standards instead of adhering to rigid standards from those who reported being objective or rigid. Students' motivation also differed based on how teachers' feedback strategies were coded. In particular, the results indicated that teachers who reported supporting students with formative feedback which focused on learning differed from teachers who reported that their feedback focused on improving performance. And finally, chi-square (X^2) tests for independence revealed that some of decisions made by teachers were associated with students' SRL scores. For instance, design and evaluation styles were associated with how students chose to attribute their failure. Overall, the results suggest that teachers' reported assessment decisions can contribute to students' approaches before and after a formal assessment task. These findings shed insights into the value teachers' intentions for formal assessments could have in promoting and sustaining SRL for science learning and performance.

4.6 Discussion

The current study aimed to broaden the scope of SRL development by understanding teachers' formal assessment decisions and examining their impact on and relation to students' SRL processes. Several important findings were identified. First, how teachers think about assessment (and its purposes) may contribute to what students are interested in and how they try to learn and perform. Second, a combination of design approaches appeared to affect students' confidence in learning and performance, their goals, and reasons they believed they were unsuccessful. Third, the results also suggested that students reported perceived competencies and interest differed between teachers who stated they instructed students to use surface learning approaches such as rehearsal and comprehension and those teachers who reported instructing students with deep learning strategies. Finally, teachers' evaluation and feedback strategies were also associated with students' SRL processes. In particular, the findings suggest that students who felt more interested and confident, and made adaptive choices were among teachers whose feedback was categorised as focused on learning.

Findings from the current study support contemporary theoretical frameworks which advocate for assessments to have a learning-focused intention for improving SRL (Andrade & Brookhart, 2020; Chen & Bonner 2020). In particular, it was found that students' mean scores for forethought (strategy selection - breadth and interest) and self-reflection (self-evaluation) sub-processes were higher among students who had a teacher who intended the assessment to support student learning when compared to students whose teachers focused on assessments on performance objectives. This finding suggests that teachers' perceptions of the assessment purpose could contribute to students’ confidence in the subject and their interest in a positive way. Teachers who have a learning-focused perception could have encouraged students to deepen their understanding of the subject using elaborative learning strategies (Andrade & Brookhart, 2020; Butler & Schnellert, 2015). Additionally, students who reported choosing mastery criteria or their past performances to evaluate their performance were among the teacher’s whose stated purpose focused on learning rather than performance (Andrade & Brookhart, 2020; Heritage, 2010; Moss et al., 2013). This approach could strengthen how students evaluate their performance and their view of assessments as valuable to learning, thus expanding the scope of assessment in their perspective, beyond performance outcomes.

The current study indicates that students' SRL processes vary based on how teachers' task design approaches were categorized, implying that intentions for design decisions about the assessment tasks could positively contribute to SRL. Teachers whose intentions were to design the assessment with a challenging and balancing approach had students who reported deep learning strategies, more interest in the subject, and evaluated their performance based on how well they responded to questions or past scores. Although the relationships reported in this study are not causal, it is assumed to have flow-on effects according to the social cognitive theory: the environment influences SRL and vice versa (Zimmerman, 2002). These approaches particularly differed from those teachers who intended to use a creative or by-the-book design approach (Lodewyk & Winne, 2006; Perry et al., 2013). At the same time, teachers design was coded as innovative in their design approach had students' report process goals that are more adaptive than performance goals, when compared to other design approaches. Furthermore, students' confidence for the subject and attributions for failures were higher for the teacher whose reported design approach was confined to the textbook. Similar to Lodewyk et al. (2005), results from this study suggest that students might feel more confident about their capacities within defined parameters. Students might have found reasons that were internal and controllable for their failures because they could have been aware of the textbook's questions. This knowledge could have encouraged them to find reasons within their control rather than focusing on task difficulty or teacher bias in grading. Thus, intentions for design approach can promote SRL. If teachers’ intentions include a combination of balance and innovation when designing questions, they could enhance students' self-efficacy, interest, and strategy selection.

This study also established differences in students' SRL processes based on teachers' reported strategy instruction. Students who reported feeling competent and interested, and valued the task were among teachers who advised them to use surface learning strategies compared to those whose advice was coded as deep learning strategies such as elaboration (e.g., finding real-life examples). Although there may be several reasons why students' scores were higher in these teacher groups, one possible reason may be that surface learning strategies give students a sense of confidence in the subject matter in terms of understanding the material (Hattie & Donoghue, 2016). In contrast, deep learning strategies may aid students in manipulating, synthesizing, and integrating knowledge it can only occur once content knowledge is acquired (Hattie & Donoghue, 2016). Accordingly, it may be possible that teachers who reported advising students to use deep learning strategies without surface-level strategies might have students feeling incompetent due to reduced knowledge acquisition that surface learning facilitates. These findings reiterate the significance of explicitly instructing students on a range of strategies that suit the context. Therefore, pre-service and in-service teachers could be supported in their effort to teach students how to learn which differs from learning, especially in the Indian context where this aspect has largely been ignored (NCET, 2020). In the current study, students' SRL processes in the forethought and self-reflection phases differed based on the evaluation strategy categories. On the one hand, students' motivation scores were lower for teachers who reported using an objective or rigid grading approach when compared to those whose responses were coded as a subjective approach. This finding might suggest that objective criteria might not be sufficient in enhancing students' motivational beliefs. It might be worth considering students' performance from a nuanced point of view. For example, teachers in India may credit student responses that adequately meet the question even if they do not include specific keywords. The new curricular framework for teacher education in India recommends that teachers take a more conscious effort to develop standards for evaluation of student work (NCET, 2020).

On the other hand, mean scores for causal attribution of failure were higher for those students whose teachers adopted a more objective approach, suggesting that students found a controllable reason for their mistakes in their performance. This finding is consistent with contemporary research wherein evaluative criteria that reflect clear expectations and recommend corrective action promote adaptive strategies and improve performance (Andrade & Brookhart, 2020). Indeed, both grading practices positively and negatively impact student learning. However, it is possible to combine the advantages of both approaches to facilitate a positive impact on student learning. For instance, clearly outlining and sharing expectations and criteria for various performance levels with students and using the same criteria could improve their overall SRL processes.

Finally, students' SRL processes were higher for those teachers who focused their feedback on learning when compared to performance or no feedback groups. These findings broaden the importance of feedback decisions from formative assessments to formal classroom assessments (Brookhart, 2016; Butler, 1987; Butler & Nissan, 1986). Teachers who focused their feedback on learning had students who reported more interest, higher self-efficacy, and made adaptive inferences such as improving a skill, knowledge, or strategy for their next learning cycle. Thus, these findings reiterate the value of qualitative and constructive feedback to enhance students' learning (Hattie & Timperley, 2007). Teachers can combine learning objectives with evaluation criteria to integrate feedback into formal classroom assessments. This approach can help students identify their areas of strengths and weaknesses in the content and their responses, and teachers can provide more specific feedback to address these areas of growth. Future research could focus on examining these practical implications and understanding how students interpret feedback and incorporate it into their SRL processes for their formal assessments.

Overall, the study provided an insight into the assessment context in Indian schools. A content analysis of teacher responses indicated that teachers' assessment decisions were based on the assessment context of the school such as policies regarding when the assessment takes place, the format (i.e., exam/test), and the aims of the assessment (i.e., measure student learning). The results form an analysis of variances, it was evident that teachers' intentions for formal assessment decisions results in differences in students' SRL processes. Importantly, teachers' intentions for formal assessments (such as exams) may be expanded to include SRL processes.

Based on the findings from this study and existing literature, I suggest a checklist that teachers could use to reflect on their intentions for formal classroom assessment design decisions and to purposefully include practices that promote SRL (Figure 16).

Figure 15

Design Guide for Teachers to Integrate SRL into Formal Assessments

Purpose and Goal-Setting	☐
Have I considered the purpose of this assessment?	☐
Have I considered ways in which the purpose of the assessment encourages a learning-focused orientation (e.g., identify learning objectives and convert them into success criteria)?	☐
Have I considered sharing the assessment purpose with my students?	☐
Have I considered ways in which I can encourage my students to use success criteria to help them set goals before the assessment?	☐
Task Design and Strategy Instruction	☐
Have I considered whether the task design is appropriate for the selected learning objectives?	☐

Have I designed the task with variability in complexity, familiarity, and types of questions?	☐
Have I considered whether questions align with the learning objectives/success criteria?	☐
Have I considered whether my students are aware of the different types of questions?	☐
Have I considered the various learning strategies required for the assessment task?	☐
Have I considered ways in which I can teach students different strategies and how to use them?	☐
Have I considered the ways in which I can encourage my students to monitor and reflect on their strategy use?	☐
Evaluation Criteria	☐
Have I set clear expectations for different types of questions (e.g., short answers, computations, diagrams)	☐
Have I set clear expectations for questions that vary in weightage (e.g., two marks, one mark)?	☐
Have I shared these expectations with students in ways that are accessible to students at all times?	☐
Feedback Strategies	☐
Have I considered how my feedback supports student learning **and** performance?	☐
Have I considered how I can encourage my students to engage in self-assessment?	☐
Have I considered how I can use the evaluation criteria to support my feedback?	☐
Have I considered ways in which I can check with my students on how they have interpreted my feedback?	☐
Have I considered ways in which I can encourage students to use this feedback to set goals and reflect on strategies?	☐

4.6.1 Limitations

The findings from the current study provide compelling links between teachers' intentions for assessment design decisions and students' SRL but must be considered with its limitations. This research project is focused on a small sample within the Indian context which does not allow for the generalisation of the findings.

Furthermore, I used interview protocols to collect the data for a single formal assessment event, thus limiting the interpretation of the results to a specific context and temporal dimension. It is possible that participants' responses were reflective of the social desirability bias. Additionally, categorical data were the primary variables because of the data collection decisions. The results would not suggest causal or correlational relationships but merely establish differences among teachers' intentions regarding assessment decisions. And finally, the study does not account for extraneous variables that could influence teachers' responses, such as professional development, personal beliefs, or prior knowledge.

And finally, the study was conducted during a global pandemic which has undoubtedly had an impact on teachers' intentions for assessment design. Teachers needed to adapt to unprecedented circumstances, which included designing and administering assessments, and supporting students through online platforms. It is likely that teachers' intentions will differ based on the support and resources they received from their school. Therefore, the results from the study are specific to this unique circumstance and cannot be generalized to other teachers or schools within the country or globally.

Although the current study did not compare teachers' intentions for assessment with the assessment they actually designed, the findings offer a glimpse into how a teacher in India considers various assessment-related decisions. In India, professional development for pre-service and in-service teachers has been inadequate. Recent changes to the policy have made teacher education a priority, particularly regarding assessment design (NCERT, 2020). Indeed, the new curricular framework emphasises that teachers need to be competent at designing comprehensive and valid assessments that can assess more than information. However, there are no guidelines on how these decisions can be made nor are there any indications of ensuring the assessment comprises elements that can support self-regulated learning processes. The findings provide suggest potential areas for further professional development for improving students' SRL processes. Future research could focus on making connections between intention for assessment and the assessment that actually used with students. In this way, teachers and researchers could work together to understand the processes and elements required for an assessment to be valid, meaningful, and comprehensive.

This research project expands the scope of supporting students' SRL with formal assessments. The results reiterate the significance of teachers' perceptions, decisions, and instruction on students' learning approaches and motivational beliefs and further support Zimmerman's social cognitive model of SRL. Even though the study has limitations, the findings from the study have implications for practice. The results from the study were used to propose a self-assessment checklist for integrating SRL into planning and decision-making processes for formal assessments. The outcomes from the study also provide directions for future research, which include integrating SRL into formal classroom assessments with intentional design principles.

Chapter 5: General Discussion and Conclusions

In this concluding Chapter, I review the findings from each study in this thesis and illustrate how they jointly contribute to the field of SRL within a classroom assessment context using novel methods. I also describe some implications for practice and research, emphasising how these findings could inform assessment decisions in the classroom. I conclude this Chapter by offering my final reflections on this thesis.

This thesis was designed to study self-regulated learning (SRL) processes for three primary reasons. First, SRL processes - when enacted effectively - improve academic outcomes such as deepening conceptual understanding of subject matter and enhancing skills such as critical thinking and problem-solving (Callan, 2018; Cleary & Kitsantas, 2017; DiBenedetto & Zimmerman, 2010; Perry et al., 2007). Second, by improving these outcomes with guidance in school, young people become strong, independent learners who can cope with and respond to problems they encounter after school (Sinatra & Taasoobshirazi, 2018; Usher & Schunk, 2018). And third, in science, strong SRL has the potential to change the downward achievement curve, particularly in India (Andrzejewski et al., 2016; NCERT, 2020; 2017). This thesis builds on prior research about SRL and its connection to formal classroom assessments.

Classroom assessments have gained a critical position in literature as a powerful context within which SRL processes can be taught, modelled, and sustained (Andrade & Brookhart, 2016; Chen & Bonner, 2020; Panadero et al., 2016). Several theoretical frameworks are proposed by scholars in the effort to link assessment practices with SRL processes (Allal, 2020; Andrade & Brookhart, 2021; Chen & Bonner, 2021). Each of these frameworks identify goal-setting, feedback, and self-assessment as critical practices that promote SRL when used in formative assessments. The significance of informal classroom assessments in improving students' SRL processes provided a strong evidentiary basis for investigating formal classroom assessments. Since SRL is dependent on context, I focused on aspects of formal assessment which include the task design and teachers' decisions regarding purpose, task, evaluation styles and feedback strategies.

Unpacking these elements of a formal assessment is particularly relevant to India because an exam-oriented system continues to dominate the formal assessment context. Research findings indicate that national exams have a strong influence on teachers' assessment tasks and decisions (Sivaraman, 2011). High school teachers in India also struggle with formative assessments and tend to conduct formative assessments as summative tasks (Brown et al., 2015). And finally, SRL processes remain unexplored in the context of formal assessments. Much of the literature in India explores students' study habits for national exams or competitive exams for professional courses. Therefore, these aspects of the formal assessment tasks are the focus of this research study, adding to our understanding of how the formative interactions occur in day-to-day teaching.

I aimed to integrate these insights from SRL literature and formative assessment frameworks into India's formal classroom assessment context. I narrowed my thesis down to the following enquiries about formal classroom assessment and its relationship to SRL:

1. What SRL processes do students demonstrate in the forethought and self-reflection phases of Zimmerman's model for a science formal classroom assessment?

2. To what extent do students' SRL processes in the forethought and self-reflection phases of Zimmerman's model relate to academic performance in science?
3. In what ways do the design features of the assessment task relate to students' SRL processes in the forethought and self-reflection phases of Zimmerman's model?
4. What are teachers' decisions regarding a formal classroom assessment for science?
5. In what ways do teachers' formal classroom assessment decisions relate to students' SRL processes in the forethought and self-reflection phases of Zimmerman's model?

These research questions were identified to explore how the relationship between a formal classroom assessment and SRL processes could further our understanding of SRL within an authentic, formal classroom assessment context and whether this understanding could be used to improve students' use of SRL processes. In the subsequent paragraphs, I review the findings followed by a discussion on how the results contribute to the existing body of literature. I conclude with limitations and implications for future research and practice.

5.1 Integrating Findings from Study 1 and Study 2

The study described in Chapter 3 investigated the first three research questions while the study in Chapter 4 examined the last two research questions. In the forethought phase, most students reported performance goals and predominantly resorted to rehearsal strategies. These findings are consistent with the results from Chapter 4, where most teachers who perceived assessments as serving a summative function advised students to use surface learning strategies (e.g., recall and rehearsal). There was only one teacher who perceived assessments with a learning-focused orientation, and they had students with higher scores on SRL processes such as strategy selection (breadth), interest, and perceived value. This outcome is critical to the formal assessment context because it suggests the potential role teachers' intentions can play in improving SRL processes through formal assessments.

Students' motivational beliefs - confidence in their abilities to learn and perform on the assessment and interest in the subject - varied across assessment design features and teachers' decisions. Students' perceived confidence in the assessment was inversely related to the cognitive complexity of the assessment. This finding was further strengthened by results in Chapter 4: students' confidence was higher for the teacher who preferred to design their assessment by-the-book. These results are expected given that more students and teachers reported rehearsal strategies than elaborative strategies and familiarity enhances self-efficacy (Lodewyk et al., 2005; Omrod, 2011).

SRL processes in the self-reflection phase of Zimmerman's model (2009) varied across assessment features and teacher decisions. Students reported evaluating their performance using the grade or score they received from their teachers. In the case of attribution scores, results suggested that students were less likely to attribute failure to controllable factors when there were more questions related to real-life contexts. It was also found that students who had high attribution scores for failure were among the teacher who relied on textbook questions. It is possible that questions relating to real-life contexts were unfamiliar to students and their learning strategies were less effective to face novel questions. In this case, students' scientific competencies and SRL skills are curbed which may lead to self-handicapping strategies (e.g., task-avoidance; Zimmerman, 2013).

And finally, students' reactions to their performance differed across assessment features and teacher decisions. Students were more likely to adapt their learning when assessment was more challenging, familiar, and presented various types of questions across several sub-topics. It is possible that students performed less than expected in challenging assessments and thus sought strategies to improve their next performance. Another explanation might be that students had more data to reflect upon when questions were diverse which encouraged them to consider specific aspects they could improve. This outcome is critical for scientific competencies because the subject often covers a range of skills (e.g., diagram drawing, problem-solving, explanation and reasoning). Students who reported higher adaptive scores were also among teachers' whose feedback practices focused improving performance. This outcome further reiterates students' and teachers' conceptions of science assessments in India are indeed rooted in performance which are expected for graded assessments (Irving et al., 2011).

In sum, these findings provide a meaningful and focused look at SRL in contemporary classroom assessment literature. They demonstrate that many of the formative assessment practices advocated in the literature for SRL development are also relevant to formal classroom assessment contexts. Specifically, the findings suggest two important implications. First, the data suggests that assessment task features do relate with students' SRL processes, but research is needed to understand how the task design can be used to promote SRL. The second finding from this thesis related to teachers' intentions regarding decisions, interactions, and presentation of the assessment. Results suggest that the established coding from the study contributes to differences in students' SRL processes, but further investigation is needed to establish how teachers' intentions and actual behaviours regarding formal assessment can potentially be used to enhance SRL among students. When these two implications are viewed together, formal classroom assessments can be a consistent and effective opportunity for students to demonstrate their SRL processes and for teachers to enhance their processes.

5.2 Contributions to the Field: Broadening the Assessment Context

In this thesis, I investigated SRL processes in a relatively under-researched context of a formal classroom assessment: India. I designed this research to contribute to the relevant knowledge and methodologies. Formal classroom assessment is typically viewed as high-stakes tests that aim to evaluate how much knowledge students have gained on a particular topic or set of topics (Gardner, 2010). Some scholars suggest that formal assessments are also opportunities for students to use effective self-regulated learning processes to perform and achieve successful learning outcomes (Chen & Bonner, 2020). Ultimately, assessments are opportunities for students to advance their understanding of what they are learning and how they are learning (Black & Wiliam, 2018; Hattie, 2007). However, the potential of formal classroom assessments to support SRL is lacking in scholarly literature. This research has attempted to fill in this gap and use the insights to offer a new perspective on students' SRL processes for a formal assessment task.

The findings from this thesis indicate that aspects of a formal classroom assessment relate to students' SRL processes. From a social cognitive perspective, this outcome is not unexpected because it is well established that SRL is the resulting interaction between the environment and the individual (Zimmerman, 2013). The findings support that formal classroom assessments are valuable opportunities to promote SRL among students. For instance, teachers' perceptions of the assessment, such as its purpose, are typically differentiated in the literature as formative or summative. There is a general agreement that formative assessment is for self-regulated learning (Clark, 2012; Panadero et al., 2018). The results from the current study indicate that students who reported adaptive SRL processes (e.g., goal-setting, perceived satisfaction) were among the teacher whose responses were coded as viewing formal assessments with a learning lens. Indeed, not all students in the classroom of that teacher reported adaptive SRL processes. Contemporary scholars encourage that assessments be viewed as opportunities to support learning and learning processes (e.g., Black & Wiliam, 2018; Broadbent et al., 2018). Moreover, a learning-focused perspective enables educators and students to consider all assessments as moments for reflection and action, thus promoting and sustaining SRL competencies for short and long-term success.

Although scholars and teachers agree that formal classroom assessments can serve both summative and formative functions, it is often much more challenging in practice. One reason for this challenge might be because summative assessments signal the end of a topic; thus, curbing opportunities for teachers and students to revisit the content. These circumstances are common across the world because formal assessments typically occur in the middle of a term or in the end to collect data on how much students have learned about a topic (Gardner, 2010; NCERT, 2020). While it is understandable that teachers may not be able to teach the content again, it is possible to reflect upon and improve SRL processes through the formal assessment process. For instance, designing and implementing varied classroom-based assessments rather than only exams/tests could help students reconsider their learning processes and attitudes.

This dynamic approach to assessment is especially useful to science education. Teachers' view of assessments can be broadened to include learning for subject mastery and is not limited to improving learning for exams alone (Brown et al., 2014). Asia-Pacific countries such as the New Zealand, Taiwan, Singapore, and Japan are moving away from an examination-dominant culture toward low-stake assessments (Koh & Luke, 2009). This move reduces the pressure of standardized assessments teachers experience in other countries such the United States, United Kingdom, and India (NCERT, 2020; NCET, 2020). With reduced pressure of performance and increased support for professional development, teachers in India can exercise their agency in the classroom to integrate SRL processes into the formal classroom assessment process.

Additionally, science education encompasses the integration of skills, knowledge, and competencies, which in turn necessitates many different types of opportunities to demonstrate their competencies. School leaders need to lend significant support to teachers as they endeavour to meet the learning needs of students. Teachers' practices and perceptions are guided and influenced by personal beliefs and their school's professional learning and development opportunities (Sahanowas & Halder, 2016). Creating new programmes for teachers to implement in the classroom is a step in the right direction but is not sufficient (Black et al., 2011). Teachers will need resources, policies, and the support of key stakeholders such as parents and principals (Black et al., 2011; NCERT, 2020). Specifically, policies that emphasize the value of assessments in acquiring mastery of the subject for students to become strong, independent learners will lay the foundation for future practices in the classroom. Although education policies aim for such learning and mastery to occur, policies such as No Child Left Behind in the United States (US) and

Continuous and Comprehensive Evaluation in India place a strong emphasis on test scores through. In contrast, a few countries in the Nordic (e.g., Finland) and Asia-Pacific regions (e.g., New Zealand) have optional standardized assessments; thus, allowing teachers to maximize their attention on learning rather than performance.

A core component of this thesis was examining the task design, teachers' decisions behind the design, and their relation to SRL processes. From the findings discussed in Chapters 3 and 4, assessment tasks that measured subject knowledge more than application and evaluation of knowledge were more likely to encourage students to report strategies that strengthened retrieval and explanatory skills (Crooks, 1988). Although content knowledge is essential for science, it is equally important that students are given the opportunity to evaluate and transfer their knowledge to develop and strengthen competencies such as problem-solving and critical thinking (Hattie & Donoghue, 2016; Sinatra & Taasoobshirazi, 2018). Earlier studies have highlighted the value of variety in a task because it encourages discernment or conditional knowledge and prepares students for unfamiliar situations or problems (Anderson et al., 2001; Smith & Smith, 2014). This was not the case in the current study, as students were more likely provided with familiar questions which might reduce the development of metacognitive strategies (Smith & Smith, 2014).

In Chapter 3, it was found that adaptive inferences were predicted by assessment design features such as cognitive complexity. Students were more likely to change their strategy or pay more attention to specific topics when the questions required deep conceptual understanding and problem-solving skills. When students are challenged about what they know, they are more likely to find ways to improve their knowledge. Developing tasks that are appropriately challenging for students with diverse individual characteristics is likely a difficult endeavour for teachers in a classroom; thus, promoting SRL is complex.

Therefore, including classroom assessment literacy in professional learning and development programmes for teachers could enhance their skills in assessment design. With these competencies, teachers may be able to consider several effective ways in which they can assess, monitor, and enhance student learning as required for their subject (Black et al., 2011). In the case of science, teachers could design assessment tasks that measure knowledge and scientific competencies such as logical thinking, problem-solving and critical thinking. Governments and educational organizations such as the IB have frameworks that outline key principles in assessment which include design, evaluation, and feedback. Scholars also recommend including students in the assessment process (Stiggins, 2005). This collaboration raises students' autonomy and ownership about learning. Such an approach can be beneficial to teachers in India, where assessment frameworks and professional workshops can encourage and empower teachers with assessment literacy that will be beneficial to student learning.

In the current study, students' motivations differed based on the task features and teachers' intentions regarding decisions for a formal classroom assessment. For instance, interest tended to pique when the task covered several subtopics, and students' efficacy was likely lower when the task demanded deeper learning processes. It was also found that students' mean scores for self-efficacy were higher for the teacher who reported using only textbook resources as a basis for their task design when compared to teachers who reported preferring a challenging style. This was also the case for teachers whose strategy advice was coded as surface learning strategies. But mean scores for interest were higher among teachers who used a tougher approach. It is important that teachers and assessment designers heed these findings because prior studies and results from this thesis indicate that motivational beliefs are key predictors of academic success and effective SRL (e.g., Lodewyk et al., 2006). While it may be difficult to motivate every student with the one assessment task, it is possible to consider ways in which autonomy and competence can be enhanced. This approach is likely to improve their motivation and attitudes toward assessments (Ryan and Deci, 2000). One way in which students can exercise agency is by having some meaningful choices in the assessment process such as the types of questions (e.g., short answer, computational problems). Teachers and students can also collaborate to identify and implement learning strategies that are task specific, thus improving their ability to learn for and perform on assessments.

These findings suggest that teachers' intentions and approaches to assessment and students' interpretations of the final task are related and could be aligned to improve learning and performance. Moreover, teachers have the autonomy to design formal classroom assessments. Yet, teachers in India and other countries experience exams as the most dominant form of assessment (NCERT, 2020; Pellegrino, 2013). Nonetheless, it is worth considering how assessment intentions and approaches can be reflected upon to include support needed for students become strong, independent learners and simultaneously measuring subject-specific competencies (Pellegrino, 2013; Perry et al., 2006; Perry et al., 2020). It is essential that teachers are provided with on-going support as they learn to design various types of assessment tasks and other stakeholders (e.g., parents, school leaders, tertiary institutions) are informed about these changes (Black et al., 2011). In this way, teachers may face lesser pressure from external sources to conform to conventional examinations and approach assessment design with more competence and assurance.

In addition to the task design, the evaluation criteria are critical to students' SRL processes and motivation. Since the assessment in the current study counted toward students' final grades, it is not surprising that they chose normative criteria for self-evaluation. Additionally, teachers' evaluation practices focused on 'cutting marks' for incomplete or incorrect answers rather than 'awarding marks' for partial or correct answers. Such an approach leaves little room for developing adaptive strategies as they do not inform students with feedback that could feed into the next cycle (Zimmerman, 2013). It was also found that students who attributed their failures to controllable factors such as effort or strategy use belonged to classrooms where teachers graded students with objective criteria. It is likely that students who were aware of why they lost marks were able to find reason in their performance or process instead of teacher bias.

The findings from this study provide reasonable grounds to consider self-assessment tools based on mastery of the subject and the assessment design to instil a learning-focused orientation, foster adaptive learning approaches, and enhance academic success for formal assessments (Butler, 2018; Harris et al., 2015; Panadero et al., 2017). Learning and advancing knowledge require self-reflective practices that help students understand their current education gaps (Hattie & Timperley, 2007). With more fine-grained self-assessment tools, students could pay more attention to their responses to a task. This would include how they plan (e.g., goal-setting and strategy selection) and reflect on the efficacy of their plan (e.g., judgement criteria, attributions). Such proactive approaches to learning would encourage students to focus on the learning process as a whole rather than only on the outcome, such as grades or scores (Andrade, 2019; DiBenedetto & Zimmerman, 2010; Panadero & Romero, 2014).

In the current study, teachers' feedback practices were more focused on performance than on learning, which is less effective in promoting SRL processes. Students who reported higher SRL processes (e.g., interest, perceived importance, and self-efficacy) were among those teachers who provided learning-focused feedback. Indeed, teachers prefer assessment feedback to be focused on learning, however when the assessment serves summative functions the focus on grades tends to take precedence over learning (Harris et al., 2009; Irving et al., 2010). Results from this study provide reasonable evidence to extend the use of formative feedback practices (e.g., self-assessment tools, descriptive feedback, and support for planning next steps) into formal assessments. It is recommended that teachers help students advance their understanding of the subject while also potentially improving performance outcomes (Broadbent et al., 2018; Kirschner & Hendrick, 2020; Pintrich, 2000). By promoting and encouraging a learner-focused perspective (e.g., purpose and feedback) on formal assessments, students also report higher levels of interest, efficacy, and perceived value of the task, which ultimately strengthen other SRL processes (Zimmerman, 2013).

It could also be useful to consider other forms of feedback in addition to teacher comments such as self-assessment and peer-assessment. Teachers can develop self- and peer-assessment tools based on Hattie and Timperley’s feedback model (2007) to encourage students to evaluate performance in relation to the task (performance), process (cognitive strategies and actions), regulation (metacognitive strategies), and self (e.g., effort, motivation). It will be necessary for teachers to guide students to evaluate their work and interpret feedback in effective ways (Harris et al., 2015). Since teachers in India emphasize on performance in their feedback, they will need on-going support to implement a wider framework for feedback.

In the current study, students' adaptive scores were statistically different between teachers whose feedback focussed on performance and teachers whose feedback focused on learning. This result is expected given that performance outcomes are more sought-after than learning outcomes. Nonetheless, taking an adaptive approach is useful for students because they begin to perceive mistakes as learning opportunities and potentially develop a growth mindset which will support them to become lifelong learners (Karlen et al., 2021). Taking this perspective is often challenging because students' inferences from feedback and evaluation strategies are more likely to be ego-protective than growth-oriented (Harris et al., 2018). Educators will need to consider how to support students to make inferences from feedback data to be adaptive rather than defensive (Bandura, 1986; Zimmerman, 1989). In the case of formal assessments, students will have repeated opportunities to make modifications to their learning processes. Encouraging self-assessment and feedforward practices in the formal assessment practice, students can learn to react to their performance with an adaptive and positive mindset. With this skill students are more likely to be better equipped to confront challenges outside the classroom, where solutions and decisions are not as straightforward as a formal assessment.

Indeed, formal assessments are typically perceived as serving a summative purpose, which refers to determining and reporting how much students have learned. Findings from the current research provide evidence that formal assessments can also serve formative functions, particularly for SRL development in students. This implication is critical to consider because formal assessments can carry high stakes, and students must possess the required competencies and attitudes to succeed in a task designed to assess knowledge on several topics within a specific time frame. SRL processes in contemporary frameworks tend to be embedded broadly in the assessment context (Andrade & Brookhart, 2020; Chen & Bonner, 2020). Insights from the current thesis show in greater detail how particular processes of SRL can be captured in the formal assessment process. These insights offer novel and unique avenues to reflect on intentional assessment design to promote SRL.

5.3 Contributions to the Field: Broadening Methodological Design

Classroom assessments encompass various practices that can be leveraged to improve SRL processes to further students' academic outcomes and performance. Given this immense potential, research must embrace multiple methodological approaches to gain a nuanced and in-depth understanding of the relationship between classroom assessments and SRL. Chapter 2 explained my methodological decisions to investigate the relationship between SRL and classroom assessments. The expansive field of SRL measurement offers a range of tools to employ, from self-report questionnaires to more fine-grained tools such as think-aloud interviews and microanalytic protocols. The microanalytic protocol appeared promising for the current dissertation because of its adaptability and specificity. Moreover, the microanalytic protocol offers immediate data for practical implications, a core objective of this doctoral research (Cleary & Callan, 2018). In this thesis, I provide further evidence supporting this tool as a method of data collection to measure several SRL processes across age, academic domains, and culture (e.g., student, teacher, task, school). The microanalytic protocol, grounded in Zimmerman's SRL model and strengthened by its structured guidelines, allowed me to modify the instrument to suit the context and student age.

Indeed, the SRL microanalysis protocol offered flexibility in its structure and administration, but the coding and scoring guidelines were challenging to implement. For instance, priori codes were not a successful approach to analysing the data reported for a formal assessment. The wide range of responses required me to modify the codes to authentically reflect students' reported SRL processes. Through this approach, I was able to develop a coding scheme with detailed descriptions for each SRL sub-process and specific instructions for establishing validity and reliability of the tool (Callan & Cleary, 2018).

Additionally, developing a scoring scheme for the codes for complex statistical analyses required thoughtful deliberation and informed decision-making which was not available in the protocol's guidelines. This limitation of the protocol required me to make crucial decisions about how the data were treated: ordinal or categorical (Cohen et al., 2018). The vast body of SRL literature allowed me to treat data as ordinal and post-hoc analyses provided reasonable support for my decision. By approaching the data in this way, I was able to offer meaningful insights regarding SRL and its relationship to performance, assessment design, and teachers' intentions for assessment design.

Furthermore, classroom assessments are rooted in context and driven by the larger society's values and priorities of the larger society thus requiring the use of diverse methods to collect and analyse its data. Chapter 2 of this thesis explained the advantages of a mixed-methods approach to address these context-specific requirements. This unique approach led to the development of an instrument to measure the characteristics of assessment tasks in Chapter 3. Scholars have identified types of academic tasks that challenge students' learning, such as questions that target higher order thinking skills such as problem solving, analytical thinking, and logical reasoning, that appear to promote SRL (Lodewyk et al., 2006; Perry et al., 2004). But there have been no precise methods to examine the design features of a task. Chapter 3 discusses the value of investigating task features concerning SRL because when SRL and task characteristics are viewed together, teachers and students can make better decisions to support learning and performance. Importantly, through the development and use of an analytical framework, I was able to quantify the features that were evident in the task. This approach is distinct from previous studies in which the researchers developed the task items (e.g., Callan, 2014; DiBenedetto & Zimmerman, 2010; Lodewyk et al., 2006) and thus provides a novel method in which a teacher-designed task can be measured that relates explicitly to SRL.

In the study described in Chapter 4, I continued to mix methodologies to determine differences in students' SRL based on teachers' assessment design decisions. A qualitative method for data collection, such as a semi-structured interview, allowed me to understand design decisions within a context. Indeed, recent assessment literacy frameworks seek to measure the multiple facets embedded in the assessment context (e.g., Classroom Assessment Inventory, Luca et al., 2016). However, many of these surveys are developed by keeping in mind the assessment standards of Western, educated, industrialized, rich and democratic (WEIRD) countries. In India, assessment standards are met with ambivalence, and the use of existing assessment literacy frameworks could be misleading (Singhal, 2012). Since there are no clear guidelines or professional development programmes focused on assessments, teachers have unique practices and interpretations of policy. Furthermore, unlike other countries (e.g., New Zealand, Canada), India's national education policy (NEP) provides a general idea about students' learning outcomes and assessment practices. For example, in New Zealand the curriculum documents clearly state the expected learning outcomes for students for each school year and there are resources for assessment (Ministry of Education, 2022). In the case of India, the NEP vaguely suggests the skills to be acquired and that assessments need to examine student learning in comprehensive ways (NCERT, 2020).

Therefore, a mixed methodological approach could draw broader insights into how teachers make assessment design decisions in culturally diverse contexts. Through this research, I lend further support to adopting a mixed-methods approach in educational psychology research because of its immense potential to broaden and deepen our understanding of classroom design decisions and their relationship to SRL for future practice and research. Findings from this research reiterate the emphasis on the environment in SRL development. Based on these results, I offer insights that teachers and students can use to unpack the task characteristics and deliberate on design decisions that support SRL.

5.4 Implications for Future Practice

SRL is a robust set of processes and beliefs that can propel academic outcomes beyond subject knowledge and into lifelong learning competencies such as problem-solving, creativity, and critical thinking. Classroom assessments are complementary to the SRL process and are advocated as practices for promoting and enhancing SRL. When viewed together, they encompass a dynamic set of strategies to boost student learning and help them tap into their immense potential. My motivation to research SRL processes in classroom assessments was to provide practical insights to empower students with strong SRL processes that can support them in becoming lifelong learners and work through challenges outside the classroom. I was encouraged by existing theoretical frameworks linking classroom assessments to SRL and was eager to understand how this into formal assessments such as examinations for a diverse context to provide evidence-based guideline to strengthen these links. My methodological decisions, backed by a pragmatic worldview, supported the data gathering and analyses to eventually help apply the findings to classroom contexts.

This thesis provides further evidentiary basis to consider how formal classroom assessments can impact SRL positively. In particular, task design features and decisions leading to its design play an integral role in how students view and respond to them, suggesting that teachers and students could deliberate carefully upon their feelings, thoughts, and language around assessments to achieve academic success. Teachers and students can be empowered through learning modules that integrate SRL strategies into formal assessments. Chapter 3 suggests that students' strategy use is limited to memorisation and comprehension, and their perceived interest in the task essentially predicts them. In Chapter 4, it was found that teachers' strategy instruction varied between surface and deep learning strategies with little instruction to students on *how* to use them. It could be worth providing pre-service and in-service teachers with professional development and resources (e.g., checklist from Chapter 4) about strategies that target different cognitive processes and how they can be used to support student learning. For example, associating rehearsal strategies with definitions and elaborative strategies such as relating subject matter to analysis could help broaden the vocabulary and repertoire of teachers and students.

Furthermore, the instrument developed for the study in Chapter 3 provides a concrete tool to support SRL processes. Teachers can help students set goals based on their efficacy and select strategies aligned with the task design. Teachers can use the same approach to prompt student reflection to meet task demands and expectations and reflect upon their assessment design. For example, fostering a mastery-focused orientation by reiterating that assessments are opportunities for deepening self-awareness about their learning processes and the subject can be helpful for students. Teachers could also use questions to help students analyse the task and respond effectively. Adopting a reflective view of assessment decisions makes allowances for context. For instance, in India, formal assessments are the most dominant assessment format, and by reimagining their assessment decisions, educators can implement the principles detailed in the national education policy. Teachers and students would be able to redirect the focus of assessments on learning and how to learn, rather than on using rote memorisation strategies to enhance performance outcome.

Typically, formal assessments in India are informed by school policy influenced by national and state examinations but do not adequately meet the expectations in the national curriculum framework (NCERT, 2005; 2020). In particular, national education policies recommend that assessments steer away from exam-oriented formats focused on outcomes because they encourage rote memorisation techniques. Instead, policymakers suggest using diverse assessment tasks with a developmental or formative approach that focuses on improving the learning processes of that subject (NCERT, 2020). For instance, science assessments should ideally capture content knowledge and thinking processes such as making observations, reasoning, and evaluating inferences. Yet, no clear guidelines help teachers design assessments to meet these recommendations. It could be helpful to refer to the instrument developed in Chapter 3 and the checklist in Chapter 4 to help teachers reflect and evaluate their assessment design strategies to meet standard policies and enhance SRL processes simultaneously. I only suggest this approach as a step toward integrating SRL into formal assessments. And perhaps with this step and on-going reflection and research, values associated with assessments shift from performance to learning. In the subsequent paragraphs, I propose a working model to describe a decision-making framework for teachers to incorporate formal assessments practices that can encourage strong SRL sub-processes among students.

5.4.1. Integrated Working Model for Formal Assessments and SRL

The findings from the current study provide reasonable empirical evidence to formulate a working model for a framework that integrates formal assessment practices and intentions to promote strong SRL skills among students. This working model (Figure 17) is distinguished into three inter-related phases along with the forethought and self-reflection phases (part of Zimmerman’s (2009) model).

In the first preparatory stage, teachers consider the why (purpose) and the what (subject knowledge and skills) of the assessment. In this stage, teachers also decide how to assess and evaluate student performance (e.g., rubric, answer key). As established in the study described in Chapter 4, students reported stronger SRL processes (e.g., goal-setting, self-confidence) among teachers who view assessments with a learning-focused intention. Therefore, at this point of the model teachers could explain to students the importance of viewing assessments a learning opportunity rather than just performance. Through this learning-focused lens, teachers could work with students to set goals and choose strategies based on the content and the process. For example, for an up-coming science formal assessment, teachers can help students understand the key learning objectives from a lesson. With this understanding, teachers and students can begin to recognize which strategies would help them achieve each objective. In this way, students can also monitor their progress and adapt their strategies to optimize their learning.

The next phase in the preparatory stage focuses on teachers' approaches to task design and questions included in the task. Based on the results from the studies described in Chapter 3 and 4, task design and design style relate to SRL processes in different ways. For instance, cognitive complexity and a challenging design style from teachers are associated with students' perceived interest in the subject. But it was also established that students felt more confident with the task when the task was less complex and had teachers use the textbook for designing the questions. These findings suggest that students feel competent and are more likely to adapt their strategies when questions are familiar to them. From this perspective, teachers could consider how they can make the formal assessment task feel familiar to students. The green dotted-arrows in Figure 17 suggest that identifying learning objectives and success criteria could help students gain a stronger awareness of what they need to learn thus providing them with an anchor for familiarity.

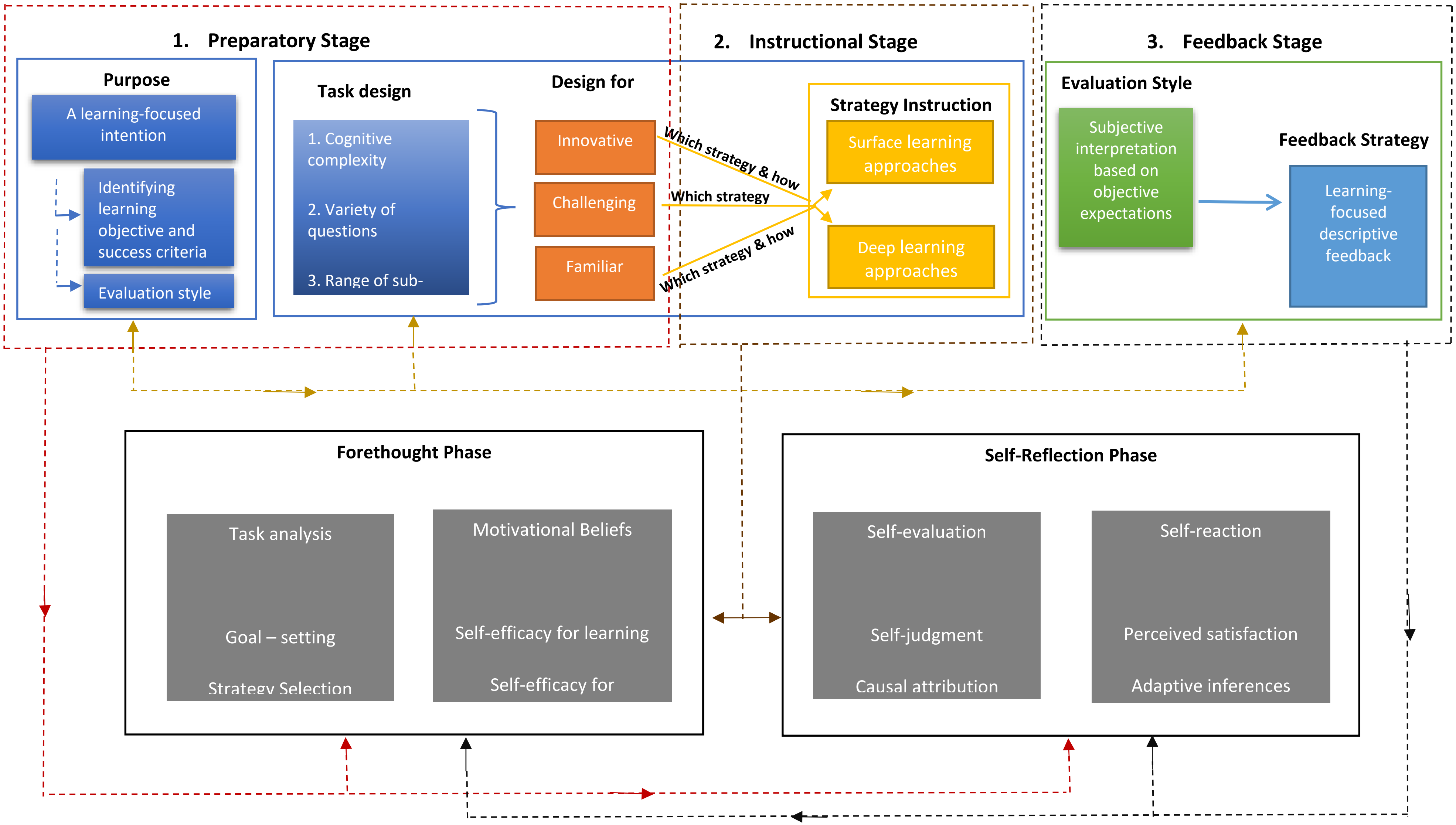
Additionally, students' goal-setting strategies, causal attributions, and adaptive strategies varied across assessment aspects. It was found that students deliberate on their learning process and set up a step-by-step approach to their goals when the assessment task comprises questions that cover several sub-topics. It might be possible that having several sub-topics to learn helps students split the content into sections which makes it manageable for them. Therefore, teachers can use the checklist suggested in Chapter 4 to reflect on their assessment task.

Another direction that can raise students' perceived efficacy for challenging and unfamiliar questions when presented in a formal assessment is for teachers to identify and teach strategies that support surface and deep learning during the second stage: the instructional stage. This instructional stage comprises explicit instruction by teachers to students regarding how to learn and acquire knowledge which is necessary for deep conceptual understanding and for enhancing students' metacognition (Greene, 2021). However, students rarely use strategies aligned with the task expectations (Lodewyk et al., 2006; Omrod, 2011). This was also seen in the current study, wherein students reported surface-level strategies that aid memorization rather than deep learning strategies required for some tasks. These findings in conjunction with existing literature imply that teachers need to instruct students on which strategies are useful for surface learning and deep learning and *how* to use and modify strategies as needed. This could be challenging because many students feel confident in their strategy-use and find it unnecessary to change their approach.

Indeed, it is well-established that strategy use is related to students' motivational beliefs (Zimmerman & Schunk, 2011). The results from this thesis indicated that interest positively correlated with students' strategic selection and predicted performance. Therefore, teachers can consider ways in which they can evaluate and accordingly promote situational interest. One way in which teachers could develop students' motivation is by increasing their autonomy. For example, teachers could design the task (e.g., setting goals, identifying learning objectives, setting evaluation criteria) in collaboration with students. Another approach in which teachers can make this change is by integrating and modelling strategy use during classroom instruction. For example, when teaching a concept, teachers can explain an appropriate strategy and ask students to practice the strategy with another concept.

Figure 16

Teachers' Decision-Making Framework to Support and Sustain Students' SRL for Formal Assessments



Finally in stage 3, the feedback stage, is comprised of decisions related to the types of feedback provided to students. A formal assessment could typically contain two types of feedback. The first type is more evaluative such as performance outcomes (grades/scores) because a function of formal assessments is to provide summative data. In this case, it is necessary for teachers to consider the standards against which student performance will be assessed. Results from the current study indicate that teachers who were less rigid and awarded marks for meeting the question requirements - in their view - had students report greater motivation and adaptive strategies. Students' causal attributions scores were also higher among teachers who used objective criteria. This meant that students could identify controllable factors as reasons for failure. When students attribute their failure to factors within their control, they can adapt their strategies to improve their learning in the next cycle. For example, if a student believes their failure was because they forgot the answer, they could try different memory strategies for retention.

The second type of feedback is more descriptive in nature and can take several different forms. Prior research as well as findings from the current study indicate the value of learning-focused feedback. Students reported stronger SRL processes among those teachers who helped students identify errors in their learning. And students reported adaptive goals among those teachers who provided descriptive feedback focused on improving performance. Thus, teachers need to consider ways in which these two types of feedback can foster strong SRL sub-processes. In the model (Figure 16), it is suggested that teachers use their subjective knowledge of the student while still adhering to objective criteria for evaluation. Teacher could also provide students with descriptive feedback that focuses on learning to strengthen students' motivation and cognitive strategies. Ultimately, this would reinforce decisions in stage 1, thus consistently strengthening the SRL process among students. This model is designed as a working model and would require further research to strengthen the links between formal assessments and SRL development.

5.5 Limitations of Current Research

There are some limitations to this thesis which should be addressed in future research. The first limitation is the sample included in this study. The sample comprise a small number of students and teachers from three different educational boards in India. As such there are variations in assessment practices and instructional strategies that were not controllable. To mitigate this limitation, I recruited students from similar regions of the country (i.e., south India) who performed on science formal assessments. Due to the small sample size the sample could not be nested for statistical analyses. There are also differences in the professional learning development offered to teachers across the school, which could have impacted teachers' assessment intentions. Therefore, findings from this study are specific to the schools included in the study and cannot be generalized to all schools and teachers belonging to these different boards. Future research in the Indian context could either recruit more students from each board or limit the study to a single educational board across the country. A larger sample size of a single board could limit the number of extraneous variables and provide fine-grain results.

The second limitation concerns the methods used in the studies described in Chapter 3 and Chapter 4. Due to limited prior research on assessment task structure, I had to develop an instrument to measure design features for study 1. This instrument was developed to measure a formal classroom assessment specific to the Indian context and could be missing key elements, thus making it less robust and reliable. In designing the instrument, I strived to ensure elements included would support SRL development and was derived from previous studies. Future studies focused on assessments could use the instrument to provide insights on its reliability and validity for similar formal assessments. It could also be useful to remove or add an element to gain causal understanding of task design on SRL.

The interviews conducted with teachers in Chapter 4 were specific to one assessment task, thus limiting teachers' opportunities to elaborate on their assessment practices. However, teachers were encouraged to elaborate on each assessment decision and relied on their general experience in their responses. Future research could use multiple methods of data collection to support interview data. For instance, classroom observations and examining teachers' evaluation practices of student performance could shed perspective on student-teacher interactions during instruction and assessment.

Furthermore, the microanalytic protocol is a fairly new methodological approach for SRL measurement and requires fluency in English, which was the second language for most participants in the current study. As a result, it might have been possible that students were limited in their vocabulary to articulate their thoughts or comprehend the question entirely. The questions for the protocol were designed to ensure the language was student-friendly and un-related examples were provided if students were still confused. It was also an option for students to use their regional language if they wanted to communicate their thoughts better. Future research could consider designing the protocol in English and the participants' preferred language.

Additionally, the protocol is a self-report format and participant responses are taken at face value. It might have been possible that students provided socially desirable responses. Participants were informed that their responses will remain confidential thus encouraging them to provide authentic responses. Future research could include more than one source of data collection for SRL (e.g., self-report questionnaires) to ensure student responses are valid and reliable. While the findings established relationships between SRL and formal classroom assessments, it did not explain any causal relationship. Future research could collect SRL data from the same sample over a period of time for several types of assessment tasks. This approach would lend a deeper understanding of the interactions between assessment structure and SRL processes.

And finally, data for this thesis was collected during a global pandemic which posed time constraints and limited access to students and teachers. All interviews were conducted online, and the temporal dimensions required for the SRL microanalysis protocol were not strictly adhered. Moreover, the whole assessment and learning process was likely influenced by the effects of the pandemic. Students and teachers adjusted to learning and performing assessment at home. This situation could have impacted student and teacher responses to the interview questions. These unusual circumstances limit the extent to which the findings from the current study can be generalized. In India, formal assessments are regularly occurring, however during the pandemic, schools cancelled most assessments to reduce the pressure on students (Dara, 2022). As a result, students had fewer opportunities to engage in learning which may have impacted their regular studying habits, and in turn their responses to the microanalytic questions. These limitations provide scope for further research.

5.6 Directions for Future Research

The thesis highlights limitations in the findings, providing further research to fully understand the complex interaction between SRL and classroom assessments. Indeed, I have already briefly outlined these limitations at the end of the two studies presented in Chapters 3 and 4. Here, I elaborate on the gaps and explore possibilities for exciting and innovative research on SRL and classroom assessments in the future.

Based on the ubiquitous practice of classroom assessments and its impact on SRL processes, the insights from this dissertation provide further evidence for the careful deliberation on how SRL can be promoted and sustained through various assessment processes. The results signify the value of broadening the scope of assessment to include decisions around design, teachers' intentions and instruction that promote strong SRL processes. In particular, teacher education for in-service and pre-service teachers and interventions for students on SRL strategies and how they are relevant to the assessment process will help maximise student learning and performance. The findings from this dissertation indicate the need for a closer examination of intentional assessment design for promoting SRL in the classroom context. Decades of research have documented the type of knowledge that teacher-designed academic tasks demand. There is also copious research on designing assessments with cognitive validity and interpretive validity for assessment data to impact learning positively. There are several recommendations in the literature regarding the theoretical links between classroom assessments and SRL, but there is limited understanding of how these links manifest themselves for formal assessments. This may be partly due to the variance in assessment formats across academic subjects and partly because of the high stakes associated with such assessments. Formal assessments are structured tasks that can contribute to SRL development, and event measures such as the microanalytic protocol are suitable methods to gather and analyse data.

Although the findings in the current study are limited to formal assessments in science, they indicate that task design is a valuable element to consider concerning SRL. Given that students vary SRL across tasks and contexts to optimize learning and performance, there is value in understanding how students respond to diverse assessment tasks with distinct characteristics. In Chapter 3, I suggest that teaching students' task-specific strategies and tracking students' SRL across assessment tasks could improve students' responses. Such a longitudinal analysis will offer deeper insights into how SRL evolves and its impact on student learning and performance. Moreover, the microanalytic protocol as a measurement tool can be modified to suit diverse types of assessments across various subjects. A longitudinal study could also offer insights into trends regarding learning strategies and motivational beliefs, how they vary across tasks and time as well as the understanding of causal influences. In this thesis perceived interest emerged as a strong motivational construct related to performance, other SRL processes, and task features. This insight warrants more research on interest across assessment tasks, subjects, and teachers' instructional practice. It would be interesting to examine the factors that promote interest, such as curiosity and creativity in relation to tasks, instruction, and SRL processes.

A significant challenge I encountered for my thesis was the scarcity of information on characteristics of assessment tasks, including formal assessments, that relate to SRL. There is scattered evidence on the types of academic tasks described by Perry et al. (2020) for primary classroom writing tasks and Lodewyk et al. (2006) for tasks in higher education. As a result, I had to synthesise insights from these studies and refer to older literature, such as the work of Doyle (1984) and Crooks (1980), to develop an integrative framework that defines task characteristics. Indeed, having criteria for assessment tasks related to SRL is helpful. It could be further expanded to include assessment characteristics such as cognitive and interpretive validity. This approach would present a comprehensive assessment framework that considers the design in relation to subject skills and knowledge and how those elements promote SRL. For example, science tasks in India are limited to evaluating students' subject knowledge (i.e., content), but do not capture the skills required for scientific thinking such as reasoning or critical thinking. Therefore, it is essential to develop a framework that helps determine the validity of the assessment task and maps to SRL processes that could promote student learning and performance.

The National Education Policy (2020) of India urges educators and schools across the country to use evidence-based practices rather than import research findings from other countries. Yet there is no substantial evidence available in the Indian context to support this endeavour (Midha, 2022). In my survey of literature, I was challenged to find empirical evidence on student learning and assessment in India to support the rationale of this research. I had to rely on studies such as teachers’ perceptions of assessment policy (Brown et al., 2015), students’ study habits for national exams (Julius & Evans, 2015), and national achievement surveys (NCERT, 2017) to make support the rationale of this research project. Indeed, teachers in India are provided with guidelines and recommendations for learning and assessment based on national policies. Nonetheless, educational research is critical to policy development, teacher professional development and empowering student learning.

There is limited research investigating the impact of teachers' instructional strategies and assessment practices on students' learning and performance outcomes. This thesis offers the first substantial evidence on these aspects in the Indian education context.

Findings from Chapter 4 highlight that classroom assessments pose a challenging and complex environment due to school policy and expectations. Nevertheless, decisions made by teachers impact students' SRL in meaningful ways. More research is needed to understand the impact of intentionally including SRL-related language and objectives in teachers' assessment decisions on students' SRL. A mixed-method design could provide insights into students' thoughts, motivation, and actions regarding teachers' assessment decisions. This approach could facilitate a collaborative process between students and teachers in which learning is a shared responsibility, thus increasing student accountability and autonomy. In Chapter 4, I suggest a checklist for formal assessments that teachers can use when making decisions for classroom assessments. This could be further expanded to include students' involvement in the classroom assessment process. The transition from students' passively reacting and responding to assessments to actively exploring and engaging with their learning processes and assessment data can be encouraged. One way to involve students in the assessment process is by using structured questioning strategies that target SRL processes and build autonomy in the learning process.

Despite its limitations, the microanalytic protocol grounded in sound theory presents itself as a valuable self-diagnosing tool for teachers and students to track learning processes and attitudes over time. I contend that by integrating microanalytic questions, with guidelines for cognitive functions into various tasks, including assessments, teachers and students can intervene as necessary to enhance self-awareness of their learning and encourage more robust SRL practices. For example, before students begin a task, they can be encouraged to identify the learning objective, and consider an appropriate strategy while reflecting on their efficacy, interest, and outcome expectations. Although it might appear as an effortful endeavour, creating time for planning and self-reflection have positive long-term consequences for learning.

5.7 Concluding Thoughts

Throughout this thesis, I have endeavoured to articulate the potential of formal classroom assessment to enhance and sustain students' SRL processes. I have argued that the task design and the decisions behind the designing process impact students' approach and responses to formal assessments. In this concluding Chapter, I have discussed how these findings make original contributions to the field of SRL and assessment literature. In light of this discussion, I draw attention to the immensity of research yet to be explored within the assessment context. The dynamic and varied nature of assessment, especially its placement within policy and the emphasis on performance outcomes, make formal assessments a worthwhile avenue for future research.

Self-regulated learning is naturally occurring for every young person. Along the way, some are in an environment that encourages and sustains effective strategies, while others are in less effective environments. Every assessment presents an opportunity to construct an environment that nurtures successful SRL strategies and attitudes. It thrills me to imagine a future in India where science assessments are less performative and more reflective. A future where every student succeeds on assessments because they know how to strategize their learning, optimize their motivation, and make adaptive decisions. A future where science is perceived as a competency rather than an ability. A future in which a younger me would have not just the fascination and interest to learn, but also the means to endure curiosity.

This thesis exemplifies the untapped potential of formal classroom assessments yet to be discovered in the self-regulated learning literature. It provides meaningful and context-driven knowledge but also hopefully prompts ideas for future research on how SRL is integrated into a broader assessment context that leads to empowering young people for the world outside the classroom.

References

Allal, L. (2020). Assessment and the co-regulation of learning in the classroom. *Assessment in Education: Principles, Policy & Practice*, *27*(4), 332-349. https://doi.org/10.1080/0969594X.2019.1609411

Anderman, E. M., Sinatra, G. M., & Gray, D. L. (2012). The challenges of teaching and learning about science in the 21st century: Exploring the abilities and constraints of adolescent learners. *Studies in Science Education*, *48* (1), 89– 117.

Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research, and Evaluation*, *10*(3). https://doi.org/10.7275/g367-ye94

Andrade, H. L., Du, Y., & Wang, X. (2008). Putting rubrics to the test: The effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing. *Educational Measurement: Issues and Practice*, *27*(2), 3-13. https://doi.org/10.1111/j.1745-3992.2008.00118.x

Andrade, H., & Brookhart, S. M. (2016). The role of classroom assessment in supporting self-regulated learning. In L. Allal & D. Laveault (Eds.), *Assessment for learning: Meeting the challenge of implementation* (pp. 293-309). Springer.

Andrade, H. L., & Brookhart, S. M. (2020). Classroom assessment as the co-regulation of learning. *Assessment in Education: Principles, Policy & Practice*, *27*(4), 350-372. https://doi.org/10.1080/0969594X.2019.1571992

Ananthakrishnan, G. (2019, June 14). *Do exams throttle India's education system?* The Hindu. Do exams throttle India's education system? - The Hindu

Andrzejewski, C. E., Davis, H. A., Shalter Bruening, P., & Poirier, R. R. (2016). Can a self-regulated strategy intervention close the achievement gap? Exploring a classroom-based intervention in 9th grade earth science. *Learning and Individual Differences*, *49*, 85-99. https://doi.org/10.1016/j.lindif.2016.05.013

Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, *36*(5), 258-267. https://doi.org/10.3102/0013189x07306523

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191. https://doi.org/10.1037/0033-295x.84.2.191

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist, 44*(9), 1175–1184. https://doi.org/10.1037/0003-066x.44.9.1175

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*(1), 1-26. https://doi.org/10.1146/annurev.psych.52.1.1

Bearman, M., Dawson, P., Boud, D., Bennett, S., Hall, M., & Molloy, E. (2016). Support for assessment practice: Developing the assessment design decisions framework. *Teaching in Higher Education*, *21*(5), 545-556. https://doi.org/10.1080/13562517.2016.1160217

Bell, B., & Cowie, B. (2001). The Characteristics of Formative Assessment. In B. Bell & B. Cowie (Eds.), *Formative assessment and science education* (pp. 62-79). Springer Dordrecht.

Berry, J., Kannan, H., Mukherji, S., & Shotland, M. (2020). Failure of frequent assessment: An evaluation of India's continuous and comprehensive evaluation program. *Journal of Development Economics*, *143*, 102 - 406. https://doi.org/10.1016/j.jdeveco.2019.102406

Biggs, J. (1998). Assessment and classroom learning: A role for summative assessment?. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 103-110. https://doi.org/10.1080/0969595980050106

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74. https://doi.org/10.1080/0969595980050102

Black, P., Harrison, C., Hodgen, J., Marshall, B., & Serret, N. (2011). Can teachers' summative assessments produce dependable results and also enhance classroom learning? *Assessment in Education: Principles, Policy & Practice*, *18*(4), 451-469. https://doi.org/10.1080/0969594x.2011.557020

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in education: Principles, Policy & Practice*, *25*(6), 551-575. https://doi.org/10.1080/0969594x.2018.1441807

Blasiman, R. N., Dunlosky, J., & Rawson, K. A. (2017). The what, how much, and when of study strategies: comparing intended versus actual study behaviour. *Memory*, *25*(6), 784-792. https://doi.org/10.1080/09658211.2016.1221974

Blumenfeld, P. C., Mergendoller, J. R., & Swarthout, D. W. (1987). Task as a heuristic for understanding student learning and motivation. *Journal of Curriculum Studies*, *19*(2), 135-148. https://doi.org/10.1080/0022027870190203

Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist*, *26*(3-4), 369-398. https://doi.org/10.1207/s15326985ep2603&4_8

Bodas, J., Ollendick, T. H., & Sovani, A. V. (2008). Test anxiety in Indian children: A cross-cultural perspective. *Anxiety, Stress, & Coping*, *21*(4), 387-404. https://doi.org/10.1080/10615800701849902

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology*, *54*(2), 199-231. https://doi.org/10.1111/j.1464-0597.2005.00205.x

Boekaerts, M. (2011). Emotions, emotion regulation, and self-regulation of learning. n B. J. Zimmerman & D. H. Schunk (eds.), *Handbook of self-regulation of learning and performance* (408–425). Routledge

Boekaerts, M. (2017). Cognitive load and self-regulation: Attempts to build a bridge. *Learning and Instruction*, *51*, 90-97. https://doi.org/10.1016/j.learninstruc.2017.07.001

Boud, D., & Falchikov, N. (2007). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education, 31*(4), 399-413. https://doi.org/10.1080/02602930600679050

Brookhart, S. M., & Durkin, D. T. (2003). Classroom assessment, student motivation, and achievement in high school social studies classes. *Applied Measurement in Education*, *16*(1), 27-54. https://doi.org/10.1207/s15324818ame1601_2

Brookhart, S. M. (2004). Classroom assessment: Tensions and intersections in theory and practice. *Teachers College Record: The Voice of Scholarship in Education*, *106*(3), 429-458. https://doi.org/10.1177/016146810410600302

Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, *30*(1), 3-12. https://doi.org/10.1111/j.1745-3992.2010.00195.x

Brookhart, S. M. (2013). The use of teacher judgement for summative assessment in the USA. *Assessment in Education: Principles, Policy & Practice*, *20*(1), 69-90. https://doi.org/10.1080/0969594x.2012.703170

Brookhart, S. M., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, *67*(3), 343-368. https://doi.org/10.1080/00131911.2014.929565

Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016). A century of grading research. *Review of Educational Research, 86*(4), 803-848. https://doi.org/10.3102/0034654316672069

Brown, G. T., Harris, L. R., & Harnett, J. (2012). Teacher beliefs about feedback within an assessment for learning environment: Endorsement of improved learning over student well-being. *Teaching and Teacher Education*, *28*(7), 968-978. https://doi.org/10.1016/j.tate.2012.05.003

Butler, D. L., Schnellert, L., & MacNeil, K. (2015). Collaborative inquiry and distributed agency in educational change: A case study of a multi-level community of inquiry. *Journal of Educational Change*, *16*(1), 1-26. https://doi.org/10.1007/s10833-014-9227-z

Butler, D. L., & Schnellert, L. (2015). Success for students with learning disabilities: What does self-regulation have to do with it? In T. Cleary (Ed.), *Self-regulated learning interventions with at-risk youth: enhancing adaptability, performance, and well-being* (pp. 89–111). American Psychological Association. https://doi.org/10.1037/14641-005

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, *65*(3), 245-281. https://doi.org/10.3102/00346543065003245

Callan, Gregory Lee, "Self-regulated Learning (SRL) Microanalysis for Mathematical Problem Solving: a Comparison of a SRL Event Measure, Questionnaires, and a Teacher Rating Scale" (2014). *Theses and Dissertations*. 557. https://dc.uwm.edu/etd/557

Case, J., & Gunstone, R. (2002). Metacognitive development as a shift in approach to learning: An in-depth study. *Studies in Higher Education*, 27(4), 459-470. https://doi.org/10.1080/0307507022000011561

Chen, P. P., & Bonner, S. M. (2020). A framework for classroom assessment, learning, and self-regulation. *Assessment in Education: Principles, Policy & Practice*, *27*(4), 373-393. https://doi.org/10.1080/0969594x.2019.1619515

Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In *The Cambridge handbook of expertise and expert performance* (pp. 21– 30). Cambridge: Cambridge University Press.

Chin, C., & Chia, L. G. (2004). Implementing project work in biology through problem-based learning. *Journal of Biological Education*, *38*(2), 69-75. https://doi.org/10.1080/00219266.2004.9655904

Cleary, T. J., & Sandars, J. (2011). Assessing self-regulatory processes during clinical skill performance: A pilot study. *Medical Teacher*, *33*(7), e368-e374. https://doi.org/10.3109/0142159x.2011.577464

Cleary, T. J. (2011). Emergence of Self-Regulated Learning Microanalysis. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 329-345). Taylor & Francis.

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*. https://doi.org/10.1155/2012/428639

Cleary, T. J., & Kitsantas, A. (2017). Motivation and self-regulated learning influences on middle school mathematics achievement. *School Psychology Review*, *46*(1), 88-107. https://doi.org/10.1080/02796015.2017.12087607

Cleary, T. J., & Callan, G. L. (2018). Assessing Self-Regulated Learning Using Microanalytic Methods. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 338-351). Routledge. https://doi.org/10.4324/9781315697048

Cleary, T. J., Slemp, J., & Pawlo, E. R. (2020). Linking student self-regulated learning profiles to achievement and engagement in mathematics. *Psychology in the Schools*, *58*(3), 443-457. https://doi.org/10.1002/pits.22456

Cleary, T. J., Slemp, J., Reddy, L. A., Alperin, A., Lui, A., Austin, A., & Cedar, T. (2020). Characteristics and Uses of SRL Microanalysis across Diverse Contexts, Tasks, and Populations: A Systematic Review. *School Psychology Review*, 1-21. https://doi.org/10.1080/2372966x.2020.1862627

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Elsevier Science & Technology.

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education*. Taylor & Francis Group.

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/10.1037/e518362013-265

Crooks, T., & Collins, E. (1986). What do 1st year university examinations assess. *New Zealand Journal of Educational Studies*, *21*(2), 123-132.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, *58*(4), 438-481. https://doi.org/10.3102/00346543058004438

Davis, D. S., & Neitzel, C. (2011). A self-regulated learning perspective on middle grades classroom assessment. *The Journal of Educational Research*, *104*(3), 202-215. https://doi.org/10.1080/00220671003690148

Deb, S., Strodl, E., & Sun, H. (2015). Academic stress, parental pressure, anxiety and mental health among Indian high school students. *International Journal of Psychology and Behavioral Science*, *5*(1), 26-34. https://doi.org/10.5923/j.ijpbs.20150501.04

Deci, E. L., & Ryan, R. M. (2000). The "What" and "Why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227-268. https://doi.org/10.1207/s15327965pli1104_01

DiBenedetto, M. K., & Zimmerman, B. J. (2010). Differences in self-regulatory processes among students studying science: A microanalytic investigation. *The International Journal of Educational and Psychological Assessment*, *5*(1), 2-24. https://doi.org/10.1016/j.lindif.2013.04.004

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., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, *3*(3), 231-264. https://doi.org/10.1007/s11409-008-9029-x

Dignath, C., & Büttner, G. (2018). Teachers' direct and indirect promotion of self-regulated learning in primary and secondary school mathematics classes – insights from video-based classroom observations and teacher interviews. *Metacognition and Learning*, *13*(2), 127-157. https://doi.org/10.1007/s11409-018-9181-x

Dignath, C., & Veenman, M. V. (2021). The role of direct strategy instruction and indirect activation of self-regulated learning—Evidence from classroom observation studies. *Educational Psychology Review*, *33*(2), 489-533. https://doi.org/10.1007/s10648-020-09534-0

Doyle, W. (1983). Academic work. *Review of Educational Research*, *53*(2), 159-199.

Doyle, W., & Carter, K. (1984). Academic tasks in classrooms. *Curriculum Inquiry*, *14*(2), 129-149. https://doi.org/10.1080/03626784.1984.11075917

Duncan, T. G., & McKeachie, W. J. (2005). The making of the motivated strategies for learning questionnaire. *Educational Psychologist*, *40*(2), 117-128. https://doi.org/10.1207/s15326985ep4002_6

Durik, A. M., Vida, M., & Eccles, J. S. (2006). Task values and ability beliefs as predictors of high school literacy choices: a developmental analysis. *Journal of Educational Psychology*, *98*(2), 382-393. https://doi.org/10.1037/0022-0663.98.2.382

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, *41*(10), 1040-1048. https://doi.org/10.1037/0003-066x.41.10.1040

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*(2), 256 - 273. https://doi/10.1037/0033-295X.95.2.256

Dwivedi, R. D. (2011). Effect of Classroom Climate and Parental Awareness on Academic Achievement of Secondary School Students. *Journal of Indian Education*, 75-84.

https://n20.ncert.org.in/pdf/publication/journalsandperiodicals/journalofindianeducation/JIE_NOV2011.pdf#page=77

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

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363-406. https://doi/10.1037/0033-295X.100.3.363

Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In A.M. Williams, A. Kozebelt, K. A. Ericsson, & R. R., Hoffman (Ed.s), *The Cambridge handbook of expertise and expert performance*, (pp 685-705). Cambridge University Press.

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs—principles and practices. *Health Services Research*, *48*(6-2), 2134-2156. https://doi.org./10.1111.1475-6773.12117

Fleming, M., & Chambers, B. A. (1983). Teacher-made tests: Windows on the classroom. *New Directions for Testing & Measurement*, 29–38.

Follmer, D. J., & Sperling, R. A. (2019). Examining the role of self-regulated learning microanalysis in the assessment of learners' regulation. *The Journal of Experimental Education*, *87*(2), 269-287. https://doi.org/10.1080/00220973.2017.1409184

Fraile, J., Panadero, E., & Pardo, R. (2017). Co-creating rubrics: The effects on self-regulated learning, self-efficacy and performance of establishing assessment criteria with students. *Studies in Educational Evaluation*, *53*, 69-76. https://doi.org/10.1016/j.stueduc.2017.03.003

Frederiksen, N. (1984). Implications of cognitive theory for instruction in problem solving. *Review of Educational Research*, *54*, 363–407. https://doi.org/10.1002/j.2330-8516.1983.tb00019.x

García-Pérez, D., Fraile, J., & Panadero, E. (2021). Learning strategies and self-regulation in context: How higher education students approach different courses, assessments, and challenges. *European Journal of Psychology of Education*, *36*(2), 533-550. https://doi.org/10.1007/s10212-020-00488-z

Gardner, J. (2010). An Introduction. In J. Gardner, W. Harlen, & L. Hayward (Eds.), *Developing teacher assessment* (pp. 1-11). McGraw-Hill Education (UK).

Geist, E. (2010). The anti-anxiety curriculum: Combating math anxiety in the classroom. *Journal of Instructional Psychology*, *37*(1), 24-31. Retrieved from: Geist.indd (andrews.edu)

Graham, S., & Harris, K. R. (2003). Students with learning disabilities and the process of writing: A meta-analysis of SRSD studies. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 323-344). The Guilford Press.

Greene, J. A., Hutchison, L. A., Costa, L. J., & Crompton, H. (2012). Investigating how college students' task definitions and plans relate to self-regulated learning processing and understanding of a complex science topic. *Contemporary Educational Psychology*, *37*(4), 307-320. https://doi.org/10.1016/j.cedpsych.2012.02.002

Greene, J.A. (2017). Self-Regulation in Education (1st ed.). Routledge.

Greene, J. A. (2020). Building upon synergies among self-regulated learning and formative assessment research and practice. *Assessment in Education: Principles, Policy & Practice*, *27*(4), 463-476. https://doi.org/10.1080/0969594x.2020.1802225

Greene, J. A. (2021). Teacher support for metacognition and self-regulated learning: A compelling story and a prototypical model. *Metacognition and Learning*, *16*(3), 651-666. https://doi.org/10.1007/s11409-021-09283-7

Guo, W. (2017). The Relationships between Chinese Secondary Teachers' Feedback and Students' Self-Regulated Learning [Unpublished doctoral dissertation]. *The Chinese University of Hongkong.*

Gupta, M., & Mehtani, D. (2017). Type of school, locality and gender as determinants of selfregulated learning among students: An empirical study. *International Journal of Research in Economics and Social Sciences (IJRESS)*, *7*(1), 37-51.

Guskey, T. R. (2009). Mastery learning. In T. L. Good (Ed.), *21st century education: A reference handbook* (pp. 194-202). Sage.

Hadwin, A. F., Winne, P. H., Stockley, D. B., Nesbit, J. C., & Woszczyna, C. (2001). Context moderates students' self-reports about how they study. *Journal of Educational Psychology*, *93*(3), 477-487. https://doi.org/10.1037/0022-0663.93.3.477

Haertel, E. (1986). The valid use of student performance measures for teacher evaluation. *Educational Evaluation and Policy Analysis*, *8*(1), 45-60. https://doi.org/10.3102/01623737008001045

Harlen, W. (2005). Teachers' summative practices and assessment for learning – tensions and synergies. *The Curriculum Journal*, *16*(2), 207-223. https://doi.org/10.1080/09585170500136093

Harris, L. R., & Brown, G. T. (2009). The complexity of teachers' conceptions of assessment: Tensions between the needs of schools and students. *Assessment in Education: Principles, Policy & Practice*, *16*(3), 365-381. https://doi.org/10.1080/09695940903319745

Harris, L. R., Brown, G. T., & Harnett, J. A. (2014). Understanding classroom feedback practices: A study of New Zealand student experiences, perceptions, and emotional responses. *Educational Assessment, Evaluation and Accountability*, *26*(2), 107-133. https://doi.org/10.1007/s11092-013-9187-5

Harris, L. R., Brown, G. T., & Harnett, J. A. (2014). Analysis of New Zealand primary and secondary student peer- and self-assessment comments: Applying Hattie and Timperley's feedback model. *Assessment in Education: Principles, Policy & Practice*, *22*(2), 265-281. https://doi.org/10.1080/0969594x.2014.976541

Harris, L. R., Brown, G. T., & Dargusch, J. (2018). Not playing the game: Student assessment resistance as a form of agency. *The Australian Educational Researcher*, *45*(1), 125-140. https://doi.org/10.1007/s13384-018-0264-0

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research*, *66*(2), 99-136. https://doi.org/10.3102/00346543066002099

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112. https://doi.org/10.3102/003465430298487

Hattie, J. A., & Donoghue, G. M. (2016). Learning strategies: A synthesis and conceptual model. *npj Science Learn*, *1, 16013*. https://doi.org/10.1038/npjscilearn.2016.13

Hayward, L. (2015). Assessment is learning: The preposition vanishes. *Assessment in Education: Principles, Policy & Practice*, *22*(1), 27-43. https://doi.org/10.1080/0969594x.2014.984656

Heritage, M. (2010). *Formative assessment and next-generation assessment systems: are we losing an opportunity?* Council of Chief State School Officers. https://files.eric.ed.gov/fulltext/ED543063.pdf

Hoffman, B., & Spatariu, A. (2008). The influence of self-efficacy and metacognitive prompting on math problem-solving efficiency. *Contemporary Educational Psychology*, *33*(4), 875-893. https://doi.org/10.1016/j.cedpsych.2007.07.002

Irving, S. E., Harris, L. R., & Peterson, E. R. (2011). 'One assessment doesn't serve all the purposes' or does it? New Zealand teachers describe assessment and feedback. *Asia Pacific Education Review*, *12*(3), 413-426. https://doi.org/10.1007/s12564-011-9145-1

Jensen, J. L., McDaniel, M. A., Woodard, S. M., & Kummer, T. A. (2014). Teaching to the test… or testing to teach: Exams requiring higher order thinking skills encourage greater conceptual understanding. *Educational Psychology Review*, *26*(2), 307-329. https://doi.org/10.1007/s10648-013-9248-9

Jonassen, D. H. (1997). Instructional design models for well-structured and III-structured problem-solving learning outcomes. *Educational Technology Research and Development*, *45*(1), 65-94. https://doi.org/10.1007/BF02299613

Julius, M., & Evans, A. S. (2015). Study of the relationship between study habits and academic achievement of students: A case of Spicer higher secondary school, India. *International Journal of Educational Administration and Policy Studies*, *7*(7), 134-141. https://doi.org/10.5897/ijeaps2015.0404

Kane, M. T., & Wools, S. (2020). Perspectives on the validity of classroom assessments. In S. M. Brookhart & J. H. McMillan (Eds.), *Classroom assessment and educational measurement* (pp. 11 - 26). Taylor and Francis.

Karpicke, J. D., Butler, A. C., & Roediger III, H. L. (2009). Metacognitive strategies in student learning: do students practise retrieval when they study on their own?. *Memory*, *17*(4), 471-479. https://doi.org/10.1080/09658210802647009

Kapur, K. (2008). Assessment for Improving Learning in Schools in India: A perspective. Unpublished manuscript, NCERT.

Kitsantas, A., & Zimmerman, B. J. (2010). Comparing self-regulatory processes among novice, non-expert, and expert volleyball players: A microanalytic study. *Journal of Applied Sport Psychology*, *14*(2), 91-105. https://doi.org/10.1080/10413200252907761

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254-284. https://doi.org/10.1037/0033-2909.119.2.254

Koh, K., & Luke, A. (2009). Authentic and conventional assessment in Singapore schools: An empirical study of teacher assignments and student work. *Assessment in Education: Principles, Policy & Practice*, *16*(3), 291-318. https://doi.org/10.1080/09695940903319703

Kuhn, D. (2009). The importance of learning about knowing: Creating a foundation for development of intellectual values. *Child Development Perspectives*, *3*(2), 112-117. https://doi.org/10.1111/j.1750-8606.2009.00089.x

Kuhn, D. (2009). Do students need to be taught how to reason?. *Educational Research Review*, *4*(1), 1-6. https://doi.org/10.1016/j.edurev.2008.11.001

Lavonen, J., & Laaksonen, S. (2009). Context of teaching and learning school science in Finland: Reflections on PISA 2006 results. *Journal of Research in Science Teaching:*

The Official Journal of the National Association for Research in Science Teaching, *46*(8), 922-944. https://doi.org/10.1002/tea.20339

Lodewyk, K. R., & Winne, P. H. (2005). Relations among the structure of learning tasks, achievement, and changes in self-efficacy in secondary students. *Journal of Educational Psychology*, *97*(1), 3 – 12. https://psycnet.apa.org/doi/10.1037/0022-0663.97.1.3

Lodewyk, K. R., Winne, P. H., & Jamieson-Noel, D. L. (2009). Implications of task structure on self-regulated learning and achievement. *Educational Psychology*, *29*(1), 1-25. https://doi.org/10.1080/01443410802447023

Lombardi, D., Sinatra, G. M., & Nussbaum, E. M. (2013). Plausibility reappraisals and shifts in middle school students' climate change conceptions. *Learning and Instruction*, *27*, 50-62. https://doi.org/10.1016/j.learninstruc.2013.03.001

Lombardi, D., Bailey, J. M., Bickel, E. S., & Burrell, S. (2018). Scaffolding scientific thinking: Students' evaluations and judgments during Earth science knowledge construction. *Contemporary Educational Psychology*, *54*, 184-198. https://doi.org/10.1016/j.cedpsych.2018.06.008

Malmberg, J., Järvelä, S., & Kirschner, P. A. (2014). Elementary school students’ strategic learning: does task-type matter?. *Metacognition and Learning*, *9*(2), 113-136. https://doi.org/10.1007/s11409-013-9108-5

McDonald, B., & Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations.

Assessment in Education: Principles, Policy & Practice, *10*(2), 209-220. https://doi.org/10.1080/0969594032000121289

McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice*, *20*(1), 20-32. https://doi.org/10.1111/j.1745-3992.2001.tb00055.x

McMillan, J. H. (2003). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice. *Educational Measurement: Issues and Practice*, *22*(4), 34-43. https://doi.org/10.1111/j.1745-3992.2003.tb00142.x

McMillan, J. H. (2005). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice. *Educational Measurement: Issues and Practice*, *22*(4), 34-43. https://doi.org/10.1111/j.1745-3992.2003.tb00142.x

Meyer, L. H., McClure, J., Walkey, F., Weir, K. F., & McKenzie, L. (2009). Secondary student motivation orientations and standards-based achievement outcomes. *British Journal of Educational Psychology*, *79*(2), 273-293. https://doi.org/10.1348/000709908x354591

Moss, C. M., Brookhart, S. M., & Long, B. A. (2013). Administrators' roles in helping teachers use formative assessment information. *Applied Measurement in Education*, *26*(3), 205-218. https://doi.org/10.1080/08957347.2013.793186

National Council of Educational Research and Training. (2005). *National curriculum framework 2005.* Retrieved from: https://ncert.nic.in/pdf/nc-framework/nf2005-english.pdf

National Council of Educational Research and Training. (2017). *National Achievement Survey: State Reports*, NCERT, New Delhi, https://ncert.nic.in/src.php?ln=

National Council of Educational Research and Training. (2020). *National education policy* (NEP 2020). Retrieved from: NEP_Final_English_0.pdf (education.gov.in)

Neuville, S., Frenay, M., & Bourgeois, E. (2007). Task value, self-efficacy and goal orientations: Impact on self-regulated learning, choice and performance among university students. *Psychologica Belgica*, *47*(1), 95-117. https://doi.org/10.5334/pb-47-1-95

OECD (2019), *PISA 2018 Results (Volume I): What Students Know and Can Do*, PISA, OECD Publishing, Paris, https://doi.org/10.1787/5f07c754-en.

Olakanmi, E. E., & Gumbo, M. T. (2017). The effects of self-regulated learning training on students' metacognition and achievement in chemistry. *International Journal of Innovation in Science and Mathematics Education*, *25*(2), 35-48.

Ormrod, J. E. (2011). *Human learning*. Prentice Hall.

Pajares, F., Schunk, D. H., & Zimmerman, B. J. (2008). Motivational role of self-efficacy beliefs in self-regulated learning. In *Motivation and self-regulated learning: Theory, research, and applications* (pp. 111 - 138). Routledge.

Pallant, J. (2007). *SPSS survival manual: A step by step guide to data analysis using SPSS.* Routledge.

Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS (7th ed.).* Routledge.

Panadero, E., & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. Assessment in *Education: Principles, Policy & Practice*, *21*(2), 133-148. https://doi.org/10.1080/0969594x.2013.877872

Panadero, E., Jonsson, A., & Strijbos, J. W. (2016). Scaffolding self-regulated learning through self-assessment and peer assessment: Guidelines for classroom implementation. In *Assessment for learning: Meeting the challenge of implementation* (pp. 311-326). Springer.

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

Panadero, E., Andrade, H., & Brookhart, S. (2018). Fusing self-regulated learning and formative assessment: a roadmap of where we are, how we got here, and where we are going. *The Australian Educational Researcher*, *45*(1), 13-31. https://doi.org/10.1007/s13384-018-0258-y

Peeters, J., De Backer, F., Kindekens, A., Triquet, K., & Lombaerts, K. (2016). Teacher differences in promoting students' self-regulated learning: Exploring the role of student characteristics. *Learning and Individual Differences*, *52*, 88-96. https://doi.org/10.1016/j.lindif.2016.10.014

Perry, N., & Drummond, L. (2002). Helping young students become self-regulated researchers and writers. *The Reading Teacher*, *56*(3), 298-310. http://www.jstor.org/stable/20205197

Perry, N., Phillips, L., & Dowler, J. (2004). Examining features of tasks and their potential to promote self-regulated learning. *Teachers College Record, 106*(9), 1864-1878.

Perry, N. E., Phillips, L., & Hutchinson, L. (2006). Mentoring student teachers to support self-regulated learning. *The Elementary School Journal*, *106*(3), 237-254. https://doi.org/10.1086/501485

Perry, N. E., Hutchinson, L., & Thauberger, C. (2008). Talking about teaching self-regulated learning: Scaffolding student teachers' development and use of practices that promote self-regulated learning. *International Journal of Educational Research, 47*(2), 97-108. https://doi.org/10.1016/j.ijer.2007.11.010

Perry, N. E., Lisaingo, S., Yee, N., Parent, N., Wan, X., & Muis, K. (2020). Collaborating with teachers to design and implement assessments for self-regulated learning in the context of authentic classroom writing tasks. *Assessment in Education: Principles, Policy & Practice*, *27*(4), 416-443. https://doi.org/10.1080/0969594x.2020.1801576

Pilcher, J. K. (1994). The value-driven meaning of grades. *Educational Assessment*, *2*(1), 69-88. https://doi.org/10.1207/s15326977ea0201_4

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*(1), 33-40. https://doi.org/10.1037/0022-0663.82.1.33

Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research*, *31*(6), 459-470. https://doi.org/10.1016/s0883-0355(99)00015-4

Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, *92*(3), 544 - 555. https://psycnet.apa.org/doi/10.1037/0022-0663.92.3.544

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekarts, P.R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451-502). Academic Press. https://doi.org/10.1016/B978-012109890-2/50043-3

Pintrich, P. R., & Zusho, A. (2002). The development of academic self-regulation: The role of cognitive and motivational factors. In A. Wigfield, J.S. Eccles (Eds.), *Development of Achievement Motivation* (pp. 249-284). Academic Press. https://doi.org/10.1016/B978-0-12-750053-9.X5000-1

Pressley, M., Borkowski, J. G., & Schneider, W. (1987). Cognitive strategies: Good strategy users coordinate metacognition and knowledge. *Annals of Child Development,* (4)*,* 89-129.

Pressley, M., & Hilden, K. (2006). Cognitive Strategies. In D. Kuhn, R. S. Siegler, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Cognition, perception, and language* (pp. 511–556). John Wiley & Sons Inc

Ratnam, T., & Tharu, J. (2018). Integrating assessment into classroom instruction to create zones of development for teachers and learners: Some perspectives from India. *Teacher Learning with Classroom Assessment*, 119-140. https://doi.org/10.1007/978-981-10-9053-0_7

Reynolds-Keefer, L. (2010). Rubric-referenced assessment in teacher preparation: An opportunity to learn by using. *Practical Assessment, Research, and Evaluation*, 15(8), 1-9. https://doi.org/10.7275/psk5-mf68

Roth, W. M. (1994). Experimenting in a constructivist high school physics laboratory. *Journal of Research in Science Teaching*, *31*(2), 197-223. https://doi.org/10.1002/tea.3660310209

Rovers, S. F., Stalmeijer, R. E., van Merriënboer, J. J., Savelberg, H. H., & De Bruin, A. B. (2018). How and why do students use learning strategies? A mixed methods study on learning strategies and desirable difficulties with effective strategy users. *Frontiers in Psychology*, *9*, 2501. https://doi.org/10.3389/fpsyg.2018.02501

Rozencwajg, P. (2003). Metacognitive factors in scientific problem-solving strategies. *European Journal of Psychology of Education*, *18* (3), 281– 294. https://doi.org/10.1007/BF03173249

Sanford, J. P. (1985). *Academic Tasks and Research in Science Teaching (R&D Rep. 6196)*. Research and Development Centre for Teacher Education. The University of Texas at Austin.

Schunk, D. H. (1989). Self-efficacy and achievement behaviors. *Educational Psychology Review*, *1*(3), 173-208. https://doi.org/10.1007/bf01320134

Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist*, *25*(1), 71-86. https://doi.org/10.1207/s15326985ep2501_6

Schunk, D. L., & Ertmer, P. A. (2000). Self-regulation and academic learning: Self-efficacy enhancing interventions. In M. Boekarts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 631-649). Academic Press.

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

Schunk, D.H., & Zimmerman (2007). Motivation. In Schunk, D.H & B.J. Zimmerman (Eds.), *Self-regulated learning: theory, research, and applications* (pp 1 – 30). Routledge. https://doi.org/10.4324/9780203831076

Schunk, D. H., & Usher, E. L. (2013). Social cognitive theory and motivation. In R. M. Ryan (Ed.), *The Oxford handbook of human motivation* (pp. 13-27). Oxford University Press.

Sebesta, A. J., & Bray Speth, E. (2017). How should I study for the exam? self-regulated learning strategies and achievement in introductory biology. *CBE—Life Sciences Education*, *16*(2), ar30. https://doi.org/10.1187/cbe.16-09-0269

Segers, M., Nijhuis, J., & Gijselaers, W. (2006). Redesigning a learning and assessment environment: The influence on students' perceptions of assessment demands and their learning strategies. *Studies in Educational Evaluation*, *32*(3), 223-242. https://doi.org/10.1016/j.stueduc.2006.08.004

Shepard, L. A., Penuel, W. R., & Pellegrino, J. W. (2018). Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. *Educational Measurement: Issues and Practice*, *37*(1), 21-34. https://doi.org/10.1111/emip.12189

Shin, N., Jonassen, D. H., & McGee, S. (2003). Predictors of well-structured and ill-structured problem solving in an astronomy simulation. *Journal of Research in Science Teaching*, *40* (1), 6– 33. https://doi.org/10.1002/tea.10058

Shute, V. J., & Kim, Y. J. (2014). Formative and stealth assessment. In *Handbook of research on educational communications and technology* (pp. 311-321). Springer.

Sinatra, G. M., & Taasoobshirazi, G. (2011). Intentional conceptual change. In *Handbook of self-regulation of learning and performance*, 203-216.

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist*, *50*(1), 1-13. https://doi.org/10.1080/00461520.2014.1002924

Sinatra, G. M., & Taasoobshirazi, G. (2018). The self-regulation of learning and conceptual change in science: Research, theory, and educational applications. In J. A Greene & D.H Schunk (Eds.), *Handbook of self-regulation of learning and performance* (2nd Ed., pp. 170 – 182). Routledge/Taylor & Francis Group.

Singhal, P. (2012). Continuous and comprehensive evaluation: A study of teachers' perception. *Delhi Business Review*, *13*(1), 81-99. dbr_v13n1g.pmd (delhibusinessreview.org)

Singhal, R., & Misra, G. (1994). Achievement goals: A situational-contextual analysis. *International Journal of Intercultural Relations*, *18*(2), 239-258. Retrieved from: PII: 0147-1767(94)90030-2 (umich.edu)

Sivaraman, U. (2012). Assessment Practices at School Stage: The Case of CCE – A School-based Assessment Implemented by the CBSE in its Affiliated Schools. *The Primary Teacher*, *37*(1–2), 25–35. Retrieved from: PTJan-Apr2012.pdf (ncert.nic.in)

Smith, S. N., & Miller, R. J. (2005). Learning approaches: Examination type, discipline of study, and gender. *Educational Psychology*, *25*(1), 43-53. https://doi.org/10.1080/0144341042000294886

Smyth, E., & Banks, J. (2012). 'There was never really any question of anything else': young people's agency, institutional habitus and the transition to higher education. *British Journal of Sociology of Education*, *33*(2), 263-281. https://doi.org/10.1080/01425692.2012.632867

Spiro, R.J., Coulson, R.L., Feltovich, P.J., & Anderson, D.K. (1988). Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains. In V. Patel & G. Groen (Eds), *Tenth annual conference of the Cognitive Science Society* (pp. 375–383). Erlbaum.

Starr, E. J., & Lovett, S. B. (2000). The ability to distinguish between comprehension and memory: Failing to succeed. *Journal of Educational Psychology*, *92*(4), 761-771. https://psycnet.apa.org/doi/10.1037/0022-0663.92.4.761

Sternberg, R. J., & Williams, W. M. (2009). *Educational Psychology* (2nd ed.). New York: Pearson.

Stiggins, R. J. (2005). The unfulfilled promise of classroom assessment. *Educational Measurement: Issues and Practice*, *20*(3), 5-15. https://doi.org/10.1111/j.1745-3992.2001.tb00065.x

Stiggins, R., & Chappuis, J. (2005). Using student-involved classroom assessment to close achievement gaps. *Theory into Practice*, *44*(1), 11-18. https://doi.org/10.1207/s15430421tip4401_3

Subramani, C., & Venkatachalam, J. (2019). Parental expectations and its relation to academic stress among school students. *International Journal of Research and Analytical Reviews* (IJRAR), *6*(2), 95-99. Retrieved from: Parental_Expectations_and_Its_Relation_to_Academic_Stress_among_School_Students20191203-39090-12j9al8-libre.pdf (d1wqtxts1xzle7.cloudfront.net)

Taasoobshirazi, G., & Carr, M. (2009). A structural equation model of expertise in college physics. *Journal of Educational Psychology*, *101* (3), 630 - 643. https://psycnet.apa.org/doi/10.1037/a0014557

Tashakkori, A., & Teddlie, C. (2009). Integrating qualitative and quantitative approaches to research. In L. Bickman & M. D. Rog (Eds.), *The SAGE handbook of applied social research methods* (2nd ed., pp. 283-317). SAGE.

Thomas, J. W., & Rohwer Jr, W. D. (1986). Academic studying: The role of learning strategies. *Educational Psychologist*, *21*(1-2), 19-41. https://doi.org/10.1080/00461520.1986.9653022

Thomas, J. W. (1993). Promoting independent learning in the middle grades: The role of instructional support practices. *The Elementary School Journal*, *93*(5), 575-591. https://doi.org/10.1086/461741

Tuckman, B. W., & Kennedy, G. J. (2011). Teaching learning strategies to increase success of first-term college students. *The Journal of Experimental Education*, *79*(4), 478-504. https://doi.org/10.1080/00220973.2010.512318

Turner, J. C. (1995). The influence of classroom contexts on young children's motivation for literacy. *Reading Research Quarterly,* 410-441. https://doi.org/10.2307/747624

Usher, E. L., & Schunk, D. L. (2018). Social cognitive theoretical perspective of self-regulation. In D. L. Schunk & J. A. Greene (Eds.), *Handbook of Self-Regulation* (2nd ed., pp. 19-35). Routledge.

Van Meter, P., Yokoi, L., & Pressley, M. (1994). College students' theory of note-taking derived from their perceptions of note-taking. *Journal of Educational Psychology*, *86*(3), 323-338. https://doi.org/10.1037/0022-0663.86.3.323

Veenman, M. V. (2011). Alternative assessment of strategy use with self-report instruments: A discussion. *Metacognition and Learning*, *6*(2), 205-211. https://doi.org/10.1007/s11409-011-9080-x

Venkatachalam, K. S. (2017, January 25). Why Does India Refuse to Participate in Global Education Rankings? The Diplomat. Retrieved from https://thediplomat.com/2017/01/why-does-india-refuse-to-participate-in-globaleducation-rankings

Verma, S., Sharma, D., & Larson, R. W. (2002). School stress in India: Effects on time and daily emotions. *International Journal of Behavioral Development*, *26*(6), 500-508. https://doi.org/10.1080/01650250143000454

Wellington, J. (2015). *Educational research: Contemporary issues and practical approaches* (2nd ed.). Bloomsbury Publishing.

Winne, P. H., & Marx, R. W. (1982). Students' and teachers' views of thinking processes for classroom learning. *The Elementary School Journal*, *82*(5), 493-518. https://doi.org/10.1086/461284

Winne, P. H. (1995). Inherent details in self-regulated learning. *Educational Psychologist*, *30*(4), 173-187. https://doi.org/10.1207/s15326985ep3004_2

Winne, P. H., and Hadwin, A. F. (1998). Studying as self-regulated engagement in learning. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in Educational Theory and Practice* (pp 277–304). Erlbaum.

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In B.J Zimmerman & D.H Schunk (Eds), *Handbook of self-regulation* (pp. 531-566). Routledge.

Winne, P. H. (2011). A cognitive and metacognitive analysis of self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp 15-32). Routledge

Winne, P. H., & Hadwin, A. F. (2013). nStudy: Tracing and supporting self-regulated learning in the Internet. In R. Azvedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 293-308). Springer.

Wolters, C. A. (2003). Regulation of motivation: Evaluating an underemphasized aspect of self-regulated learning. *Educational Psychologist*, *38* (4), 189– 205. https://doi.org/10.1207/S15326985EP3804_1

Yagnamurthy, S. (2017). Continuous and comprehensive evaluation (CCE): Policy and practice at the national level. *The Curriculum Journal*, *28*(3), 421-441. https://doi.org/10.1080/09585176.2016.1275725

Zepeda, C. D., Richey, J. E., Ronevich, P., & Nokes-Malach, T. J. (2015). Direct instruction of metacognition benefits adolescent science learning, transfer, and motivation: An in vivo study. *Journal of Educational Psychology*, *107*(4), 954 - 970. https://doi.org/10.1037/edu0000022

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology*, *81*(3), 329-339. https://doi/10.1037/0022-0663.81.3.329

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal*, *29*(3), 663-676. https://doi.org/10.3102/00028312029003663

Zimmerman, B. J. (1995). Self-regulation involves more than metacognition: A social cognitive perspective. *Educational Psychologist*, *30*(4), 217-221. https://doi.org/10.1207/s15326985ep3004_8

Zimmerman, B. J., & Kitsantas, A. (1996). Self-regulated learning of a motoric skill: The role of goal setting and self-monitoring. *Journal of Applied Sport Psychology*, *8*(1), 60-75. https://doi.org/10.1080/10413209608406308

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekarts, P.R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-39). Academic press. https://doi.org/10.1016/B978-0-12-109890-2.X5027-6

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, *41*(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2

Zimmerman, B. J., & Campillo, M. (2003). Motivating self-regulated problem solvers. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 233-262). Cambridge University Press.

Zimmerman, B. J., & Cleary, T. J. (2009). Motives to Self-Regulate Learning. In K. Wentzel, A. Wigfield, & D. Miele (Eds.), *Handbook of Motivation at School* (pp. 247-267). Routledge.

Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 311-328). Routledge.

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist*, *48*(3), 135-147. https://doi.org/10.1080/00461520.2013.794676

Appendix A: SRL Microanalysis Coding and Scoring for a classroom assessment

The coding and scoring guide is developed for this study which aims to measure students' SRL for a classroom science assessment. Each SRL process is categorized using the SRL Microanalysis protocol that is based on Zimmerman's SRL model (Cleary, 2011). Every SRL process is clearly defined with example responses. Following the description of the process is the coding guide, each of which is grounded in sound theoretical frameworks. Every SRL process is also allocated a specific value.

Phase 1: Forethought

This phase includes the following forethought processes: goal-setting, strategy selection

Phase 2: Self-Reflection

This phase will include the following SRL processes: self-judgement, perceived satisfaction, causal attributions (success and failure), and adaptive inferences.

Goal-setting

Question: As you begin getting ready for your assessment, do you have a goal in mind?

Definition. Goals are defined as the 'the object of a person's ambition or effort, an aim or desired result.' Please use the numerical value assigned next to each category as the code.

Process goals*.* Statement indicates a focus on the execution of procedures, or the processes involved in studying for an assessment

Process specific definition (1). Statements that focus on the process of studying for the assessment and also identify the use of specific study strategies, tactics, or procedure as the primary focus of the studying sessions.

- I'll probably solve the problems <u>first and then go to the theory</u>
- I'll <u>make a timetable</u> to study
- I'll <u>mark</u> the important points
- I will <u>read</u> the entire syllabus

Process general definition (2). Statements indicating a focus on a process in general but does not identify any particular procedures. <u>**DO NOT**</u> code **Process General Goals** if the participant has also indicated a <u>**process specific**</u> goal.

- I'll finish the portions
- I want to study well
- I will TRY my best
- I will understand the concepts

- I'll give it my best
- I'll work hard
- I'll work on the numericals/problems

Outcome goals definition. Statement that indicates a focus on performance or an outcome for the assessment

Outcome specific goals definition (3). Statements that identify a clear and measurable outcome as the focus of the studying session

- I want to get to an A/green card/excellence
- I want to get at least 65+; 75+; 80+; etc. on the exam
- I want to get <u>all</u> the numericals/problems correct

Outcome general goals definition (4). Statement identifies as an outcome that is unclear, not quantifiable, or not directly measurable as the focus of the study session.

- I want to do the best I can
- I want to get the highest I can
- I want to do better than last time
- I want a good grade
- I want to get good marks
- I want to do well

Non-task goal definition (5). Statements that indicate a goal that is so incongruent with the current assessment task that the goal reflects inadequate understanding of the task.

- To become a doctor
- To get into college

No goal definition (6). Statement indicates that the student doesn't have a goal for the assessment.

- No
- I don't know
- Not really
- I don't really have a goal

Goal Setting	Process specific	Process general	Outcome specific	Outcome general	Non-task goal	No goal
Ordinal scale	3	3	2	2	1	1

Strategy selection

Question: As you begin getting ready for the assessment, do you have any plans or strategies to help you prepare?

Definition: A range of statements that describes a behaviour and/or cognitive operation that is used for learning. Strategies examined using a two-part rubric: processing level and the breadth level. The first part comprises statements that reflect four levels of processing: rehearsal, comprehension, organization, and integration. The second part of the rubric indicates the number of strategies reported.

Part 1. Processing-level. This section describes the four levels of processing.

Rehearsal: statements that describe encoding learned content to memory. Ultimately, the aim of the learning strategy is to understand or retrieve information.

- Highlighting, Flash cards, mnemonics
- Rehearsal and memorization
- Test practice

Comprehension. Statements that describe learning of subject matter which focuses on understanding the concepts.

- Reading for comprehension
- Note-taking
- Summarizing and paraphrasing

Organization. A statement describing organization or consolidation of ideas or checking of conceptual understanding

- Concept maps, mind maps,
- self-explanation
- Self-questioning

*Integration***.** Statements that describe approaches to learning that deepen their understanding of the concept such as elaboration

- relating content to prior knowledge,
- relating content to real-life situations
- Becoming the teacher

Strategy Selection	Rehearsal	Comprehension	Organization	Integration
Ordinal scale	+1	+2	+3	+2

Breadth of learning strategies. The total number of strategies students reported

Self-evaluation

Definition. A statement reflecting the criteria used in judging of one's own performance. The four categories include: (1) Mastery-focused, (2) Normative comparison, and (3) Social Comparison.

Mastery-focused Definition (1). Statements that describe mastery-focused standards, such as skill, knowledge, or competency

Normative self-comparison Definition (2). Statements that describe evaluating one's performance to a numerical value or grade, but based on how they performed on previous assessments/exams/tests

Normative Social Comparison Definition (3). Statements that describe evaluating one's performance to their peer or classmate.

Normative Comparison Definition (4). Statements that describe evaluating one's performance only the score, regardless of their past performance, or the performance of their peers.

No comparison Definition (5). Statements that describe **no** evaluation criteria. **Only code this if no other category applies.** E.g**., "**I don't care how I do," "It doesn't matter to me," "I don't think about it"

Self-Judgement	Mastery-based	Self-comparison	Normative comparison	Social comparison	No judgement
Ordinal Score	3	3	2	1	1

Perceived Satisfaction

Definition. Statements that reflect whether students are content with their performance scores. Participant responses will be coded into three categories: (a) yes, (b) no, and (c) neutral (e.g., somewhat, and kind of). The value of each are as follows:

Perceived Satisfaction	Yes	Somewhat	No
Ordinal Score	+3	+2	+1

Causal attribution for success and failure

Question: Why do you think you did well (success) or did not do well (failure) in some test items?

Definition. A statement that indicates a plausible reason for success and failure on the questions in the assessment. Responses are coded into 8 categories, based on three levels: locus, controllability, and stability.

Locus. The locus refers to the location of the causal factor: internal (to the individual) or external (others; environment).

Controllability. This construct refers to the volitional control that the individual can exert on the causal factor. In other words, controllability indicates whether the causal factor was subject to the actor's control (controllable) or beyond it (uncontrollable).

Stability. This dimension refers to whether the causal factor remains stable over time or changes (i.e., variable).

Codes. Based on the above three dimensions, eight codes were established for classification. The description of each code is given below. There is one additional code for when students do not report a causal attribution. The code is the numerical value assigned beside each category.

Internal, stable, and controllable (1). Statements that describe causes that are stable over time, ascribed to the individual and are controllable by the individual. E.g., long-term effort

Internal, unstable, controllable (2). Statements that describe the cause to be attributed to the individual are controllable and but unstable over time. E.g., situational/temporary effort, skills/knowledge

Internal, stable, uncontrollable (3). Statements that ascribe the cause to the individual, are stable over time, however uncontrollable. E.g., ability, aptitude

Internal, uncontrollable, and unstable (4). Statements that ascribe cause to the individual, are uncontrollable and unstable over time. E.g., health on the day of the exam, mood

*External, Controllable, and Stable (5)***.** Statements that ascribe cause to sources outside the individual, are controllable, and stable over time. E.g., teacher bias, favouritism

External, controllable, unstable (6). Statements that ascribe cause to sources outside the individual, are controllable, but unstable over time. E.g., help from friends/ teacher

*External, uncontrollable, stable (7)***.** Statements that ascribe cause to sources outside the individual, are uncontrollable, but stable over time. E.g., ease/difficulty of course or exam.

External, uncontrollable, unstable (8). Statements that ascribe cause to sources outside the individual, are uncontrollable, and unstable over time. E.g., luck, chance.

*No attribution (9)***.** Statements that describe no reason for success and failure on the task items. Only code if no other category fits. E.g., I don't know, or I don't think about it

Causal Attribution	Internal, controllable, stable, and unstable	External controllable stable and instable	Internal, uncontrollable, and unstable	External uncontrollable Stable and unstable	No attribution
Ordinal scale	3	2	1	1	1

Adaptive Inferences

Question: If you had another chance to do this assessment, would you do anything differently?

Definition. Statements reflecting an inference based on current performance for a future course of action. Responses may be coded into the four categories: specific inferences, general inferences, situational inferences, and no inference/I don't know.

Specific inferences. Statements that describe a particular change to their study plan or strategies. Sample responses include:

- I'll work on my word problems
- I'll do more practice tests

General inferences. These statements reflect general adaptations. Examples include:

- I'll study harder
- I'll put in more effort

No inferences. Statements that indicate no inferences. For example, "I don't know" or "I won't change anything"

Adaptive Inferences	Specific Inferences	General inferences	No inference	I will not change
Score	3	2	1	1

Appendix B: Coding scheme for Assessment Task

SOLO taxonomy Levels

The SOLO taxonomy classifies learning outcomes into four levels: unistructural, multistructural, relational, and extended abstract. Biggs and Collins use the depth of cognitive processing to distinguish these levels. To code the questions into one of the levels, the coder analyses the verb which indicates the response needed from the student. The following four levels are described in the table below.

Table 44

Description of SOLO Levels with Examples

SOLO Level	Description	Example
Unistructural Level	Focuses on a single element of a concept	Define Match Label Name
Multistructural Level	Refers to multiple elements of a concept	Classify Describe Explain
Relational Level	Multiple aspects of the concept are understood in relation to each other	Predict Compare Contrast Distinguish Apply
Extended Abstract	Refers to the transfer of knowledge, such that the student uses knowledge to extend understanding beyond what is presented on the topic	Reflect Theorize Generate Hypothesize Imagine Evaluate

Nature of Questions

The criterion measuring nature of question comprises specific types of questions: Close-ended questions, open-ended questions, and subject-specific questions.

Table 45

Description for Nature of Questions with Examples

Nature of Question	Description	Example
Close-ended	These questions require students to respond in a single word or phrase	Name, Match, Choose
Open-ended	These questions require students to construct a response in the form of a paragraph	Explain, describe
Subject-specific	These questions differ based on the subject	Draw a diagram, compute

Appendix B: Analytical Framework applied to Assessment F (Science – CBSE)

Q. No.	Topic	Learning outcome	Question Type	SOLO	Familiarity	Weightage
1.	Friction	Identify the relation between mass and friction	MCS/Close-ended	Relational	Unfamiliar	1
2.	Friction	Recognize the type of friction present in a car	MCQ/Close-ended	Relational	Unfamiliar	1
3.	Friction	Recognize that drag refers to friction produced by liquids and gases	MCQ/Close-ended	Unistructural	Familiar	1
4.	Friction	Identify the friction present in a book lying on the table	Short answer/Close ended – Real-life	Relational	Unfamiliar	1
5.	Friction	Identify the friction present in a rock rolling down the hill	Short answer/Close ended – Real-life	Relational	Unfamiliar	1
6.	Friction	Identify the type of friction that helps ice-skating.	Short answer/Close ended – Real-life	Relational	Unfamiliar	1
7.	Friction	Describe two advantages of friction	Short answer/open-ended	Multistructural	Familiar	2
8.	Friction	Explain why vehicle tires have treaded design	Short answer/open-ended – real-life	Relational	Familiar	2
9.	Friction	Explain why bodies of aeroplanes and birds are streamlined	Short answer/open-ended – real-life	Relational	Familiar	2
10.	Pressure	Recognize the term for pressure exerted by air	MCQ/Close-ended	Unistructural	Familiar	1
11	Pressure	Recognize the relation between pressure and depth/height	MCQ/Close-ended	Relational	Familiar	1
12.	Pressure	Recognize the locations in which pressure is least	MCQ/Close-ended	Relational	Unfamiliar	1
13	Pressure	Recall the SI unit of pressure	Short answer/Close ended	Unistructural	Familiar	1
14	Pressure	Recall the name of the scientist who presented the hydraulic principle	Short answer/Close ended	Unistructural	Familiar	1
15	Pressure	Recall the instrument used to measure atmospheric pressure	Short answer/Close ended	Unistructural	Familiar	1
16	Pressure	Distinguish atmospheric pressure based on locations	Short answer/Close ended	Relational	Unfamiliar	1

17	Pressure	Explain the difference in pressure as exerted by water and air.	Short answer/open-ended	Relational	Familiar	2
18	Pressure	Compute pressure using a formula that relates force and area	Computational	Multistructural	Familiar	2
19	Adolescence	Recognize the period of adolescence	MCQ/Close-ended	Unistructural	Familiar	1
20	Adolescence	Recall that sex of the baby is dependent on the chromosomal combination	MCQ/Close-ended	Relational	Familiar	1
21	Adolescence	Recognize the hormones secreted by testes (male reproductive organs)	MCQ/Close-ended	Multistructural	Familiar	1
22	Adolescence	Recognize the process of metamorphosis	MCQ/Close-ended	Unistructural	Familiar	1
23	Adolescence	Recall the number of chromosomes present in the nuclei	True/false – close ended	Multistructural	Familiar	1
24	Adolescence	Apply the understanding of chromosomal combination to determine whose chromosome is responsible for sex determination	True/false – close ended	Relational	Familiar	1
25	Adolescence	Recall that sex of the unborn child depends on the whether the zygote has XX or XY chromosome	Short answer/ close ended	Relational	Familiar	1
26	Adolescence	Explain puberty	Short answer/open -ended	Multistructural	Familiar	2
27	Adolescence	Describe the differences between menarche and menopause	Short answer/open – ended	Multistructural	Familiar	2
28	Adolescence	Explain why young people have acne and pimples during adolescence	Short answer/open – ended (real-life)	Relational	Familiar	4
						40

Rubric: Whole Assessment

Criterion	1	2	3	4	5

Distribution of SOLO	Overall, the items on this assessments are focused on the unistructural and multistructural levels	There are more unistructural and multistructural items than relational and extended abstract	There is a moderate mixture of all four levels of SOLO	There are more relational and extended abstract than unistructural and multistructural items	The majority of items in the assessment are focused on the relational and extended-abstract levels
Familiarity of items	Overall, the items on the items are familiar to the students	There are more familiar items than unfamiliar items	There is a moderate mix of familiar and unfamiliar items in the assessment	There are more unfamiliar items than familiar items	Overall, the items on the items are unfamiliar to the students.
Types of Questions	The items in the assessment are all close-ended requiring one word or phrase	There are more close-ended items than open-ended items	There is a moderate mixture of close and open-ended questions	There are more open-ended items than close-ended items	The items in the assessment are all open-ended requiring students to construct a response
Distribution of sub-topics outcomes	The items in the assessment focus on very few sub-topics from each unit	The items on this assessment unit focus on a few sub-topics from the unit	The items on this assessment focus on some sub-topics from each unit	The items on this assessment focus on most sub-topics from each unit.	The items in this assessment have a good distribution of all sub-topics from each unit
Relevance to rea-life/ application of knowledge to real-life situations	Students are never provided with opportunities to apply their learning to real-life situations	Students are rarely provided with opportunities to apply their learning to real-life situations	Students are occasionally provided with opportunities to apply their learning to real-life situations	Students are often provided opportunities to apply their learning to real-life situations	Students are frequently provided with opportunities to apply their learning to real-life situations

Appendix C: Teacher Interview

Demographical Information:

Name:

Gender:

Ethnicity:

Researcher: Hi *teacher's name*. Thank you for participating in this study. I understand that as a teacher you have a lot of responsibilities, so I appreciate you taking the time to speak with me. Before I begin, this is the information sheet regarding the study. Please let me know if you have any questions. I would like to remind you that this interview will be audio-recorded. Also please be assured that your responses will not be shared with anyone except my supervisors. Do you have any questions for me? *After the teacher reads the information sheet and questions are answered* Great, if you could please just go through this consent form, and sign with your name below? *After teacher hands back the consent form*. Shall we get started? Alright then!

Researcher: Ok to start off with, I'd like to know a little about your teaching experience.

Ask q. 1, 2

1. How many years have you been teaching? How many years have you been teaching at this school? For how long have you been teaching science?

2. Does the school regularly offer professional development workshops for teachers? If yes, how frequently, and what are topics that are usually covered?

Researcher: Perfect. Thank you for sharing that with me. Now, the rest of my questions are going to be around assessments. Please take your time to answer, there is no rush or time limit. If any of the questions are not clear, please feel free to let me know and I can help clarify. Does that sound good?

3. Do you develop/create assessments for your subject? Can you please describe the assessment process for science?
    a. What is the purpose for this assessment?
    b. Do colleagues/peers work with you to develop this assessment?
    c. Does the assessment need approval from head of school/academic co-ordinator?
    d. Why did you choose this format for the assessment?
    e. How did you choose the questions/tasks?
4. Are there strategies/tips you share with your students that help them prepare for their assessments? If yes, please explain.
    a. Is there any reason for these specific tips/suggestions?
    b. How do you think these suggestions/tips help students prepare?
5. How are student responses marked?
    a. What criteria are used to mark responses? Is there a rubric or answer key? Is it always rubric/answer key or does it vary based on the questions?
    b. Is this shared with the students?

   c. Who marks the responses? Are there opportunities for students to self-assess and/or peer assessment?

6. How do students find out the outcome/results of their performance?
   a. Is there a reason for this practice?
7. What do you think a student should do to best prepare for this assessment? or how do you think a student should prepare for this assessment?
8. What does achievement mean to you? What would achievement look like for you on this assessment?

Researcher: That brings us to the end of the interview. Thank you for sharing your all of that information with me. I appreciate your time and effort. I'd just like to remind you that I will send you the transcripts of this conversation via email, and you can check it and let me know if I got anything incorrect. Additionally, if you decide that you no longer want your responses to be a part of this project, please don't hesitate to let me know before <TBD>. After this date it will become difficult for me to remove your data as I will begin my analysis. Once again, thank you so much for being a part of this project.

Appendix D: Teacher Assessment Decisions Coding Scheme

Assessment Purpose

Learning – 1

- Maintain connectivity with the subject matter
- Opportunity to go through the content before moving on

Performance -2

- Evaluate how much students have learned
- Evaluate what they have understood
- Future exam training

Learning and performance – 3

- A combination of the goals mentioned above

Assessment Designer - Type of designer:

By the book designer - 1

The teacher who chooses questions that are from the textbook, or creates questions based on examples from the textbook. This teacher also uses standard resources to select questions, such as old question papers and question banks.

Creatively restricted designer – 2

The teacher who develops their own questions that meet the requirements of the test format and academic standards set forth by the educational board.

Challenging but balanced designer – 3

This teacher creates challenging questions that increase in cognitive depth and complexity. Also, this teacher considers the abilities of all students and therefore strives to achieve a balance between straightforward questions (e.g., define, explain, describe) and higher-order thinking questions (e.g., application, reasoning, problem-solving).

Evaluation Practices

Evaluation/Scoring

The rigid marker – 1

This marker evaluates student work with strict expectations and no leniency.

The subjective marker – 2

This marker evaluates student work based on their past experience and grades student work based on personal expectations for each question.

The 'objective' marker – 3

This marker develops an answer key to evaluate student work. The answer key derived from past experience and based on standards followed in national standardized exams is more objective than the subjective marker because the key is explicit and is used against each student paper.

Feedback practices

Feedback for Performance – 1

The teacher focuses on where and why marks were lost.

Feedback for learning – 2

The teacher focuses on provide + clarify answers.

No feedback provided - 3

Instructions for assessment preparation

Cognitive strategies

- Surface only (Rehearsal and comprehension) - 1
- Deep only (Organization & Elaboration) - 2
- Combination of Surface learning & Deep learning strategies – 3

Behavioural strategies

- Long-term: regular attendance, time management - 1
- Short-term: Sit upright, study in the morning, reduce distractions - 2
- Combination of Long- & Short-term strategies – 3
- No behaviour strategies – 4

Performance strategies – 1

If performance strategies are absent, category – 0

Appendix E: Ethics Approval

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

HUMAN ETHICS COMMITTEE
Secretary, Rebecca Robinson
Telephone: +64 03 369 4588, Extn 94588
Email: human-ethics@canterbury.ac.nz

Ref: 2020/07/ERHEC

7 April 2020

Maansa Bajaj Prakash
School of Educational Studies and Leadership
UNIVERSITY OF CANTERBURY

Dear Maansa

Thank you for providing the revised documents in support of your application to the Educational Research Human Ethics Committee. I am very pleased to inform you that your research proposal "Self-Regulated Learning in the Indian Middle School Context" has been granted ethical approval.

Please note that this approval is subject to the incorporation of the amendments you have provided in your emails of 11^{th} March and 3^{rd}, 4^{th} and 7^{th} April 2020.

Should circumstances relevant to this current application change you are required to reapply for ethical approval.

If you have any questions regarding this approval, please let me know.

We wish you well for your research.

Yours sincerely

pp. R. Robinson

Dr Patrick Shepherd
Chair
Educational Research Human Ethics Committee

Please note that ethical approval relates only to the ethical elements of the relationship between the researcher, research participants and other stakeholders. The granting of approval by the Educational Research Human Ethics Committee should not be interpreted as comment on the methodology, legality, value or any other matters relating to this research.

F E S