TOWARDS AN INTERACTION THEORY OF STUDENT SELF-ASSESSMENT: MAPPING THE COGNITIVE AND AFFECTIVE, METACOGNITIVE, AND EMOTIONAL DIMENSIONS VIA MULTIMODAL DATA

by

NATHAN HARRY RICKEY

A thesis submitted to the Faculty of Education

In conformity with the requirements for

the degree of Doctor of Philosophy

Queen's University

Kingston, Ontario, Canada

(May, 2025)

Abstract

Facilitating student self-assessment across education systems is imperative for enhancing learning for all students, yet there is currently no theoretical framework that articulates how students' actions, cognition and affect, metacognition, and emotions interact within the self-assessment process—a critical barrier to understanding the phenomenon of self-assessment and how it can be supported effectively in classrooms. Building on scholarship which has mapped the self-assessment process as a series of actions and practices, the purpose of this dissertation was to examine and provide initial empirical evidence of the cognitive and affective, metacognitive, and emotional dimensions of the student self-assessment process—and how these interact with each other and students' actions. This dissertation is predicated on multimodal data (i.e., eye tracking, think-alouds) collected while 25 university students engaged in a self-assessment activity. The first study in this multiple manuscript dissertation focused on the cognitive and affective dimension, drawing on joint display integration of eye tracking and think-aloud data to contribute an understanding of the unique cognitive and affective processes students engage in when using resources that are commonly used to facilitate student self-assessment (i.e., rubrics, exemplars, their work). The second study focused on the metacognitive dimension of student self-assessment via inductive analysis of think aloud data, elucidating the metacognitive experiences and metacognitive skills whereby students monitor and control their self-assessment processes. The third study, an investigation of the emotional dimension of student self-assessment, leveraged think-aloud and questionnaire data to contribute initial empirical evidence of the types and intensities of emotions that emerge during self-assessment. State-transition analysis offered insight into how certain types of emotions may be likely to provoke certain self-assessment actions. The dissertation concludes with the advancement of a

novel interaction theory of student self-assessment. This interaction theory advances the first theoretical framework of the student self-assessment process that articulates the interrelations between students' self-assessment actions, cognitive and affective processes, metacognition, and emotions, generating implications for future research and education, providing an initial comprehensive framework of the phenomenon of student self-assessment to guide research and practice.

Co-Authorship

This dissertation follows a manuscript-style format in accordance with the guidelines of the Faculty of Education and the School of Graduate Studies and Postdoctoral Affairs at Queen's University. I would like to acknowledge the contributors for studies in Chapters 2, 3, and 4.

Chapter 2: Rickey, N. H., Panadero, E., DeLuca, C., & Chahine, S. (In preparation). Using rubrics and exemplars to support student self-assessment: Investigating the cognitive and affective processes via joint display integration.

Chapter 3: Rickey, N. H., Panadero, E., & DeLuca, C. (In review). How do students self-assess? Examining the metacognitive processes of student self-assessment.

Chapter 4: Rickey, N. H., Panadero, E., DeLuca, C., & Muis, K. R. (In preparation). Student self-assessment emotions: Investigating the role of emotions in self-assessment.

Acknowledgements

To my supervisor, Chris DeLuca—I am grateful and honoured to work with you. Your mentorship, guidance, and care have made me the researcher I am today.

To my committee members, Saad Chahine and Krista Muis—I could not have asked for a more supportive committee. Thank you for inspiring me to aim high.

To my examining committee members, Yan Zi, Brian Frank, and Ben Kutsyuruba—thank you for your thoughtful engagement with my work and a compelling defence which will guide the next steps in my career.

To my co-authors on the manuscripts included in this dissertation, Chris DeLuca, Ernesto Panadero, Saad Chahine, and Krista Muis—thank you for collaborating with me and making this dissertation possible. Our collaborations continue to hone my research skills.

To my host supervisor in Spain, Ernesto Panadero—your generosity, guidance, and friendship have greatly influenced my research career. Thank you for your ongoing support.

To my mentor, Andrew Coombs—your support, especially early in my research career, was fundamental to me developing key skills to support my career.

To the Social Sciences and Humanities Research Council of Canada—thank you for recognizing the value of my research and for your financial support which made it possible.

To the graduate students who I had the honour of working with during my graduate studies, including members of the Queen's Assessment and Evaluation Group and the Education, Regulated Learning, and Assessment research group—our collaborations make research joyful and continue to teach me how to be part of a supportive research community.

To my family—you inspire me and continue to guide me through life's decisions.

To Beata—you are my center. Thank you for your unwavering love and support.

Table of Contents

List of Figures

List of Tables

Chapter 1

General Introduction

Across education contexts, assessment researchers and policymakers have called on educators to facilitate student self-assessment—i.e., processes whereby students describe and evaluate their own performance to generate self-feedback—in classrooms to enhance student learning in K-12 and post-secondary education (Andrade, 2010; Birenbaum et al., 2015; Black & Wiliam, 2018; British Columbia Ministry of Education, 2023; Nicol & MacFarlane-Dick, 2006; OECD, 2023). These calls are driven by empirically documented benefits of student self-assessment which have been advanced via discourses emphasizing cognitive, socio-cultural, affective and societal benefits. Students' capacity to self-assess has been identified as a top educational factor influencing student learning (Hattie, 2023). Multiple meta-analyses have identified moderate to strong effects of student self-assessment on academic achievement (Lee et al., 2020; Yan et al., 2021; Yan, Lao, et al., 2022). Meta-analyses have also implicated self-assessment as a driver of self-regulated learning, a process whereby students manage their own learning to achieve their goals (Brown & Harris, 2013; Panadero et al., 2017). Additionally, researchers have reported empirical evidence that self-assessment can facilitate creativity and creative thinking (Bolden & DeLuca, 2022; Yan, Lee, et al., 2022). There is strong empirical evidence that student self-assessment has fundamental cognitive benefits for students' learning.

In addition to these cognitive benefits, student self-assessment has been promoted from a socio-cultural framing. Researchers have argued that student self-assessment can promote classroom inclusion by forefronting students' own values in learning and assessment (Bourke & Mentis, 2013). Moreover, researchers have noted that self-assessment which values students'

own learning goals can enable them to take control of their learning and become agentic, lifelong learners (Bourke, 2016; Tan, 2007). Student self-assessment is a powerful mechanism for promoting student voice (Ahmadi, 2022) and can reduce power imbalances between educators and students associated with grading (Nieminen, 2022). Disrupting these power imbalances, which are rooted in colonial education systems, is necessary for decolonizing assessment in Canadian schools: by shifting the central agent of learning from the teacher to the student, student self-assessment can challenge colonial worldviews of education and make space for Indigenous ways of learning (Claypool & Preston, 2011). Given that assessment shapes students' identities (Nieminen, 2024), student self-assessment presents an opportunity for students to make sense of how their performance in schools influences their self-concept.

Self-assessment may also promote essential affective learning outcomes. A meta-analysis in medical education found that self-assessment was more strongly linked with affective outcomes, such as self-efficacy and learning satisfaction, than cognitive ones (Sitzmann et al., 2010). By showing students that they have control of their learning, self-assessment may enhance their self-efficacy which is needed to motivate students to mobilize cognitive resources in future learning tasks (Ross, 2006). Additionally, Panadero et al. (2023) found that providing students with rubrics during self-assessment elicited more positive emotional responses during self-assessment than instructor feedback, indicating that self-assessment can facilitate positive emotional experiences of learning. Affect and emotions fundamentally shape self-regulated learning, student behaviour, and wellbeing (Boekaerts, 2010; Fiedler & Beier, 2014; Muis et al., 2018), so the affective benefits of student self-assessment are critical.

More recently, advocacy to cultivate student self-assessment in education systems has been embedded in international societal improvement discourses advanced by the Organization

for Economic Cooperation and Development (OECD). These discourses, which aim to map the future of education to address international societal crises, have promoted the development of students' global competencies, such as self-regulated learning, creativity, and communication. The OECD identifies these global competencies as essential priorities for education systems because students need them to address rapidly developing global crises, such as climate change and societal inequities (OECD, 2023). Self-assessment, as a fundamental subprocess of self-regulated learning and learner agency (Bourke, 2016; Panadero, 2017), is essential for fostering the global competencies, such as creativity which is predicated on agency and self-regulation (OECD, 2018, 2021). In particular, the OECD identifies "Taking Responsibility" as a fundamental competency that includes the "capacity to consider the consequences of one's actions, evaluate risk and reward, and accept accountability" (OECD, 2019, p. 86)—a competency that necessitates self-assessment, going beyond mere self-appraisal and requiring students to deeply reflect on their actions and their outcomes. This societal discourse is supported by empirical evidence that exemplary integrated Science Technology Engineering and Mathematics (STEAM) teachers rely on student self-assessment to facilitate transdisciplinary learning targeting the global competencies (DeLuca et al., 2024; Dubek et al., 2021; Rickey et al., 2023).

Given the cognitive, socio-cultural, affective, and potential societal benefits of student self-assessment, it is unsurprising that educators across contexts have long been called on to facilitate student self-assessment in their classrooms. Facilitating student self-assessment is a fundamental goal for education systems globally to support students' success and wellbeing in K-12 and postsecondary education—and to address the dynamic and rapidly developing challenges facing society today. This dissertation addresses a fundamental challenge for student

self-assessment theory and practice with the aim of promoting the use of student self-assessment in schools.

The Challenge for Facilitating Student Self-Assessment in Schools

Despite international calls to facilitate student self-assessment in schools, researchers across education contexts have repeatedly found that student self-assessment is among the least used assessment practices by both educators and students (Brandmo et al., 2020; Johnson et al., 2019; Kippers et al., 2018; Veugen et al., 2021; Volante & Beckett, 2011). In these studies, researchers have identified a lack of guidance on how to support student self-assessment as a primary reason educators and students are not widely engaging student self-assessment in classrooms. Empirical studies have indicated that this trend is pronounced in university education, where instructors' lack of training in using self-assessment and negative prior experiences with student self-assessment have been identified as primary barriers (Panadero et al., 2014; Panadero, Fraile, et al., 2019)—despite self-assessment being an essential skill for succeeding in university education (Tai et al., 2018; Usher, 2020). These studies signal the need for guidance which can inform pedagogical approaches to fostering student self-assessment across education contexts and developmental stages (Brown & Harris, 2014).

With the aim of developing such guidance, researchers have studied the effects of various ways educators facilitate student self-assessment. In her critical review of the self-assessment literature, Andrade (2019) pointed out that these efforts have largely been conducted from behaviorist paradigms, where researchers have studied how certain inputs (e.g., providing students with rubrics) relate to certain learning outputs (e.g., academic achievement, self-assessment accuracy, self-regulated learning) primarily via (quasi-)experimental research designs. These behaviourist approaches have identified widely varied effects of similar

approaches to student self-assessment across studies—from strong to negative effect sizes—as documented in various meta-analyses and reviews (Andrade, 2019; Brown & Harris, 2013; To et al., 2021; Yan, Lao, et al., 2022). For example, in their meta-review of self-assessment in K-12 contexts, Brown and Harris (2013) found that using rubrics to support student self-assessment had both the highest and lowest effect sizes on academic achievement of the studied self-assessment interventions. The authors concluded that students' self-assessment training—i.e., how teachers pedagogically supported students in using rubrics to self-assess—likely explained much of this variance. The wide variance in these learning effects offers little guidance to teachers because the impacts of the studied approaches are unpredictable. However, these studies point to the need for research that moves beyond behaviourist paradigms and comprehensively describes how students are engaging in self-assessment to inform pedagogical and tailored approaches.

Currently, such guidance is lacking because of a fundamental gap in the research literature: the self-assessment process is not yet fully understood. Specifically, the ways in which students are thinking and feeling during self-assessment—the primary mechanisms whereby self-assessment inputs (e.g., providing rubrics) are translated into learning outcomes (e.g., enhanced academic achievement)—are understudied. These mechanisms may be the key to making sense of variable effects of self-assessment on learning, illuminating why the similar self-assessment interventions work in some instances and not others (Andrade, 2019). Researchers have argued that understanding what students do, think, and feel when they are self-assessing is essential for informing pedagogical approaches and tailored interventions (Brown & Harris, 2014; Panadero et al., 2016). Mapping the cognitive and affective process that comprises student self-assessment can elucidate key challenges students are facing and opportunities for educators and the students

themselves to intervene. Perhaps most importantly, understanding the student self-assessment process is requisite to understanding how to develop students' self-assessment processes and skills so that they can self-assess in ways that progress their learning. This dissertation therefore builds on emerging lines of research which are addressing these calls (Panadero et al., 2024; Yan & Brown, 2017) to develop comprehensive descriptions of the student self-assessment process to inform practice and future research.

Dissertation Purpose, Structure, and Research Questions

The purpose of this dissertation was to examine and provide initial empirical evidence of the cognitive and affective, metacognitive, and emotional dimensions of the student self-assessment process. This multiple manuscript dissertation includes three interrelated research studies based on the same dataset that address this purpose. These study manuscripts comprise Chapters 2-4 and each includes its own introduction, literature review, and method sections; Chapter 1 is a general introduction that outlines the overall context and purpose underpinning the dissertation. Chapters 2-4 each focus on one dimension of the self-assessment process.

Chapter 2 focuses on the cognitive and affective dimension, contributing an understanding of the unique cognitive and affective processes students engage in when using resources that are commonly used to facilitate student self-assessment. This study addressed one research question:

RQ2.1. What are the cognitive and affective processes of students using rubrics, exemplars, and their essays to self-assess their writing?

Chapter 3 presents a study of the metacognitive dimension of student self-assessment. This study contributes the first empirical evidence of self-assessment metacognition, which

refers to the ways students monitor and control their own self-assessment processes. This study addressed two research questions:

RQ3.1. What metacognitive experiences shape students' self-assessment processes?

RQ3.2. What metacognitive skills do students employ during self-assessment?

Chapter 4 focuses on the emotional dimension of student self-assessment, contributing initial empirical evidence of the types and intensities of emotions that emerge during self-assessment, and how certain types of emotions may be likely to provoke certain self-assessment actions. This study addressed three research questions:

RQ4.1. What self-assessment actions do students employ during self-assessment?

RQ4.2. What types of emotions do students experience during self-assessment, and how frequently and intensely do they experience these emotions?

RQ4.3. Do certain emotions increase the likelihood of a transition into a subsequent self-assessment action?

The dissertation concludes with the advancement of a novel interaction theory of student self-assessment in Chapter 5. This interaction theory advances the first theoretical framework of the student self-assessment process that articulates the interrelations between students' self-assessment actions, cognitive and affective processes, metacognition, and emotions, generating implications for future research and education. This framework is predicated on a synthesis of the contributions of the studies included in this dissertation and relevant theoretical and empirical literature. It builds directly on Panadero et al.'s (2024) self-feedback, model which articulates the primary actions that comprise student self-assessment, by elucidating interactions between self-assessment actions, cognition and affect, metacognition, and emotions, providing an initial

comprehensive framework of the phenomenon of student self-assessment to guide research and practice.

Overall Research Design and Data Collection Procedure

The three studies in this dissertation were based on the same data sources and overall research design and data collection procedure. Data were analyzed in different ways and using different theoretical lenses to address the aims and research questions of each study. The methods used for each study are described in detail in each of the chapters. Here, I provide a broad overview of the data collection procedure which underpinned the three studies and explain some key methodological choices to orient the reader to the data sources.

To collect data for this dissertation, I used a concurrent parallel mixed methods design to collect qualitative and quantitative concurrently and then mix them (Creswell & Clark, 2018). I leveraged a multimodal approach to data collection, wherein I collected multiple channels of online data (i.e., collected during learning) to inform analyses of each other, with each channel accessing a different dimension of the self-assessment experience (Järvelä et al., 2021). Online data were collected while students engaged in a self-assessment activity: eye tracking (i.e., observed dimension), gaze-cued retrospective think-alouds (i.e., reported dimension), and electrodermal activity (i.e., physiological or felt dimension). I also collected offline data (i.e., generated outside of, before, or after learning) in the form of online questionnaires, administered via Qualtrics, and a semi-structured interviews.

Participants were English-speaking international students (N=25) studying at a private university in Spain. Full demographic information is provided in each of the dissertation chapters. I recruited the participants and collected data for this dissertation while completing a Mitacs-funded research internship (i.e., Mitacs Globalink Graduate Research Award) under joint

supervision of Professors Christopher DeLuca and Ernesto Panadero. I recruited international university students for three reasons. First, they represented an extremely diverse group of students, allowing me to develop self-assessment initial theory based on a wide range of academic and socio-cultural backgrounds. Second, as university students living and studying abroad, they had demonstrated strong prior academic achievement and, as effective learners, were likely to have the reflective skills to support an ideal model of the student self-assessment process. Third, the international students represented a convenient population as they were accessible through institutional connections and had strong English speaking and writing skills needed to participate in the study. Because the sample was small and diverse, with low numbers of students in each cultural and academic group, I did not examine between-participant differences in this dissertation. Rather, the three studies and final theoretical chapter are exploratory in nature, contributing an initial interaction theory of student self-assessment supported by a diverse participant sample and rich data. Further research is needed to examine key between-participant differences, such as how culture influences students' self-assessment processes. Additionally, I recruited a relatively small sample size because I used a wide variety of rich data sources to understand students' self-assessment processes in depth. Mixed multimodal data are incredibly fine grained and time-consuming to analyze. The focus on depth of understanding and theory development outweighed the need to generalize at this stage of my research agenda, pointing me to a modest sample size.

The general data collection procedure is illustrated in Figure 1. This procedure involved participants participating in two data collection sessions. The first session, i.e., Writing, was focused on the participants producing a reflective essay which would be used as the object of self-assessment for the study. In a study lab, participants first reviewed and signed the Letter of

Information/Consent Form. I reminded the participants about the purpose of the study: to understand their self-assessment processes and emotions they experienced during self-assessment. Participants then completed an online demographics questionnaire. Collecting demographics information during this first session reduced the chances that emotions brought up as a result of the demographics questionnaire would influence the emotions felt in the self-assessment activity, which occurred at a later date and time. Last, working on a computer in a Microsoft Word document, participants completed a short reflective essay on their experiences living and studying abroad. The essay prompt, which is stated in the chapter manuscripts, was provided on a written instruction sheet so that the essay instructions were consistent across participants. The instruction sheet also included a list of writing criteria to indicate to participants criteria against which they could self-assess their essays in the self-assessment activity, to be completed in the second session. Participants spent 20-60 minutes writing their essays.

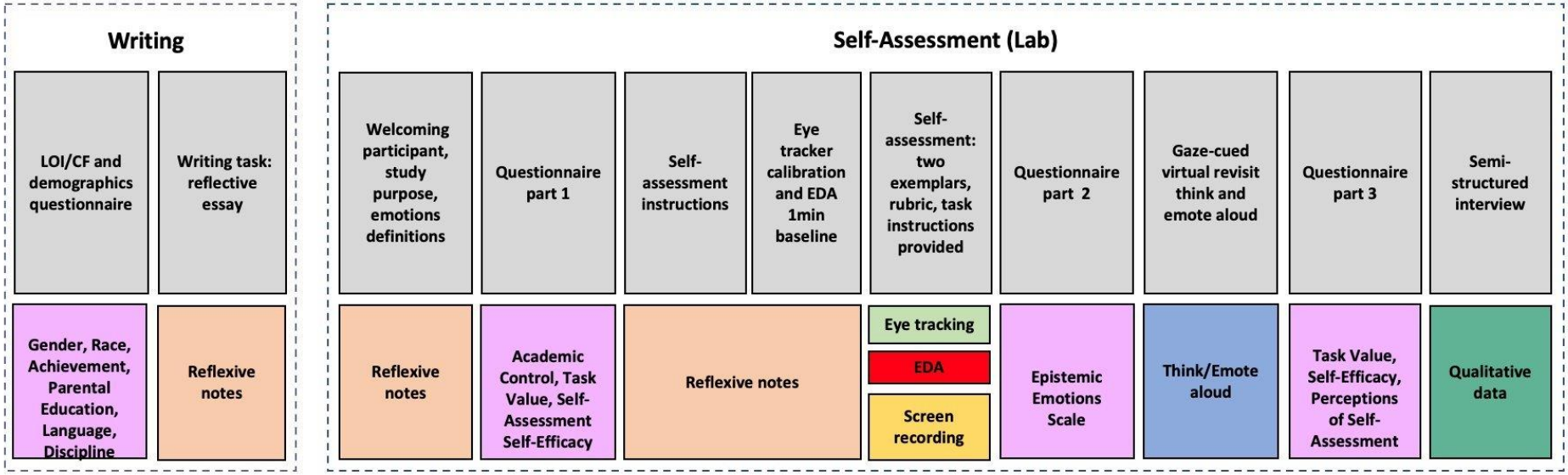
In the second data collection session, i.e., Self-Assessment, each participant returned to the study lab at a later date and time to engage in a self-assessment activity based on their reflective essays. I reminded them of the study purpose, provided definitions of key epistemic emotions to facilitate consistent understandings, and trained participants to think aloud. Working on the lab computer, participants then completed the first part of the second online questionnaire, providing data on their initial states such as their perceptions of academic control, self-assessment task value, and self-assessment self-efficacy. Next, I provided the instructions for the self-assessment activity and then calibrated the eye tracker and established a baseline electrodermal activity measure (i.e., one minute at rest). The self-assessment activity is described in detail in each study chapter. Importantly, it focused on participants deeply engaging with their work with the aim of generating self-feedback and optionally revising their work. Participants

could open numerous resources onscreen to support this process: their essay (with tracked-changes enabled to capture revisions), a validated analytic writing rubric, two contrasting writing exemplars, a blank document for generating written self-feedback, and the essay task instruction sheet. After completing the self-assessment activity, participants returned to the questionnaire and completed the Epistemic Emotions Scale (Pekrun et al., 2017). Next, they engaged in a gaze-cued think-aloud protocol, watching the entire screen recording of their self-assessment process including integrated eye tracking (i.e., real-time gazepath) and verbalized everything they were doing, thinking, and feeling (see study chapters for full details). Participants then completed the third and final part of the online questionnaire, reporting on their perceptions of self-assessment task value, self-assessment self-efficacy, and outcomes of the self-assessment task itself. Last, participants engaged in a semi-structured interview focused on generating deeper insights into emotions and ideas that emerged during the think-alouds.

Importantly, not all the data sources described here were used in this dissertation. I only selected and used the data needed to address each of the research questions. The semi-structured interview, screen recording, and electrodermal activity data will be analyzed and reported on in future studies. In this dissertation, the interview data were used to clarify and understand the think-aloud data.

Figure 1

Data Collection Procedure Overview



Significance

Understanding the cognitive and effective, metacognitive, and emotional dimensions of student self-assessment is critical for facilitating student self-assessment in K-12 and university education. As explained in the beginning of this chapter, self-assessment can be a powerful pedagogy with essential cognitive, affective, socio-cultural, and societal benefits. This dissertation builds on process-oriented conceptions of student self-assessment which are uniquely positioned to inform pedagogical approaches and tailored interventions for supporting student self-assessment in classrooms: by understanding self-assessment as a process, researchers and educators can develop pedagogical approaches to support productive student self-assessment processes and interventions at appropriate points in the process. I understand student self-assessment as a process comprised of actions, thoughts, and feelings which students engage to generate self-feedback on their own performance (Panadero et al., 2016). While self-assessment can be initiated independently by students in and outside of classrooms, the process can also be explicitly facilitated and taught by educators in classrooms (Panadero & Alonso-Tapia, 2013). Research examining the process of student self-assessment has elucidated the practices, actions, and criteria students use to self-assess (Panadero et al., 2024; Yan & Brown, 2017), providing a strong foundation for this dissertation to advance self-assessment theory and practice.

This dissertation builds on this research foundation by examining dimensions of the self-assessment process that have not yet been fully explored. Previous studies have highlighted the importance of a particular self-assessment action—i.e., using external resources (e.g., rubrics, exemplars, practice assessments) to generate internal feedback—as central to effective self-assessment processes (Nicol, 2021; Panadero et al., 2024; Yan, 2020); however, how students

process these resources is not yet understood, meaning that generating guidance for supporting students in using them effectively is implausible. Examining the cognitive and affective processes of students as they use rubrics and exemplars advances self-assessment theory by providing a more comprehensive understanding of a particularly important self-assessment action, and supports practice by highlighting opportunities for educators to intervene and scaffold productive uses of rubrics and exemplars (see Chapter 2). Additionally, a metacognitive dimension of student self-assessment has been theorized (Panadero, Lipnevich, et al., 2019), but research has yet to generate empirical understandings of the metacognitive experiences and skills that shape the self-assessment process. Illuminating these metacognitive experiences and skills articulates the ways in which students monitor and control their own self-assessment processes—an essential step for supporting students in developing their own self-assessment processes (see Chapter 3). Moreover, self-assessment has been theorized to be a highly emotional process (Baca & Smith, 2018; Belkaid et al., 2017; Sitzmann et al., 2010), but researchers have yet to empirically identify the emotions students experience during self-assessment and how these may influence students' actions. Elucidating the types, intensities, and consequences of emotions during self-assessment can expand understandings of a core aspect of self-assessment and enable educators to help students regulate and respond to emotions during self-assessment (see Chapter 4). Synthesizing these dimensions in a novel theoretical framework of the self-assessment process can articulate the interrelations between them, providing a foundation for future research to expand and refine understandings of the phenomenon of self-assessment and the generation of educational implications (see Chapter 5). Together, the studies in this dissertation contribute some of the first empirical evidence illuminating critical elements of cognitive and affective,

metacognitive, and emotional dimensions of student self-assessment, generating implications for theory and practice.

Key Terms

Throughout this dissertation, my co-authors and I use several key terms to situate this research. Here, I provide definitions for these key terms as they are used in the context of this dissertation to signpost fundamental concepts. We expand on their definitions and theoretical foundations in each of the following chapters.

Student Self-Assessment: A psychosocial process involving actions, thoughts, and feelings whereby students generate self-feedback on their own performance (Panadero et al., 2016)

Classroom Assessment: Assessment processes and practices initiated at the classroom level—i.e., by teachers and students—with the primary purpose of advancing student learning (McMillan, 2013)

Cognitive and Affective Processes: Distinct patterns of thinking that may involve an affective dimension (e.g., feelings, emotions, mood) and that require cognitive resources (Fiedler & Beier, 2014)

Metacognition: A form of cognition that involves a learner thinking about their own thinking (Winne, 2017)

Emotions: Multi-componential coordinated processes of physiological subsystems including affective, cognitive, motivational, and physiological experiences (Pekrun, 2006)

Self-Regulated Learning: A process whereby learners manage their thinking, emotions, behaviours, and environments in pursuit of goals (Panadero et al., 2017)

References

Ahmadi, R. (2022). Students' perceptions of student voice in assessment within the context of Iran: The dynamics of culture, power relations, and student knowledge. *Higher Education Research and Development*, *41*(2), 211–225. https://doi.org/10.1080/07294360.2021.1882401

Andrade, H. L. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In H. L. Andrade & G. Cizek (Eds.), *Handbook of formative assessment* (pp. 90–105). Routledge.

Andrade, H. L. (2019). A critical review of research on student self-assessment. *Frontiers in Education*, *4*, 87. https://doi.org/10.3389/feduc.2019.00087

Baca, L., & Smith, S. R. (2018). The role of self-reflection and self-assessment in the psychological assessment process. In S. R. Smith & R. Krishnamurthy (Eds.), *Diversity-sensitive personality assessment* (pp. 3–24). Routledge.

Belkaid, M., Cuperlier, N., & Gaussier, P. (2017). Emotional metacontrol of attention: Top-down modulation of sensorimotor processes in a robotic visual search task. *PLoS ONE*, *12*(9). https://doi.org/10.1371/journal.pone.0184960

Birenbaum, M., DeLuca, C., Earl, L., Heritage, M., Klenowski, V., Looney, A., Smith, K., Timperley, H., Volante, L., & Wyatt-Smith, C. (2015). International trends in the implementation of assessment for learning: Implications for policy and practice. *Policy Futures in Education*, *13*(1), 117–140. https://doi.org/10.1177/1478210314566733

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

Boekaerts, M. (2010). The crucial role of motivation and emotion in classroom learning. In H. Dumont, D. Istance, & F. Benevides (Eds.), *The nature of learning: Using research to inspire practice* (pp. 91–111). The Center for Educational Research and Innovation, Organization for Economic Co-operation and Development.

Bolden, B., & DeLuca, C. (2022). Nurturing student creativity through assessment for learning in music classrooms. *Research Studies in Music Education*, *44*(1), 273–289. https://doi.org/10.1177/1321103X211054793

Bourke, R. (2016). Liberating the learner through self-assessment. *Cambridge Journal of Education*, *46*(1), 97–111. https://doi.org/10.1080/0305764X.2015.1015963

Bourke, R., & Mentis, M. (2013). Self-assessment as a process for inclusion. *International Journal of Inclusive Education*, *17*(8), 854–867. https://doi.org/10.1080/13603116.2011.602288

Brandmo, C., Panadero, E., & Hopfenbeck, T. N. (2020). Bridging classroom assessment and self-regulated learning. *Assessment in Education: Principles, Policy and Practice*, *27*(4), 319–331. https://doi.org/10.1080/0969594X.2020.1803589

British Columbia Ministry of Education. (2023). *A Framework for Classroom Assessment*.

Brown, G. T. L., & Harris, L. (2013). Student self-assessment. In J. H. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 367–393). Sage.

Brown, G. T. L., & Harris, L. (2014). The future of self-assessment in classroom practice: Reframing self- assessment as a core competency. *Frontline Learning Research*, *2*(1), 22–30. https://doi.org/10.14786/flr.v2i1.24

Claypool, T. R., & Preston, J. P. (2011). Redefining learning and assessment practices impacting Aboriginal students: Considering Aboriginal priorities via Aboriginal and Western worldviews. *In Education*, *17*(3), 84–95.

Creswell, J. W., & Clark, V. L. P. (2018). Choosing a mixed methods design. In J. W. Creswell (Ed.), *Designing and conducting mixed methods research* (3rd ed., pp. 53–106). Sage.

DeLuca, C., Dubek, M., & Rickey, N. H. (2024). How hermeneutics can guide grading in integrated STEAM education: An evidence-informed perspective. *British Educational Research Journal*, *50*, 2163-1280. https://doi.org/10.1002/berj.3979

Dubek, M., DeLuca, C., & Rickey, N. (2021). Unlocking the potential of STEAM education: How exemplary teachers navigate assessment challenges. *Journal of Educational Research*, *114*(6), 513–525. https://doi.org/10.1080/00220671.2021.1990002

Fiedler, K., & Beier, S. (2014). Affect and cognitive processes in educational contexts. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 36–55). Routledge.

Hattie, J. (2023). *Visible learning: The sequel: A synthesis of over 2,100 meta-analyses relating to achievement*. Routledge.

Järvelä, S., Malmberg, J., Haataja, E., Sobocinski, M., & Kirschner, P. A. (2021). What multimodal data can tell us about the students' regulation of their learning process? *Learning and Instruction*, *72*, 101203. https://doi.org/10.1016/j.learninstruc.2019.04.004

Johnson, C. C., Sondergeld, T. A., & Walton, J. B. (2019). A study of the implementation of formative assessment in three large urban districts. *American Educational Research Journal*, *56*(6), 2408–2438. https://doi.org/10.3102/0002831219842347

Kippers, W. B., Wolterinck, C. H. D., Schildkamp, K., Poortman, C. L., & Visscher, A. J. (2018). Teachers' views on the use of assessment for learning and data-based decision making in classroom practice. *Teaching and Teacher Education*, *75*, 199–213. https://doi.org/10.1016/j.tate.2018.06.015

Lee, H., Chung, H. Q., Zhang, Y., Abedi, J., & Warschauer, M. (2020). The effectiveness and features of formative assessment in US K-12 education: A systematic review. *Applied Measurement in Education*, *33*(2), 124–140. https://doi.org/10.1080/08957347.2020.1732383

McMillan, J. (2013). Why we need research on classroom assessment. In J. McMillan (Ed.), *SAGE Handbook of Research on Classroom Assessment* (pp. 3–16). Sage.

Muis, K. R., Chevrier, M., & Singh, C. A. (2018). The role of epistemic emotions in personal epistemology and self-regulated learning. *Educational Psychologist*, *53*(3), 165–184. https://doi.org/10.1080/00461520.2017.1421465

Nicol, D. (2021). The power of internal feedback: Exploiting natural comparison processes. *Assessment and Evaluation in Higher Education*, *46*(5), 756–778. https://doi.org/10.1080/02602938.2020.1823314

Nicol, D., & MacFarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. https://doi.org/10.1080/03075070600572090

Nieminen, J. H. (2022). Disrupting the power relations of grading in higher education through summative self-assessment. *Teaching in Higher Education*, *27*(7), 892–907. https://doi.org/10.1080/13562517.2020.1753687

Nieminen, J. H. (2024). How does assessment shape student identities? An integrative review. *Studies in Higher Education*. https://doi.org/10.1080/03075079.2024.2334844

OECD. (2018). *Preparing our youth for an inclusive and sustainable world: The OECD PISA global competence framework*. https://www.oecd.org/pisa/Handbook-PISA-2018-Global-Competence.pdf

OECD. (2019). *OECD future of education and skills 2030: OECD learning compass 2030: A series of concept notes*. https://www.oecd.org/education/2030-project/contact/OECD_Learning_Compass_2030_Concept_Note_Series.pdf

OECD. (2021). *PISA 2021 Creative Thinking Framework*.

OECD. (2023). *Building the future of education 2023*.

Panadero, E., & Alonso-Tapia, J. (2013). Student self-assessment: Theoretical and practical connotations. When it happens, how it is acquired, and how to develop it in our students. *Electronic Journal of Research in Education Psychology*, *11*(30), 551–576. https://doi.org/10.14204/ejrep.30.12200

Panadero, E., Brown, G., & Courtney, M. (2014). Teachers' reasons for using self-assessment: A survey self-report of Spanish teachers. *Assessment in Education: Principles, Policy and Practice*, *21*(4), 365–383. https://doi.org/10.1080/0969594X.2014.919247

Panadero, E., Brown, G. T. L., & Strijbos, J. W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review*, *28*(4), 803–830. https://doi.org/10.1007/s10648-015-9350-2

Panadero, E., Fernández, J., Pinedo, L., Sánchez, I., & García-Pérez, D. (2024). A self-feedback model (SEFEMO): secondary and higher education students' self-assessment profiles.

Assessment in Education: Principles, Policy and Practice, *31*(3–4), 221–253. https://doi.org/10.1080/0969594X.2024.2367027

Panadero, E., Fraile, J., Fernández Ruiz, J., Castilla-Estévez, D., & Ruiz, M. A. (2019). Spanish university assessment practices: Examination tradition with diversity by faculty. *Assessment and Evaluation in Higher Education*, *44*(3), 379–397. https://doi.org/10.1080/02602938.2018.1512553

Panadero, E., García-Pérez, D., Fernández Ruiz, J., Fraile, J., Sánchez-Iglesias, I., & Brown, G. T. L. (2023). Feedback and year level effects on university students' self-efficacy and emotions during self-assessment: Positive impact of rubrics vs. instructor feedback. *Educational Psychology*, 1–24. https://doi.org/10.1080/01443410.2023.2254015

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, *22*, 74–98. https://doi.org/10.1016/j.edurev.2017.08.004

Panadero, E., Lipnevich, A., & Broadbent, J. (2019). Turning self-assessment into self-feedback. In *The Impact of Feedback in Higher education: Improving assessment outcomes for learners* (pp. 147–163). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-25112-3_9

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Pekrun, R., Vogl, E., Muis, K. R., & Sinatra, G. M. (2017). Measuring emotions during epistemic activities: The Epistemically-Related Emotion Scales. *Cognition and Emotion*, *31*(6), 1268–1276. https://doi.org/10.1080/02699931.2016.1204989

Rickey, N., Dubek, M., & Deluca, C. (2023). Toward a praxis-oriented understanding of student self-assessment in STEAM Education: How exemplary educators leverage self-assessment. *Cambridge Journal of Education*, *53*(5), 605–625. https://doi.org/10.1080/0305764X.2023.2196245

Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research, and Evaluation*, *11*, 10. https://doi.org/10.7275/9wph-vv65

Sitzmann, T., Ely, K., Brown, K. G., & Bauer, K. N. (2010). *Self-Assessment of knowledge: A cognitive learning or affective measure? 9*(2), 169–191.

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education*, *76*(3), 467–481. https://doi.org/10.1007/s10734-017-0220-3

Tan, K. (2007). Conceptions of self-assessment: What is needed for long-term learning? In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education* (pp. 114–127). Taylor & Francis Group. http://ebookcentral.proquest.com/lib/queen-ebooks/detail.action?docID=292915

To, J., Panadero, E., & Carless, D. (2021). A systematic review of the educational uses and effects of exemplars. *Assessment and Evaluation in Higher Education*. https://doi.org/10.1080/02602938.2021.2011134

Usher, N. (2020). Developing evaluative judgement in higher education. Assessment for knowing and producing quality work. *Assessment in Education: Principles, Policy & Practice*, *27*(6), 707–710. https://doi.org/10.1080/0969594X.2020.1850419

Veugen, M. J., Gulikers, J. T. M., & den Brok, P. (2021). We agree on what we see: Teacher and student perceptions of formative assessment practice. *Studies in Educational Evaluation*, *70*, 101027. https://doi.org/10.1016/j.stueduc.2021.101027

Volante, L., & Beckett, D. (2011). *Formative assessment and the contemporary classroom: Synergies and tensions between research and practice*. *34*(2), 239–255. https://doi.org/10.2307/canajeducrevucan.34.2.239

Winne, P. H. (2017). Cognition and metacognition within self-regulated learning. In P. A. Alexander, D. H. Schunk, & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 36–48). Routledge. https://doi.org/10.4324/9781315697048.ch3

Yan, Z. (2020). Self-assessment in the process of self-regulated learning and its relationship with academic achievement. *Assessment and Evaluation in Higher Education*, *45*(2), 224–238. https://doi.org/10.1080/02602938.2019.1629390

Yan, Z., & Brown, G. T. L. (2017). A cyclical self-assessment process: Towards a model of how students engage in self-assessment. *Assessment and Evaluation in Higher Education*, *42*(8), 1247–1262. https://doi.org/10.1080/02602938.2016.1260091

Yan, Z., Lao, H., Panadero, E., Fernández-Castilla, B., Yang, L., & Yang, M. (2022). Effects of self-assessment and peer-assessment interventions on academic performance: A pairwise and network meta-analysis. *Educational Research Review*, *37*, 100484. https://doi.org/10.1016/j.edurev.2022.100484

Yan, Z., Lee, J. C.-K., Hui, S. K. F., & Lao, H. (2022). Enhancing students' self-efficacy in creativity and learning performance in the context of English learning: The use of self-

assessment mind maps. *Frontiers in Psychology*, *13*, 871781. https://doi.org/10.3389/fpsyg.2022.871781

Yan, Z., Wang, X., Boud, D., & Lao, H. (2021). The effect of self-assessment on academic performance and the role of explicitness: A meta-analysis. *Assessment & Evaluation in Higher Education*, *48*(1), 1–15. https://doi.org/10.1080/02602938.2021.2012644

Chapter 2

Using Rubrics and Exemplars to Support Student Self-Assessment: Investigating the Cognitive and Affective Processes via Joint Display Integration

Abstract

Despite being a global priority for education, cultivating self-assessment remains a challenge for educators and students across contexts. The purpose of this study was to advance self-assessment theory by investigating students' cognitive and affective processes while they self-assessed their writing with the support of rubrics and exemplars. We collected eye tracking data from 25 university students as they engaged in a self-assessment task and optionally revised their reflective essays on a computer, with on-screen access to a writing rubric and two exemplar essays of contrasting quality. Immediately after the self-assessment task, participants engaged in a gaze-cued think and emote aloud protocol, generating data on their cognitive and affective processes throughout the self-assessment process which we analyzed qualitatively using a hybrid approach. Using joint display, we integrated fixation durations across five self-assessment resources (rubric, two exemplars, essay, self-feedback) with codes from the think aloud data—i.e., cognitive and affective processes—that emerged when participants were looking at each resource. Rubrics and exemplars provoked distinct types of comparison and affect. When reviewing their essays, participants prioritized making important revisions while being distracted by minor writing errors. The self-feedback space elicited self-forgiveness and reflections on the role of bias. This study advances self-assessment theory and practice by elucidating what students are thinking and feeling when they use typical resources to self-assess and revise their writing. Our findings support implications for practice by signaling pedagogical strategies to help students manage and develop their self-assessment processes effectively.

Introduction

The capacity to self-assess and revise performance is foundational to self-regulated learning and learner agency (Andrade, 2010; Bourke, 2016). Students' ability to self-assess can therefore support their academic achievement throughout their schooling and help them set and pursue their personal and professional goals beyond school (Tai et al., 2018; Yan et al., 2023). Formative assessment researchers have long highlighted the importance of cultivating student self-assessment across all years of education to support self-regulated learning and achievement (e.g., Black & Wiliam, 2018; Nicol & MacFarlane-Dick, 2006; Stobart, 2008). However, despite being a global priority for education, studies have continually shown that student self-assessment is one of the least used formative assessment practices across education contexts (Brandmo et al., 2020; Johnson et al., 2019; Panadero, Fraile, et al., 2019; Veugen et al., 2021). Teachers and students in these studies consistently report a lack of guidance on supporting students in developing their self-assessment processes, reflecting a key gap in the literature: researchers have yet to fully illuminate what students are doing, thinking, and feeling while self-assessing (Andrade, 2019; Brown & Harris, 2014; Panadero et al., 2016), meaning that supporting students in developing their self-assessment processes via tailored classroom interventions and pedagogical approaches is implausible.

The present study sought to address a key gap in self-assessment research by examining the cognitive and affective processes of students using rubrics and exemplars to self-assess and revise their writing. Using external indicators of success, namely rubrics and exemplars, is a common self-assessment practice (Lipnevich et al., 2023). However, how students process these resources is understudied. Investigating these cognitive and affective processes can inform guidance for educators on supporting students with using rubrics and exemplars to self-assess

and improve their performance. With the aim of advancing self-assessment theory, this study addressed one research question: *What are the cognitive and affective processes of students using rubrics, exemplars, and their essays to self-assess their writing?*

Defining Student Self-Assessment: A Process View

While self-assessment has been described as a skill underpinning self-regulated learning—i.e., a processes wherein learners manage their cognition, affect, motivation, and behaviour in pursuit of their goals (e.g., Boekaerts, 1991)—it is increasingly also viewed as a process involving thoughts, emotions, and actions that learners use to describe, and possibly evaluate, the quality of their performance (Panadero et al., 2016). This process-oriented view is especially helpful for developing pedagogical approaches to supporting self-assessment, as illuminating the self-assessment process can enable the development of tailored interventions and pedagogical approaches for improving how students self-assess. With the aim of informing pedagogical interventions, we adopt this understanding of student self-assessment as a process. Our research is predicated on the assumption that understanding the cognitive and affective mechanisms involved in this process can elucidate why and how self-assessment can have positive learning benefits (Andrade, 2019), supporting the development of tailored approaches to teaching student self-assessment.

Recognising the importance of illuminating the self-assessment process, researchers have begun to empirically identify the practices, strategies, and criteria that students use during self-assessment (Panadero et al., 2024; Yan & Brown, 2017). Based on qualitative interviews with Master of Education students focused on how they self-assess, Yan and Brown (2017) introduced a cyclical model of the self-assessment process comprised of four self-assessment practices: seeking external feedback via monitoring (e.g., completing a practice test, comparing

performance to a rubric or exemplar), seeking external feedback through inquiry (e.g., asking a peer or teacher for feedback or help), seeking internal feedback (e.g., referring to gut reactions or emotions), and self-reflection (e.g., reflecting on the product or learning process to inform a judgment about strengths or weaknesses). Central to this model were self-directed feedback seeking practices via internal (e.g., gut reactions, emotions) and external (e.g., practice assessments, rubrics, exemplars) resources which inform self-reflection and calibration of performance. Yan and colleagues have expanded and validated this model using questionnaire methods in various education contexts (Yan, 2018, 2020a; Yan et al., 2020; Yang et al., 2023). In a study of the relationships between these self-assessment practices and phases of self-regulated learning, Yan (2020b) found, of the four self-assessment practices, students' use of external indicators of success (e.g., rubrics, exemplars) in the performance phase of task completion was the highest predictor of achievement on the task, signaling the importance of students referring to external indicators to support self-assessment during learning tasks. Using external indicators of quality during the performance phase of task completion appeared to be a particularly powerful self-assessment practice, likely because it enabled students to revise and adjust their performance before submitting their work.

Panadero and colleagues, taking another approach, have sought to map how self-assessment unfolds as a process in real time by capturing think aloud data from K-12 and university students engaged in self-assessment (Panadero et al., 2020, 2022, 2024). Their research culminated in a self-feedback model involving three key phases. First, information processing involves students reading resources (e.g., their work, the task instructions, rubrics, exemplars) and recalling details about their performance. Second, analysis involves drawing on internal (e.g., gut reactions) and external (e.g., teacher) criteria, comparing performance to

indicators of quality (e.g., rubrics, exemplars, past performance), and redoing or revising work (e.g., redoing a math problem to confirm a response, revising an essay to see if it improves the essay). Third, self-assessment judgement involves making an overall evaluation of performance, possibly by drawing on the previous phases. This process emphasizes the important role that students' own work and external indicators of quality, such as rubrics and exemplars, play in enabling students to process information (e.g., reading their essays, reading a rubric) and analyze their work (e.g., drawing on external criteria from a rubric) to inform self-judgements. Further, by identifying and analyzing underlying profiles of self-assessment, Panadero et al. (2024) argued that the primary distinguishing feature of advanced self-assessment processes—the most sophisticated profile—was students' deep engagement in the analysis phase, namely by comparing their work to external criteria and redoing or revising work to (dis)confirm their performance and inform self-feedback. For example, redoing a math problem allowed students to confirm the correctness of their answer, and thinking about possible revisions an essay enabled students to understand how their work could be improved. Redoing or revising work may engage students in thinking more deeply about their work leading to enhanced self-assessment and learning and is therefore an important part of the process (Andrade, 2019).

Developed via different methodologies, these lines of research both implicate drawing on external feedback in the form of indicators of quality performance (e.g., rubrics, exemplars) as a particularly important action within the self-assessment process. Additionally, they implicate redoing or revising as an important part of the self-assessment process, rather than an end goal or separate action, that can be effectively supported using external indicators of quality. Last, these self-assessment process models highlight the importance of students having access to their work (e.g., given the prevalence of reading; Panadero et al., 2024) and generating self-feedback as part

of the process. However, research on facilitating student self-assessment with resources, such as rubrics and exemplars, indicates that how students engage with these resources shapes the outcomes of self-assessment, warranting further investigation of how students engage in these self-assessment actions.

Supporting Student Self-Assessment via Rubrics and Exemplars

While students may initiate self-assessment independently, educators can help students develop their self-assessment processes by explicitly guiding them in self-assessment and providing them with tools to help them self-assess (Panadero & Alonso-Tapia, 2013). A common way that educators help students self-assess is to provide them with external indicators of success criteria, namely rubrics and/or exemplars, to help them contrast their work with these criteria to accurately assess and revise their work (Lin-Siegler et al., 2015; To et al., 2021). Rubrics can provide students with disaggregated descriptors of performance related to success criteria across various levels of performance. Exemplars can provide students with a holistic understanding of levels of performance and include concrete exemplars of success criteria, thereby supplementing rubrics. However, providing students with rubrics or exemplars to support self-assessment does not guarantee positive learning benefits. Meta-analytic reviews have identified widely varying effects on achievement associated with the use of rubrics or exemplars to support self-assessment (Andrade, 2019). The use of rubrics to support self-assessment in K-12 contexts has been associated with both the highest and lowest effect sizes on achievement (Brown & Harris, 2013), and a similar trend has been observed for exemplar use in higher education (To et al., 2021). Much of this variation is likely explained by teachers' pedagogical approaches to engaging students with rubrics and exemplars (e.g., students' training; Brown & Harris, 2013). Therefore, to inform pedagogical approaches that facilitate effective use of rubrics and exemplars during

self-assessment, researchers must first elucidate the cognitive and affective processes students are engaged in when using these resources to self-assess (Andrade, 2019).

Supporting this notion, research has indicated that rubrics and exemplars influence the magnitude of cognitive load that students experience during self-assessment, suggesting that self-assessment with rubrics and exemplars is fundamentally shaped by cognitive and affective processes because such processes require cognitive resources and therefore increase cognitive load. Students who used success criteria to self-assess have reported higher cognitive load than students who self-assessed without criteria, possibly because interpreting success criteria requires deeper cognitive processing (Fastré et al., 2012). Contrasting this finding, another experimental study found that students who used rubrics to self-assess experienced lower cognitive load and were more accurate than those who self-assessed without rubrics (Krebs et al., 2022). Similarly, students who received feedback in the form of a rubric during self-assessment have reported significantly increased positive emotions and decreased negative emotions compared to those who received only instructor feedback or a combination of rubric and instructor feedback (Panadero et al., 2023). Supporting the notion that multiple self-assessment resources may enhance cognitive load via deeper cognitive processing, an experimental study involving 200 high school students found that the combined use of rubrics and exemplars was less beneficial for students' writing performance than rubrics or exemplars alone, a finding the authors attributed to the high cognitive load needed to use multiple self-assessment resources (Lipnevich et al., 2023). The observed variance in cognitive load—which may shape the learning outcomes of self-assessment—across studies signals the need to understand the cognitive and affective processes students may engage in when using rubrics and exemplars in order to inform ways to support effective rubric and exemplar use.

Cognitive and Affective Processes

These findings about cognitive load are important because, together, they indicate that the use of rubrics and exemplars provoke cognitive and affective processes which require and direct cognitive resources, thereby influencing cognitive load (Efklides, 2017). This cognitive load may in turn affect the learning outcomes of engaging in self-assessment using rubrics and exemplars (e.g., Fastré et al., 2012; Lipnevich et al., 2023). However, what students are thinking and feeling as they use rubrics and exemplars to self-assess—i.e., the cognitive and affective processes themselves—are understudied, meaning that supporting students in using rubrics and exemplars to self-assess and revise their performance is implausible.

In this study, we intentionally examined cognitive and affective processes rather than just cognitive processes for two reasons. First, although cognition and affect are determined by distinct brain architectures, they are highly linked in operation (Altmann et al., 2012; Muja & Appelbaum, 2012). Affective states (e.g., emotions, moods) are multicomponential and include a (meta)cognitive component. For example, a student experiencing anxiety experiences the physiological effects of anxiety alongside thoughts about the anxiety and the reasons they are experiencing it. There is a bi-directional relationship between cognition and affect, where thoughts can provoke affective states, and vice-versa (Fiedler & Beier, 2014). Metacognitive experiences, such as feelings of difficulty, often follow or precede an affective state (e.g., frustration) and provoke cognition about the feeling of difficulty (e.g., noting processing disfluency as a heuristic for self-assessment; Efklides, 2017; Reber & Greifeneder, 2017). Therefore, it can be difficult to operationally separate cognition and affect as these processes are deeply interrelated. Second, in terms of self-assessment, affect can enforce biases about oneself via perceptual distortion and illogical interpretation (Muja & Appelbaum, 2012). Additionally,

Sitzmann et al. (2010) found, via a meta-analysis, that self-assessment was more closely associated with affective than cognitive outcomes, reinforcing that there is a strong connection between self-assessment and affect. Affective states may therefore be just as important as cognitive ones in shaping students' self-assessment processes and should be examined.

Multimodal Data

Elucidating students' cognitive and affective processes as they use rubrics and exemplars requires capturing students' thinking as they are engaged in self-assessment. Researchers have therefore explicitly called for the use of online measures and multimodal data to examine self-assessment processes (To et al., 2021; Yan et al., 2020). Online measures collect data when learning is happening; examples include think aloud protocols, eye tracking, digital trace data, and physiological measures. They contrast offline measures which collect data before or after learning (e.g., interviews, questionnaires; Chevrier et al., 2019). Online measures are uniquely capable of capturing fleeting thoughts and feelings that characterize a learning process—thoughts and feelings that learners often forget or which elude learners' awareness (Winne et al., 2019). Research has demonstrated that online measures can inform more granular and veridical understandings of learning process (Beheshitha et al., 2015; Callan & Cleary, 2018; Schellings et al., 2013). Therefore, in this study, we used online measures to capture students' cognitive processes unobtrusively and in real time.

When using online measures, researchers have stressed the value of using multimodal data to inform justified interpretations of cognitive processes (Hadwin et al., 2007; Winne, 2018). Multimodal approaches combine multiple channels of online data, each of which can inform interpretations of the others. Eye tracking can illuminate what visual resources a learner is cognitively processing; think alouds can reveal what learners are thinking when processing

specific visual resources (Järvelä et al., 2021). In the present study, we combined eye tracking and think aloud data using joint display integration to investigate what visual resources students attend to during self-assessment and what they are thinking (i.e., cognitive processes) while looking at them. Joint display is an approach to data integration which aligns and quantitative and qualitative data to inform meta-inferences—i.e., insights generated from comparing and analyzing qualitative and quantitative findings together (Creswell, 2021). By displaying quantitative eye tracking data (i.e., gaze durations) with think aloud data, we elucidated the differential cognitive and affective processes underpinning students' use of rubrics, exemplars, their essays, and self-generated feedback.

Method

To examine students' cognitive processes as they used resources to self-assess, we employed a convergent parallel mixed methods design. This approach involved collecting qualitative and quantitative data concurrently and then subsequently mixing them primarily through joint displays to inform inferences (Creswell & Clark, 2018). Ethics clearance for this study was granted by institutional research ethics boards. All participants actively consented to participate in the study prior to data collection.

Participants

We recruited 25 English-speaking international students from multiple courses at a university in Spain to participate in this study. The majority of participants identified as female (*n*=16), followed by males (*n*=8) and one was another category. Most (*n*=18) of the participants' ages ranged from 20-30, with the remaining being older than 30. Participants were studying in a range of disciplines, with largest representations from the social sciences (*n*=6), engineering (*n*=4), humanities (*n*=3), and business (*n*=3); other disciplines represented were sciences,

languages, communications, design, and energy. The sample was diverse, including participants from 15 different countries across four continents: Australia, Belgium, Canada, Croatia, France, Georgia, Germany, Ghana, Hungary, Iran, Netherlands, Nigeria, Pakistan, Philippines, South Korea, Turkey, Ukraine, United Kingdom, United States of America, and Puerto Rico. Most participants were white (*n*=15) and the rest were Black (*n*=2), Asian/Asian American (*n*=2), Latinx (*n*=2), or did not know or preferred not to identify (*n*=4). Ten of the participants (40%) spoke English as a first language—the largest language group—while the rest had various other first languages but were proficient and studying in English.

Capturing Cognitive and Affective Processes

All data were collected in a controlled lab setting to standardize the conditions of data collection. We captured participants' cognitive and affective processes while they self-assessed their own writing using eye tracking (Stark et al., 2018) and a gaze-cued combined think and emote aloud protocol (Chevrier et al., 2019; Craig et al., 2008; McIntyre et al., 2022). Eye tracking provided insights into what visual resources participants attended to and for how long, indicating what resources they were cognitively processing (Genc Aksaray & Ozcelik, 2022; Holmqvist et al., 2023). We used a Tobii Pro Fusion screen-based eye tacking system (sampling rate 250Hz) to collect eye tracking data, including gaze coordinates, fixations, gazepaths, saccades, and pupil dilation (Tobii Pro AB, 2014b). To optimize the validity of interpretations of eye tracking data, we instructed participants to wear no makeup during the data collection and we ensured the lab environment, including the lighting, equipment positions, and monitor temporal and spatial resolution, were consistent across participants.

While eye tracking can indicate what visual resources participants are likely processing, think alouds can provide insights into what participants are thinking about and feeling when

looking at visual resources (Chevrier et al., 2019). We used a retrospective think aloud protocol, despite their reliance on participants' ability to recall information, to avoid imposing extraneous cognitive load on research participants. Researchers have found that, when participants are engaged in complex cognitive tasks such as self-assessment, verbalizations produced by concurrent think alouds were of lower quality and lower frequency than to those elicited by cued retrospective ones (Beach & Willows, 2017; van Gog et al., 2005). Since we were interested in understanding the sources of students' cognitive load when self-assessing, it was important to avoid creating extra cognitive load by asking students to think aloud concurrently while self-assessing. We included a gaze cue (i.e., participants' real time gazepath) to provide a fine-grained visual reminder of what aspects of each resource participants were looking at. Additionally, retrospective think alouds which included a gaze cue have been shown to be the optimal form of think aloud for helping effective learners understand their own self-assessment process (Kostons et al., 2009).

Data Collection Procedure

Pre-Writing and Demographics

The data collection procedure occurred in two parts. First, participants came individually to a study lab to prepare a writing sample to self-assess. Participants completed a demographics questionnaire which included variables related to eye tracking (e.g., use of eyeglasses). We provided participants with a standardized instruction sheet to ensure participants received consistent instructions and allotted time. Participants were asked to respond to the prompt: "Write an essay in English about your experiences living and studying in a new country." The instruction sheet included suggested writing prompts and a list of criteria participants would later use to self-assess their writing. Participants were given 20-30 minutes to write their essays.

Student Self-Assessment Task Instructions

Later, participants returned to the study lab to self-assess their essays for the second part of the data collection procedure. First, we reminded participants of the purpose of the study and provided training on the think aloud protocol. Participants then completed a questionnaire to provide data on their initial states (e.g., academic control, task value, self-assessment self-efficacy). Data from the questionnaire were used for a separate study and are not reported on here. Next, we provided instructions for the self-assessment task. We instructed participants to, using a study computer, review their essays and try to provide themselves with written self-feedback and/or revise their essays. We showed participants where they could open on the computer screen several resources to help them self-assess and revise their writing: a validated writing analytic rubric, two contrasting exemplars (i.e., mastering and developing levels), the task instructions, a blank document for writing self-feedback or making notes, and their own essay (tracked changes enabled). The writing rubric was adapted from a validated writing assessment included in the Diagnostic Assessment and Achievement of College Skills (Bryer et al., 2025). Participants received scripted instructions for the self-assessment to ensure consistency.

Since researchers have identified multiple different self-assessment profiles (Panadero et al., 2024) and a wide variety of ways that students and educators can engage student self-assessment (e.g., Brown & Harris, 2013; Panadero et al., 2016), we did not standardize the self-assessment process. For example, we did not standardize what resources participants should use and for how long. Instead, we viewed participants as agentic learners (Winne et al., 2019) and allowed them to self-assess however they liked, using whatever resources they wanted to for any amount of time they wanted in order to more closely replicate an authentic learning context and

enhance ecological validity, given our aim to inform practice.

Equipment Calibration and Eye Tracking

Once participants' questions were answered, we calibrated the eye tracking system to optimize the validity of eye tracking data. We repeated calibration until we achieved acceptable average validation results across all calibration targets (i.e., > 80%). During calibration, we helped the participants establish a comfortable and sustainable head position and asked them to maintain that head position throughout the self-assessment process once the eye tracker was calibrated. Tobii's I-VT Fixation Filter algorithm accounts for natural head movements in identifying fixations; however, we asked and reminded participants to maintain a consistent head position throughout the self-assessment to minimize data loss resulting from head movements (Tobii Pro AB, 2014b).

Once the eye tracker was calibrated, we reminded participants of the self-assessment instructions and participants then completed the self-assessment task. Participants' self-assessment processes ranged in length from approximately 7 to 45 minutes. Each participant's self-assessment process was screen-recorded and eye tracking was recorded using the Tobii Pro Fusion screen-based eye tracking system. Eye tracking data were concurrently integrated into the screen recoding files using Tobii Pro Lab software (version 1.181; Tobii Pro AB, 2014a), allowing the gaze path to be visualized in the screen recordings for the gaze cued retrospective think alouds.

Think Aloud Protocol

Immediately after completing the self-assessment task, participants engaged in a gaze-cued retrospective think and emote aloud protocol where they watched the complete screen recording of their self-assessment process and verbalized what they were doing, thinking, and

feeling (Beach & Willows, 2014; Chevrier et al., 2019). We instructed participants to view the entire recording and continually share aloud what they were doing, thinking, and feeling. We asked participants to try to talk continuously and we prompted them to keep talking if they were silent for 30 seconds, following Panadero et al. (2024). Think alouds were audio-recorded superimposed in the screen recording files. After completing the think alouds, participants engaged in a brief semi-structured interview where we asked them questions about ideas that emerged during their think alouds to help us more accurately interpret the think aloud data.

Quantitative Data Analysis

To investigate what resources participants attended to during self-assessment, and for how long, we calculated fixation durations for five areas of interest (AOIs): rubric, mastering exemplar, developing exemplar, essay, and self-feedback (see Figures 1 and 2). Eye tracking data were analyzed using Tobii Pro Lab software (version 1.181; Tobii Pro AB, 2014a). We defined dynamic AOIs in each participant's screen recording, which can be moved to follow an object's location in a screen recording, and calculated the total fixation durations in milliseconds for each of the AOIs. Fixations were identified using the Tobii I-VT Fixation Filter algorithm (Tobii Pro AB, 2014b).

Figure 1

Example Dynamic Areas of Interest (AOIs): Rubric

	Developing	Emerging	Mastering
Content	Ideas are general statements about the writer's experience without specific, personalized insights.	Ideas are general statements with some examples or anecdotes to illustrate the ideas.	Ideas are clear and compelling. The writer offers a well-developed and highly personal interpretation.
Organization & Structure	Individual ideas are clear, but the progression of one idea to the next is unclear.	The essay features a short introduction, at least one body paragraph, and a short conclusion.	The essay is well-organized, with an order and structure that present the discussion effectively.
Paragraphs	Paragraphs abruptly shift from one idea to the next, or there is only one paragraph. Paragraphs have several topics. Sentences contain independent ideas.	Transitions between paragraphs are simple. Paragraphs focus on one or two main ideas. A limited range of linking words are used to connect sentences within paragraphs.	Transitions between paragraphs strengthen the progression of the essay. Paragraphs are consistently and clearly focused on a main idea. Sentences are seamlessly linked using a variety of types of words.
Sentences	Sentences are often very short or long without a clear effect. Sentences follow a basic subject-verb-object pattern. Sentence lengths are all similar.	Sentences are generally correct. There may be minor agreement errors (e.g., "the information in each section are …"). One type of sentence is primarily used throughout. Fragments are used for effect.	Consistent and appropriate use of a variety of sentence structures and lengths to achieve clear effects (e.g., complex sentences to create a sense of rushing).
Conventions	Capital letter errors and/or spelling errors make meaning unclear. Full stops/period errors cause the reader to pause to decipher meaning.	Minor spelling, capital letter and/or punctuation errors are noticeable but do not distract from the meaning. Punctuation is used to ensure that meaning is clear.	Spellings and capital letters are used correctly. Punctuation is used to convey subtle meanings (e.g., semicolon used to contrast two parallel ideas).

Figure 2

Example Dynamic Areas of Interest (AOIs): Mastering Exemplar

Qualitative Data Analysis

We employed inductive thematic analysis (Thomas, 2006) to identify cognitive and affective processes in the think aloud data. Qualitative analysis of think aloud data is well suited to capturing individual differences in thinking and supporting theory development (Charters, 2003), two benefits that aligned with the objectives of this research. Coding was primarily conducted by the first author with guidance in the form of thematic review and refinement from the other authors and self-assessment researchers. Additionally, a small panel of three self-assessment researchers, comprised of graduate students and professors with experience using think aloud data to analyze students' self-assessment processes, reviewed the codebook during development and provided feedback and guidance.

The analytic process was conducted in 4 main parts. First, the first author read the complete think-aloud data to develop a holistic understanding of the data. Second, by cross-referencing timestamps in the eye tracking AOI data and think aloud recordings, the first author sorted all the text according to the active AOI. Third, they identified utterances related to the study objective and delineated units of text that contained one distinct idea (i.e., units varied in length), following a coding procedure developed by Campbell et al. (2013). Fourth, the first author coded each unit according to the idea expressed and categorized units into codes based on logical association. Multiple iterations of this coding process and discussion among the co-authors were conducted until a consensual interpretation of the data emerged which informed a final codebook (Thomas, 2006). The final codebook was then presented to a small review panel of three expert self-assessment researchers, comprised of graduate students and professors with experience using think aloud data to analyze students' self-assessment processes, who provided feedback to inform final refinement of the codebook. The first author then recoded all the data

with the finalized codebook.

To assess interrater reliability, we calculated Cohen's kappa (K) between the first author and a trained rater who independently coded 10% of the think aloud data. The first author trained the rater to use the coding protocol and provided the data with units of meaning delineated, following Campbell et al. (2013). The rater then coded the data independently using the final codebook to code each unit of meaning. The Cohen's kappa (K) value was 0.77 ($p<.001$), indicating a substantial level of agreement between the raters.

Data Mixing via Joint Display Integration

To mix the quantitate eye tracking and qualitative think aloud data, we used joint display integration to present codes—i.e., cognitive and affective processes—participants expressed when looking at each AOI. In mixed methods research, joint display is a method of data mixing which typically involves a graph or table that includes results from both the quantitative and qualitative databases which have been analyzed separately. Joint display allows mixed methods researchers to move beyond sequential discussions of results from each type of data. By visualizing the quantitative and qualitative findings together, joint displays allow researchers to draw meta-inferences about the relationship between the two databases (Creswell, 2021). Joint displays have been used to support advances in psychology, health sciences, and education research (Bustamante, 2019; McCrudden et al., 2021).

Researchers have documented various joint display integration approaches that have been used in research (Guetterman et al., 2015). Joint displays can be used to support data collection or data analysis (McCrudden et al., 2021). McCrudden et al. (2021) identified five distinct mixing purposes which joint display may address: triangulation, to investigate different aspects of the same phenomenon; complementarity, to use findings from one form of data to clarify or

enhance findings from another; development, to inform decisions about another form of data; initiation, to identify contradictions between findings from two forms; and expansion, to use findings from both to investigate two different phenomena. In our study, we mixed data for a complimentary purpose: we wanted to display think aloud themes alongside AOI fixation durations to generate insights about what participants were thinking when looking at each AOI. In this way, eye tracking and think aloud data complemented each other, with eye tracking indicating what visual resources participants thought about, and for how long, and think aloud data illuminating participants' latent cognitive and affective processes associated with each visual resource.

Different mixing purposes may be supported by different types of joint displays. To address a complementary mixing purpose, we created an integrated results matrix, which researchers use to show quantitative and qualitative results side by side to compare findings and generate meta-inferences (McCrudden et al., 2021). Meta-inferences go beyond inferences that can be made from findings from a single methodical tradition; they are generated and substantiated through comparing data from different methodological traditions (Creswell, 2021). Using an integrated results matrix to generate meta-inferences supported our complementarity purpose, enabling eye tracking and think aloud data to inform interpretations of the other that would be unsupported by only one of these forms alone (e.g., what participants were thinking about when looking at the rubric).

Results

Table 1 presents the integrated results matrix of our quantitative eye tracking and qualitative think aloud findings. The three leftmost columns indicate the AOI and quantitative fixation durations. We included both absolute fixation durations for each AOI and weighted

fixation durations, which we calculated by dividing the total fixation duration for each participant by the number of words on each AOI and then totalling those values. In essence, the weighted durations reflect the amount of time participants spent fixating on each AOI distributed across the number of words. The weighted fixation durations account for the differing number of words in each AOI, which may have influenced the time participants spent looking at each AOI. As shown in the integrated results matrix, participants spent the most time looking at the essay (97.13 mins), followed by the self-feedback space (32.86 mins), the rubric (15.39 mins), the mastering exemplar (13.81 mins), and the developing exemplar (4.23 mins). This pattern held true in the weighted durations, with the exception that participants spent more time on the developing exemplar (2.33) than the developing exemplar (1.72) when accounting for the word counts. Fixation durations from eye tracking data suggest that participants spent the most time cognitively processing their essays and self-feedback. In terms of self-assessment resources, fixation duration was highest for the rubric and lowest for the developing exemplar.

In the two rightmost columns, we present the codes that emerged from our inductive analysis of the think aloud data. We identified 17 codes in total. These codes articulate participants' cognitive and affective processes when they were fixating on each resource, as indicated by eye tracking data. We included representative quotations to illustrate each cognitive and affective process.

Table 1

Integrated Results Matrix of Eye Tracking and Think Aloud Findings

AOI	Fixation Duration (min)	Weighted Fixation Duration (s)	Cognitive and Affective Processes	Representative Quotations
Essay	97.13	36.76	Searching for mistakes or areas for improvement	"I was finding parts that can improve, going through all the essay."
				"I tried to find some letter errors and spelling mistakes. I was reading this time as a like a teacher."
			Remembering their work	"I'm reading the essay that I wrote first because I want to remember what I have written."
			Managing distractions	"And it's slightly frustrating because I wasn't able to read it through without stopping to edit something, I would've rather have just been able to read and read it through in one pass."
			Confusion, enjoyment, and frustration	"I was a little confused or unknowing if this was right."
				"I felt good about editing this because I felt like this was one of the better paragraphs."
				"And here I was very frustrated by the way I phrased that."
Self-Feedback	32.86	28.17	Actualizing their own self-assessment processes	"From my experience, I need to write down all my feedback because otherwise I'm coming up with my next thought and then I forget that and then I'm really upset."
			Reflecting on the role of bias	"Maybe I should be more critical of myself, maybe it's not well structured and I'm just saying this because it is the work that I did and I'm being kind of biased."
			Self-forgiveness	"Yeah, and I tried to tell myself at the end, it doesn't really matter that much, so I don't know, I tried to do something."
Rubric	15.39	6.81	Focusing on relevant information	"I focus more on the emerging and mastering sections because I believe my work will probably be between either one of those sections of the rubric."

			Remembering criteria	"To make sure that I did not miss any important points that I wanted to touch upon in my feedback."
			Comparison: systematic consideration of each criterion	"So I would just look at the category piece and then read through my text quickly with a critical eye to see where I would put it."
			Overwhelm	"I was a little bit overwhelmed to be honest with this table."
Mastering Exemplar	13.81	1.72	Curiosity	"I found it very interesting to kind of compare and to see different ways of approaching it and how different people had written about."
				"The way it was written made me feel better because it was easier to understand and I was feeling more interested."
			Resonance	"It was more like reading a book. It was better to read easier. And interesting too. I was way more interested in this one. And I was, I know I was nodding. I was agreeing."
			Comparison: adopting or discarding approaches with consideration of personal style	"And then with reading that paper, I felt pretty sure that mine was also well written, even though the approaches were I would say completely different. And then I would say the middle part of this paper, I noticed that this person just has I would say a different writing style or maybe a different background in academia."
Developing Exemplar	4.23	2.33	Comparison: confidence building	"I thought to myself, it has pretty common basic mistakes it has. But I know that this is not my writing."
				"And here I knew it was a lot better than the developing text."

Our analysis indicates that, when processing the rubric, participants engaged four distinct cognitive and affective processes: *focusing on relevant information*; *remembering criteria*; *comparison: systematic consideration of each criterion;* and *overwhelm*. Focusing on relevant information involved participants making decisions about which criteria or descriptors were most relevant and thinking about those. At the same time, participants' cognitive processing was directed at remembering criteria as described in the rubric so that they could apply the criteria to their own essays. Beyond merely holding criteria in working memory, participants also compared their essay to the criteria in the rubric by systematically considering each criterion. For example, one participant shared, "I decided to first write like every dimension on how to grade the text down" and proceeded to evaluate their essay against each criterion one at a time. Overall, these codes highlighted the large amount of information contained in the rubric to process, which participants also expressed in a code related to affect: *overwhelm*. Participants expressed feeling overwhelmed with the amount of information in the rubric, and their cognitive processes reflected their efforts to focus on and extract the most helpful information to manage overwhelm. Overall, when looking at the rubric, participants engaged in a range of cognitive and affective processes aimed at consolidating and making sense of disaggregated information within the analytic rubric.

Cognitive and affective processes participants experienced when viewing the mastering exemplar included *curiosity*, *resonance*, and *comparison: adopting or discarding approaches with consideration of personal style*. Participants verbalized high levels of interest when reading the mastering exemplar, which we coded as curiosity—an epistemic emotion (Pekrun & Perry, 2014). Curiosity was provoked by the content of the exemplar essay, e.g., "well, more of curiosity really to see what had been written," and when participants were interested in how

another person approached the essay task, e.g., “it’s interesting to see how my perception of the conclusion is highly different.” Participants experienced resonance when they felt the exemplar described experiences similar to ones they had while studying abroad. When viewing the mastering exemplar, participants also engaged in comparison processes; however, the comparison process was concerned with considering which approaches they might adopt in their own writing that would maintain their own personal writing style. For example, when fixating on the mastering exemplar, one participant shared, “Yeah, so here I’m thinking what to compare with, then just trying to capture in general this style or approach of moderate level.” The exemplar, which synthesized and exemplified the elements in the rubric, provided a holistic indicator of performance that provoked participants to think of alternative ways to achieve mastering writing.

During the relatively short time participants spent processing the developing exemplar, they expressed a cognitive and affective process of *comparison: confidence building*. Participants perceived that their essays were stronger than this exemplar, which was a source of confidence for them. For example, one participant fixating on the developing exemplar shared, “I don’t think that my work was quite like this, I feel like it was a little bit better than that.” While we only coded one cognitive and affective process associated with the developing exemplar, building confidence via favourable comparison supported positive evaluations of performance and helped participants more precisely situate their work relative to the available indicators of performance (i.e., rubric, mastering and developing exemplars).

While viewing their own essays, participants verbalized six cognitive and affective processes: *searching for mistakes or areas for improvement*, *remembering their work*, *managing distractions, enjoyment, confusion*, and *frustration*. Participants read their essays with specific

goals in mind, such as searching for errors to correct or identifying ways to develop their writing. For example, as one read their essay, they shared, “And then I think I’m just having one more read through to check with anything that I wanted to change.” As participants read their essays or searched for ways to improve them, they became distracted by correcting minor errors (e.g., typos) or thinking about ways to reword sentences they disliked. One participant articulated noticing and managing distractions when they shared, “Here I was time after time re-reading the text and realizing this word stuff, and I would lose focus.” Participants experienced confusion and frustration when they detected that they were being distracted from prioritizing high-impact revisions and feedback, and when they were unable to resolve a perceived issue in their essay. They also experienced enjoyment when they read about their own positive experiences and when they could resolve an issue in their essay and, e.g., “I was happy about that to be able to give that kind of feedback.” The cognition and affect provoked by the essay were highly interrelated and predicated on a tension between their initial goals and perceived issues that arose as they read their essays.

Last, when viewing their self-feedback, participants expressed three cognitive and affective processes: *actualizing their own self-assessment processes*, *reflecting on the role of bias*, and *self-forgiveness*. Participants reflected on how they normally self-assess, or their level of experience with self-assessment, when approaching their self-feedback, and selected strategies that they normally use (e.g., writing down feedback immediately when they think of it, creating a bullet point list). Participants also reflected on their biases, considering whether they were being too harsh or too lenient in their self-feedback and questioning their accuracy, e.g., “there were no major grammar vocabulary errors, but this is also my, like, assessment, so I don’t know.” Participants also considered the full context of the essay writing process to practice self-

forgiveness: “I accepted that I’ll never be perfect. I’m not trying to achieve perfect. I just focused on bettering my writing.” The cognitive and affective processes associated with the self-feedback space appeared to center on metacognitive expressions about the self-assessment process itself which helped students to construct and emotionally process self-feedback.

Discussion

Research has begun to unpack what learners do when self-assessing. This study makes an important contribution to theories of the self-assessment process (Panadero et al., 2024; Yan & Brown, 2017) by elucidating what students are thinking about and feeling when they use typical resources to self-assess and revise their writing—i.e., rubrics and contrasting exemplars, in addition to their own reflective essays and self-feedback. Joint display integration of eye tracking (i.e., AOI fixation durations) and think aloud data generated key meta-inferences about the cognitive and affective processes participants engaged in when looking at self-assessment resources, supporting implications for research and practice. In this discussion, we address our research question and discuss implications for self-assessment, internal feedback, and cognitive load theory. We then conclude with a focus on key implications for practice and directions for future study.

What are the Cognitive and affective Processes of Students using Rubrics, Exemplars, and their Essays to Self-Assess their Writing?

Our findings contribute to self-assessment theory by elucidating the cognitive and affective processes students may engage in when using common resources to self-assess. Our joint display indicated that participants engaged in distinct comparison processes when using the rubric and both exemplars: the rubric provoked systematic comparison between each criterion and participants’ essays; the mastering exemplar provoked a focus on differences in writing

styles or approaches; and the developing exemplar provoked feelings of confidence through positive caparisons to participants' essays. Participants employed strategies (e.g., focusing on relevant information) to make sense of disaggregated information in the rubric while navigating feelings of overwhelm. Processing their essays was characterized by participants trying to focus on important revisions while being distracted by minor errors. When processing their self-feedback, participants were concerned with generating fair and balanced feedback, considering the context of their writing (e.g., limited time, a new genre) to forgive perceived shortcomings.

These findings build directly on current self-assessment process theories (Panadero et al., 2024; Yan & Brown, 2017). By elucidating the cognitive and affective processes that underpin the use of key self-assessment resources, our findings provide insight into how students engage in important self-assessment practices and actions. Since we examined how students generated self-feedback and revisions using rubrics and exemplars in real time, our findings offer particular insights into the cognition and affect underpinning seeking external feedback via monitoring, a self-assessment practice that has been identified as a predictor of academic achievement when used in the performance phase of task completion (Yan, 2020b). This practice appears to be primarily driven by distinct types of comparison processes that are provoked by specific types of resources. Educators may do well, therefore, to focus on developing students' comparison processes, helping them select and use specific processes depending on the type of resource. For example, educators could help students using high-quality exemplars of writing to brainstorm ways that they can approximate the quality of the exemplar in their own writing style. Supporting these underlying comparison processes will require educators to help students make their thinking explicit so they can intervene and support as needed (Nicol, 2021; Panadero, Lipnevich, et al., 2019).

Further, we situate our findings within Panadero et al.'s (2024) self-feedback model to offer insights into the cognitive and affective processing underlying the three phases of the self-assessment process: information processing, analysis, and self-assessment judgements. In the information processing phase, when students are bringing together information from their essays, rubrics, and exemplars that they can use to analyze and generate self-feedback about their work, students may be engaged in the cognitive and affective processes we identified in relation to reading and recalling information about their work. When processing information about their essays, these processes include searching for areas for improvement, remembering their work, managing distractions, and feelings of confusion, enjoyment, and frustration. When processing information in rubrics, students may try to focus on relevant information, remember the criteria, and manage feelings of overwhelm. When processing information about high quality exemplars, they may experience feelings of resonance and curiosity. Comparison is a key action in the analysis phase, and when students use rubrics to provoke comparisons, they may engage in systematic comparison between their essays and each individual criterion. Conversely, when using high-quality exemplars, they may attempt to compare their work holistically and consider ways to match the quality while maintaining their writing style. Comparisons with low-quality exemplars may provoke positive appraisals of their work, possibly enhancing their self-efficacy and persistence with self-assessment (Panadero et al., 2017; Ross, 2006). Last, some of the cognitive and affective processes we observed relate to the self-judgement phase, where students arrive at an overall judgement about their work based on the information the processed and analyzed. When participants generated self-feedback in our study, they aimed to actualize their own personal self-assessment processes, reflected on the role of bias in their self-feedback, and practiced self-forgiveness, considering the context and relative importance of their essays to keep

their feedback in perspective. By understanding these underlying cognitive and affective processes across the self-feedback model, educators can help students engage in these processes effectively to support students in processing information, analyzing, and generating feedback about their work.

The Role of Comparison in Internal Self-Feedback

The meta-inferences from our joint display integration illuminate important cognitive and affective processes that drive internal feedback, signaling that rubrics and exemplars provoked distinct types of comparison. Researchers have theorized that a core mechanism of internal feedback is learners making comparisons between their work and external feedback (Butler & Winne, 1995; Nicol, 2021). Our findings suggest distinct types of comparison that may be driven by different resources. When using rubrics, students might compare their work against one criterion at a time. Exemplars of high-quality performance may encourage students to balance achieving the same quality of the exemplar while maintaining personal style. Exemplars of low-quality performance, conversely, may inspire confidence in students, which could support motivation. The distinct comparison processes make sense given that analytic rubrics disaggregate elements of successful performance while exemplars provide a holistic sense of successful performance.

Researchers have theorized that exemplars can provide students with concrete examples of rubric descriptors, which can be abstract, and that using contrasting exemplars can help students more accurately self-assess by demonstrating concrete differences between performance levels (Lin-Siegler et al., 2015). Notably, we found little evidence that participants used the exemplars to understand abstract criteria in the rubric or compare the two contrasting exemplars to orient their understanding of different levels of performance. We only found two instances

where participants explicitly compared the two exemplars, and one instance where a participant explicitly compared the rubric and the exemplars. Considering the participants in this study were international university students and therefore likely highly effective and self-regulated learners, these findings signal that it could be beneficial for educators at all levels to guide students in using exemplars to identify concrete examples of abstract criteria articulated in rubrics, and to compare contrasting exemplars to support self-assessment accuracy.

Sources of Cognitive Load in Self-Assessment

Our joint display integration facilitated meta-inferences about possible sources of cognitive load when using rubrics and contrasting exemplars to self-assess writing. Findings may therefore help explain variance in cognitive load students experience when using different types of resources to self-assess (Fastré et al., 2012) and how rubrics can provoke strong emotional reactions (Panadero et al., 2023). Managing feelings of overwhelm and curiosity while grappling with demanding comparative processes—e.g., systematically comparing each rubric criterion to the essay; matching the level of performance of an exemplar without comprising personal writing style—could explain the increased cognitive load students have reported when using both rubrics and exemplars (Lipnevich et al., 2023). Additionally, when reviewing their essays, navigating a tension between achieving self-assessment goals (e.g., to identify areas for improvement) and addressing surprising things they encounter as they read (e.g., typos) could increase cognitive load students experience when self-assessing their writing. When students are generating and reviewing self-feedback, they may be engaged in metacognition about the influence of bias or managing emotions by practicing self-forgiveness, both cognitive and affective processes which may increase cognitive load.

Being aware of these sources of cognitive load during self-assessment can help educators

support students in using rubrics and exemplars effectively. While cognitive load can reflect beneficial deep thinking (Fastré et al., 2012), excessive cognitive load can lead to frustration and disengagement, limit students' self-assessment accuracy, and hinder their capacity to actualize their goals in self-assessment (Collins et al., 2019; Eva & Regehr, 2011; Krebs et al., 2022). Our findings highlight the value of educators in supporting students in achieving optimal levels of cognitive load to support beneficial self-assessment. Simply helping students recognize that self-assessment is a cognitively demanding activity and helping them monitor their emotions and cognitive effort during self-assessment could be a starting point for providing students with tools to regulate their self-assessment processes.

Importantly, only 40% of the participants in our study spoke English as a first language. The remaining participants were highly fluent in English and were completing undergraduate and graduate programs of study primarily in English. However, writing an essay and engaging in the self-assessment task in English could still have introduced extraneous cognitive load for those participants whose first language was not English. When we separated data from English first language participants and the others, we did not see any different themes emerging between the two groups. Yet, more research with participants completing the self-assessment task in their first language could help verify our findings related to cognitive load.

Conclusion and Future Research

Overall, our findings indicate that self-assessment requires substantial cognitive and emotional resources—supporting the conception of student self-assessment as a complex psychosocial process (Andrade & Brown, 2016; Boud & Brew, 1995). Recognizing the complexity of this process can be helpful for educators by highlighting the need for curricular, scaffolded, and tailored approaches to facilitating student self-assessment. It may not be enough

to simply invite students to self-assess and provide rubrics and exemplars; students can benefit from guidance which helps them monitor and control their self-assessment processes when using such resources—and develop these skills over time. Understanding the cognitive processes and emotions students are grappling with when using specific resources can help educators guide students in becoming aware of these processes and select those that are most helpful. For example, when supporting students in using rubrics, educators can provide students with strategies to mitigate feelings of overwhelm and process the information efficiently, such as pre-estimating performance to focus on the most salient parts of the rubric. When providing students with exemplars, educators could help students consider ways that they could match the performance level while keeping their personal writing style consistent—e.g., brainstorming alternative ways to demonstrate key success criteria. When students are reviewing their own work, educators can help students review their essays various times for specific reasons—e.g., reviewing once to correct all typos, then once to identify confusing sentences; then once to identify areas that can be developed. Such co-regulation of the self-assessment process could help students develop their capacity to self-assess and revise their own work effectively.

We situate this research within descriptive aspects of theory development which, rather an explaining why phenomena occur, provide comprehensive portrayals of phenomena and for which mixed methods approaches are well suited (Greene, 2022). Contributing to comprehensive descriptions of phenomena, such as student self-assessment, is critical for ensuring future research considers all aspects of phenomena and does not focus on aspects which may be construct irrelevant or confounding. Our study lays some initial theoretical groundwork and suggests pathways for future research in student self-assessment. First, a descriptive study with a small sample size, our findings cannot be generalized. Future research would do well to

investigate and test these cognitive processes and emotions in different education and socio-cultural contexts, as self-assessment processes may be shaped by socio-cultural influences (Andrade & Brown, 2016). Second, since our study pointed to the value of curricular approaches to supporting student self-assessment, we see value in research which can inform developmental perspectives on student self-assessment processes, echoing previous calls for developmental approaches to teaching self-assessment (Brown & Harris, 2014). For example, studies which examine how learners at different developmental stages think and feel when self-assessing in various subjects could inform developmentally appropriate pedagogies of student self-assessment which facilitate students' development of their self-assessment processes throughout their schooling. Third, as our study focused on students' use of rubrics, exemplars, essays, and self-feedback, future research is needed to examine the cognitive and affective processes associated with students' use of other common self-assessment resources, such as checklists of success criteria (Fastré et al., 2012).

References

Altmann, U., Bohrn, I. C., Lubrich, O., Menninghaus, W., & Jacobs, A. M. (2012). The power of emotional valence—from cognitive to affective processes in reading. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00192

Andrade, H. L. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In H. L. Andrade & G. Cizek (Eds.), *Handbook of formative assessment* (pp. 90–105). Routledge.

Andrade, H. L. (2019). A critical review of research on student self-assessment. *Frontiers in Education*, *4*, 87. https://doi.org/10.3389/feduc.2019.00087

Andrade, H. L., & Brown, G. T. L. (2016). Student self-assessment in the classroom. In G. T. L. Brown & L. Harris (Eds.), *Handbook of human and social conditions in assessment* (pp. 319–334). Routledge.

Beach, P., & Willows, D. (2014). Investigating teachers' exploration of a professional development website: An innovative approach to understanding the factors that motivate teachers to use internet based resources. *Canadian Journal of Learning and Technology*, *43*(3), 1–16.

Beach, P., & Willows, D. (2017). Understanding teachers' cognitive processes during online professional learning: A methodological comparison. *Online Learning Journal*, *21*(1), 60–84. https://doi.org/10.24059/olj.v21i1.949

Beheshitha, S. S., Gašević, D., & Hatala, M. (2015). A process mining approach to linking the study of aptitude and event facets of self-regulated learning. *ACM International Conference Proceeding Series*, *16-20-March-2015*, 265–269. https://doi.org/10.1145/2723576.2723628

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

Boekaerts, M. (1991). Subjective competence, appraisals and self-assessment. *Learning and Instruction*, *1*, 1–17.

Boud, D., & Brew, A. (1995). Developing a typology for learner self-assessment practices. *Research and Development in Higher Education*, *18*, 130–135.

Bourke, R. (2016). Liberating the learner through self-assessment. *Cambridge Journal of Education*, *46*(1), 97–111. https://doi.org/10.1080/0305764X.2015.1015963

Brandmo, C., Panadero, E., & Hopfenbeck, T. N. (2020). Bridging classroom assessment and self-regulated learning. *Assessment in Education: Principles, Policy and Practice*, *27*(4), 319–331. https://doi.org/10.1080/0969594X.2020.1803589

Brown, G. T. L., & Harris, L. (2013). Student self-assessment. In J. H. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 367–393). Sage.

Brown, G. T. L., & Harris, L. (2014). The future of self-assessment in classroom practice: Reframing self- assessment as a core competency. *Frontline Learning Research*, *2*(1), 22–30. https://doi.org/10.14786/flr.v2i1.24

Bryer, J., Andrade, H. L., & Cleary, T. (2025). *Writing*. https://docs.daacs.net/overview/assessments/writing

Bustamante, C. (2019). TPACK and teachers of Spanish: Development of a theory-based joint display in a mixed methods research case study. *Journal of Mixed Methods Research*, *13*(2), 163–178. https://doi.org/10.1177/1558689817712119

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

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

Campbell, J. L., Quincy, C., Osserman, J., & Pedersen, O. K. (2013). Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods and Research*, *42*(3), 294–320. https://doi.org/10.1177/0049124113500475

Charters, E. (2003). The use of think-aloud methods in qualitative research: An introduction to think-aloud methods. *Brock Education*, *12*(2), 68–82.

Chevrier, M., Muis, K. R., Trevors, G. J., Pekrun, R., & Sinatra, G. M. (2019). Exploring the antecedents and consequences of epistemic emotions. *Learning and Instruction*, *63*, 101209. https://doi.org/10.1016/j.learninstruc.2019.05.006

Collins, J., Regenbrecht, H., Langlotz, T., Can, Y. S., Ersoy, C., & Butson, R. (2019). Measuring cognitive load and insight: A methodology exemplified in a virtual reality learning context. *Proceedings - 2019 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2019*, 351–362. https://doi.org/10.1109/ISMAR.2019.00033

Craig, S. D., D'Mello, S., Witherspoon, A., & Graesser, A. (2008). Emote aloud during learning with AutoTutor: Applying the Facial Action Coding System to cognitive - affective states during learning. *Cognition and Emotion*, *22*(5), 777–788. https://doi.org/10.1080/02699930701516759

Creswell, J. W. (2021). Sampling, integration, and metainferences. In J. W. Creswell (Ed.), *A concise introduction to mixed methods research* (2nd ed., pp. 101–121). Sage.

Creswell, J. W., & Clark, V. L. P. (2018). Choosing a mixed methods design. In J. W. Creswell (Ed.), *Designing and conducting mixed methods research* (3rd ed., pp. 53–106). Sage.

Efklides, A. (2017). Affect, epistemic emotions, metacognition, and self-regulated learning. *Teachers College Record*, *119*, 130305.

Eva, K. W., & Regehr, G. (2011). Exploring the divergence between self-assessment and self-monitoring. *Advances in Health Sciences Education*, *16*(3), 311–329. https://doi.org/10.1007/s10459-010-9263-2

Fastré, G. M. J., van der Klink, M. R., Sluijsmans, D., & van Merriënboer, J. J. G. (2012). Drawing students' attention to relevant assessment criteria: Effects on self-assessment skills and performance. *Journal of Vocational Education and Training*, *64*(2), 185–198. https://doi.org/10.1080/13636820.2011.630537

Fiedler, K., & Beier, S. (2014). Affect and cognitive processes in educational contexts. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 36–55). Routledge.

Genc Aksaray, S., & Ozcelik, E. (2022). Attention mediates the effect of emotional arousal on learning outcomes in multimedia learning: an eye-tracking study. *Educational Psychology*. https://doi.org/10.1080/01443410.2022.2154317

Greene, J. A. (2022). What can educational psychology learn from, and contribute to, theory development scholarship? *Educational Psychology Review*, *34*(4), 3011–3035. https://doi.org/10.1007/s10648-022-09682-5

Guetterman, T. C., Fetters, M. D., & Creswell, J. W. (2015). Integrating quantitative and qualitative results in health science mixed methods research through joint displays. *Annals of Family Medicine*, *13*(6), 554–561. https://doi.org/10.1370/afm.1865

Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning*, *2*, 107–124. https://doi.org/10.1007/s11409-007-9016-7

Holmqvist, K., Örbom, S. L., Hooge, I. T. C., Niehorster, D. C., Alexander, R. G., Andersson, R., Benjamins, J. S., Blignaut, P., Brouwer, A. M., Chuang, L. L., Dalrymple, K. A., Drieghe, D., Dunn, M. J., Ettinger, U., Fiedler, S., Foulsham, T., van der Geest, J. N., Hansen, D. W., Hutton, S. B., … Hessels, R. S. (2023). Eye tracking: Empirical foundations for a minimal reporting guideline. *Behavior Research Methods*, *55*(1), 364–416. https://doi.org/10.3758/s13428-021-01762-8

Järvelä, S., Malmberg, J., Haataja, E., Sobocinski, M., & Kirschner, P. A. (2021). What multimodal data can tell us about the students' regulation of their learning process? *Learning and Instruction*, *72*, 101203. https://doi.org/10.1016/j.learninstruc.2019.04.004

Johnson, C. C., Sondergeld, T. A., & Walton, J. B. (2019). A study of the implementation of formative assessment in three large urban districts. *American Educational Research Journal*, *56*(6), 2408–2438. https://doi.org/10.3102/0002831219842347

Kostons, D., Van Gog, T., & Paas, F. (2009). How do I do? Investigating effects of expertise and performance-process records on self-assessment. *Applied Cognitive Psychology*, *23*(9), 1256–1265. https://doi.org/10.1002/acp.1528

Krebs, R., Rothstein, B., & Roelle, J. (2022). Rubrics enhance accuracy and reduce cognitive load in self-assessment. *Metacognition and Learning*, *17*, 627–650. https://doi.org/10.1007/s11409-022-09302-1

Lin-Siegler, X., Shaenfield, D., & Elder, A. D. (2015). Contrasting case instruction can improve self-assessment of writing. *Educational Technology Research and Development*, *63*(4), 517–537. https://doi.org/10.1007/s11423-015-9390-9

Lipnevich, A. A., Panadero, E., & Calistro, T. (2023). Unraveling the effects of rubrics and exemplars on student writing performance. *Journal of Experimental Psychology: Applied*, *29*(1), 136–148. https://doi.org/10.1037/xap0000434

McCrudden, M. T., Marchand, G., & Schutz, P. A. (2021). Joint displays for mixed methods research in psychology. *Methods in Psychology*, *5*. https://doi.org/10.1016/j.metip.2021.100067

McIntyre, N. A., Draycott, B., & Wolff, C. E. (2022). Keeping track of expert teachers: Comparing the affordances of think-aloud elicited by two different video perspectives. *Learning and Instruction*, *80*, 101563. https://doi.org/10.1016/j.learninstruc.2021.101563

Muja, N., & Appelbaum, S. H. (2012). Cognitive and affective processes underlying career change. *Career Development International*, *17*(7), 683–701. https://doi.org/10.1108/13620431211283814

Nicol, D. (2021). The power of internal feedback: Exploiting natural comparison processes. *Assessment and Evaluation in Higher Education*, *46*(5), 756–778. https://doi.org/10.1080/02602938.2020.1823314

Nicol, D., & MacFarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. https://doi.org/10.1080/03075070600572090

Panadero, E., & Alonso-Tapia, J. (2013). Student self-assessment: Theoretical and practical connotations. When it happens, how it is acquired, and how to develop it in our students. *Electronic Journal of Research in Education Psychology*, *11*(30), 551–576. https://doi.org/10.14204/ejrep.30.12200

Panadero, E., Brown, G. T. L., & Strijbos, J. W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review*, *28*(4), 803–830. https://doi.org/10.1007/s10648-015-9350-2

Panadero, E., Fernández, J., Pinedo, L., Sánchez, I., & García-Pérez, D. (2024). A self-feedback model (SEFEMO): secondary and higher education students' self-assessment profiles. *Assessment in Education: Principles, Policy and Practice*, *31*(3–4), 221–253. https://doi.org/10.1080/0969594X.2024.2367027

Panadero, E., Fernández-Ruiz, J., & Sánchez-Iglesias, I. (2020). Secondary education students' self-assessment: The effects of feedback, subject matter, year level, and gender. *Assessment in Education: Principles, Policy & Practice*, *27*(6), 607–634. https://doi.org/10.1080/0969594x.2020.1835823

Panadero, E., Fraile, J., Fernández Ruiz, J., Castilla-Estévez, D., & Ruiz, M. A. (2019). Spanish university assessment practices: Examination tradition with diversity by faculty. *Assessment and Evaluation in Higher Education*, *44*(3), 379–397. https://doi.org/10.1080/02602938.2018.1512553

Panadero, E., García-Pérez, D., Fernández Ruiz, J., Fraile, J., Sánchez-Iglesias, I., & Brown, G. T. L. (2023). Feedback and year level effects on university students' self-efficacy and emotions during self-assessment: Positive impact of rubrics vs. instructor feedback. *Educational Psychology*, 1–24. https://doi.org/10.1080/01443410.2023.2254015

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, *22*, 74–98. https://doi.org/10.1016/j.edurev.2017.08.004

Panadero, E., Lipnevich, A., & Broadbent, J. (2019). Turning self-assessment into self-feedback. In *The Impact of Feedback in Higher education: Improving assessment outcomes for learners* (pp. 147–163). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-25112-3_9

Panadero, E., Pérez, D. G., Ruiz, J. F., Fraile, J., Sánchez-Iglesias, I., & Brown, G. T. L. (2022). University students' strategies and criteria during self-assessment: Instructor's feedback, rubrics, and year level effects. *European Journal of Psychology of Education*, *58*, 1031–1051. https://doi.org/10.1007/s10212-022-00639-4

Pekrun, R., & Perry, R. P. (2014). Control-value theory of achievement emotions. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 120–141). Routledge.

Reber, R., & Greifeneder, R. (2017). Processing fluency in education: How metacognitive feelings shape learning, belief formation, and affect. *Educational Psychologist*, *52*(2), 84–103. https://doi.org/10.1080/00461520.2016.1258173

Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research, and Evaluation*, *11*, 10. https://doi.org/10.7275/9wph-vv65

Schellings, G. L. M., Van Hout-Wolters, B. H. A. M., Veenman, M. V. J., & Meijer, J. (2013). Assessing metacognitive activities: The in-depth comparison of a task-specific questionnaire with think-aloud protocols. *European Journal of Psychology of Education*, *28*(3), 963–990. https://doi.org/10.1007/s10212-012-0149-y

Sitzmann, T., Ely, K., Brown, K. G., & Bauer, K. N. (2010). *Self-Assessment of knowledge: A cognitive learning or affective measure? 9*(2), 169–191.

Stark, L., Brünken, R., & Park, B. (2018). Emotional text design in multimedia learning: A mixed-methods study using eye tracking. *Computers and Education*, *120*, 185–196. https://doi.org/10.1016/j.compedu.2018.02.003

Stobart, G. (2008). *Testing times: The uses and abuses of assessment*. Routledge.

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education*, *76*(3), 467–481. https://doi.org/10.1007/s10734-017-0220-3

Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, *27*(2), 237–246. https://doi.org/10.1177/1098214005283748

To, J., Panadero, E., & Carless, D. (2021). A systematic review of the educational uses and effects of exemplars. *Assessment and Evaluation in Higher Education*. https://doi.org/10.1080/02602938.2021.2011134

Tobii Pro AB (2014a). Tobii Pro Lab (Version 1.181) [Computer software]. Tobii Pro AB.

Tobii Pro AB (2014b). Tobii Pro Lab User Manual (Version 1.181). Tobii Pro AB, Danderyd, Sweden.

van Gog, T., Paas, F., van Merriënboer, J. J. G., & Witte, P. (2005). Uncovering the problem-solving process: cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology. Applied*, *11*(4), 237–244. https://doi.org/10.1037/1076-898X.11.4.237

Veugen, M. J., Gulikers, J. T. M., & den Brok, P. (2021). We agree on what we see: Teacher and student perceptions of formative assessment practice. *Studies in Educational Evaluation*, *70*, 101027. https://doi.org/10.1016/j.stueduc.2021.101027

Winne, P. H. (2018). Discussion paradigmatic issues in state-of-the-art research using process data. *Frontline Learning Research*, *6*(3), 250–258. https://doi.org/10.14786/flr.v6i3.551

Winne, P. H., Teng, K., Chang, D., Lin, M. P. C., Marzouk, Z., Nesbit, J. C., Patzak, A., Rakovic, M., Samadi, D., & Vytasek, J. (2019). NStudy: Software for learning analytics about learning processes and self-regulated learning. *Journal of Learning Analytics*, *6*(2), 95–106. https://doi.org/10.18608/jla.2019.62.7

Yan, Z. (2018). Student self-assessment practices: The role of gender, school level and goal orientation. *Assessment in Education: Principles, Policy and Practice*, *25*(2), 183–199. https://doi.org/10.1080/0969594X.2016.1218324

Yan, Z. (2020a). Developing a short form of the self-assessment practices scale: Psychometric evidence. *Frontiers in Education*, *4*, 153. https://doi.org/10.3389/feduc.2019.00153

Yan, Z. (2020b). Self-assessment in the process of self-regulated learning and its relationship with academic achievement. *Assessment and Evaluation in Higher Education*, *45*(2), 224–238. https://doi.org/10.1080/02602938.2019.1629390

Yan, Z., & Brown, G. T. L. (2017). A cyclical self-assessment process: Towards a model of how students engage in self-assessment. *Assessment and Evaluation in Higher Education*, *42*(8), 1247–1262. https://doi.org/10.1080/02602938.2016.1260091

Yan, Z., Brown, G. T. L., Lee, J. C. K., & Qiu, X. L. (2020). Student self-assessment: Why do they do it? *Educational Psychology*, *40*(4), 509–532. https://doi.org/10.1080/01443410.2019.1672038

Yan, Z., Brubacher, S., Boud, D., & Powell, M. (2023). The influence of self-assessment practices on online training for investigative interviewing skills. *Studies in Educational Evaluation*, *76*, 101236. https://doi.org/10.1016/j.stueduc.2023.101236

Yang, L., Yan, Z., Zhang, D., Boud, D., & Datu, J. A. (2023). Exploring the roles of academic self-concept and perseverance of effort in self-assessment practices. *Assessment in Education: Principles, Policy & Practice*, 1–26. https://doi.org/10.1080/0969594X.2023.2191161

Chapter 3

How do Students Self-Assess?

Examining the Metacognitive Processes of Student Self-Assessment

Abstract

Self-assessment is foundational to self-regulated learning and academic achievement. As such, education systems globally are called to foster student self-assessment; however, research continues to indicate that educators and students struggle to cultivate student self-assessment. A core barrier to supporting student self-assessment is that the metacognitive processes that shape self-assessment—i.e., how students monitor and control their self-assessment processes—are not yet understood, meaning that supporting students in developing their self-assessment processes is implausible. This study aimed to investigate the metacognitive processes of students engaged in self-assessment. We recruited English-speaking university students (N=25) to participate. In a lab, participants self-assessed and revised their own essays on a computer with access to on-screen self-assessment resources (e.g., rubric, exemplars). Participants' thoughts, emotions, and actions were captured via gaze-cued retrospective think alouds. Using a general inductive approach, we identified two overarching themes in the think aloud data: *metacognitive experiences* and *metacognitive skills* (Flavell, 1979). Metacognitive experiences, especially feelings of difficulty, helped participants monitor their self-assessment processes. Metacognitive skills (e.g. goal setting/reasoning, selecting a strategy) expressed the ways that participants controlled their self-assessment processes and indicate that students are continuously updating their goals to direct their self-assessment processes and drawing on metacognitive self-knowledge. Findings elucidate mechanisms whereby self-assessment may foster metacognition. By explicitly teaching students the metacognitive processes of self-assessment and supporting

students in making their self-assessment metacognition explicit, educators may use self-assessment to foster co-regulated learning and develop students' self-regulated learning. Findings advance descriptive aspects of self-assessment theory, providing an initial empirical understanding of the metacognitive dimension of student self-assessment.

Introduction

A core imperative of education systems globally is the development of learners who can self-assess their own performance and learning (OECD, 2019). Self-assessment is catalytic to self-regulated learning (i.e., a process whereby students manage their thinking, emotions, and behaviours in pursuit of goals) as it enables learners to mobilize appropriate prior knowledge, monitor their performance, and reflect on their learning to inform next steps (Andrade, 2013; Boekaerts, 1991). Meta-analyses focused on the contexts of both K-12 and higher education have implicated self-assessment as a driver of self-regulated learning (Panadero et al., 2017), creativity and creative self-efficacy (Yan, Lee, et al., 2022), and academic achievement (Lee et al., 2020; Yan, Lao, et al., 2022). Given that self-assessment appears foundational for promoting educational outcomes of self-regulated learning, creativity, and achievement, it is unsurprising that researchers and policymakers globally have called on educators to facilitate student self-assessment in K-12 (e.g., Bourke, 2016; British Columbia Ministry of Education, 2023) and higher education classrooms (e.g., Malecka et al., 2022; Nicol, 2021) across education contexts.

Despite these calls, researchers across education contexts have repeatedly raised concerns that student self-assessment is not being effectively fostered in classrooms. Studies have shown that facilitating self-assessment is underrepresented among teachers' formative assessment practices, a finding reported by both teachers and students (Brandmo et al., 2020; Johnson et al., 2019; Veugen et al., 2021; Volante & Beckett, 2011). This problem may be particularly

persistent in higher education (Panadero, Fraile, et al., 2019). Research has indicated that university teachers are less likely than secondary or primary teachers to facilitate student self-assessment due to a lack of formative assessment training and previous experiences with self-assessment (Panadero et al., 2014). A primary challenge to supporting student self-assessment is that researchers and educators do not fully understand the cognitive processes that comprise self-assessment (Andrade, 2019; Brown & Harris, 2014; Panadero et al., 2016). Without understanding these processes, students may be unprepared to make decisions to regulate their self-assessment processes, and teachers may be unable to help students learn how to self-assess. Research is needed that maps the cognitive processes of self-assessment, illuminating what students do when they are self-assessing, to actualize the benefits of student self-assessment in schools.

One particularly productive area would be to examine how students monitor and control their self-assessment processes in order to make decisions and guide their actions. Researchers have begun to identify the specific actions that students employ to self-assess (Panadero et al., 2024; Yan & Brown, 2017); however, students' awareness of their actions and the decisions that guide their actions are understudied. These monitoring and controlling functions can be understood as self-assessment metacognition—i.e., how students think about their thinking during self-assessment. Understanding the metacognitive processes that underpin self-assessment can illuminate how metacognition shapes self-assessment processes and ways that students can enhance their capacity to self-assess and direct their own learning. With the aim of supporting educator and student practice, the objective of this study was to investigate the metacognitive processes of students engaged in self-assessment.

Student Self-Assessment: Definitions, Effects, and Two Object Foci

Self-assessment is comprised of various actions whereby learners describe, and possibly evaluate, the quality of their performance (Panadero et al., 2016). Self-assessment can be initiated by learners independently, inside or outside of schools, and student self-assessment can be explicitly facilitated in formal educational contexts in pursuit of educational aims (Panadero & Alonso-Tapia, 2013). Our study is focused on how self-assessment occurs in or in relation to formal educational contexts, whether initiated by students or teachers. When facilitated in classrooms with the aim of supporting student learning, self-assessment appears to be most effective when the process also includes the opportunity for students to revise their work or performance based on their evaluation (Andrade, 2019). Educators can facilitate student self-assessment by providing students with resources (e.g., rubrics, exemplars) to make success criteria explicit and clear. A common way educators support self-assessment, providing students with analytic rubrics, which include qualitative descriptors of performance at various levels across success criteria, have been shown to enhance student achievement (Panadero et al., 2023) and are suited to supporting self-assessment of writing (Andrade et al., 2009). Another common way to support student self-assessment is by providing students with exemplars of performance (To et al., 2021).

Student self-assessment can be understood as a process comprised of actions that students employ with the aim of evaluating and revising their performance. We explicitly include revisions as part of the process following the work of Panadero et al. (2024) who empirically demonstrated that revising or redoing a task can inform students' self-assessments—students may redo a task to (dis)confirm their initial response, or revise a task to improve it and reflect on the process, leading to an updated self-assessment of their performance. Self-assessment as a

process is therefore not separate from revisions, in our view; it can inform revisions but is also informed by revisions—a view further supported by Yan (2020b) who found that the self-assessment practice of seeking external feedback via monitoring (i.e., referring to indicators of quality such as rubrics, exemplars, practice assessments) during the performance phase of task completion was the highest predictor of achievement on the task among all self-assessment practices. This finding suggests that students may perform better on a task if they seek external feedback using resources during the performance phase when they can use the feedback to inform how they complete and adjust their work. Since we understand learners as agentic (Winne et al., 2019), students may select and use or discard any number of actions and feedback during self-assessment; so while learners may make revisions as part of the process, other learners or the same learner in different instances of self-assessment may choose not to make revisions as part of the process. During self-assessment, students select and use available internal and external resources to generate internal feedback (Butler & Winne, 1995; Nicol, 2021; Yan & Brown, 2017). Internal resources include intuition, memory, and emotions, while external resources include exemplars, rubrics, and teacher or peer feedback, among others.

The actions that comprise self-assessment may be closely related to the subprocesses which support self-regulated learning. Researchers have theorized that self-assessment is a fundamental subprocess of self-regulated learning in all three phases of Zimmerman's foundational self-regulation model (2013): forethought (e.g., planning, prior knowledge activation, self-efficacy judgements before a learning task), performance (e.g., self-monitoring during a learning task), and reflection (e.g., appraisals of performance and strategies after a learning task; Boekaerts, 1991; Panadero, 2017; Panadero & Alonso-Tapia, 2013). Supporting this notion, in a meta-analytic review of practices used to support student self-assessment in K-

12 contexts, Brown and Harris (2013) found varying effects of self-assessment on achievement and concluded that self-assessment practices support learning when they engage students in subprocesses associated with self-regulated learning. Acting on the forethought phase, self-assessment may enhance students' self-efficacy to successfully complete a task or learn something—a factor which can motivate students to mobilize resources needed to engage in a task—over time, as indicated by a meta-analysis of self-assessment's effects on self-regulated learning (Panadero et al., 2017), possibly because criteria-referenced self-assessment can help students increase their understanding of success criteria and recognize their capacity to improve their performance (Ross, 2006). Alternatively, criteria-referenced self-assessment could refine the accuracy of students' self-efficacy via more accurate metacognitive monitoring, supporting students in mobilizing the right resources (Hartelt & Martens, 2024; Talsma et al., 2019). Acting on the performance phase, student self-assessment is theorized to facilitate self-monitoring towards goals, an integral component of self-regulated learning, by helping students interpret and adjust their behaviour and thinking (e.g., usage of learning strategies) as necessary during learning to improve their performance (Andrade, 2013). This theorized mechanism is supported by empirical evidence that students' use of external indicators of success (e.g., rubrics, exemplars) to self-assess during the performance phase of task completion was the highest predictor of achievement on the task (Yan, 2020b). Acting on the reflection phase, self-assessment has been shown to guide students' capacity to reflect on their performance and calibrate their future learning goals accordingly (Yan & Brown, 2017). Taken together, these findings indicate that student self-assessment actions may mirror the cognitive subprocesses of self-regulated learning and help students internalize these subprocesses.

Researchers have theorized two levels of self-assessment actions that have different object foci (Panadero et al., 2016). The first level are actions that have the students' performance or work as the object focus. For example, a student may compare their essay to a scoring rubric to determine their current level of performance and identify ways to improve their work in relation to specific criteria (Nicol, 2021). This comparison action is focused on the student's work. The second level are actions that have the self-assessment processes itself as the object focus. For example, the student self-assessing their essay may realize that they do not understand a criterion in the rubric and are therefore struggling to feel confident in their self-assessment. They may consider looking for an exemplar essay to elucidate the criterion. This action is directed at the student's self-assessment process, rather than the student's essay. Therefore, during self-assessment, students can theoretically employ actions to evaluate and revise their own (a) work and (b) self-assessment process, allowing researchers to categorize self-assessment actions according to object focus.

Researchers have begun to elucidate the actions that have the student's work as their object focus. These actions have emerged primarily from two lines of research. The first was predicated on a cyclical model of students' self-assessment practices developed by Yan and colleagues using interview (Yan & Brown, 2017) and survey methods (Yan, 2020a, 2020b; Yan et al., 2020). Their model included four self-assessment practices, each of which represent a type of specific self-assessment action: seeking external feedback through monitoring (e.g., completing a practice test, comparing performance to a rubric or exemplar), seeking external feedback through inquiry (e.g., asking a peer or teacher for feedback or help), seeking internal feedback (e.g., referring to gut reactions or emotions), and self-reflection (e.g., reflecting on the product or learning process to inform a judgment about strengths or weaknesses).

The second line of research, developed by Panadero and colleagues, has examined self-assessment strategies and criteria. This research has been advanced using concurrent think aloud methods to capture self-assessment actions as secondary school (Panadero et al., 2020) and university students (Panadero et al., 2022) were engaged in self-assessment. The researchers developed a coding framework for self-assessment strategies and criteria based on an inductive analysis of the think aloud data (Panadero et al., 2024). Strategies refer to specific actions that learners employ during self-assessment which may be subject-specific or general (e.g., read the essay, compare the essay to the question, check the mathematics equation). These strategies are categorized into six overarching processes: read, recall, compare, rate, assess, and redo. Criteria are categorized as internal (e.g., intuition, hindsight, experience or self-efficacy) or external (e.g., from a rubric, from a teacher). This research culminated in the recent publication of a model of self-feedback articulating the self-assessment process as comprised of three phases. The first, information processing, involves students compiling information about their work and success criteria that they can use to self-assess through two primary actions—i.e., read (e.g., their work, rubrics, feedback) and recall (e.g., their work). In the second phase, analysis, students make sense of their performance by drawing on internal and external criteria, comparing performance to indicators of success, or revising or redoing work to (dis)confirm their performance or try a new approach. Last, students draw on their information processing and analysis to arrive at a self-assessment judgement (i.e., rating performance; Panadero et al., 2024). Importantly, the think aloud data revealed that revising or redoing their work was integral to the self-assessment process as participants used revision and redoing as tools to (dis)confirm their appraisals of their work or to re-calibrate their goals—in addition to making improvements to their work. This work

indicated that revision is not always merely an end goal of student self-assessment but also part of the self-feedback process.

Both these lines of research have focused on self-assessment actions that have, as their focus, the student's work. However, researchers have yet to examine the actions that have the self-assessment process as the object focus. This gap in the literature means that how students monitor and control their self-assessment processes is not yet understood. Since this object focus involves monitoring and control of cognition, i.e., metacognition, it can be understood as self-assessment metacognition. In this study, we therefore used metacognition theory as a lens to identify how students monitor and control their own self-assessment processes.

Theorizing the Metacognitive Component of Self-Assessment

Metacognition, a dimension of self-regulated learning, is a form of cognition that involves thinking about information that describes features of cognition (Winne & Azevedo, 2014). In other words, metacognition occurs when a learner thinks about their own thinking (Winne, 2017). Nelson and Narens (1990) defined metacognition according to three principles. First, cognitive processes are divided into two interrelated levels: the object-level and the meta-level. Second, the meta-level contains a dynamic model of the object-level. The meta-level involves a perception of the object-level. Third, there are two relations between these levels which are characterized by the direction of the flow of information between the levels. Monitoring refers to the flow of information from the object-level to the meta-level. The object-level informs the meta-level and can update the learner's mental model of the object-level. Control refers to the opposite flow of information, from the meta-level to the object-level. Here, the meta-level can support modifications to the object-level whereby the learner may initiate, continue, or terminate an object-level action. Monitoring elicits information about the object-

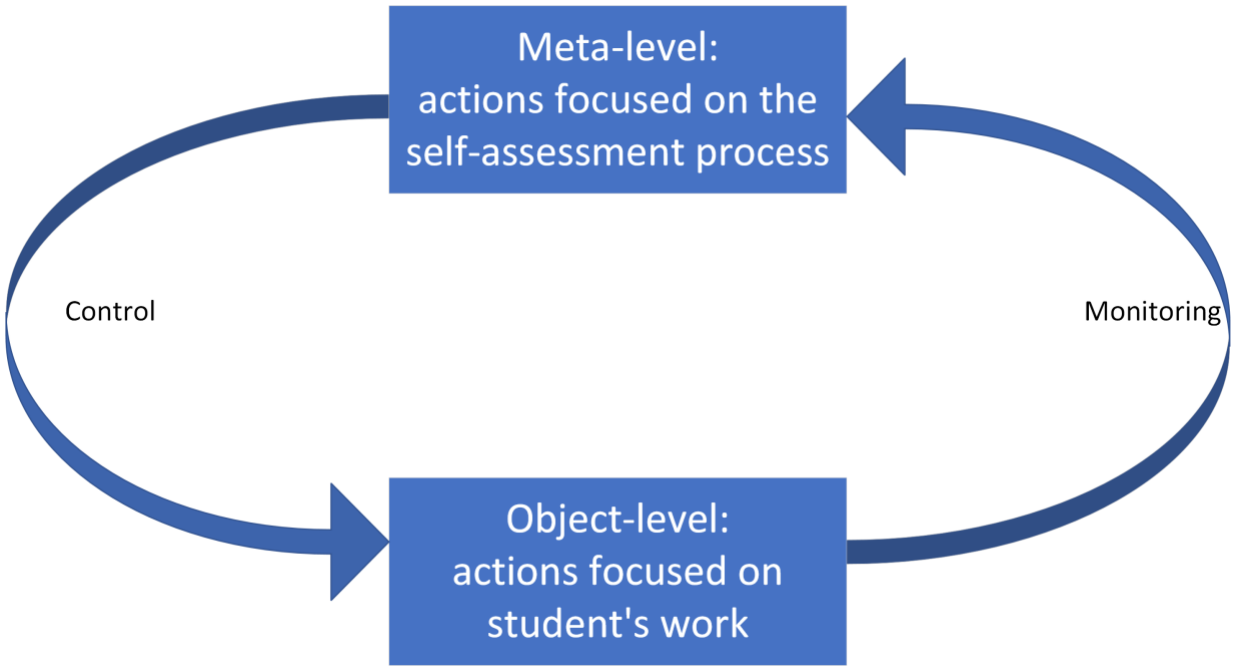
level to inform the dynamic model at the meta-level; control involves the learner using that information to modify the object-level. These definitions imply that students can have an awareness of their cognition and experiences during learning (Efklides, 2008). Metacognition plays a fundamental role in all phases of self-regulated learning, enabling students to (a) draw on relevant knowledge about similar tasks and themselves in the forethought phase; (b) monitor how their current states relate to desired goals and sustain or adapt their thinking, emotions, or behaviour accordingly; and simultaneously (c) (re)construct knowledge about the task and themselves through experience (Greene et al., 2024).

Applying this framework to student self-assessment, actions students apply with their work as the object focus are aligned with the object-level. These are the actions which have been profiled in previous research, e.g., read the essay, compare the response to feedback, refer to intuition (Panadero et al., 2024; Yan & Brown, 2017). The meta-level, in contrast, reflects students' thinking about their self-assessment actions, centered on their perceptions (i.e., dynamic model) of their self-assessment processes. Students can monitor their self-assessment actions (i.e., object-level) to elicit information about their nature and effectiveness. Processing this information to update their dynamic model (i.e., meta-level), students can modify their self-assessment actions and exert control over the self-assessment process (see Figure 1). Note that self-assessment metacognition could be conceptualized as meta-metacognition in a nested model, with students' work as the object level and monitoring and self-evaluations as the meta-level which initiates a third cycle (i.e., meta-meta-level) of cognitions about the self-assessment process. However, to align with our view of the relationship between self-assessment and self-regulated learning in our literature review—where self-assessment is integrated with monitoring and evaluations made based on monitoring—we see monitoring and evaluations *of students'*

work as the object-level, and not a separate meta-level, while monitoring and evaluation *of the self-assessment process* itself comprises the meta-level. What distinguishes the level of monitoring and evaluations is whether they are focused on their work or their self-assessment process. Moreover, rather than understanding how self-assessment fits into cognition, our aim is to conceptualize the metacognitive dimension of the self-assessment process; as such, we consider actions with the self-assessment process as the object of focus as pertaining to the meta-level.

Figure 1

Theorizing the Metacognitive Dimension of Student Self-Assessment



Note. Adapted from Nelson and Narens (1990).

The existence of self-assessment actions focused on the self-assessment process have been theorized (Nicol, 2021; Panadero, Lipnevich, et al., 2019). Panadero et al. (2019) called for educators to help students reflect on their self-assessment processes, making these implicit processes explicit, so that they can support students in developing their self-assessment processes

and model effective self-assessment processes. However, researchers have not yet elicited empirical evidence of thinking that occurs at the meta-level during self-assessment, nor have they used metacognition theory to identify and understand these processes. The focus of this study, therefore, was on eliciting evidence of self-assessment metacognition to generate insights into the actions whereby students monitor and control their cognition during self-assessment.

Metacognition theory provides an architecture for understanding how students may monitor and control their self-assessment processes. Metacognition can occur at any stage of a cognitive event: before (e.g., a learner preparing for a challenging learning task), during (e.g., a learner monitoring their performance and considering strategies to enhance it), or after (e.g., a learner reflecting on how effective their cognitive strategies were; Winne & Azevedo, 2014). Flavell (1979) envisioned a fundamental framework for metacognition which has underpinned developments in metacognition and its role in self-regulated learning. This framework characterizes three facets of metacognition. First, metacognitive knowledge refers to declarative knowledge, stored in memory, of cognitive processes, as well as information about the self, others, tasks, strategies, and goals. Metacognitive knowledge includes task-, strategy-, and goal-specific knowledge. Second, metacognitive experiences are metacognitive feelings, judgements, and online task-specific knowledge that a learner is aware of during a task as they process information related to it. This facet is both cognitive and affective in nature. Metacognitive feelings include feelings of knowing, familiarity, confidence, fluency, and difficulty. Metacognitive judgements include judgements of learning, estimates of effort, and estimates of time required. Online task-specific knowledge refers to task-specific knowledge learners are attending to at any time during a task and metacognitive knowledge retrieved from memory to help with processing information related to a task. Third, metacognitive skills encompass

deliberate use of strategies to control cognition. Metacognitive skills (i.e., strategy use) are driven by metacognitive knowledge and metacognitive experiences (Efklides, 2008).

Together, these facets provide a framework for studying how students engage monitoring and control of their self-assessment processes. In the context of self-assessment, metacognitive knowledge includes knowledge about self-assessment; metacognitive experiences are the feelings, judgements, and knowledge that students are aware of during self-assessment; and metacognitive skills are strategies students consciously deploy during self-assessment to control the process.

Elucidating Self-Assessment via Online Measures

As we were interested in capturing instances of self-assessment metacognition while students were engaged in self-assessment, we employed online measures in this study. Online measures collect data during learning (Chevrier et al., 2019; Winne & Perry, 2000). Examples of online measures include think alouds, eye tracking, and digital trace data. Interviews and questionnaires (i.e., offline measures) rely on decontextualized and often misremembered recollections of generalized learning events or hypothetical scenarios (Winne, 2020), offering "a composite measure of users' recollections of their overall experience" (Collins et al., 2019, p. 352). In contrast, online measures are uniquely capable of capturing fleeting thoughts, feelings, oculomotor actions, and physical sensations that occur during learning and escape participants' recollection or awareness—all of which are needed to fully understand a learning process (Winne et al., 2019). Empirical evidence has repeatedly demonstrated that online measures can inform more granular and comprehensive understandings of metacognition, self-regulated learning, and studying processes (Beheshitha et al., 2015; Callan & Cleary, 2018; Hadwin et al.,

2007; Schellings et al., 2013). Online measures are ideally suited to developing fine-grained understandings of how metacognition unfolds throughout the self-assessment process.

The use of online measures necessitates examining self-assessment within a specific task and subject context, as data are collected during the self-assessment process. Since self-assessment actions have been shown to vary across subjects (Panadero et al., 2020), defining the subject context is particularly important. We decided to focus on self-assessment of writing for four reasons. First, research has suggested that self-assessment is central to the process of learning to write (Andrade et al., 2009; Graham & Harris, 2018; Lin-Siegler et al., 2015). Second, writing has been theorized as applied metacognition whereby students monitor and control their thinking and externalize the metacognitive process through writing (Hacker, 2018), suggesting that a writing task may provoke evidence of self-assessment metacognition. Third, the first author of this paper has experience as an English Language Arts teacher and therefore has experience assessing writing and supporting self-assessment of writing. Fourth, writing is an important skill across a wide variety of contexts and professions; therefore, examining how self-assessment operates in writing contexts can support promotion of an important transdisciplinary skill.

Study Aim and Research Questions

The aim of this study was to investigate the metacognitive processes of students engaged in self-assessment. We were specifically interested in the metacognition students deployed during self-assessment—i.e., thoughts that had the self-assessment process as the object of focus. The cognition level of self-assessment (i.e., actions that have the student's work as the object of focus) was beyond the scope of this study and has been explored in other studies (Panadero et al., 2024; Yan & Brown, 2017). While we are examining the cognition level and other dimensions of

self-assessment in other studies, this study exclusively isolated the metacognition dimension of self-assessment in order to uncover how students monitor and control their own self-assessment processes.

Since we intended to capture the metacognitive processes that occur during self-assessment, our analysis focused primarily on metacognitive experiences and skills—the facets of metacognition at play during a learning process (Efklides, 2017). Students call upon metacognitive knowledge stored in memory to inform monitoring and control (Efklides, 2008), so we aimed to capture metacognitive knowledge as it arose during the process. Rather than trying to elicit participants' full range of metacognitive knowledge about self-assessment, we wanted to capture the metacognitive knowledge that students applied during the process.

Two research questions (RQ) guided this study:

RQ1. What metacognitive experiences shape students' self-assessment processes?

RQ2. What metacognitive skills do students employ during self-assessment?

Methods

In this study, we examined students' self-assessment metacognition using a qualitative research design by adopting a qualitative orientation to think aloud data (Charters, 2003). Eye tracking served solely as a gaze cue for retrospective think alouds and eye tracking data are not presented in this study. The eye tracking data will be analyzed in another study. University institutional ethics boards granted approval for this research.

Participants

Using a convenience sampling strategy, we recruited 25 English-speaking international students completing their studies at a university in Spain. Recruiting international students provided a socio-culturally diverse sample. Participants were from 15 countries: Australia,

Belgium, Canada, Croatia, France, Georgia, Germany, Ghana, Hungary, Iran, Netherlands, Nigeria, Pakistan, Philippines, South Korea, Turkey, Ukraine, United Kingdom, United States of America, and Puerto Rico. Most were white (*n*=15; 60%) and the rest were Black (*n*=2), Asian/Asian American (*n*=2), Latinx (*n*=2), or did not know or preferred not to identify (*n*=4). The largest group had English as their first language (*n*=10; 40%). While the remaining participants had various other first languages, all were proficient in English and studying primarily in English. Most participants were female (*n*=16) and most were 20-30 years old (*n*=18), with the remaining being older than 30. Participants were a mixture of undergraduate and graduate students studying in a range of disciplines, including the social sciences (*n*=6), engineering (*n*=4), humanities (*n*=3), and business *(n*=3), sciences, languages, communications, design, and energy. For detailed information on the participants, see the Supplemental Material.

Measures

Students' self-assessment metacognition was captured via gaze-cued retrospective think alouds (McIntyre et al., 2022). After self-assessing their writing on a computer equipped with screen-based eye tracking, participants watched the entire screen recording of their self-assessment process, which included the real-time gaze path and all screen-based actions (e.g., scrolling, typing, navigating resources), and verbalized what they were doing, thinking, and feeling. The researcher instructed the participants to say everything that came to mind as they watched the recording and to try to keep talking constantly. They were prompted to keep talking if they were silent for more than 30 seconds (Panadero et al., 2024). We selected cued think aloud protocols because they have been shown to elicit more verbalizations related to metacognition than concurrent think alouds or un-cued retrospective think alouds (Beach & Willows, 2017; McIntyre et al., 2022; van Gog et al., 2005). Moreover, gaze-cued retrospective think alouds have been shown to help more effective

learners better evaluate their own learning processes (Kostons et al., 2009). Gaze-cued retrospective think alouds are theorized to support greater utterances related to metacognition than concurrent think alouds because they provide fine-grained visual cues to inform participants' recollections, while concurrent think alouds may reduce metacognitive utterances because participants' focus is divided between thinking aloud and engaging in the task (van Gog et al., 2005). Participants received training on the think aloud protocol at the start of the data collection session and were reminded of the instructions immediately before the gaze cued think alouds.

Think alouds were audio recorded and transcribed. Demographic data were collected via an online questionnaire.

Study Procedure

Prior to data collection, participants met with the first author to complete a reflective essay, responding to the prompt, "Write an essay in English about your experiences living and studying in a new country." The essay was completed for the study and was not a part of students' course work. Participants were given a handout with the full instructions, including suggested writing prompts and a list of writing criteria, and they typed their essays on a computer. Participants also completed a short demographics questionnaire.

At a later time, participants came to the study lab for data collection. Working on the lab computer, participants reviewed their essay, generated written self-feedback, and optionally revised their essays. Their goal was to generate self-feedback and possibly revise their essays, not grade or rate their work, though they were advised that they could if they wanted to. Participants could open several on-screen resources throughout the self-assessment task: an analytic writing rubric, two contrasting exemplar essays, the essay instructions, and the essay. Participants were offered several resources—namely rubrics and exemplars—to use freely while

self-assessing. This was done to increase the ecological validity as this solution emulated more a real classroom situation than having a standardized implementation of rubrics and exemplars (e.g., standardizing the reading or use of the resources). Our aim was to capture self-assessment metacognition across a range of self-assessment processes and approaches, rather than a specific and standardized one, as the ways in which self-assessment is operationalized in classrooms and in students' independent learning vary drastically (e.g., Panadero et al., 2016). Additionally, research on self-assessment has identified at least four distinct self-assessment profiles (Panadero et al., 2024) which may evoke different types or levels of metacognition, so we did not standardize a self-assessment process to capture metacognition from the full range of profiles in our sample. The analytic writing rubric was adapted from a validated writing rubric developed for the Diagnostic Assessment and Achievement of College Skills (DAACS), an assessment and research tool which collects data and provides assessment takers with feedback on key skills needed to succeed in university (e.g., writing, self-regulated learning; see Bryer et al., 2025). Because such analytic rubrics disaggregate qualitative descriptions of performance at different levels across criteria, they provide a helpful way resource for students to self-assess and improve specific aspects of their writing (Andrade et al., 2009). The exemplars were developed by the first author to align with the performance levels articulated in the rubric and were refined based on feedback from three rubric researchers. See the Supplemental Material for the rubric and exemplars that students could open on the computer screen.

Eye tracking was recorded using a Tobii Pro Fusion screen-based eye tracking system (sampling rate 250Hz) and integrated with screen recordings of the entire process. Immediately after self-assessing, participants engaged in a gaze-cued retrospective think aloud protocol (McIntyre et al., 2022), wherein they were instructed to watch their complete screen recordings

with integrated gaze paths and verbalize everything they were doing, thinking, and feeling throughout.

Data Analysis

We interpreted think aloud data using a qualitative lens. While think aloud studies are often analyzed quantitatively, using a qualitative analysis approach can support the development of novel theories and elucidate individual differences in thinking processes (Charters, 2003)—benefits which aligned with our aim to theorize self-assessment metacognition. Additionally, metacognition researchers have recently used and called for qualitative think aloud approaches to support understandings of intersections between metacognition and self-regulated learning (Marra et al., 2022). We analyzed the data primarily using a general inductive approach which is suitable for qualitative studies with clear research objectives and for which a guiding framework is lacking (Thomas, 2006). This approach was ideal in the absence of an empirically-informed framework for self-assessment metacognition, and it allowed us to capture the full range of metacognitive approaches that appeared in the data. However, drawing on our theoretical framework predicated on Flavell's framework (1979), we also employed some deductive coding to denote the overarching themes, categorizing codes as either *metacognitive experiences* or *metacognitive skills*.

Coding was primarily conducted by the first author with guidance from the other co-authors. The first author inductively coded utterances related to metacognition (i.e., that had self-assessment process as the object focus) and then categorized these codes into overarching themes based on logical association. Utterances were defined as single idea units, so they varied in length. Utterances were only assigned a single code; none were double coded. We used text surrounding each utterance and participants' screen recorded actions (i.e., context) to interpret

the core meaning of each utterance when assigning codes (Winne et al., 2019). Following a general inductive approach, we engaged in multiple rounds of coding and interpretation of the data. Interrater agreement was establishing following a procedure developed by Campbell et al. (2013). The first author met with the co-authors, as well as other self-assessment researchers, after each round of coding to discuss the utterances (i.e., agree on the lengths of idea units) and codes. These discussions involved raters coding samples from multiple transcriptions and then discussing disagreements and, once consensus emerged, we adjusted the coding scheme as needed. This process was repeated until all authors agreed on the codes, after which the first author re-coded all the data using the finalized codebook. We then assessed interrater reliability of the coding by calculating Cohen's kappa (K) between the first author and a trained rater who independently coded 10% of the think aloud data. Following Campbell et al. (2013), the first author delineated pre-defined units of meaning (i.e., utterances containing or expressing a single idea) in the sample data and trained the rater to use the coding protocol. The rater coded the sample data independently, using the final codebook to code each unit of meaning. The Cohen's kappa (K) value was 0.78 ($p<.001$), indicating substantial agreement between the raters.

Through this iterative process, we grouped codes into two overarching themes, *metacognitive experiences* and *metacognitive skills*, which were aligned with research questions 1 and 2, respectively. These overarching themes were named according to Flavell's (1979) metacognition framework; hence, while our fine coding was driven by inductive analysis, we used a hybrid approach guided in part by Flavell's framework to arrive at the larger, overarching themes. These themes and their codes articulated the metacognitive processes of self-assessment.

Results

Metacognitive Experiences (RQ1)

Metacognitive experiences included thoughts or feelings participants had about their self-assessment processes that emerged during self-assessment. We observed 9 metacognitive experiences in the data: *abduction*, *efficacy judgements*, *feelings of difficulty*, *feelings of (dis)satisfaction*, *judgements of learning*, *monitoring focus or concentration*, *noting the time or time management*, *realizing their self-assessment process*, and *reflecting on task value*. See Table 1 for codes, representative quotations, and frequencies.

Table 1

Metacognitive Experiences (RQ1): Codes and Representative Quotations

Code	Description	Representative Quotation	*n*
Metacognitive Experiences (RQ1)			
Abduction	Coming to an important realization, a moment of insight, about their thinking in the moment	"… and I realized it made more sense to kind of connect those two points."	14
Feelings of difficulty	Expressing a feeling of difficulty or challenge	"I couldn't find it. I think I was a bit stressed."	77
Feelings of (dis)satisfaction	Metacognitive feeling of (dis)satisfaction with something they have done or thought	"So at that time, I thought my feedback is a bit weird."	44
Efficacy judgements	Making a judgement about their perceived capacity to successfully engage in self-assessment or apply a relevant skill	"I have some weakness some … some bad, not habits, like bad skills for this self-assessment"	20
Judgements of learning	Making a judgement about their learning from the self-assessment	"And then after, yeah, in the middle, I said okay, I think I've got the structure of it."	11
Monitoring focus or concentration	Explicitly noticing their level of focus or concentration	"Yeah, I think because my eyes are making such weird patterns, it's because I'm not that focused. I'm reading, but start … I'm losing my concentration"	14
Noting the time or time management	Commenting on time and/or time management	"oh, I'm running out of time"	34
Realizing their self-assessment process	Realizing or learning something new about how they self-assessed, or self-assessment in general	"It's crazy because it's like all over the place what I did"	46
Reflecting on task value	Reflecting on the value of self-assessment	"I feel like if you go back, you can give yourself feedback first on like in a separate document. Then you can go back to your work and edit or revise what you've already written."	15

The most frequent metacognitive experience was feelings of difficulty (*n*=77). This metacognitive experience involved participants articulating a challenge they were experiencing during self-assessment. Participants experienced feelings of difficulty for various reasons. For example, feelings of difficulty arose when participants felt compelled to correct technical errors while reading through their work. Participants expressed that noticing technical errors (e.g., typos) in their essay distracted them from reading it entirely. One shared, "I would've rather have just been able to read and read it through in one pass, well, that's great, and then think of it. But no, I had to make edits along way where I thought I didn't have to." This participant felt that, while correcting technical errors was necessary, it was also not the most efficient use of the self-assessment time, suggesting that they would have preferred to focus on more substantial concerns (e.g., structure, originality, diction). Another participant expressed a similar sentiment when they said, "And this [overthinking] sometimes, I guess, could interfere with the general performance, like overthinking about the use of language rather than actually paying attention to the contents." Participants recognized, via metacognitive experiences, that being distracted by technical errors could limit their capacity to make substantial revisions to improve their essay.

A common source of feelings of difficulty was in participants' doubts about the accuracy or objectivity of their self-assessment. When experiencing feelings of difficulty, participants expressed concerns that they were not qualified to assess their own writing, suggesting they felt they lacked the skills or technical knowledge about writing needed to self-assess effectively. Participants also expressed concerns about holding positive biases about their own work. For example, one shared, "maybe it's not well structured and I'm just saying this because it is the work that I did and I'm being kind of biased or something." One the other hand, participants also worried that their judgements were overly critical. While thinking about a transition they had

written, one said, “I didn’t know if it was a good transition or if it was just me being a bit too harsh.” Participants raised doubts about judgements they made about their own work, a tension that manifested in feelings of difficulty.

Feelings of difficulty were also provoked when participants struggled to think of how they might revise a part of their essay or correct a mistake. One participant was attempting to evaluate their essay as too colloquial, but they could not remember the word colloquial or an appropriate synonym. They re-read the passage several times and said, “but still I can’t remember the English word, but I mean the language you use in your daily life and not in academic context.” Participants at times simply expressed that they were struggling to think or write, e.g., “I was struggling to compose.” Uncertainty and doubt were expressed in many feelings of difficulty, e.g., “Um, but I don’t even know if I should include that”, “But I think I was doubting myself again a little bit.” Participants also experienced feelings of difficulty when they were not sure what action to take next in their self-assessment process. For example, one said, “What am I looking for? I wasn’t sure what I was thinking here.”

Interestingly, one participant highlighted the topic of her essay as emotionally challenging which created a feeling of difficulty. The participant, an international student, was going to be returning to their home country shortly after data collection occurred, and they found it difficult to re-read and think about their essay reflecting on their experiences in Spain: “It’s hard to write about something that is emotionally, you’re very emotionally attached to when you’re also leaving it.” Re-engaging with a challenging topic during self-assessment interrupted the flow for this participant and caused them to reflect on a feeling of difficulty.

The second most frequent metacognitive experience was realizing self-assessment processes (*n*=46), which occurred when participants experienced a moment of insight about their

self-assessment process. These insights tended to relate to specific actions within the self-assessment, or broad realizations about how they self-assessed after the self-assessment had occurred. To exemplify the former type of realization, one participant shared, "I now realized that I was more critical for this one paragraph than I was earlier on." The participant then checked on their other paragraphs to ensure they were being equally critical across their whole essay. These in-the-moment realizations held potential to shape participants' self-assessment processes. As an example of the latter, broader type of realization about self-assessment, one participant shared, "But yeah, I think, during the process also I realized that self-assessment requires quite a lot of self-organization, too." By engaging in self-assessment and explicitly reflecting on the process, participants generated broad insights about self-assessment with the potential to shape how they approach it in the future.

The third most frequent metacognitive experience was feelings of (dis)satisfaction (n=44), which involved participants explicitly sharing their level of satisfaction with something they thought or did as part of the self-assessment process. In some instances, participants were clearly aware of why they experienced feelings of (dis)satisfaction. They reflected experienced feelings of (dis)satisfaction towards revisions they made to their work or about the self-feedback they had written. For example, one shared, "I'm happy with that, the edits I had made I guess," and another shared, "So at that time, I thought my feedback is a bit weird." Regarding their self-feedback, one participant reflected, "I felt like my thoughts weren't really expressed the way that I wanted them to be." Participants also expressed (dis)satisfaction regarding their approach to self-assessment. For example, after skimming an exemplar essay and realizing that they did not learn anything from skimming it, one participant said, "It was a stupid approach." These participants clearly identified sources of (dis)satisfaction. On the other hand, feelings of

(dis)satisfaction were at times not explicitly linked to sources. For example, after making a revision, one participant said "That is good. I like it." Feelings of (dis)satisfaction with their self-assessment process held potential for participants to direct their self-assessment processes, making changes to their approach when they felt unsatisfied or moving on to a new action when they felt satisfied.

Metacognitive Skills (RQ2)

Metacognitive skills comprised strategies that participants used to deliberately shape their self-assessment processes. Importantly, metacognitive skills were strategies participants employed with the self-assessment process as the object, rather than their essays. We clustered the metacognitive strategies into 7 codes: *goal setting/reasoning*; *considering a strategy or approach*; *considering if they are finished self-assessing*; *selecting a strategy*; *reflecting on initial states or personal predispositions*; *reflecting on their personal self-assessment process*; and *reflecting on the writing process*. See Table 2 for codes, representative quotations, and frequencies.

The most frequent metacognitive skill by far was selecting a strategy (n=477), wherein participants verbalized a specific strategy they were consciously using to manage the self-assessment process. Participants articulated a wide variety of strategies. One participant, after reviewing the mastering and developing exemplars, made notes to summarize features of each one:

> and then I decided to just make a little feedback on the examples I had read … typos, of stuff that I didn't want to see in my essay … and then a little feedback on the mastering example of stuff that I would want to see in my essay.

Table 2

Metacognitive Skills (RQ2): Codes and Representative Quotations

Code	Description	Representative Quotation	*n*
Goal setting/reasoning	Articulating a short-term goal or reason for doing something	"… to make sure that I was being as fair as possible to myself."	157
Considering a strategy or approach	Weighing different ways they might approach all or part of the self-assessment task	"… you know sometimes if you are asking someone to explain, it's like a bad thing. I was supposed to ask you but I didn't ask you."	8
Considering if they are finished self-assessing	Explicitly considering if they are finished self-assessing and revising their essay or if they want to continue self-assessing and revising	"And then now I thought I was thinking, 'Okay, what do I do now? Do I continue writing the feedback or do I stop?'	21
Selecting a strategy	Naming their strategy for regulating or managing the self-assessment process without explaining why	"Okay, so I thought first I was going to compare my work with the criteria."	447
Reflecting on initial states or personal predispositions	Explicitly reflecting on their psycho-social state and possibly how it might impact their self-assessment	"… and so I feel like I have a certain amount of perfectionism."	69
Reflecting on their personal self-assessment process	Explicitly reflecting on how they usually or like to approach self-assessment, and possibly how that is influencing their current self-assessment process	"… usually for me it's easier to self-assess myself when I have something as a comparison. I think that is a general understanding of a lot of people."	25
Reflecting on the writing process	Reflects on or remembers details about the process of writing the essay	"… because when I wrote this, the other time I really did not put much effort in the commas, I mean the grammatical structure of the content."	35

This example represented a relatively well-developed articulation of a strategy. Most examples were shorter verbalizations, where participants articulated in-the-moment decisions they made about the self-assessment process, e.g., "now I'm starting with [writing] the feedback. So I wanted to do bullet points." Such verbalizations operated as connecting points between self-assessment actions. In the previous example, selecting a strategy connected reading their essay with assessing their essay (i.e., self-generating feedback), two self-assessment actions which had the student's work as the object of focus (i.e., object-level). In some instances of strategy selection, participants paused to summarize their overall approach to the self-assessment, e.g., "And I think my approach with it was I changed any kind of grammatical issues or like if I kept reading where something didn't seem quite right." These instances involved participants reflecting on their strategies more broadly.

Goal setting/reasoning was the second most frequently used metacognitive skill and it was closely connected to selecting a strategy: it involved participants explaining why they employed certain actions or approaches. Goal setting/reasoning therefore often co-occurred with selecting a strategy. For example, one participant said, "So, I'm going to read it [the essay] again to be more objective." The first part of the statement was coded as selecting a strategy (i.e., the decision to re-read the essay) and the latter part was coded as goal setting/reasoning (i.e., the participant explaining why they decided to re-read the essay). Participants reported a variety of goals that guided their approaches and actions, including to:

- confirm they were satisfied with a judgement or revision (e.g., "to see if I agreed with my changes"),
- generate fair and unbiased feedback (e.g., "like, how do I be fair on myself here?"),

- locate a specific type or piece of information ("I tried to find some letter errors and spelling mistakes"),
- remember something (e.g., "to remind myself what I wrote about"),
- orient themselves to the spectrum of performance ("So yeah, I was trying to follow the first two stuff [criteria] to understand the categories how I need to understand myself"), and
- improve the essay ("because it makes the sentence nice").

These short-term goals or reasons for doing and thinking things played an important role in the self-assessment actions and approaches participants undertook.

While they occurred relatively less frequently, instances of participants reflecting on their initial states, personal self-assessment process, or the writing process were important because they involved participants drawing on metacognitive knowledge about themselves, self-assessment, and writing. Some participants commented on having little experience with self-assessment, or not having self-assessed in a long time, highlighting their awareness of how this initial state might shape their self-assessment. One participant shared,

> I decided to start with rubric first [*selecting a strategy*] to kind of structure my feedback on my piece of writing and yeah just to get the general idea of what rubric said [*goal setting/reasoning*] because I don't have any experience with self-assessment [*reflecting on initial states or personal predispositions*].

Here, the participant considered their initial state (i.e., inexperienced with self-assessment) which informed their reasoning about why they selected a certain strategy. Participants' metacognitive knowledge of themselves also guided their processes. One noted that, when self-assessing, they often have a hard time starting, but once they get started they are able to self-assess effectively:

"Yeah, usually I, when I write anything, including, for example, self-assessment, I have always hard time starting with things … But once I start, then it's much easier to continue." This metacognitive self-knowledge helped them to overcome a feeling of confusion about where to start during the self-assessment. Another participant reflected on their initial state in the self-assessment as being creative and energized, helping them to improve their essay: "So I could tell that my brain just functioned, I would say a bit more, I was feeling more productive and more creative". Participants reflected on the writing process (of writing the essay), drawing on metacognitive knowledge about writing and the writing process, to prioritize what they attended to in the self-assessment process. For example, one participant recalled that they were tired when they wrote their essay, saying, "but I think I was too tired to use my actual vocabulary skills, my better vocabulary skills," which informed their subsequent judgement about their work: "Yeah, I think like regarding transition signals, I think it's fine." Actively remembering the writing process was a metacognitive skill whereby participants controlled their self-assessment processes.

Discussion

Previous research has begun to uncover the actions and practices that comprise student self-assessment (Panadero et al., 2024; Yan & Brown, 2017). However, while a metacognitive dimension of self-assessment has been theorized, there has been limited empirical research examining how students think about their self-assessment processes while engaged in self-assessment. The findings from the present study provide initial empirical evidence of the metacognitive process that shape how students self-assess. Overall, our findings suggest that metacognitive experiences served an important monitoring function in self-assessment, and served to support students in goal setting/reasoning and selecting a strategy to facilitate control.

The identification of metacognitive processes holds important implications for researchers studying self-assessment and for educators supporting students in developing their self-assessment processes.

RQ1: What metacognitive experiences shape students' self-assessment processes?

As explained in the theoretical framework, metacognitive experiences refer to metacognitive feelings, judgements, and online task-specific knowledge that a learner is aware of during a task (Flavell, 1979). During the think alouds, participants verbalized a range of metacognitive experiences during self-assessment: *abduction*; *feelings of difficulty*; *feelings of (dis)satisfaction*; *efficacy judgements*; *judgements of learning*; *monitoring focus or concentration*; *noting the time or time management*; *realizing their self-assessment process*; and *reflecting on task value*. Together, these metacognitive experiences served a monitoring function during self-assessment, helping students to re-evaluate their self-assessment process. For example, feelings of (dis)satisfaction represented participants' evaluations of aspects of their own self-assessment processes (e.g., self-feedback, revisions). Efficacy judgements, conversely, were a mechanism whereby participants calibrated their capacity to successfully assess and improve their work based on their developing perceptions of their self-assessment process. The metacognitive experiences identified in this study thereby held the potential for participants to be metacognitively aware of and evaluate their self-assessment processes (i.e., monitoring) so they could act on these evaluations (i.e., control).

While previous research has not examined self-assessment metacognition specifically, metacognitive experiences have been linked to self-assessment in theories of assessment as the co-regulation of learning. Researchers have theorized that classroom assessment is a mechanism of co-regulation (i.e., regulation with the support of peers and teachers) whereby teachers,

students, and classroom resources interact to regulate learning and instruction (Allal, 2020). In a systematic review of evidence of assessment as co-regulation, Andrade and Brookhart (2020) noted that student self-assessment may be an effective vehicle for promoting metacognitive monitoring by engaging students in understanding success criteria and reflecting on the extent to which their performance aligns with criteria. Our findings provide empirical evidence to support the link between student self-assessment and metacognitive monitoring: our participants' self-assessment processes elicited a range of metacognitive experiences, indicating that student self-assessment can operate as a co-regulator of learning and help students practice and internalize metacognition. Andrade and Brookhart (2020) also found evidence that formative and summative assessments can be used to elicit students' judgements of learning, a metacognitive experience that we also found as part of the self-assessment process.

In our study, feelings of difficulty were reported relatively much more frequently than other metacognitive experiences, suggesting they may play a particularly important role in self-assessment. The prevalence of feelings of difficulty in our participants' self-assessment processes highlight the documented connection between affect and metacognition, whereby metacognitive experiences have an affective or emotional component (Efklides, 2017). The affective component associated with feelings of difficulty could require students to dedicate most of their cognitive resources to attending to these experiences. Studies have reported varying levels of cognitive load during self-assessment which may be associated with the types of resources students use to self-assessment (e.g., a list of criteria, analytic rubrics, exemplars; Fastré et al., 2012; Krebs et al., 2022; Lipnevich et al., 2023). Our study suggests that feelings of difficulty may be an important source of cognitive load, given that it was reported so frequently by participants. Feelings of difficulty may explain some of the variance in the cognitive load

students experience during self-assessment, which would in turn shape the outcome of self-assessment on performance and learning.

Research on processing fluency can also support interesting insights related to this finding. In a review highlighting implications of processing fluency research for educational psychology, Reber and Greifeneder (2017) described how learners' perceptions of processing fluency inform metacognitive experiences, beliefs, and affect. Processing fluency can operate as a heuristic for judgements of learning and feelings of knowing, saving learners time and cognitive load required to consider evidence or retrieve information required to support these metacognitive experiences. Instead, a feeling of processing fluency can quickly signal to a learner how well they are learning or how much they know. An amount of processing disfluency can also be beneficial for learning. Disfluency in encoding has been associated with enhanced retrieval, and confusion due to contradictory information can enhance learning. In terms of affect, processing fluency may enhance interest in a task. Processing fluency has also been shown to increase the perceived truth of a statement. Relatedly, confusion has been shown to be beneficial for learning when students are able to engage self-regulated learning strategies that resolve the confusion (D'Mello et al., 2014), implicating confusion—which is related to feelings of difficulty and processing disfluency—as a heuristic that can facilitate strategy selection. Confusion has been shown to provoke metacognitive processing in elementary school students engaged in math problem solving (Munzar et al., 2021). These findings offer support to the important role that metacognitive feelings of difficulty may play as an instantiation of processing disfluency which may provoke confusion and support strategy selection.

In our study, feelings of difficulty could be interpreted as instances of processing disfluency, and our findings therefore illuminate how processing fluency may operate in self-

assessment. Participants experienced feelings of difficulty when they detected that increased cognitive resources and time were needed to process information—i.e., when they were unsure how to resolve a perceived impasse in their work, or when they realized their attention was being distracted by compulsions to correct typos. Additionally, participants experienced feelings of difficulty when reading about difficult topics in their essays, demonstrating a link between affect and processing difficulty. Processing fluency also seemed to shape our participants' beliefs about their work, as they attributed feelings of difficulty to doubt about their own self-assessment and objectivity; processing disfluency may have undermined participants' confidence in the truth of their own self-feedback. Students' doubts about their self-judgements of performance have been empirically documented in research studying second-order judgements (i.e., confidence ratings about previous self-judgements of performance) and provide information about students' metacognitive awareness (Händel & Fritzsche, 2016), supporting our findings that self-doubt may be an important part of the metacognitive dimension of self-assessment. Overall, processing disfluency seemed to shape each of the domains identified by Reber and Greifeneder (2017)—participants' metacognition, affect, and beliefs—during self-assessment. Feelings of difficulty informed participants' understanding of their self-assessment process (e.g., struggling to stay focused, confused about technical issues) as they rationalized why they were experiencing disfluency. Processing fluency and disfluency may both have an important role in guiding self-assessment, supporting efficient decisions and directing students' cognitive and affective resources to areas that students sense require further consideration (e.g., objectivity, sensing a source of distraction), respectively.

RQ2: What metacognitive skills do students employ during self-assessment?

As noted in our theoretical framework, metacognitive skills refer to the strategies that learners use to control cognition and are driven by metacognitive knowledge and metacognitive experiences (Flavell, 1979). Participants expressed seven distinct metacognitive skills while self-assessing: *goal setting/reasoning*; *considering a strategy or approach*; *considering if they are finished self-assessing*; *selecting a strategy*; *reflecting on initial states or personal predispositions*; *reflecting on their personal self-assessment process*; and *reflecting on the writing process*. These metacognitive skills articulate mechanisms whereby participants deliberately shaped their self-assessment processes. For example, considering a strategy or approach involved participants thinking about how a certain strategy (e.g., asking for clarification, highlighting text) might be helpful. Participants reflected on their personal self-assessment processes to inform decisions about how to proceed during the self-assessment.

Our findings highlighted the particular importance of goal setting/reasoning and selecting a strategy, the most frequently verbalized metacognitive skills. Metacognitive goal setting/reasoning involved participants setting short-term goals or explaining why they did or thought something related to their self-assessment process specifically. Selecting a strategy was a similar metacognitive skill, involving participants expressing their self-assessment strategy or approach but without explicitly stating why they were doing it. While research specifically focused on goal setting in self-assessment is lacking, assessment as co-regulation theory has posited that assessment activities, such as assessment portfolios, can forefront goal setting as a learning skill for students (Andrade & Brookhart, 2020). What our findings add to this discussion is insight into the highly recursive nature of metacognitive goal setting/reasoning during self-assessment. While assessment portfolios may help students and teachers co-construct goal setting

before or between learning tasks, our think aloud data revealed that students are continuously involved in setting and updating their goals to direct their self-assessment processes. The recursive nature of self-assessment processes has also been identified in a K-12 context using digital trace data (Rickey et al., 2023). With the addition of the present study, the highly recursive nature of self-assessment is corroborated now with two different online methods, suggesting the importance of online measures in detecting recursion. This recursive goal setting and updating may be an instantiation of volatile data, thoughts and feelings that occur quickly during learning which offline measures are not well suited to capturing because participants forget about or are not aware of them (Winne et al., 2019). Notably, utterances wherein participants expressed goal setting/reasoning and selecting a strategy were often very short, with participants quickly moving on to their next thought, implicating them as possible instances of volatile data.

Overall, the prevalence of goal setting/reasoning and selecting a strategy throughout participants' self-assessment processes highlights the importance of students being able to articulate their self-assessment processes in order to control them, as participants were asked to verbalize what they were doing, thinking, and feeling during the self-assessment task. In other words, being intentionally aware of their actions, decision-making, and reasoning during self-assessment is an important part of being able to control and direct their own self-assessment process. This finding is supported by self-regulated learning research: a foundational meta-analysis found that self-regulated learning training interventions were most effective when they included metacognitive reflection (Dignath & Büttner, 2008).

Participants' metacognitive knowledge about self-assessment emerged through reflecting on initial states or personal predispositions; reflecting on their personal self-assessment process;

and reflecting on the writing process. In these instances, participants drew on their metacognitive knowledge about themselves as learners, how they usually engage in self-assessment, and their recollection of the essay writing process to make decisions about or contextualize how they self-assessed and revised their essays. Our findings reinforce that students draw on metacognitive knowledge about themselves and their prior cognitive processing to inform metacognitive skills (Efklides, 2008). Our data also suggest that students may have conceptions about their personal approaches to self-assessment which they rely on to inform their self-assessment processes. Notably missing in our findings was evidence that participants applied task-specific knowledge about self-assessment, apart from knowledge about how they have approached self-assessment in the past. This finding is unsurprising; given that researchers have yet to fully theorize the cognitive, metacognitive, and affective processes that comprise self-assessment (Brown & Harris, 2014; Panadero et al., 2024), we would not expect the university students who participated in this study to hold much task-specific knowledge about self-assessment, beyond that accurate self-assessment requires awareness of personal biases.

Implications

The findings from this study have implications for theory and practice. Theory development is primarily comprised of descriptive and explanatory aspects (Greene, 2022). Explanatory aspects articulate why a phenomenon in the world manifests, including relationships between, predictions about, and ways to control phenomena. They tend to be informed by experimental research and quantitative analyses of relationships between constructs. Less frequently discussed, descriptive aspects provide an accurate, useful, and comprehensive portrayal of a phenomenon. Descriptive aspects, based on observations of a phenomenon, are needed to precede explanatory ones: incomplete understandings of phenomena leave aspects of

the phenomena unexamined, leading to gaps in research and practice. Qualitative and mixed methods research are well suited to developing descriptive aspects of theory by capturing comprehensive and contextualized observations. However, in psychology, theory development has tended to focus on explanatory aspects of theory development, with insufficient attention given to descriptive aspects (Greene, 2022).

The present study specifically advances the descriptive aspects of self-assessment theory. Leveraging a qualitative orientation to online think aloud data, we elucidated the metacognitive processes of self-assessment, a dimension of self-assessment that has been theorized but not empirically investigated. Our findings provide an initial empirical understanding of the underlying metacognitive experiences and skills which comprise self-assessment metacognition. Greene (2022) argued that "descriptive aspects of theories are particularly useful when they elucidate less-commonly-known phenomena or aspects of those phenomena" (p. 3014). Perspectives on self-assessment have evolved from seeing self-assessment as a skill to understanding it as a multidimensional process comprised of cognitive, affective, and metacognitive actions (Panadero et al., 2016, 2024). By providing an initial articulation of the metacognitive dimension of self-assessment, our findings have expanded previous understandings of the actions and criteria that comprise the process of self-assessment. This articulation can inform future research and prevent explanatory aspects from overlooking the role of metacognition in how self-assessment relates to self-regulated learning.

This study also holds implications for metacognition theory. Assessment policies (British Columbia Ministry of Education, 2023; Manitoba Education, 2006; Ontario Ministry of Education, 2010) and teachers (Braund & DeLuca, 2018) have theorized that engaging students in self-assessment is an important part of fostering student metacognition. However, the

mechanisms linking student self-assessment and metacognition are not yet empirically understood. Findings from this study may introduce some specific mechanisms whereby self-assessment may help students engage in and internalize metacognition over time. For example, engaging in self-assessment seems to provoke a range of metacognitive experiences, especially feelings of difficulty, which provides students with an opportunity to learn to recognize and respond to metacognitive experiences in other learning situations. Self-assessment also affords students an opportunity to apply metacognitive skills, such as goal setting and selecting a strategy, reflecting self-assessment's capacity to catalyze students' metacognitive control and application of metacognitive knowledge (e.g., reflecting on initial states/personal predispositions). Self-assessment can therefore help students practice and internalize the metacognitive processes identified in this study.

In terms of practice, findings from this study highlight the importance of supporting students' metacognitive knowledge about self-assessment, an aspect of metacognition that seemed to be lacking in participants' think alouds. It was clear that participants had perceptions of their personal self-assessment processes that they drew on to inform how they approached self-assessment in this study. However, it was not evident that they had learned strategies or approaches to self-assessment beyond those that they had personally developed. Moreover, reflections on their personal self-assessment processes tended to elicit superficial strategies, such as organizing self-feedback using bullet points, rather than high-impact approaches to self-assessment, such as redo (making or thinking of revisions; Panadero et al., 2024). The high frequency of feelings of difficulty that participants reported emphasize that they found self-assessment challenging: they reported feeling unsure about how to regulate the process, manage their time effectively, or respond to biases that may have shaped their self-assessments—despite

being university students who have spent many years in formal education and who are presumably relatively highly self-regulated. In previous studies, both students and teachers have identified a perceived lack of student self-knowledge as a limitation to supporting self-assessment (Kippers et al., 2018; Volante & Beckett, 2011). Our findings would suggest that, for university students, a lack of self-knowledge may not be the primary issue. Rather, a lack of metacognitive knowledge about self-assessment may be a more substantial limitation to reaping the benefits of student self-assessment. Our findings highlight the importance of teaching students to self-assess to help them control their own self-assessment processes.

Our findings also illuminate some of the metacognitive skills and knowledge that could support self-assessment. Educators could support student self-assessment by explicitly teaching students about the metacognitive experiences, skills, and knowledge identified in this study to help them monitor, interpret, and control their self-assessment processes. Self-assessment is a core aspect of feedback literacy, positioned as a space where students draw on and interpret internal and external feedback to evaluate their own performance and make decisions about learning (e.g., eliciting more feedback, making revisions; Yan & Carless, 2021). By teaching the processes identified in this study, educators can help students monitor and control their own thinking as they process and consolidate internal and external feedback to make such decisions and thereby foster students' feedback literacy.

Last, findings from this study have implications for using assessment to support self-regulated learning and co-regulation. Researchers have theorized that self-assessment supports learning when it engages students in subprocesses of self-regulated learning (Andrade, 2010; Brown & Harris, 2013). Theories and empirical findings have indicated several mechanisms whereby self-assessment may facilitate self-regulated learning. Self-assessment may foster self-

efficacy, a driver of self-regulated learning (Panadero et al., 2017; Ross, 2006), goal setting (Andrade & Brookhart, 2020), self-monitoring (Andrade, 2013), and knowledge of success criteria (Stobart, 2008). When educators invite students to self-assess and support them using guidance and resources, self-assessment can be seen as supporting co-regulated learning (Andrade & Brookhart, 2020). Self-assessment can also co-regulate teaching and learning when it affords opportunities for students to shape what is valued in a classroom, e.g., via criteria co-construction or feedback to teachers (Bourke & Mentis, 2013; Tanner & Jones, 1994). Findings from our study provide initial empirical insights into a mechanism whereby self-assessment may facilitate self-regulated learning and co-regulation via metacognition. Metacognition is a core dimension of self-regulated learning (Greene et al., 2024), and our findings indicate that self-assessment activates metacognitive monitoring of the self-assessment process via metacognitive experiences, especially feelings of difficulty, and metacognitive control of the self-assessment process via metacognitive skills, especially goal setting and articulation of the self-assessment process. By explicitly supporting these metacognitive processes when facilitating student self-assessment, teachers may enhance the co-regulatory effects of student self-assessment and students' development of self-regulated learning. For self-assessment to support the metacognitive dimension of co-regulation, teachers must also help students make their self-assessment metacognition visible. Students could do so using reflection journals, video diaries, or thinking aloud about their work for teachers. Making these processes visible is needed for teachers to be able to offer tailored supports (Nicol, 2021; Panadero, Lipnevich, et al., 2019) and can provide educators with concrete indicators of when students are taking an active role in assessment, a requisite for cultivating student-centered assessment (Gulikers et al., 2021).

Limitations and Future Research

As this is a qualitative study based on a small, albeit diverse, sample of university students, findings cannot be widely generalized. Our aim was not to identify generalizable trends; rather, with a focus on developing descriptive aspects of theory (Greene, 2022), we adopted a qualitative approach to capture rich data and account for individual differences (Charters, 2003). Future research with larger sample sizes and across education contexts is needed to validate and extend our findings. Given that our participants were university students studying abroad, many of whom were graduate students, they may exhibit relatively well-developed self-regulation and metacognition compared to other groups of students. In the think alouds, participants expressed varied amounts of experience and training with self-assessment which is unsurprising especially given how they came from a diverse range of education contexts and disciplines. Studies examining self-assessment metacognition in K-12 contexts could help illuminate the metacognitive processes of self-assessment in younger students and those with less experience in education or varied levels of self-assessment experience which all may shape self-assessment metacognition (see Andrade & Brown, 2016; Brown & Harris, 2014).

The context of self-assessment in this study also necessitates important considerations when interpreting our findings. How we conceptualized self-assessment is one such consideration. We adopted a formative perspective, where the purpose of self-assessment is to support learning and especially revisions to performance, in contrast to a summative perspective on self-assessment that only involves quantitative self-rating or self-grading (Andrade, 2019). Self-rating or self-grading, without the need to revise or develop qualitative self-feedback, might elicit different metacognitive processes. Research has indicated that low achieving students tend to be overconfident in their self-rating, while high achieving students tend to be underconfident

(Andrade, 2019; Hacker et al., 2000; Hacker & Bol, 2019). Future study of the metacognitive processes of self-assessment for summative purposes is therefore warranted to understand how self-assessment accuracy and confidence can be supported. However, we focused on formative self-assessment as we are primarily interested in how students generate self-feedback for improvement and because, pedagogically, self-rating may be less helpful for students' learning: it can rush students to a judgement without deep, generative analysis of their own performance (Andrade, 2019; Boud & Brew, 1995).

As a second contextual consideration, participants in our study had access to various onscreen resources to help them self-assess (i.e., rubric, contrasting exemplars, essay instructions, self-feedback space). The different resources students have access to could also shape the metacognitive processes that emerge. These contextual considerations highlight that self-assessment happens within a socio-cultural context that intrinsically shapes it (Andrade & Brown, 2016). We examined self-assessment metacognition primarily from a psychological perspective, collecting data via think alouds in a lab setting to uncover cognitive processes. While such psychological perspectives are important, they are limited to what occurs inside students' minds (Shepard & Penuel, 2018); they therefore need to be supplemented with socio-cultural perspectives to elucidate how social and cultural environments shape psychological phenomena. Ethnographic and longitudinal observational studies are needed to supplement the findings from this research and illuminate situative perspectives (Nolen, 2024) on how students and educators co-construct self-assessment processes over time using different resources and across diverse learning contexts. Writing has been conceptualized as a fundamentally metacognitive process—a view supported by the wide variety of metacognitive experiences and skills our participants articulated—that is subject to developmental stages (Hacker, 2018).

Research is therefore needed to verify if the metacognitive processes of self-assessment we identified are transferrable across subject contexts (e.g., math, arts) and developmental stages. Such research would continue to extend the descriptive aspects of self-assessment theory which are needed to inform explanatory aspects.

Conclusion

Despite these limitations, this study provides an initial understanding of the metacognitive processes of student self-assessment via the first empirical data of metacognitive experiences and skills—i.e., self-assessment actions which have the students' own self-assessment process as the object of focus—captured during the self-assessment process. It thereby makes an important contribution to descriptive aspects of self-assessment theory, elucidating the metacognitive dimension of self-assessment for the first time and highlighting areas of future research, and contributing to ongoing efforts to open the black box of self-assessment (Panadero et al., 2024). Our findings highlight the particular importance of feelings of difficulty, selecting a strategy, and goal setting/reasoning as metacognitive processes shaping self-assessment. While self-assessment's role as a subprocess of self-regulated learning has been empirically examined (Panadero et al., 2017; Yan, 2020b), this study offers an initial bridge between self-assessment and metacognition research—an underexplored link that stands to support more comprehensive understandings of how self-assessment relates to learning. To support self-assessment in education, mapping the various dimensions of the process of self-assessment remains central to the research agenda.

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

Andrade, H. L. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In H. L. Andrade & G. Cizek (Eds.), *Handbook of formative assessment* (pp. 90–105). Routledge.

Andrade, H. L. (2013). Classroom assessment in the context of learning theory and research. In J. H. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 17–34). Sage. http://ebookcentral.proquest.com/lib/queen-ebooks/detail.action?docID=1207749.

Andrade, H. L. (2019). A critical review of research on student self-assessment. *Frontiers in Education*, *4*, 87. https://doi.org/10.3389/feduc.2019.00087

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

Andrade, H. L., & Brown, G. T. L. (2016). Student self-assessment in the classroom. In G. T. L. Brown & L. Harris (Eds.), *Handbook of human and social conditions in assessment* (pp. 319–334). Routledge.

Andrade, H. L., Wang, X., Du, Y., & Akawi, R. (2009). Rubric-referenced self-assessment and self-efficacy for writing. *Journal of Educational Research*, *102*(4), 287–302. https://doi.org/10.3200/JOER.102.4.287-302

Beach, P., & Willows, D. (2017). Understanding teachers' cognitive processes during online professional learning: A methodological comparison. *Online Learning Journal*, *21*(1), 60–84. https://doi.org/10.24059/olj.v21i1.949

Beheshitha, S. S., Gašević, D., & Hatala, M. (2015). A process mining approach to linking the study of aptitude and event facets of self-regulated learning. *ACM International Conference Proceeding Series*, *16-20-March-2015*, 265–269. https://doi.org/10.1145/2723576.2723628

Boekaerts, M. (1991). Subjective competence, appraisals and self-assessment. *Learning and Instruction*, *1*, 1–17.

Boud, D., & Brew, A. (1995). Developing a typology for learner self-assessment practices. *Research and Development in Higher Education*, *18*, 130–135.

Bourke, R. (2016). Liberating the learner through self-assessment. *Cambridge Journal of Education*, *46*(1), 97–111. https://doi.org/10.1080/0305764X.2015.1015963

Bourke, R., & Mentis, M. (2013). Self-assessment as a process for inclusion. *International Journal of Inclusive Education*, *17*(8), 854–867. https://doi.org/10.1080/13603116.2011.602288

Brandmo, C., Panadero, E., & Hopfenbeck, T. N. (2020). Bridging classroom assessment and self-regulated learning. *Assessment in Education: Principles, Policy and Practice*, *27*(4), 319–331. https://doi.org/10.1080/0969594X.2020.1803589

Braund, H., & DeLuca, C. (2018). Elementary students as active agents in their learning: An empirical study of the connections between assessment practices and student metacognition. *Australian Educational Researcher*, *45*(1), 65–85. https://doi.org/10.1007/s13384-018-0265-z

British Columbia Ministry of Education. (2023). *A Framework for Classroom Assessment*.

Brown, G. T. L., & Harris, L. (2013). Student self-assessment. In J. H. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 367–393). Sage.

Brown, G. T. L., & Harris, L. (2014). The future of self-assessment in classroom practice: Reframing self- assessment as a core competency. *Frontline Learning Research*, *2*(1), 22–30. https://doi.org/10.14786/flr.v2i1.24

Bryer, J., Andrade, H. L., & Cleary, T. (2025). *Writing*. https://docs.daacs.net/overview/assessments/writing

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

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

Campbell, J. L., Quincy, C., Osserman, J., & Pedersen, O. K. (2013). Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods and Research*, *42*(3), 294–320. https://doi.org/10.1177/0049124113500475

Charters, E. (2003). The use of think-aloud methods in qualitative research: An introduction to think-aloud methods. *Brock Education*, *12*(2), 68–82.

Chevrier, M., Muis, K. R., Trevors, G. J., Pekrun, R., & Sinatra, G. M. (2019). Exploring the antecedents and consequences of epistemic emotions. *Learning and Instruction*, *63*, 101209. https://doi.org/10.1016/j.learninstruc.2019.05.006

Collins, J., Regenbrecht, H., Langlotz, T., Can, Y. S., Ersoy, C., & Butson, R. (2019). Measuring cognitive load and insight: A methodology exemplified in a virtual reality learning context. *Proceedings - 2019 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2019*, 351–362. https://doi.org/10.1109/ISMAR.2019.00033

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

D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction*, *29*, 153–170. https://doi.org/10.1016/j.learninstruc.2012.05.003

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist*, *13*(4), 277–287. https://doi.org/10.1027/1016-9040.13.4.277

Efklides, A. (2017). Affect, epistemic emotions, metacognition, and self-regulated learning. *Teachers College Record*, *119*, 130305.

Fastré, G. M. J., van der Klink, M. R., Sluijsmans, D., & van Merriënboer, J. J. G. (2012). Drawing students' attention to relevant assessment criteria: Effects on self-assessment skills and performance. *Journal of Vocational Education and Training*, *64*(2), 185–198. https://doi.org/10.1080/13636820.2011.630537

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, *34*(10), 906–911.

Graham, S., & Harris, K. R. (2018). Evidence-based writing practices: A meta-analysis of existing meta-analyses. In R. Fidalgo, T. Olive, K. R. Harris, & M. Braaksma (Eds.),

Studies in writing series: design principles for teaching effective writing (Vol. 34, pp. 13–37). Brill Academic Publishers. https://doi.org/10.1163/9789004270480_003

Greene, J. A. (2022). What can educational psychology learn from, and contribute to, theory development scholarship? *Educational Psychology Review*, *34*(4), 3011–3035. https://doi.org/10.1007/s10648-022-09682-5

Greene, J. A., Bernacki, M. L., & Hadwin, A. F. (2024). Self-regulation. In P. A. Schultz & K. R. Muis (Eds.), *Handbook of educational psychology* (4th ed., pp. 314–334). Routledge. https://doi.org/10.4324/9780429433726-17

Gulikers, J., Veugen, M., Baartman, L., Beekman, K., Brookhart, S. M., & Tay, H. Y. (2021). What are we really aiming for? Identifying concrete student behavior in co-regulatory formative assessment processes in the classroom. *Frontiers in Education*, *6*, 750281. https://doi.org/10.3389/feduc.2021.750281

Hacker, D. J. (2018). A Metacognitive Model of Writing: An Update From a Developmental Perspective. *Educational Psychologist*, *53*(4), 220–237. https://doi.org/10.1080/00461520.2018.1480373

Hacker, D. J., & Bol, L. (2019). Calibration and self-regulated learning: Making the connections. In J. Dunlosky & K. Rawson (Eds.), *Cambridge handbook on cognition and education* (pp. 647–677). Cambridge University Press.

Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology*, *92*(1), 160–170. https://doi.org/10.1037/0022-0663.92.1.160

Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning*, *2*, 107–124. https://doi.org/10.1007/s11409-007-9016-7

Händel, M., & Fritzsche, E. S. (2016). Unskilled but subjectively aware: Metacognitive monitoring ability and respective awareness in low-performing students. *Memory and Cognition*, *44*(2), 229–241. https://doi.org/10.3758/s13421-015-0552-0

Hartelt, T., & Martens, H. (2024). Self-regulatory and metacognitive instruction regarding student conceptions: Influence on students' self-efficacy and cognitive load. *Frontiers in Psychology*, *15*. https://doi.org/10.3389/fpsyg.2024.1450947

Johnson, C. C., Sondergeld, T. A., & Walton, J. B. (2019). A study of the implementation of formative assessment in three large urban districts. *American Educational Research Journal*, *56*(6), 2408–2438. https://doi.org/10.3102/0002831219842347

Kippers, W. B., Wolterinck, C. H. D., Schildkamp, K., Poortman, C. L., & Visscher, A. J. (2018). Teachers' views on the use of assessment for learning and data-based decision making in classroom practice. *Teaching and Teacher Education*, *75*, 199–213. https://doi.org/10.1016/j.tate.2018.06.015

Kostons, D., Van Gog, T., & Paas, F. (2009). How do I do? Investigating effects of expertise and performance-process records on self-assessment. *Applied Cognitive Psychology*, *23*(9), 1256–1265. https://doi.org/10.1002/acp.1528

Krebs, R., Rothstein, B., & Roelle, J. (2022). Rubrics enhance accuracy and reduce cognitive load in self-assessment. *Metacognition and Learning*, *17*, 627–650. https://doi.org/10.1007/s11409-022-09302-1

Lee, H., Chung, H. Q., Zhang, Y., Abedi, J., & Warschauer, M. (2020). The effectiveness and features of formative assessment in US K-12 education: A systematic review. *Applied Measurement in Education*, *33*(2), 124–140. https://doi.org/10.1080/08957347.2020.1732383

Lin-Siegler, X., Shaenfield, D., & Elder, A. D. (2015). Contrasting case instruction can improve self-assessment of writing. *Educational Technology Research and Development*, *63*(4), 517–537. https://doi.org/10.1007/s11423-015-9390-9

Lipnevich, A. A., Panadero, E., & Calistro, T. (2023). Unraveling the effects of rubrics and exemplars on student writing performance. *Journal of Experimental Psychology: Applied*, *29*(1), 136–148. https://doi.org/10.1037/xap0000434

Malecka, B., Boud, D., & Carless, D. (2022). Eliciting, processing and enacting feedback: Mechanisms for embedding student feedback literacy within the curriculum. *Teaching in Higher Education*, *27*(7), 908–922. https://doi.org/10.1080/13562517.2020.1754784

Manitoba Education. (2006). *Rethinking classroom assessment with purpose in mind*.

Marra, R. M., Hacker, D. J., & Plumb, C. (2022). Metacognition and the development of self-directed learning in a problem-based engineering curriculum. *Journal of Engineering Education*, *111*(1), 137–161. https://doi.org/10.1002/jee.20437

McIntyre, N. A., Draycott, B., & Wolff, C. E. (2022). Keeping track of expert teachers: Comparing the affordances of think-aloud elicited by two different video perspectives. *Learning and Instruction*, *80*, 101563. https://doi.org/10.1016/j.learninstruc.2021.101563

Munzar, B., Muis, K. R., Denton, C. A., & Losenno, K. (2021). Elementary students' cognitive and affective responses to impasses during mathematics problem solving. *Journal of Educational Psychology*, *113*(1), 104–124. https://doi.org/10.1037/edu0000460

Nelson, T., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125–173). Academic Press.

Nicol, D. (2021). The power of internal feedback: Exploiting natural comparison processes. *Assessment and Evaluation in Higher Education*, *46*(5), 756–778. https://doi.org/10.1080/02602938.2020.1823314

Nolen, S. B. (2024). Taking a situative approach to research in educational psychology. In P. A. Schultz & K. R. Muis (Eds.), *Handbook of educational psychology* (4th ed., pp. 119–138). Routledge. https://doi.org/10.4324/9780429433726-8

OECD. (2019). *OECD future of education and skills 2030: OECD learning compass 2030: A series of concept notes*. https://www.oecd.org/education/2030-project/contact/OECD_Learning_Compass_2030_Concept_Note_Series.pdf

Ontario Ministry of Education. (2010). *Growing success: Assessment, evaluation, and reporting in Ontario schools*.

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., & Alonso-Tapia, J. (2013). Student self-assessment: Theoretical and practical connotations. When it happens, how it is acquired, and how to develop it in our students. *Electronic Journal of Research in Education Psychology*, *11*(30), 551–576. https://doi.org/10.14204/ejrep.30.12200

Panadero, E., Brown, G., & Courtney, M. (2014). Teachers' reasons for using self-assessment: A survey self-report of Spanish teachers. *Assessment in Education: Principles, Policy and Practice*, *21*(4), 365–383. https://doi.org/10.1080/0969594X.2014.919247

Panadero, E., Brown, G. T. L., & Strijbos, J. W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review*, *28*(4), 803–830. https://doi.org/10.1007/s10648-015-9350-2

Panadero, E., Fernández, J., Pinedo, L., Sánchez, I., & García-Pérez, D. (2024). A self-feedback model (SEFEMO): secondary and higher education students' self-assessment profiles. *Assessment in Education: Principles, Policy and Practice*, *31*(3–4), 221–253. https://doi.org/10.1080/0969594X.2024.2367027

Panadero, E., Fernández-Ruiz, J., & Sánchez-Iglesias, I. (2020). Secondary education students' self-assessment: The effects of feedback, subject matter, year level, and gender. *Assessment in Education: Principles, Policy & Practice*, *27*(6), 607–634. https://doi.org/10.1080/0969594x.2020.1835823

Panadero, E., Fraile, J., Fernández Ruiz, J., Castilla-Estévez, D., & Ruiz, M. A. (2019). Spanish university assessment practices: Examination tradition with diversity by faculty. *Assessment and Evaluation in Higher Education*, *44*(3), 379–397. https://doi.org/10.1080/02602938.2018.1512553

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, *22*, 74–98. https://doi.org/10.1016/j.edurev.2017.08.004

Panadero, E., Jonsson, A., Pinedo, L., & Fernández-Castilla, B. (2023). Effects of rubrics on academic performance, self-regulated learning, and self-efficacy: A meta-analytic review. *Educational Psychology Review*, *35*(4). https://doi.org/10.1007/s10648-023-09823-4

Panadero, E., Lipnevich, A., & Broadbent, J. (2019). Turning self-assessment into self-feedback. In *The Impact of Feedback in Higher education: Improving assessment outcomes for*

learners (pp. 147–163). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-25112-3_9

Panadero, E., Pérez, D. G., Ruiz, J. F., Fraile, J., Sánchez-Iglesias, I., & Brown, G. T. L. (2022). University students' strategies and criteria during self-assessment: Instructor's feedback, rubrics, and year level effects. *European Journal of Psychology of Education*, *58*, 1031–1051. https://doi.org/10.1007/s10212-022-00639-4

Reber, R., & Greifeneder, R. (2017). Processing fluency in education: How metacognitive feelings shape learning, belief formation, and affect. *Educational Psychologist*, *52*(2), 84–103. https://doi.org/10.1080/00461520.2016.1258173

Rickey, N., DeLuca, C., & Beach, P. (2023). Towards a new theory of student self-assessment: Tracing learners' cognitive and affective processes. *Metacognition and Learning*, *18*, 945–981. https://doi.org/10.1007/s11409-023-09359-6

Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research, and Evaluation*, *11*, 10. https://doi.org/10.7275/9wph-vv65

Schellings, G. L. M., Van Hout-Wolters, B. H. A. M., Veenman, M. V. J., & Meijer, J. (2013). Assessing metacognitive activities: The in-depth comparison of a task-specific questionnaire with think-aloud protocols. *European Journal of Psychology of Education*, *28*(3), 963–990. https://doi.org/10.1007/s10212-012-0149-y

Shepard, L. A., & Penuel, W. R. (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.

Stobart, G. (2008). *Testing times: The uses and abuses of assessment*. Routledge.

Talsma, K., Schüz, B., & Norris, K. (2019). Miscalibration of self-efficacy and academic performance: Self-efficacy ≠ self-fulfilling prophecy. *Learning and Individual Differences*, *69*, 182–195. https://doi.org/10.1016/j.lindif.2018.11.002

Tanner, H., & Jones, S. (1994). Using peer and self-assessment to develop modelling skills with students aged 11 to 16: A socio-constructive view. *Educational Studies in Mathematics*, *27*(4), 413–431.

Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, *27*(2), 237–246. https://doi.org/10.1177/1098214005283748

To, J., Panadero, E., & Carless, D. (2021). A systematic review of the educational uses and effects of exemplars. *Assessment and Evaluation in Higher Education*. https://doi.org/10.1080/02602938.2021.2011134

van Gog, T., Paas, F., van Merriënboer, J. J. G., & Witte, P. (2005). Uncovering the problem-solving process: cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology. Applied*, *11*(4), 237–244. https://doi.org/10.1037/1076-898X.11.4.237

Veugen, M. J., Gulikers, J. T. M., & den Brok, P. (2021). We agree on what we see: Teacher and student perceptions of formative assessment practice. *Studies in Educational Evaluation*, *70*, 101027. https://doi.org/10.1016/j.stueduc.2021.101027

Volante, L., & Beckett, D. (2011). *Formative assessment and the contemporary classroom: Synergies and tensions between research and practice*. *34*(2), 239–255. https://doi.org/10.2307/canajeducrevucan.34.2.239

Winne, P. H. (2017). Cognition and metacognition within self-regulated learning. In P. A. Alexander, D. H. Schunk, & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 36–48). Routledge. https://doi.org/10.4324/9781315697048.ch3

Winne, P. H. (2020). Commentary: A proposed remedy for grievances about self-report methodologies. *Frontline Learning Research*, *8*(3), 164–173. https://doi.org/10.14786/flr.v8i3.625

Winne, P. H., & Azevedo, R. (2014). Metacognition. In *The Cambridge Handbook of the Learning Sciences, Second Edition* (pp. 63–87). Cambridge University Press. https://doi.org/10.1017/CBO9781139519526.006

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

Winne, P. H., Teng, K., Chang, D., Lin, M. P. C., Marzouk, Z., Nesbit, J. C., Patzak, A., Rakovic, M., Samadi, D., & Vytasek, J. (2019). NStudy: Software for learning analytics about learning processes and self-regulated learning. *Journal of Learning Analytics*, *6*(2), 95–106. https://doi.org/10.18608/jla.2019.62.7

Yan, Z. (2020a). Developing a short form of the self-assessment practices scale: Psychometric evidence. *Frontiers in Education*, *4*, 153. https://doi.org/10.3389/feduc.2019.00153

Yan, Z. (2020b). Self-assessment in the process of self-regulated learning and its relationship with academic achievement. *Assessment and Evaluation in Higher Education*, *45*(2), 224–238. https://doi.org/10.1080/02602938.2019.1629390

Yan, Z., & Brown, G. T. L. (2017). A cyclical self-assessment process: Towards a model of how students engage in self-assessment. *Assessment and Evaluation in Higher Education*, *42*(8), 1247–1262. https://doi.org/10.1080/02602938.2016.1260091

Yan, Z., Brown, G. T. L., Lee, J. C. K., & Qiu, X. L. (2020). Student self-assessment: Why do they do it? *Educational Psychology*, *40*(4), 509–532. https://doi.org/10.1080/01443410.2019.1672038

Yan, Z., & Carless, D. (2021). Self-assessment is about more than self: The enabling role of feedback literacy. *Assessment & Evaluation in Higher Education*, *47*(7), 1116–1128. https://doi.org/10.1080/02602938.2021.2001431

Yan, Z., Lao, H., Panadero, E., Fernández-Castilla, B., Yang, L., & Yang, M. (2022). Effects of self-assessment and peer-assessment interventions on academic performance: A pairwise and network meta-analysis. *Educational Research Review*, *37*, 100484. https://doi.org/10.1016/j.edurev.2022.100484

Yan, Z., Lee, J. C.-K., Hui, S. K. F., & Lao, H. (2022). Enhancing students' self-efficacy in creativity and learning performance in the context of English learning: The use of self-assessment mind maps. *Frontiers in Psychology*, *13*, 871781. https://doi.org/10.3389/fpsyg.2022.871781

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

Supplemental Material

Participant Demographics

Variable	Frequency (/25)	Percentage	Variable	Frequency (/25)	Percentage
Age			**Gender**		
20	1	4.0	Female	16	64.0
21	6	24.0	Male	8	32.0
22	4	16.0	Genderqueer	1	4.0
23	2	8.0	**Discipline**		
24	1	4.0	Humanities	3	12.0
25	1	4.0	Sciences	1	4.0
26	1	4.0	Engineering	4	16.0
28	2	8.0	Social Sciences	6	24.0
31	1	4.0	Languages	2	8.0
32	2	8.0	Business	3	12.0
33	1	4.0	Communications	1	4.0
34	1	4.0	Design	1	4.0
37	1	4.0	Energy	1	4.0
47	1	4.0	Graphic Design	1	4.0
Home Country			**First Language**		
Australia	1	4.0	English	10	40.0
Belgium	1	4.0	French	1	4.0
Canada	2	8.0	Korean	1	4.0
Croatia	1	4.0	Urdu	1	4.0
France	1	4.0	Turkish	1	4.0

Georgia	1	4.0	Hungarian	1	4.0
Germany	3	12.0	Dutch	1	4.0
Ghana	1	4.0	German	3	12.0
Hungary	1	4.0	Georgian	1	4.0
Iran	1	4.0	Croatian	1	4.0
Netherlands	1	4.0	Cebuano	1	4.0
Nigeria	1	4.0	Persian	1	4.0
Pakistan	1	4.0	Ukrainian	1	4.0
Philippines	1	4.0	Yoruba	1	4.0
South Korea	1	4.0	**Race**		
Turkey	1	4.0	Asian/Asian American	2	8.0
Ukraine	1	4.0	Black/African American	2	8.0
United Kingdom	2	8.0	Latinx	2	8.0
United States of America	2	8.0	White/European	15	60.0
Puerto Rico	1	4.0	Prefer not to say/Don’t know	4	16.0

Analytic Writing Rubric Which Participants Could Open On The Screen

	Developing	Emerging	Mastering
Content	Ideas are general statements about the writer's experience without specific, personalized insights.	Ideas are general statements with some examples or anecdotes to illustrate the ideas.	Ideas are clear and compelling. The writer offers a well-developed and highly personal interpretation.
Organization & Structure	Individual ideas are clear, but the progression of one idea to the next is unclear.	The essay features a short introduction, at least one body paragraph, and a short conclusion.	The essay is well-organized, with an order and structure that present the discussion effectively.
Paragraphs	Paragraphs abruptly shift from one idea to the next, or there is only one paragraph. Paragraphs have several topics. Sentences contain independent ideas.	Transitions between paragraphs are simple. Paragraphs focus on one or two main ideas. A limited range of linking words are used to connect sentences within paragraphs.	Transitions between paragraphs strengthen the progression of the essay. Paragraphs are consistently and clearly focused on a main idea. Sentences are seamlessly linked using a variety of types of words.
Sentences	Sentences are often very short or long without a clear effect. Sentences follow a basic subject-verb-object pattern. Sentence lengths are all similar.	Sentences are generally correct. There may be minor agreement errors (e.g., "the information in each section are …"). One type of sentence is primarily used throughout. Fragments are used for effect.	Consistent and appropriate use of a variety of sentence structures and lengths to achieve clear effects (e.g., complex sentences to create a sense of rushing).
Conventions	Capital letter errors and/or spelling errors make meaning unclear. Full stops/period errors cause the reader to pause to decipher meaning.	Minor spelling, capital letter and/or punctuation errors are noticeable but do not distract from the meaning. Punctuation is used to ensure that meaning is clear.	Spellings and capital letters are used correctly. Punctuation is used to convey subtle meanings (e.g., semicolon used to contrast two parallel ideas).

Mastering Exemplar Which Participants Could Open On The Screen

Writing Exemplar: Mastering

Living in a new country was, for me, a life-changing experience. In 2014, I moved from my home country of Canada to teach secondary school in the United Kingdom. Working as a teacher afforded lots of opportunities for travel: teachers in England get a 1-2 week holiday every six weeks, as well as a 5-6-week summer. I travelled extensively during my time living in England, visiting many countries in Europe (e.g., Iceland, Switzerland, Slovenia), North Africa (e.g., Morocco, Egypt), and Asia (e.g., Singapore, Indonesia). While traveling and living abroad was incredibly fun, it came with certain challenges that only those who have lived abroad can fully understand, such as acculturating to new belief and behaviour systems. I am very grateful, however, for these challenges. Navigating them has helped me better understand the world—and the people in it—and provided me with invaluable life lessons.

One of the first things I learned about living abroad was that subtle differences make a big difference. Moving to another English-speaking country—and one that established the original colonies in Canada—I didn't expect to experience much culture shock. However, from day one, the subtlest of differences were glaringly obvious. In Canada, we drive on the right side of the road; in the UK, we drive on the left. In Canada, there are lots of coniferous trees; in the UK, the trees are mainly deciduous. The sun rises slightly later and sets slightly earlier in the UK than in Canada, casting a different shade of light than I was used to. I often felt like I was living in an alternate, parallel universe. While these differences fostered an initial sense of shock, they ultimately caused me to slow down and pay attention to the little things. I noticed the types of trees that surrounded me; their scents, heights, and leaves, things that I would not have noticed had I stayed in Canada. The differences cultivated presence. I got to deeply notice and experience the beauty of a new place.

Acculturating to the UK proved more challenging in some ways […]

Ultimately, the differences I experienced between my home and my adopted home were at the core of my experiences living abroad. The minor differences highlighted, for me, the important similarities, i.e., core values of fun and kindness. Additionally, subtle differences in the natural world gave me a new appreciation for the beauty of Ontario—the unique way the light hits the lakes in the summer, the rugged beauty of the Canadian shield, the density of Ontario forests, all features which became more beautiful when contrasted with the rolling hills, jagged sea cliffs, and gentle yet breathtaking landscapes of England. Living abroad helped me understand other lands and people better, which in turn helped me understand myself and my home better. Living abroad enriched my life and is something I highly recommend to others.

Developing Exemplar Which Participants Could Open On The Screen

Writing Exemplar: Developing

I learned alot from my time living abroad. I went from Canada to England in 2014. I experienced challenges and wonderful experiences in England. the challenges made me develop key skills that have served me well. I noticed that the type of humour was different in the UK there they liked to wind people up. one skill I learned there was how to be patient with people to understand their perspectives, this was an important skill because it made me kinder and more understanding. I had lots of fun while living abroad I made friends and tried good food.

I traveled alot when I lived in Englad, flights to places in Europe are very cheap there. I learned about different cultures when I lived abroad. I made friends in England and still visit them sometimes. the weather there is rainy compared to Canada. But Canada is worse in the winter. in the future, I will do more research to know what to pack when I move somewhere. overall, I really enjoyed my time in England and learned alot it served me well in my life for lots of reasons.

Chapter 4

Student Self-Assessment Emotions:

Investigating the Role of Emotions in Self-Assessment

Abstract

Self-assessment is fundamental to student success. However, the emotions that learners experience while engaged in self-assessment are understudied. This study investigated (a) the self-assessment actions students employed during self-assessment, (b) the types, frequencies, and intensity of emotions students experienced while self-assessing, and (c) if certain emotions were likely to transition to certain self-assessment actions. To capture learners' self-assessment actions and emotions, we engaged international university students (N=25) in a self-assessment activity followed by a retrospective, gaze-cued virtual revisit think and emote aloud protocol with emotions questionnaire. Think-aloud data were analyzed using a hybrid approach and questionnaire data were analyzed descriptively. We used a table to triangulate questionnaire and think-aloud data to illuminate the types, frequencies, and intensities of emotions participants experienced during self-assessment. We used state-transition analysis to identify emotions to self-assessment action transitions that occurred statistically more frequently than chance. Through our analysis, we identified 11 self-assessment actions and eight types of emotions. Participants experienced a range of emotions at low to moderate intensities. Enjoyment and confusion were the most frequent emotions, while curiosity and enjoyment were the most intensely felt emotions. State-transition analysis revealed significant transitions between enjoyment and surprise and assessing with internal criteria (e.g., gut reactions), and between anxiety and frustration and metacognitive experiences (e.g., feelings of difficulty). The

emotional dynamics of self-assessment emerged as a cycle of detecting and resolving impasses, which has practical implications for supporting student self-assessment in classrooms.

Introduction

Students' capacity to self-assess their performance is one of the top educational factors influencing achievement in schools (Hattie, 2023) and meta-analyses have repeatedly demonstrated positive effects of facilitating student self-assessment in classrooms on students' academic achievement (Lee et al., 2020; Yan et al., 2022). Despite researchers' and policymakers' calls for educators to cultivate student self-assessment in classrooms (e.g., Birenbaum et al., 2015; Malecka et al., 2022; Nicol & MacFarlane-Dick, 2006), research continues to show that fostering and supporting student self-assessment in classrooms is uncommon across education contexts and teaching divisions (Brandmo et al., 2020; Brooks et al., 2021; Johnson et al., 2019; Kippers et al., 2018; Veugen et al., 2021). One key barrier to facilitating student self-assessment is that research has yet to fully examine the emotions students experience during self-assessment and how those emotions shape the self-assessment process (Andrade, 2019; Nicol, 2021). Understanding the emotions that may arise during self-assessment and the responses these emotions are likely to provoke can help teachers—and students themselves—develop students' self-assessment processes through tailored guidance, interventions, and feedback.

The aim of this study was to advance current theories of how students self-assess (Panadero et al., 2024) by investigating emotions students experienced when engaged in self-assessment and patterns in how these emotions provoked follow up self-assessment actions. This research enhances the relevance of self-assessment theories to real-world learning situations, which are inherently shaped by emotions, and generates practical implications for educators.

Defining and Mapping the Self-Assessment Process

Student self-assessment "involves a wide variety of mechanisms and techniques through which students describe (i.e., assess) and possibly assign merit or worth to (i.e., evaluate) the qualities of their own learning processes and products" (Panadero et al., 2016, p. 804). From a psychological perspective, self-assessment is thereby conceptualized as a process that involves internal cognition and behavioural actions. Self-assessment is a prerequisite for self-regulated learning—i.e., the process of students managing their own cognition, emotions, motivation, and behaviour to achieve goals—as it enables subprocesses of forethought (e.g., goal setting, efficacy judgements), performance (e.g., self-monitoring, control), and reflection (e.g., self-reflection, goal calibration; Boekaerts, 1991). Self-assessment can be initiated by students independently or explicitly facilitated by educators to support student metacognition and cultivate co-regulation of learning between teachers, peers, and learning resources (Andrade & Brookhart, 2020; Panadero & Alonso-Tapia, 2013). When engaged to support learning, an important part of the self-assessment process involves the opportunity for students to revise their work or performance which activates and reinforces deeper self-regulated learning (Andrade, 2019). In addition to these psychological benefits, facilitating student self-assessment can also have important socio-cultural benefits, such as giving students more control of their learning and what is valued in assessment, supporting more inclusive classrooms and agentic learners (Bourke & Mentis, 2013; Tan, 2007).

Given these important learning benefits, there has been an increasing interest in mapping the self-assessment process to inform classroom interventions that can support effective student self-assessment (Panadero et al., 2020, 2022, 2024; Yan & Brown, 2017). Panadero et al. (2020, 2022), drawing on extensive analysis of think-aloud and video recording data from students in

secondary and university education, developed a coding framework of self-assessment strategies and criteria. Strategies are specific, content-dependent actions students deploy during self-assessment (e.g., compare essay to rubric); criteria are standards participants refer to when evaluating their work which can range in complexity (e.g., intuition, rubric criteria). Strategies are grouped in six actions that can elucidate distinct profiles of self-assessment: Read, Recall, Compare, Rate, Assess, and Redo. Assess and Redo can be performed at two levels of complexity. In Assess, students explicitly draw on criteria to inform a judgement about their work. At the first level (i.e., Assess Level 1) students draw on internal criteria, like intuition and hindsight, while at the second level (i.e., Assess Level 2) they use external criteria, such as those in the rubric or shared with them by a teacher. Redo involves participants revising their work or redoing a task, either to improve their work or check their answer. At the first level (i.e., Redo Level 1), they describe changes they would make without making them, whereas at the second level (i.e., Redo Level 2) they actually make the changes or redo their work.

This research recently culminated in the publication of the self-feedback model (SEFEMO; Panadero et al., 2024) which, based on analysis of hundreds of secondary and university students' self-assessment processes in three different subjects (Spanish, mathematics, writing), conceptualized the self-assessment process as comprised of three overarching phases. Information processing involves the student activating available information to help them self-assess (i.e., Read and Recall), including readily accessible materials (e.g., [re]read their essay, [re]read the rubric]) and those stored in memory (e.g., recall their essay). Analysis involves the student analyzing different sources of information about their work (i.e., Compare, Assess, and Redo). They may compare their work to a rubric or exemplar, or explicitly refer criteria which may be external (e.g., in a rubric) or internal (e.g., intuition). The analysis phase may involve the

student redoing or revising their work, either by thinking about how they would approach the task different next time or making actual changes. Panadero et al. (2024) argue that Redo "represents the most effective form of self-feedback: creating a new performance while simultaneously comparing and assessing the previous one" (p. 27) and note that engaging in the redo/revision process is more important than the revisions made—e.g., a student may redo a math exercise to confirm the original answer was correct. The last phase, self-assessment judgement, involves the student synthesizing everything they learned in the previous two phases to make a reasoned judgement about their performance which ideally includes decisions about the next steps in their learning (i.e., Rate). Predicated on process data collected in real-time while students self-assessed, this framework provides a fine-grained understanding of how students self-assess within writing contexts.

Self-Assessment Emotions and Control Value Theory

While students are learning, the emotions they experience influence the strategies they employ (Boekaerts, 2010) and the information they store in memory (Genc Aksaray & Ozcelik, 2022; Pekrun & Perry, 2014; Winne, 2018). Self-assessment may provoke a wide range of emotions because it has consequences for learners' self-concept (Baca & Smith, 2018). Emotions triggered by self-assessment are theorized to direct a learner's attention (Belkaid et al., 2017), which supports cognitive processing of information (Stark et al., 2018). Moreover, a meta-analysis from the discipline of management indicated that self-assessment in educational and professional contexts was more closely associated with affective learning outcomes (i.e., motivation and learner satisfaction) than cognitive ones (i.e., academic achievement), even in environments where self-assessment was scaffolded via explicit guidance and feedback (Sitzmann et al., 2010)—further signaling a connection between self-assessment and emotions.

Together, these perspectives implicate emotions as an important part of the self-assessment process that fundamentally shape how learners self-assess.

We see control-value theory (Pekrun, 2006; Pekrun & Linnenbrink-Garcia, 2014) as a valuable lens for interpreting emotions during self-assessment because it has been used to examine how emotions shape self-regulated learning (Chevrier et al., 2019; Di Leo et al., 2019; D'Mello et al., 2014), a process self-assessment is embedded throughout (Andrade, 2010; Boekaerts, 1991). Control-value theory can therefore offer insights into how emotions shape self-assessment that crosscut self-regulated learning research. Control-value theory proposes that learners' perceptions of control (i.e., over an achievement task and its outcomes) and task value (i.e., the subjective value of the achievement task and its outcomes, including the perceived importance of success) shape the emotions that arise during learning. It categorizes emotions based on valence (positive or negative) and activation (activating or deactivating). Activating emotions (e.g., curiosity, anxiety) may provoke strategy use while deactivating emotions (e.g., contentment, boredom) may provoke disengagement. Control-value theory further distinguishes emotions' objects of focus: social emotions are focused on other learners (e.g., envy, admiration); topic emotions relate to the topic being learned about; achievement emotions arise from learners' sense of success or failure in an achievement context; and epistemic emotions arise from information-oriented appraisals about the alignment or misalignment between new information and existing beliefs, existing knowledge structures, or recently processed information (Muis et al., 2018) as the object focus is on knowledge and the generation of knowledge and related activities (Pekrun et al., 2017). These emotions have antecedents (e.g., epistemic beliefs, perceptions of control, perceptions of task value, epistemic aims) and

consequences (e.g., use of a particular strategy, metacognitive monitoring of information; disengagement from a task; Muis et al., 2018).

Although there is a rich literature on the role of emotions during learning (see Frenzel et al., 2024; Pekrun, 2024) the emotions learners experience during self-assessment—and the roles these play in the self-assessment process—are underexplored (Andrade, 2019; Nicol, 2021). Very few studies have empirically examined the relationship between self-assessment and emotions. Panadero et al. (2023) investigated the effects of feedback occasion and type and year level on university students' emotions when they self-assessed. Based on participants' responses to an emotions questionnaire before self-assessment, after self-assessment without feedback, and after self-assessment with feedback, the researchers found that feedback in the form of a rubric had the strongest positive effect on emotions—i.e., reduced negative emotions and increased positive emotions. Additionally, Panadero et al., (2014) found that rubrics were particularly effective in reducing students' anxiety and negative thoughts during self-assessment, likely due to their structured and transparent nature, which contrasted with scripts that occasionally triggered performance-related stress. While these findings offer valuable insights for educators, they are predicated on questionnaire data collected when self-assessment was not occurring, meaning they offer limited insight into how emotions emerge within the process of self-assessment and how they may influence the process. Understanding the type of emotions that emerge during self-assessment and if certain emotions are likely to transition to certain self-assessment actions can generate needed guidance for supporting students' self-assessment processes in classrooms. In the present study, we leverage data collected during the self-assessment process—i.e., think-aloud data—to extend and supplement previous research.

Online Measures

Understandings of self-assessment (Panadero et al., 2020) and emotions (Chevrier et al., 2019) informing the present study are substantiated via online measures. Online measures collect process data during learning processes (e.g., think-aloud protocols, eye tracking, electrodermal activity), in contrast to offline measures which collect data when learning is not happening (e.g., questionnaires, interviews; Chevrier et al., 2019). Online measures have consistently been shown to inform more granular and veridical understandings of learning processes than offline ones (e.g., Callan & Cleary, 2018; Hadwin et al., 2007; Schellings et al., 2013). Collecting data during the learning process may capture volatile data that offline measures cannot—i.e., thoughts, feelings, and actions that are immediately forgotten or escape participants' consciousness (Winne et al., 2019). Predicated on these perspectives, the present study drew primarily on think-aloud data, an online measure, which has been used effectively to investigate how learners self-assess (Panadero et al., 2024).

Study Aim and Research Questions

To contribute to the reviewed literature and inform such practical implications, this study leveraged real-time multimodal online measures—i.e., gaze-cued think-aloud protocols—to investigate the type and intensity of emotions that students experience while self-assessing, and how these emotions may lead to subsequent self-assessment actions. This research was guided by two research questions (RQ):

RQ1. What self-assessment actions do students employ during self-assessment?

RQ2. What types of emotions do students experience during self-assessment, and how frequently and intensely do they experience these emotions?

RQ3. Do certain emotions increase the likelihood of a transition into a subsequent self-assessment action?

We included RQ1 primarily so that we could examine how emotions related to the self-assessment actions that comprised participants' self-assessment processes. Drawing on psychological perspectives on self-assessment and control-value theory of emotions, this research is predicated on the fundamental assumptions that emotions learners experience during self-assessment shape the self-assessment process and that identifying these emotions and patterns in how they influence students' self-assessment processes can ultimately generate valuable implications for supporting student self-assessment in classrooms.

Method

Following a convergent parallel mixed methods design (Creswell & Clark, 2018), we concurrently collected quantitative data, i.e., eye tracking and questionnaires, and qualitative data, i.e., semi-structured interviews, and integrated them to examine the emotional dimension of 25 students' self-assessment processes. Our primary data source was gaze-cued retrospective think-alouds, which we treated quantitatively (i.e., frequencies, state-transition analysis; Chevrier et al., 2019) and qualitatively (i.e., inductive thematic analysis; Charters, 2003). Eye tracking data were collected primarily to support the gaze-cued retrospective think-aloud protocols and are not reported in this study. University institutional ethics boards granted approval for this research.

Participants

Participants (n=25) were English-speaking international students from a variety of countries completing exchange programs at a private university in northern Spain. Leveraging a

convenience sampling strategy, the lead author visited international students' classes to present the study and recruit participants. This strategy resulted in a diverse sample of participants.

Most participants were female (*n*=16). The majority participants were 20-30 years old (*n*=18), with the remaining being older than 30. Participants were a mixture of undergraduate and graduate students studying in a range of disciplines, including the social sciences (*n*=6), engineering (*n*=4), humanities (*n*=3), and business (*n*=3), sciences, languages, communications, design, and energy. Participants were from 15 different countries across four continents: Australia, Belgium, Canada, Croatia, France, Georgia, Germany, Ghana, Hungary, Iran, Netherlands, Nigeria, Pakistan, Philippines, South Korea, Turkey, Ukraine, United Kingdom, United States of America, and Puerto Rico. The majority of participants were white (*n*=15; 60%) and the rest were Black (*n*=2), Asian/Asian American (*n*=2), Latinx (*n*=2), or did not know or preferred not to identify (*n*=4). Participants with English as their first language were the largest language group (*n*=10; 40%). The remaining participants had various other first languages but were proficient in English and studying primarily in English.

Measures

Self-Assessment Actions

Participants' self-assessment actions (i.e., Read, Recall, Compare, Rate, Assess Level 1, Assess Level 2, Redo Level 1, Redo Level 2) were captured via a gaze-cued retrospective think-aloud protocol (McIntyre et al., 2022) guided by the think-aloud and coding procedures published in Panadero et al.'s (2023) *Manual for the data collection and analysis of self- and peer assessment via think-aloud protocols*. This protocol built on a virtual revisit think-aloud protocol, where participants watch a screen recording of their on-screen actions during an online learning task and verbalize what they were doing and thinking. Virtual revisit think-aloud have

been shown to elicit more utterances related to cognitive processing than concurrent think-alouds or uncued retrospective think-aloud, likely because it reduces the extraneous cognitive load imposed by completing a complex task and thinking aloud concurrently while still supporting participants' recall via a fine-grained, process-oriented visual cue (Beach & Willows, 2017). We added a gaze-cue to this protocol, so that the screen recordings participants watched while thinking aloud showed the real-time gaze path as captured by eye tracking, offering a visual cue of what the participant was looking at throughout the recording. Including a gaze cue in retrospective think-aloud protocols has been shown to help effective learners, such as the participants in our study (i.e., university students successfully completing higher education programs), better evaluate their own learning processes (Kostons et al., 2009). This gaze-cued protocol was, therefore, ideally suited to eliciting the most cognitive utterances on the self-assessment process to generate fine-grained insights.

Emotions: Type, Frequency, and Intensity

We measured emotions in two ways. First, we used online data (i.e., think-alouds) to capture the types and frequencies of emotions during the self-assessment process. Second, we used an emotions questionnaire administered immediately after participants self-assessed to measure the intensity of emotions participants experienced.

The types and frequencies of emotions were captured as part of the gaze-cued retrospective think-alouds. To gaze-cued think-aloud protocol, described above, included an emote-aloud component where participants were encouraged and trained to express their emotions in addition to their actions and thoughts (Craig et al., 2008). Participants received a printed list of seven emotions, surprise, enjoyment, curiosity, confusion, anxiety, frustration, and boredom, with definitions, and a researcher explained the meaning of each emotion and

answered participants' clarifying questions to ensure that the researcher and participant had a shared understanding of these emotions. Participants were asked not to say emotions if they did not truly experience them and were encouraged to express any emotions they felt, including those beyond this core list.

We explored the intensity of emotions participants experienced during self-assessment using an adapted version of the Epistemic Emotions Scales (EES; Pekrun et al., 2017). This scale is comprised of Likert-style items which ask participants to rate the intensity of 21 emotions they experienced during a learning event (e.g., intrigued, puzzled, worried) on a scale of 0 (not at all) to 4 (very). These items are divided into seven three-item subscales, each of which measures the intensity of a core academic emotion shown to influence learning: surprise, curiosity, confusion, enjoyment, anxiety, frustration, and boredom (Pekrun et al., 2017). In the questionnaire used in the present study, the scale prompt was adapted to ask participants to rate the degree to which they experienced emotions during the self-assessment activity. The adapted prompt read, "For each emotion, please indicate the strength of that emotion by selecting the option that best describes the intensity of your emotional response during *the self-assessment*" (adapted part in italics). We administered this questionnaire immediately after participants self-assessed. We calculated Cronbach's alpha (α) to estimate the reliability of each of the seven scales: curiosity (α = .56), boredom (α = .23), confusion (α = .81), surprise (α = .52), anxiety (α = .88), frustration (α = .80), and enjoyment (α = .88). Reliability estimates were generally high. The reliability estimates for curiosity, boredom, and surprise were low, possibly due to the small sample size.

Initial States: Academic Control, Task Value, and Self-Assessment Self-Efficacy

To contextualize and inform our interpretation of the findings, we measured academic control, task value, and self-assessment self-efficacy via a questionnaire participants completed

just before the self-assessment activity. Control-value theory holds that learners' appraisals of their control over the outcome of a task and the value of the task shape the emotions they experience while engaging in the task. Therefore, the questionnaire included validated academic control and task value Likert-style scales to capture participants' appraisals of academic control (8 items; e.g., I have a great deal of control over my academic performance in my courses; There is little I can do about my performance in university [reverse coded]) and task value (7 items; e.g., In general, I find self-assessing my work very interesting; The amount of effort it takes to self-assess is worthwhile to me; Di Leo et al., 2019). Items were adapted to apply to self-assessment as a learning task. Additionally, research has implicated students' self-efficacy for self-assessment as an important factor shaping their intention to self-assess and self-assessment practices (Yan et al., 2020). Therefore, the initial questionnaire also captured students' self-efficacy for self-assessment using a scale developed and validated by Yan et al. (2020) comprised of 5 Likert-style items (e.g., I know how to implement self-assessment; I can find reference materials (e.g., rubrics, examples) against which to assess myself). For these scales, participants were instructed to rate their level of agreement with each of the statements, on a scale of 1 (strongly disagree) to 5 (strongly agree). To estimate the reliability of these scales, we calculated Chronbach's alpha (α) and found sufficient reliability estimates for academic control ($\alpha = .90$), task value ($\alpha = .83$), and self-assessment self-efficacy ($\alpha = .73$).

Procedure

Data were collected in two sessions with each participant in a study lab. The primary purpose of the first session was for the participant to generate writing which they could later self-assess. In a blank Microsoft Word document on a study computer, participants responded to the prompt "Write an essay in English about your experiences living and studying in a new country."

Participants were provided with a handout with the full essay instructions, including writing prompts and a list of writing criteria based on the rubric participants would later be given to help them self-assess. Participants took 20-60 minutes to write their essays. During this first session, participants also completed a short demographics questionnaire which was administered online using Qualtrics, an online survey platform.

The second session, which occurred at a separate time, focused on collecting data while participants self-assessed and optionally revised their reflective essays. This longer session typically took 90-120 minutes and involved various forms of data collection, including questionnaires, eye tracking, and think-alouds. Participants worked on a study computer equipped with a with a screen-based eye-tracking system, and all questionnaires were administered online via Qualtrics.

First, participants completed a questionnaire to capture data on key variables needed to contextualize findings—i.e., academic control, task value, and self-assessment self-efficacy. Next, participants engaged in the self-assessment activity. In the activity, participants were asked to review a piece of writing they had completed prior to data collection (i.e., a reflective essay on their experiences living and studying abroad) and provide themselves with written feedback in a separate Microsoft Word document. Participants were advised that they could optionally revise their essay using tracked changes. They were also shown several resources which they could use to help them self-assess and revise their essays. These resources, which could be opened by clicking thumbnails located at the bottom of the screen, included a validated writing rubric, two exemplar essays, the task instructions, and the participants' essays in a Microsoft Word Document with tracked changes enabled. Participants could make revisions directly in their essay files using the tracked changes feature in Microsoft Word. Each participants' self-

assessment process was screen-recorded and their eye movements were tracked in real-time. Eye tracking was recorded using a Tobii Pro Fusion screen-based eye tracking system (sampling rate 250Hz) and each participant's complete gaze-path was embedded in the screen recording file. Participants' self-assessment processes varied in duration from 7 to 40 minutes.

Immediately after self-assessing, participants completed an adapted version of the EES to capture how intensely they experienced surprise, curiosity, confusion, enjoyment, anxiety, frustration, and boredom during the self-assessment activity. Then, they engaged in a gaze-cued retrospective think-aloud protocol (McIntyre et al., 2022) to generate data on their self-assessment actions and emotions. In this protocol, participants watched the full screen recordings including embedded real-time gaze-paths, as captured by eye tracking, and verbalized everything they were doing, thinking, and feeling. Participants received training on the think-aloud protocol, including reviewing definitions of core academic emotions, at the start of the data collection session before data collection began. During the think-alouds, participants were prompted to keep talking if they were silent for 30 seconds (Panadero et al., 2024). A randomly selected subsample of participants (n=4) completed concurrent think-alouds instead of retrospective think-alouds to assess if concurrent think-alouds captured more utterances related to emotions. All think-alouds were recorded, with recordings embedded in the screen recordings, and transcribed verbatim.

At the end of the data collection session, participants engaged in a short (10-20-minute) semi-structured interview to gather more data on their emotions and self-assessment actions. During these interviews, the researchers asked participants to expand on actions and emotions they verbalized in the think-alouds to provide more context to interpret the think-aloud data.

Data Analysis

Before analyzing the data, the first author read the data and viewed the screen recordings to develop a holistic understanding of the data (Thomas, 2006). We then coded the think-aloud data using a hybrid approach to elucidate participants' self-assessment processes, including the actions, emotions, and metacognition that comprised these processes. We used deductive coding, following the coding manual developed by Panadero et al. (2023), to identify participants' self-assessment actions (RQ1), and the approach developed by Chevrier et al. (2019) to identify emotions that arose during self-assessment (RQ2). Following Panadero et al. (2024), we coded specific self-assessment strategies (e.g., read the essay, compare essay to mastering exemplar) and criteria (e.g., intuition, external criteria from rubric) and then clustered these into eight self-assessment actions: read, recall, compare, rate, assess (levels 1 and 2), and redo (levels 1 and 2).

Through our coding process, we also identified metacognitive utterances which articulated actions that had the self-assessment process itself as the object of focus (e.g., monitoring their self-assessment process). Our coding also revealed a few instances of explicit emotional regulation where participants described strategies for managing or responding to emotions (e.g., practicing self-forgiveness). We coded metacognitive and emotional regulation utterances using a general inductive approach (Thomas, 2006) to capture the full range of these actions in the context of self-assessment. We engaged in iterative cycles of coding and discussion among the co-authors until a consensual and consistent interpretation of the data emerged. This iterative process of establishing interrater agreement was supported by assessment scholars in the [removed for peer review] research group who provided expertise on coding self-assessment processes. This process also involved adapting Panadero et al.'s (2024) coding framework to match our present study design. For example, we added "reading the mastering

exemplar" as a code within the read self-assessment action because previous self-assessment studies using the framework have not involved exemplars. Once we had established a final codebook, we used it to recode all the data.

To assess interrater reliability of our coding, we calculated Cohen's kappa (K) between the first author and a trained rater who independently coded 10% of the think-aloud data. The first author trained the rater to use the coding protocol and, following a procedure introduced by Campbell et al. (2013), pre-defined the units of meaning (i.e., utterances containing a single idea) in the sample data. The rater coded the sample data independently using the final codebook to code each unit of meaning. The Cohen's kappa (K) value was 0.83 ($p<.001$), indicating almost perfect agreement between the raters.

Triangulation of Data

Based on our coding, we calculated the frequencies of think-aloud actions (Table 1) and emotions (Table 2). We used Table 2 to triangulate data on the types, frequencies, and intensities of emotions. In this table, we presented the types of emotions alongside the number of times each type of emotion was coded across all participants in the think-aloud data (N). In the rightmost column, we presented the mean ratings of the intensity of each type of emotions from the EES (Pekrun et al., 2017) which participants completed immediately after self-assessing. Triangulating the data using this table allowed us to interpret the relative frequencies and intensities of each type of emotion (RQ2).

To examine the how emotions shaped self-assessment actions, we conducted a state-transition analysis to identify the likelihood of transitioning from an emotion to a subsequent self-assessment action. State-transition analysis is a quantitative method that has been used to study how emotions lead to particular self-regulated learning strategies (e.g., rehearsal, critical

thinking, elaboration) and metacognitive states during learning (Chevrier et al., 2019; Di Leo et al., 2019; D'Mello & Graesser, 2012). To conduct state-transition analysis, we first created a transition matrix including all possible states—i.e., emotions, self-assessment actions (including metacognition), and emotional regulation—to calculate the frequencies of transitions participants made between these states, based on the coded think-aloud data. We then used this state-transition matrix to identify if any transitions from an emotion to a self-assessment action occurred more frequently than chance. State-transition analysis is predicated on the assumption that transitions are not independent if they occur statistically more frequently than if they were randomly distributed. To identify transitions occurring more frequently than by chance, we created 2x2 contingency tables to compare the frequency distribution of specific transitions to a prior random distribution of variables. Using these tables, we conducted chi-squared tests to identify transitions which occurred at a statistically higher rate than chance as reflected by significant chi-squared tests (RQ3).

Results

Our analysis of the think-aloud data resulted in 19 codes in total. Eight of these codes reflected distinct emotions, and 11 reflected student self-assessment actions. Just before completing the self-assessment activity, participants reported high feelings of academic control (M=4.20; SD=.77; scale 1-5), moderate task value (M=3.9; SD=.52; scale 1-5), and moderate self-assessment self-efficacy (M=3.6; SD=.68; scale 1-5).

RQ1. What self-assessment actions do students employ during self-assessment?

Our analysis of the think-aloud data identified 11 student self-assessment actions. See Table 1 for codes, descriptions, representative quotations, and frequencies.

Table 1

Student Self-Assessment Actions from Think-Aloud Data

Code	Description	Representative Quotations	N
Read	Reading a self-assessment resource (e.g., essay, rubric, exemplars, task instructions, self-feedback)	"I started to read [the essay], sentence by sentence."	341
Recall	Remembering or describing (e.g., the essay, the exemplars, the rubric, the instructions) without making a judgement of quality or explicitly reading	"And I kind of spoke chronologically. So, like my first few days, and then, yes, and then challenges I experienced, then the highlights, and then kind of a conclusion at the end."	88
Compare	Explicitly comparing their essay to an indicator of writing quality (e.g., rubric, exemplars) or explicitly comparing two self-assessment resources (e.g., rubric and exemplars, two contrasting exemplars)	"This essay, it used a lot more adjectives and it just seemed much more personal than mine."	50
Rate	Making a judgement about their work or part of their work without explicitly referring to or indicating any criteria	"It's not a good conclusion."	108
Assess Level 1	Making a judgement on their work with reference to gut reaction; peripherally relevant, low-impact criteria (e.g., spelling); or unclear indicators	"But I think I was also feeling satisfied with how this was written."	341
Assess Level 2	Making a judgement on their work with reference to highly relevant and/or high impact external indicators (e.g., tone, style, audience, structure)	"I realized I missed really examples to keep my experience more engaging."	97

Redo Level 1	Explicitly indicating what changes they would make without actually making the changes	"I think if I probably wanted to spend a lot more time, I would have really edited more of the paper in general to be more interesting and kind of captivate the reader instead of just emotionally dumping information."	208
Redo Level 2	Making substantial changes to the content of the essay, structure changes, or high impact sentence level changes that improve the quality of the essay (e.g., adding or revising multiple sentences, elevating the vocabulary throughout, revising punctuation to add shades of meaning)	"We can just—I can make it more personal, as I should. [typing and speaking:] 'As in Germany, people are really materialistic and focus a lot on work. People will choose work over enjoyment or leisure time. In Spain, it does feel like there is a different focus for people, which really inspired me to re-evaluate if I want to live a really stressful life in the business world, which society seems to, seems to want from me, which society seems to expect from me.' Okay, this is just one example."	47
Metacognitive Experiences	Explicitly monitoring or evaluating their self-assessment process or skills (e.g., feelings of difficulty, judgements of learning, efficacy judgements, task value judgements)	"I have some weakness some … some bad, not habits, like bad skills for this self-assessment"	275
Metacognitive Skills	Explicitly articulating how they are controlling or making decisions about their self-assessment process (e.g., articulating their approach, reasoning/goal setting)	"… to make sure that I was being as fair as possible to myself"	762
Emotional Regulation	Identifying a strategy with the explicit aim of addressing an emotion (e.g., practicing self-forgiveness, pushing through negative affect)	"And nevertheless, I summoned courage"	44
Total self-assessment actions			**2360**

Note. The rightmost column (i.e., N) presents the frequencies of each self-assessment action across participants, calculated from the think-aloud data.

Eight of these actions were identified using the adapted coding protocol developed by Panadero et al. (2024). These actions had the participants' work (i.e., reflective essay) or a self-assessment resource (e.g., mastering exemplar, rubric) as the object of focus: Read, Recall, Compare, Rate, Assess (Levels 1 and 2), and Redo (Levels 1 and 2). For example, if a participant said "I was reading my essay" or "my essay was better than this one [developing exemplar]", the essay was the object of focus of these actions (i.e., Read and Compare, respectively). In contrast, two codes represented actions that had the self-assessment process itself as the object of focus—i.e., self-assessment metacognition. Such actions were coded as metacognitive experiences, where participants monitored their self-assessment process, or metacognitive skills, where participants explicitly controlled or made decisions about their self-assessment process. Last, the code emotional regulation represented actions participants explicitly took to address emotions. Participants verbalized a total of 2360 self-assessment actions.

RQ2. What types of emotions do students experience during self-assessment, and how frequently and intensely do they experience these emotions?

To examine the emotions students experience during self-assessment, we calculated the frequency and intensity of emotions participants reported during self-assessment (i.e., via think-alouds) and immediately after the self-assessment (i.e., via a questionnaire). Our analysis of the think-aloud data identified eight emotions participants experienced during self-assessment: surprise, curiosity, enjoyment, confusion, anxiety, frustration, boredom, and sadness. See Table 2 for codes, descriptions, representative quotations, and frequencies and intensities of each emotion.

Table 2

Types of Emotions Codes from Think-Aloud Data and Intensities of Emotions from the EES

Emotion	Description	Representative Quotation	N from think-aloud data	Mean(SD) Intensities from EES, scale 0-4
Anxiety	Worried, nervous, stressed	"It's kind of vulnerable reading my own writing, especially about something that my time in Spain, which is important."	31	.99(1.04)
Boredom	Dull, monotonous	"And then probably there, I got bored."	5	.47(.35)
Confusion	Uncertain, puzzled, unsure	"I was just a bit confused because the way was written was not super clear"	76	1.21(.84)
Curiosity	Interested, intrigued	"honestly it's very interesting for me"	42	2.60(.75)
Enjoyment	Happy, excited	"at this point I felt pretty content with the work I had done and the edits I had made to this"	95	1.99(.93)
Frustration	Irritated, annoyed, dissatisfied	"Here I was wondering if I needed to add something. I was feeling frustrated because I didn't like the way it was phrased, but I couldn't come up with another way of saying it."	27	.60(.72)
Sadness	Upset, unhappy	"Also made me kind of sad to read this because I know I'm going to be leaving so soon"	13	-
Surprise	Amazed, astonished	"I was surprised that I didn't catch initially the mistakes or grammatical things"	18	1.19(.76)
Positive			119	2.29(.63)
Negative			152	.82(.58)

Note. The frequencies of each type of emotion (i.e., column N) were calculated based on the think-aloud data. The mean intensities of each emotion (i.e., the rightmost column) were calculated via the EES (Pekrun et al., 2017) data. There is no intensity data for sadness because sadness was not measured in the Epistemic Emotions Questionnaire. Sadness emerged as a code in the think-aloud data, so only frequency data are available.

The most frequently verbalized emotion was enjoyment (N=95) and the emotion verbalized least often was boredom (N=5). Participants experienced negative emotions (confusion, anxiety, frustration, boredom, sadness; N=152) relatively more frequently than positive emotions (curiosity, enjoyment; N=119). Note that surprise is omitted from frequency counts of positive or negative emotions because surprise is considered a neutral emotion (Chevrier et al., 2019). Participants verbalized emotions a total of 289 times.

Descriptive analysis of the EES (Pekrun et al., 2017) data indicated that the intensity of emotions participants felt was relatively low. The most intensely felt emotion was curiosity (M=2.60; SD=.75; scale: 0=4) and the least intensely felt emotion was boredom (M=.47; SD=.35). Participants experienced positive emotions relatively more intensely (M=2.29; SD=.63) than negative ones (M=.82; SD=.58). Surprise, a neutral emotion (Chevrier et al., 2019), was omitted from mean calculations for positive and negative emotions. We produced a correlation matrix (Pearson's *r*) to assess the relationships between task value, academic control, and self-assessment self-efficacy and the intensities of the emotions participants experienced (see Table 3). There were no significant correlations between task value, academic control, and self-assessment self-efficacy and the intensities of any of the emotions participants rated on the EES, suggesting task value, academic control, and self-assessment self-efficacy were not associated with how intensely participants experienced emotions. However, this finding could be due to the small sample size and must be interpreted with caution. We did not calculate correlations between task value, academic control, and self-assessment self-efficacy and the frequency of types of emotions participants experienced because we did not control for verbosity. Correlations between emotions indicated that surprise was moderately positively correlated with curiosity, $r(24) = .61$ $p < .001$; enjoyment, $r(24) = .43$ $p < .005$; and confusion, $r(24) = .45$ $p < .005$.

Confusion was moderately positively correlated with anxiety, $r(24) = .59$ $p < .001$, and frustration, $r(24) = .60$ $p < .001$.These correlations suggest that the intensity at which participants experienced surprise was related to the intensity at which they experienced positive emotions of enjoyment and curiosity, and the negative emotion of confusion. The correlations also indicate that how strongly participants felt confusion was related to how strongly they felt negative emotions of anxiety and frustration.

Table 3

Pearson (r) Correlation Matrix: Task Value, Academic Control, Self-Assessment Self-Efficacy, and the Intensities of Emotions

	Task Value	Academic Control	Self-Assessment Self-Efficacy	Surprise	Curiosity	Enjoyment	Confusion	Anxiety	Frustration	Boredom
Task Value	1	0.22	**.47***	0.05	-0.04	0.06	-0.07	0.08	0.07	-0.03
Academic Control	0.22	1	**.48***	0.17	0.08	-0.25	0.18	0.33	-0.01	0.16
Self-Assessment Self-Efficacy	**.47***	**.48***	1	-0.03	-0.23	-0.23	0.27	0.16	0.06	0.22
Surprise	0.05	0.17	-0.03	1	**.61****	**.43***	**.45***	0.38	0.33	0.26
Curiosity	-0.04	0.08	-0.23	**.61****	1	0.12	0.27	0.31	0.16	0.06
Enjoyment	0.06	-0.25	-0.23	**.43***	0.12	1	0.06	0.13	-0.17	-0.02
Confusion	-0.07	0.18	0.27	**.45***	0.27	0.06	1	**.59****	**.60****	0.14
Anxiety	0.08	0.33	0.16	0.38	0.31	0.13	**.59****	1	**.49***	0.24
Frustration	0.07	-0.01	0.06	0.33	0.16	-0.17	**.60****	**.49***	1	0.28
Boredom	-0.03	0.16	0.22	0.26	0.06	-0.02	0.14	0.24	0.28	1

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

Participants experienced topic enjoyment due to feelings of nostalgia, e.g., "This paragraph was mainly about my roommate, which I really liked writing about this paragraph because I really liked my roommate and I liked editing this." Participants also experienced enjoyment when they perceived that they had made improvements to their work or resolved an impasse, which would be epistemic enjoyment. One said, "I didn't have an objective … I was happy about realizing that." Feelings of curiosity about how others approached the essay influenced students to read the exemplars: "now I just wanted to read over the mastering example to … well, more of curiosity really, kind of see what had been written." States of enjoyment, or "flow", were interrupted when participants noticed mistakes or writing that did not match their personal voice. These interruptions caused frustration; for example, one participant noted, "It's slightly frustrating because I wasn't able to read it through without stopping to edit something," or confusion when they did not immediately know how to resolve the impasse: "here I felt confused about what I wanted to say." Participants attempted to resolve achievement frustration and confusion by thinking of different ways to revise their writing; i.e., frustration and confusion provoked Redo Level 1. Redo Level 1, therefore, represents a unique self-assessment action that participants sometimes used as an emotion regulation strategy. Inability to quickly make revision decisions exacerbated frustration: "I was feeling frustrated because I didn't like the way it was phrased, but I couldn't come up with another way." When participants experienced sadness, which occurred relatively infrequently (N=13), it tended to be topic sadness as participants felt sad about something they read in their essays (e.g., having to leave Spain soon).

RQ3. Do emotions increase the likelihood of a transition into a subsequent self-assessment action?

To address research question 3, we conducted a state-transition analysis to identify transitions from emotions to subsequent self-assessment actions that occurred at a greater rate than chance in the think-aloud data. Table 4 is the transition matrix which presents the frequencies of transitions from emotions to all possible subsequent states. The numbers in each cell indicate the number of times we observed a transition from an emotion, indicated in the corresponding row, to a subsequent self-assessment action, indicated in the corresponding column. We then created multiple 2x2 contingency tables to test for statistical differences between observed and expected frequencies using chi-squared tests. Chi-squared tests assume at least five cases per cell, so we created tables for and tested all transitions that occurred five times or more. Note that, since the focus of this study is on how emotions provoke self-assessment actions, we did not examine transitions to emotional regulation or other emotions in our chi-square analyses. Our chi-squared tests indicated four statistically significant transitions: surprise to Assess Level 1; enjoyment to Assess Level 1; anxiety to metacognitive experiences; and frustration to metacognitive experiences. Tables 5-8 are the contingency tables for these transitions.

Table 4

Frequency of transition from emotion (left column) to self-assessment action, emotional regulation, or other emotion (top row)

	Read	Recall	Compare	Rate	Assess L1	Assess L2	Redo L1	Redo L2	Metacognitive Experiences	Metacognitive Skills	Emotional Regulation	Other Emotion
Surprise	1				8	1	3		1	1	1	1
Enjoyment	6	3		2	19	2	5		10	26	2	13
Curiosity	4	3	1	3	4		1		7	8		8
Confusion	12	1		1	9		9	1	8	21	3	6
Anxiety	1			2	2		1		6	11	1	2
Frustration	5	1		2	4				7	3	3	2
Boredom	2				1					2		
Sadness		2					2			1	2	4

Note. The number in each cell represents the number of transitions from an emotion, denoted in the column, to a subsequent self-assessment action, denoted in the row. Blank cells indicate 0 transitions.

Table 5

Frequency of Surprise to Assess Level 1 Transitions

	Following Surprise	Total
Assess L1	8	341
Other Self-Assessment Actions	7	1975
Self-Assessment Actions	15	2316

Table 6

Frequency of Enjoyment to Assess Level 1 Transitions

	Following Enjoyment	Total
Assess L1	19	341
Other Self-Assessment Actions	54	1975
Self-Assessment Actions	73	2316

Table 7

Frequency of Anxiety to Metacognitive Experiences Transitions

	Following Anxiety	Total
Metacognitive Experiences	6	275
Other Self-Assessment Actions	17	2041
Self-Assessment Actions	23	2316

Table 8

Frequency of Frustration to Metacognitive Experiences Transitions

	Following Frustration	Total
Metacognitive Experiences	7	275
Other Self-Assessment Actions	15	2041
Self-Assessment Actions	22	2316

Surprise to Assess Level 1

For transitions from surprise to subsequent self-assessment actions, we found that Assess Level 1 followed surprise significantly more often than expected, $\chi^2(1) = 17.45$, $p < .001$ (Table 5). Examples of this sequence included: "Yes, I was surprised [*surprise*] how they actually make the text sound much better and much more … not complex, but yeah just a little bit more profound [*Assess Level 1*]" and "And yeah, I did not expect that. And this kind of surprised me [*surprise*] because for some reason, I thought I wrote about them equally, but clearly not [*Assess Level 1*]." Overall, these transitions tended to involve participants expressing a feeling of surprise and then considering the reason for their surprise by attributing it to an evaluation of their work based on internal criteria (e.g., intuition, hindsight).

Enjoyment to Assess Level 1

For transitions from enjoyment to subsequent self-assessment actions, we found that Assess Level 1 followed enjoyment significantly more often than expected, $\chi^2(1) = 7.07$, $p = .008$ (Table 6). Examples of this sequence included: "I felt satisfied [*enjoyment*], but yeah, could have improved, could have been improved, the transition part [*Assess Level 1*]" and "but it was so emotional to meet her and I felt my mom's presence being there. It was really cool [*enjoyment*]. And that's a spelling mistake [*Assess Level 1*]." Transitions from enjoyment to Assess Level 1 either involved participants experiencing enjoyment and then explaining why with a positive evaluation of their work, or participants' feelings of enjoyment or flow being interrupted by a negative evaluation of their work.

Anxiety to Metacognitive Experiences

For transitions from anxiety to subsequent self-assessment actions, we found that metacognitive experiences followed anxiety significantly more often than expected, $\chi^2(1) = 4.35$,

p = .04 (Table 7). Examples of this sequence included: "And here I think, I don't know. I think I got flustered with the sentence before [*anxiety*], so then it took me a while to get back into reading my work again [*metacognitive experience*]" and "Here I remember being a bit not anxious but nervous about what I would put it in [*anxiety*] as I was rethinking my choices because I wasn't sure if I was actually good enough for mastering [*metacognitive experience*]." Metacognitive experiences that followed anxiety tended to involve feelings of difficulty which were provoked by or explained the feeling of anxiety.

Frustration to Metacognitive Experiences

For transitions from frustration to subsequent self-assessment actions, we found that metacognitive experiences followed frustration significantly more often than expected, $\chi^2(1) = 8.17$, $p < .001$ (Table 8). Examples of this sequence included: "And it's slightly frustrating because I wasn't able to read it through without stopping to edit something [*frustration*], I would've rather have just been able to read and read it through in one pass—well, that's great—and then think of it. But no … I had to make edits along way where I thought I didn't have to [*metacognitive experience*]" and "I got a little frustrated, I guess, irritated [*frustration*]. The fact that I couldn't put it well, put the sentence well together I was missing some words or a way to say it. So I think I got stuck here for a bit [*metacognitive experience*]". Metacognitive experiences that followed frustration, like those that followed anxiety, articulated feelings of difficulty.

Discussion

Leveraging gaze-cued retrospective think-aloud data, this study provides initial evidence of what emotions students experience while self-assessing and revising their writing, how intensely they experience these emotions, and how these emotions may shape the subsequent

actions that students employ as part of their self-assessment processes. Findings from this study offer important insights into the role of emotions in student self-assessment which have theoretical and practical implications.

RQ1. What self-assessment actions do students employ during self-assessment?

We analyzed the think-aloud data to identify the types and frequencies of student self-assessment actions in order to relate them to emotions participants experienced during the self-assessment process. We found that the think-aloud coding protocols developed by Panadero et al. (2024) were effective for identifying self-assessment actions. We identified the same primary actions in our data as have been observed in previous think-aloud studies of student self-assessment: i.e., Read, Recall, Compare, Rate, Assess (Levels 1 and 2), and Redo (Levels 1 and 2; Panadero et al., 2024). Using inductive coding, we also identified utterances which expressed self-assessment metacognition—i.e., actions which had the self-assessment process, rather than the participants' essay, as the object of focus (see Panadero et al., 2019)—and explicit emotional regulation (e.g., pushing through negative affect, practicing self-forgiveness).

RQ2. What types of emotions do students experience during self-assessment, and how frequently and intensely do they experience these emotions?

Overall, this study suggested that students experience a range of emotions while self-assessing their writing. We observed, in the think-aloud data, utterances expressing eight distinct emotions: surprise, curiosity, enjoyment, confusion, anxiety, frustration, boredom, and sadness. Participants' post-self-assessment ratings of emotions indicated that they experienced these at low to moderate intensities. While negative emotions were experienced more frequently, positive ones were more intensely felt. Participants experienced this range of emotions despite the low-stakes assessment context of this study. The essays they were assessing were not part of their

actual coursework, nor would the quality of their essays have any material consequences on their education or GPAs. The only potential consequences were for their self-concept or identity, as assessments such as the self-assessment in this study can shape students' understandings of themselves (Nieminen, 2024). When the work students are self-assessing will have consequences for students (e.g., grades, program admission), emotions are likely to be more intense, as frustration and anxiety have been associated with the feeling of being assessed in higher education (Wass et al., 2020). Negative emotions, such as frustration and anxiety, may be felt more frequently and intensely when students are self-assessing work or learning which will significantly affect their lives, such as assessments which will determine their selection into subsequent educational programs—and the intensity and frequency of these emotions may have variable or amplified effects on the self-assessment process.

Our analysis of the think-aloud data revealed that enjoyment and confusion were the most frequently expressed emotions, suggesting that they play an important role in the self-assessment process. Enjoyment and confusion fundamentally shape how students engage in self-regulated learning (Muis et al., 2018), a process predicated on student self-assessment, supporting our findings. Learners experience enjoyment when they perceive that the learning task supports their learning goals and that they have sufficient control to achieve their goals (Obergriesser & Stoeger, 2020). Our findings support this conception, as our participants reported high global academic control and moderate task value for self-assessment immediately before self-assessing. As university students across various levels of postsecondary education studying in another country, our participants are likely highly self-regulated learners who would be likely to experience enjoyment during self-assessment. This finding also supports previous research which showed that providing students with a rubric can increase their positive emotions and

decrease their negative emotions during self-assessment (Panadero et al., 2023), as our design included providing students with a rubric. In addition to enhancing students' perceptions of control and task value for self-assessment, providing students with rubrics may also be a productive way to foster enjoyment in self-assessment.

Enjoyment plays an important role in learning. It has been shown to have positive motivational consequences, helping students to persist with learning tasks (Chevrier et al., 2019). Given that self-assessment can be a challenging activity for many students to persist with—as it requires them to confront negative aspects of their performance (Baca & Smith, 2018) and to apply higher order thinking skills (Boud & Brew, 1995)—cultivating enjoyment during self-assessment may be critical to its effectiveness. Classroom interventions that support students' appraisals of academic control and task value for self-assessment could therefore be important prerequisites for fostering student self-assessment. Interestingly, engaging in self-assessment can help build students' self-efficacy for tasks they self-assess by helping them demonstrate success criteria and showing them that they have the capacity to improve their own performance (Bourke, 2016; Ross, 2006)—especially when teachers provide explicit opportunities for students to use their self-assessments to revise and improve their work (Andrade & Brookhart, 2020). Self-assessment and reflective thinking have been associated with learner enjoyment in the context of language learning (Aldosari et al., 2023). There appears to be a cyclical relationship between self-assessment and enjoyment, where enjoyment can sustain persistence with self-assessment and self-assessment, when supported as an empowering experience for students, can enhance enjoyment in self-assessment via enhanced task value and control.

Confusion arises during complex learning tasks when a learner detects discrepant information or disequilibrium (e.g., a conflict between prior knowledge and new learning;

conflicting pieces of new information) which causes a physiological response and the learner's attention to shift towards resolving the discrepancy (Muis et al., 2018). Although confusion is a negative emotion, like enjoyment it serves an important role in learning by leading to the activation of regulatory processes and strategies aimed at resolving the confusion (Munzar et al., 2021). When confusion is resolved, returning the learner to an affective state of equilibrium or enjoyment, it has been shown to promote learning (D'Mello et al., 2014). On the other hand, if the learner does not have the resources to resolve confusion in a timely manner, it can lead to frustration, boredom, and ultimately disengagement (Chevrier et al., 2019; D'Mello & Graesser, 2012). In our study, we did not find any statistically significant transitions from confusion to self-assessment actions. However, confusion to metacognitive skills was the second most frequently observed transition (N=21), supporting these previous studies' findings that confusion can provoke higher-level cognitive processing and theory that delineates the role of emotions in self-regulated learning (Muis et al., 2018). Together, our findings and the previous research signal the important role of confusion in the self-assessment process and the need for teachers and students to be able to respond effectively when confusion arises. Teachers could support students in recognizing when and why they are confused during self-assessment, and to seek help from teachers, peers, or resources (e.g., rubric, exemplar) before confusion persists into frustration or boredom.

RQ3. Do certain emotions increase the likelihood of a transition into a subsequent self-assessment action?

The transition matrix in Table 4 provides a profile of transitions that aligns with theorized connections between epistemic emotions and self-regulated learning strategies, with relatively frequent transitions between positive emotions, like enjoyment and curiosity, and sophisticated

self-assessment strategies, like metacognitive processing and Assess Level 1 (Muis et al., 2018). Moreover, confusion, a negative emotion theorized to benefit learning when it provokes self-regulated learning strategies that resolve the confusion (D'Mello et al., 2014), transitioned relatively frequently to a wide range of actions, including Read, Redo, and metacognitive processing, possibly because participants activated information processing, analysis (Panadero et al., 2024), and metacognition to address feelings of confusion related to the success criteria, their performance, and their self-assessment approach. The consequences of emotions we observed while students engaged in self-assessment reflect those identified in self-regulated learning, which is expected given self-assessment is understood as a subprocess of self-regulated learning (Boekaerts, 1991; Panadero, 2017).

Our state-transition analysis indicated that surprise and enjoyment were likely to transition to Assess Level 1, implicating these emotions are more likely than others to provoke learners to draw on gut reactions to assess their work. Enjoyment has been linked to the application of cognitive strategies (Obergriesser & Stoeger, 2020), which is reflective of this transition. In our think-aloud data, enjoyment to Assess Level 1 transitions tended to involve participants either expressing enjoyment and then explaining why, as a result of something they liked about their work, or expressing enjoyment and then having that feeling interrupted by an unrelated, negative evaluation of their work. Surprise arises when learners detect a discrepancy between new information and prior expectations or beliefs (Chevrier et al., 2019). During self-assessment, our participants expressed surprise and explained the nature of their surprise as being related to an evaluative statement about their work which was guided by gut reactions (e.g., intuition). This finding suggests that surprise may lead to Assess Level 1 as participants make sense of their feelings of surprise by relating it to their work. However, surprise may be an

immediate experience that quickly provokes follow up emotions, such as confusion or curiosity, as the nature of the surprising information comes to light for the learner (see Muis et al., 2018). The intensity of surprise, or the emotion that immediately follows surprise, may mediate the strategies that learners apply as a result (Chevrier et al., 2019). Our think-alouds could not detect the intensity of each instance of surprise during the self-assessment process, and it is possible that participants did not express intermediate emotions between surprise and Assess Level 1, so further research is needed to verify and examine the role of surprise in provoking Assess Level 1.

Overall, the link between two emotions and Assess Level 1 could be due to the affective nature of Assess Level 1 itself, which is evaluation predicated on gut reactions and internal feelings. Our study indicates a strong connection between emotions and drawing on gut reactions which provokes an area for future research. Specifically, understanding the extent to which emotions themselves operate as heuristics for learners to quickly assess their work during self-assessment would be helpful, as emotions can be shaped by various extraneous factors beyond the work itself and therefore could confound the accuracy of students' self-feedback, an important dimension in the validity of formative self-assessment (Brown et al., 2015). A practical implication of these transitions is that teachers could support students in self-assessment by directing them to also use Assess Level 2, drawing on external criteria (e.g., from rubrics or teachers), to make clearer sense of their work when they experience emotions like enjoyment or surprise, rather than relying on gut reactions. Additionally, redoing or revising aspects of their work can help them better understand their work in response to emotions. Assess Level 2 and redo are important actions associated with sophisticated approaches to self-assessment by providing learners with higher level thinking skills that they can apply to make sense of their performance (Panadero et al., 2024).

Our state-transition analysis also indicated that anxiety and frustration were likely to transition to metacognitive experiences, implicating this transition as playing an important role in student self-assessment. Anxiety and frustration are both negative activating emotions. Anxiety involves a moderate fear of failure to achieve a goal resulting from a perceived lack of control (Pekrun, 2006). Frustration tends to arise when learners perceive a persistent lack of progress towards their goals and is more intense than anxiety; it can overwhelm a learner's cognitive resources (Chevrier et al., 2019; Park & Ramirez, 2022). Anxiety and frustration tend to lead to application of shallow learning strategies (e.g., rehearsal) instead of more effective regulation (e.g., critical thinking, elaboration; Di Leo et al., 2019). However, we found that they led to metacognitive experiences during self-assessment. Our findings are supported by evidence that frustration predicted metacognitive processing which predicted math problem-solving achievement for grades 3 to 6 students (Munzar et al., 2021).Collectively, these findings offer a reminder that there are no universal consequences of emotions and that context matters (e.g., simple versus complex learning tasks; Muis et al., 2018). In some contexts, these negative emotions may provoke shallow processing, while in others, they may provoke metacognitive experiences. For our participants, anxiety- and frustration-metacognitive experiences transitions tended to involve expressing either anxiety or frustration and then articulating a feeling of difficulty, making sense of a perceived challenge in their self-assessment process (e.g., balancing correcting typos with reflecting deeply on their work; feeling unsure about how to revise a section of their writing). The link between these emotions and metacognitive experiences, especially feelings of difficulty, is commensurate with metacognition theory which posits a bidirectional relationship whereby negative emotions or moods lead to feelings of difficulty, and metacognitive experiences inform feelings of (dis)fluency towards goals to provoke positive or

negative affect (Efklides, 2017). This bidirectional relationship further implicates context as a critical factor in the emotions that may arise in relation to metacognition and their consequences.

In the context of self-assessment, our findings suggested that negative emotions indeed fostered metacognitive experiences, especially feelings of difficulty, which could be interpreted as perceptions of disfluency. From the perspective of processing fluency, these metacognitive experiences may lead students to make negative appraisals of their self-assessment processes, which could provoke them to exercise metacognitive control to regulate the process or, if the sense of disfluency persists unresolved, disengage from the self-assessment process (Reber & Greifeneder, 2017). Transitions from negative emotions to metacognitive experiences could be interpreted as negative emotions provoking deeper processing and potential regulation, or a state of paralysis that hinders the self-assessment process. Therefore, it is unclear whether anxiety and frustration, and their subsequent transitions to metacognitive experiences, were helpful for or detrimental to our participants' self-assessment processes. Research in clinical education has found that anxiety increased women's self-assessment accuracy (Colbert-Getz et al., 2013), potentially supporting a positive interpretation of anxiety-metacognitive experiences whereby the deeper processing triggered by the anxiety ultimately led to a more accurate self-assessment for women. However, more research is needed to unpack what happens following anxiety- and frustration-metacognitive experiences transitions during self-assessment. Sequential pattern analysis (Chen et al., 2017), for example, could illuminate the pathways that follow these transitions to identify when and how they might provoke more effective self-assessment. Being aware that students may experience metacognitive feelings of difficulty after feeling anxiety or frustration during self-assessment can be helpful information for teachers, who can help students

make sense of and address feelings of difficulty when they experience these negative emotions during classroom self-assessment activities.

Implications

This exploratory research contributes to the descriptive aspects of student self-assessment theory by providing the first empirical evidence of the emotional dimension. This descriptive aspect is critical, as comprehensive descriptions of psycho-social phenomena, predicated on mixed-methods studies like this one, delineate what dimensions phenomena comprise to guide explanatory research (Greene, 2022). Overall, our study suggests that there is value for theory and practice in conceptualizing the emotional dimension of student self-assessment in line with D'Mello and Graesser's (2012) framework of the dynamics of affective states during complex learning. We found that enjoyment was the most frequent emotion, and curiosity the most intense emotion, and that enjoyment and surprise tended to lead to use of Assess Level 1, a productive assessment action. We also found that negative emotions of anxiety and frustration led to metacognitive experiences, where participants articulated difficulties they were experiencing in their self-assessment process. Together, these findings characterize the emotional dynamics of student self-assessment as a process of "flow" or enjoyment (i.e., enjoyment, curiosity), as the default state, interrupted when impasses are detected (i.e., surprise, anxiety, frustration) and resolutions attempted. Practically, this understanding can be helpful for teachers in supporting student self-assessment. Knowing the consequences of emotions such as surprise, anxiety, and frustration can help teachers support students in responding productively to these emotions and return to states of enjoyment, which can help them persist in self-assessment and continue to use it productively in the future.

Limitations and Future Directions

Findings from this study need to be interpreted with consideration of the study's limitations and research context. First, findings from this study are not widely generalizable and future research is needed across education contexts to understand the role of emotions in self-assessment. As self-assessment is socio-culturally situated, emotions students experience during self-assessment likely vary across and even within socio-cultural contexts. Additional mixed methods studies with different populations could validate, situate, and extend our findings across education contexts. Importantly, future research on the emotional dynamics of students self-assessing work that has real consequences for their grades and lives could support understandings rooted in more authentic self-assessment scenarios. Second, as this was an initial study into the emotional dimension of student self-assessment, we focused on transitions between emotions and self-assessment actions. However, emotions research in the context of self-regulated learning has implicated many factors that shape emotions and how emotions relate to learning, such as epistemic cognition (e.g., Chevrier et al., 2019).

With the initial picture provided by this study, future research could expand on self-assessment affect theory by examining the antecedents of emotions during self-assessment and utilize analyses to identify longer sequences of transitions between emotions and self-assessment actions that serve important functions. Third, while our study leveraged primarily online measures of think-alouds to capture emotions as they arose within the self-assessment process, we relied primarily on self-reports. The think-alouds allowed us to identify what emotions learners experienced, and our questionnaire data provided insight into overall estimates of the intensities of these emotions; however, the use of multimodal approaches combining physiological measures, such as electrodermal activity and heart rate (Collins et al., 2019; Järvelä

et al., 2021), could offer insights into the intensity of each instance of an emotion as it occurs in real time, offering deeper insight into the impacts these emotions may have on the self-assessment process.

Conclusion

Despite these limitations, this exploratory research contributes the first empirical evidence of the types and intensity of emotions students may experience during self-assessment, and how certain emotions may be likely to provoke certain self-assessment actions. Our findings characterize the emotional dynamics of student self-assessment as a process of "flow" or enjoyment (i.e., enjoyment, curiosity), as the default state, interrupted when impasses are detected (i.e., surprise, anxiety, frustration) and resolutions attempted. Future research is needed to understand where students' self-assessment actions which arise as a result of emotions lead, and to verify and extend this framework to other learning contexts—especially higher stakes ones, where the emotions may be more intense and impactful. This study further implicates emotions as a consequential dimension of the student self-assessment process that remains at the heart of the self-assessment research agenda.

References

Aldosari, M., Heydarnejad, T., Hashemifardnia, A., & Abdalgane, M. (2023). The interplay among self-assessment, using reflection for assessment, classroom enjoyment, and immunity: into prospects of effective language learning. *Language Testing in Asia*, *13*(1). https://doi.org/10.1186/s40468-022-00213-1

Andrade, H. L. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In H. L. Andrade & G. Cizek (Eds.), *Handbook of formative assessment* (pp. 90–105). Routledge.

Andrade, H. L. (2019). A critical review of research on student self-assessment. *Frontiers in Education*, *4*, 87. https://doi.org/10.3389/feduc.2019.00087

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

Baca, L., & Smith, S. R. (2018). The role of self-reflection and self-assessment in the psychological assessment process. In S. R. Smith & R. Krishnamurthy (Eds.), *Diversity-sensitive personality assessment* (pp. 3–24). Routledge.

Beach, P., & Willows, D. (2017). Understanding teachers' cognitive processes during online professional learning: A methodological comparison. *Online Learning Journal*, *21*(1), 60–84. https://doi.org/10.24059/olj.v21i1.949

Belkaid, M., Cuperlier, N., & Gaussier, P. (2017). Emotional metacontrol of attention: Top-down modulation of sensorimotor processes in a robotic visual search task. *PLoS ONE*, *12*(9). https://doi.org/10.1371/journal.pone.0184960

Birenbaum, M., DeLuca, C., Earl, L., Heritage, M., Klenowski, V., Looney, A., Smith, K., Timperley, H., Volante, L., & Wyatt-Smith, C. (2015). International trends in the implementation of assessment for learning: Implications for policy and practice. *Policy Futures in Education*, *13*(1), 117–140. https://doi.org/10.1177/1478210314566733

Boekaerts, M. (1991). Subjective competence, appraisals and self-assessment. *Learning and Instruction*, *1*, 1–17.

Boekaerts, M. (2010). The crucial role of motivation and emotion in classroom learning. In H. Dumont, D. Istance, & F. Benevides (Eds.), *The nature of learning: Using research to inspire practice* (pp. 91–111). The Center for Educational Research and Innovation, Organization for Economic Co-operation and Development.

Boud, D., & Brew, A. (1995). Developing a typology for learner self-assessment practices. *Research and Development in Higher Education*, *18*, 130–135.

Bourke, R. (2016). Liberating the learner through self-assessment. *Cambridge Journal of Education*, *46*(1), 97–111. https://doi.org/10.1080/0305764X.2015.1015963

Bourke, R., & Mentis, M. (2013). Self-assessment as a process for inclusion. *International Journal of Inclusive Education*, *17*(8), 854–867. https://doi.org/10.1080/13603116.2011.602288

Brandmo, C., Panadero, E., & Hopfenbeck, T. N. (2020). Bridging classroom assessment and self-regulated learning. *Assessment in Education: Principles, Policy and Practice*, *27*(4), 319–331. https://doi.org/10.1080/0969594X.2020.1803589

Brooks, C., Burton, R., van der Kleij, F., Ablaza, C., Carroll, A., Hattie, J., & Neill, S. (2021). Teachers activating learners: The effects of a student-centred feedback approach on

writing achievement. *Teaching and Teacher Education*, *105*, 103387. https://doi.org/10.1016/j.tate.2021.103387

Brown, G. T. L., Andrade, H. L., & Chen, F. (2015). Accuracy in student self-assessment: Directions and cautions for research. *Assessment in Education: Principles, Policy and Practice*, *22*(4), 444–457. https://doi.org/10.1080/0969594X.2014.996523

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

Campbell, J. L., Quincy, C., Osserman, J., & Pedersen, O. K. (2013). Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods and Research*, *42*(3), 294–320. https://doi.org/10.1177/0049124113500475

Charters, E. (2003). The use of think-aloud methods in qualitative research: An introduction to think-aloud methods. *Brock Education*, *12*(2), 68–82.

Chen, B., Resendes, M., Chai, C. S., & Hong, H. Y. (2017). Two tales of time: Uncovering the significance of sequential patterns among contribution types in knowledge-building discourse. *Interactive Learning Environments*, *25*(2), 162–175. https://doi.org/10.1080/10494820.2016.1276081

Chevrier, M., Muis, K. R., Trevors, G. J., Pekrun, R., & Sinatra, G. M. (2019). Exploring the antecedents and consequences of epistemic emotions. *Learning and Instruction*, *63*, 101209. https://doi.org/10.1016/j.learninstruc.2019.05.006

Colbert-Getz, J. M., Fleishman, C., Jung, J., & Shilkofski, N. (2013). How do gender and anxiety affect students' self-assessment and actual performance on a high-stakes clinical skills

examination? *Academic Medicine*, *88*(1), 44–48. https://doi.org/10.1097/ACM.0b013e318276bcc4

Collins, J., Regenbrecht, H., Langlotz, T., Can, Y. S., Ersoy, C., & Butson, R. (2019). Measuring cognitive load and insight: A methodology exemplified in a virtual reality learning context. *Proceedings - 2019 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2019*, 351–362. https://doi.org/10.1109/ISMAR.2019.00033

Craig, S. D., D'Mello, S., Witherspoon, A., & Graesser, A. (2008). Emote aloud during learning with AutoTutor: Applying the Facial Action Coding System to cognitive - affective states during learning. *Cognition and Emotion*, *22*(5), 777–788. https://doi.org/10.1080/02699930701516759

Creswell, J. W., & Clark, V. L. P. (2018). Choosing a mixed methods design. In J. W. Creswell (Ed.), *Designing and conducting mixed methods research* (3rd ed., pp. 53–106). Sage.

Di Leo, I., Muis, K. R., Singh, C. A., & Psaradellis, C. (2019). Curiosity… Confusion? Frustration! The role and sequencing of emotions during mathematics problem solving. *Contemporary Educational Psychology*, *58*, 121–137. https://doi.org/10.1016/j.cedpsych.2019.03.001

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*(2), 145–157. https://doi.org/10.1016/j.learninstruc.2011.10.001

D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction*, *29*, 153–170. https://doi.org/10.1016/j.learninstruc.2012.05.003

Efklides, A. (2017). Affect, epistemic emotions, metacognition, and self-regulated learning. *Teachers College Record*, *119*, 130305.

Frenzel, A. C., Goetz, T., & Stockinger, K. (2024). Emotions and emotion regulation. In P. A. Schultz & K. R. Muis (Eds.), *Handbook of Educational Psychology* (pp. 219–244). Taylor and Francis. https://doi.org/10.4324/9780429433726-13

Genc Aksaray, S., & Ozcelik, E. (2022). Attention mediates the effect of emotional arousal on learning outcomes in multimedia learning: an eye-tracking study. *Educational Psychology*. https://doi.org/10.1080/01443410.2022.2154317

Greene, J. A. (2022). What can educational psychology learn from, and contribute to, theory development scholarship? *Educational Psychology Review*, *34*(4), 3011–3035. https://doi.org/10.1007/s10648-022-09682-5

Hattie, J. (2023). *Visible learning: The sequel: A synthesis of over 2,100 meta-analyses relating to achievement*. Routledge.

Järvelä, S., Malmberg, J., Haataja, E., Sobocinski, M., & Kirschner, P. A. (2021). What multimodal data can tell us about the students' regulation of their learning process? *Learning and Instruction*, *72*, 101203. https://doi.org/10.1016/j.learninstruc.2019.04.004

Johnson, C. C., Sondergeld, T. A., & Walton, J. B. (2019). A study of the implementation of formative assessment in three large urban districts. *American Educational Research Journal*, *56*(6), 2408–2438. https://doi.org/10.3102/0002831219842347

Kippers, W. B., Wolterinck, C. H. D., Schildkamp, K., Poortman, C. L., & Visscher, A. J. (2018). Teachers' views on the use of assessment for learning and data-based decision making in classroom practice. *Teaching and Teacher Education*, *75*, 199–213. https://doi.org/10.1016/j.tate.2018.06.015

Kostons, D., Van Gog, T., & Paas, F. (2009). How do I do? Investigating effects of expertise and performance-process records on self-assessment. *Applied Cognitive Psychology*, *23*(9), 1256–1265. https://doi.org/10.1002/acp.1528

Lee, H., Chung, H. Q., Zhang, Y., Abedi, J., & Warschauer, M. (2020). The effectiveness and features of formative assessment in US K-12 education: A systematic review. *Applied Measurement in Education*, *33*(2), 124–140. https://doi.org/10.1080/08957347.2020.1732383

Malecka, B., Boud, D., & Carless, D. (2022). Eliciting, processing and enacting feedback: Mechanisms for embedding student feedback literacy within the curriculum. *Teaching in Higher Education*, *27*(7), 908–922. https://doi.org/10.1080/13562517.2020.1754784

McIntyre, N. A., Draycott, B., & Wolff, C. E. (2022). Keeping track of expert teachers: Comparing the affordances of think-aloud elicited by two different video perspectives. *Learning and Instruction*, *80*, 101563. https://doi.org/10.1016/j.learninstruc.2021.101563

Muis, K. R., Chevrier, M., & Singh, C. A. (2018). The role of epistemic emotions in personal epistemology and self-regulated learning. *Educational Psychologist*, *53*(3), 165–184. https://doi.org/10.1080/00461520.2017.1421465

Munzar, B., Muis, K. R., Denton, C. A., & Losenno, K. (2021). Elementary students' cognitive and affective responses to impasses during mathematics problem solving. *Journal of Educational Psychology*, *113*(1), 104–124. https://doi.org/10.1037/edu0000460

Nicol, D. (2021). The power of internal feedback: Exploiting natural comparison processes. *Assessment and Evaluation in Higher Education*, *46*(5), 756–778. https://doi.org/10.1080/02602938.2020.1823314

Nicol, D., & MacFarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. https://doi.org/10.1080/03075070600572090

Nieminen, J. H. (2024). How does assessment shape student identities? An integrative review. *Studies in Higher Education*. https://doi.org/10.1080/03075079.2024.2334844

Obergriesser, S., & Stoeger, H. (2020). Students' emotions of enjoyment and boredom and their use of cognitive learning strategies – How do they affect one another? *Learning and Instruction*, *66*. https://doi.org/10.1016/j.learninstruc.2019.101285

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., & Alonso-Tapia, J. (2013). Student self-assessment: Theoretical and practical connotations. When it happens, how it is acquired, and how to develop it in our students. *Electronic Journal of Research in Education Psychology*, *11*(30), 551–576. https://doi.org/10.14204/ejrep.30.12200

Panadero, E., Brown, G. T. L., & Strijbos, J. W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review*, *28*(4), 803–830. https://doi.org/10.1007/s10648-015-9350-2

Panadero, E., Fernández, J., Pinedo, L., Sánchez, I., & García-Pérez, D. (2024). A self-feedback model (SEFEMO): secondary and higher education students' self-assessment profiles. *Assessment in Education: Principles, Policy and Practice*, *31*(3–4), 221–253. https://doi.org/10.1080/0969594X.2024.2367027

Panadero, E., Fernández-Ruiz, J., & Sánchez-Iglesias, I. (2020). Secondary education students' self-assessment: The effects of feedback, subject matter, year level, and gender.

Assessment in Education: Principles, Policy & Practice, *27*(6), 607–634. https://doi.org/10.1080/0969594x.2020.1835823

Panadero, E., García-Pérez, D., Fernández Ruiz, J., Fraile, J., Sánchez-Iglesias, I., & Brown, G. T. L. (2023). Feedback and year level effects on university students' self-efficacy and emotions during self-assessment: Positive impact of rubrics vs. instructor feedback. *Educational Psychology*, 1–24. https://doi.org/10.1080/01443410.2023.2254015

Panadero, E., Lipnevich, A., & Broadbent, J. (2019). Turning self-assessment into self-feedback. In *The Impact of Feedback in Higher education: Improving assessment outcomes for learners* (pp. 147–163). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-25112-3_9

Panadero, E., Pérez, D. G., Ruiz, J. F., Fraile, J., Sánchez-Iglesias, I., & Brown, G. T. L. (2022). University students' strategies and criteria during self-assessment: Instructor's feedback, rubrics, and year level effects. *European Journal of Psychology of Education*, *58*, 1031–1051. https://doi.org/10.1007/s10212-022-00639-4

Park, D., & Ramirez, G. (2022). Frustration in the classroom: Causes and strategies to help teachers cope productively. *Educational Psychology Review*, *34*(4), 1955–1983. https://doi.org/10.1007/s10648-022-09707-z

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Pekrun, R. (2024). Control-value theory: From achievement emotion to a general theory of human emotions. *Educational Psychology Review*, *36*(3). https://doi.org/10.1007/s10648-024-09909-7

Pekrun, R., & Linnenbrink-Garcia, L. (2014). *International handbook of emotions in education* (R. Pekrun & L. Linnenbrink-Garcia, Eds.). Routledge.

Pekrun, R., & Perry, R. P. (2014). Control-value theory of achievement emotions. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 120–141). Routledge.

Pekrun, R., Vogl, E., Muis, K. R., & Sinatra, G. M. (2017). Measuring emotions during epistemic activities: The Epistemically-Related Emotion Scales. *Cognition and Emotion*, *31*(6), 1268–1276. https://doi.org/10.1080/02699931.2016.1204989

Reber, R., & Greifeneder, R. (2017). Processing fluency in education: How metacognitive feelings shape learning, belief formation, and affect. *Educational Psychologist*, *52*(2), 84–103. https://doi.org/10.1080/00461520.2016.1258173

Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research, and Evaluation*, *11*, 10. https://doi.org/10.7275/9wph-vv65

Schellings, G. L. M., Van Hout-Wolters, B. H. A. M., Veenman, M. V. J., & Meijer, J. (2013). Assessing metacognitive activities: The in-depth comparison of a task-specific questionnaire with think-aloud protocols. *European Journal of Psychology of Education*, *28*(3), 963–990. https://doi.org/10.1007/s10212-012-0149-y

Sitzmann, T., Ely, K., Brown, K. G., & Bauer, K. N. (2010). *Self-Assessment of knowledge: A cognitive learning or affective measure? 9*(2), 169–191.

Stark, L., Brünken, R., & Park, B. (2018). Emotional text design in multimedia learning: A mixed-methods study using eye tracking. *Computers and Education*, *120*, 185–196. https://doi.org/10.1016/j.compedu.2018.02.003

Tan, K. (2007). Conceptions of self-assessment: What is needed for long-term learning? In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education* (pp. 114–127). Taylor & Francis Group. http://ebookcentral.proquest.com/lib/queen-ebooks/detail.action?docID=292915

Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, *27*(2), 237–246. https://doi.org/10.1177/1098214005283748

Veugen, M. J., Gulikers, J. T. M., & den Brok, P. (2021). We agree on what we see: Teacher and student perceptions of formative assessment practice. *Studies in Educational Evaluation*, *70*, 101027. https://doi.org/10.1016/j.stueduc.2021.101027

Wass, R., Timmermans, J., Harland, T., & McLean, A. (2020). Annoyance and frustration: Emotional responses to being assessed in higher education. *Active Learning in Higher Education*, *21*(3), 189–201. https://doi.org/10.1177/1469787418762462

Winne, P. H. (2018). Discussion paradigmatic issues in state-of-the-art research using process data. *Frontline Learning Research*, *6*(3), 250–258. https://doi.org/10.14786/flr.v6i3.551

Winne, P. H., Teng, K., Chang, D., Lin, M. P. C., Marzouk, Z., Nesbit, J. C., Patzak, A., Rakovic, M., Samadi, D., & Vytasek, J. (2019). NStudy: Software for learning analytics about learning processes and self-regulated learning. *Journal of Learning Analytics*, *6*(2), 95–106. https://doi.org/10.18608/jla.2019.62.7

Yan, Z., & Brown, G. T. L. (2017). A cyclical self-assessment process: Towards a model of how students engage in self-assessment. *Assessment and Evaluation in Higher Education*, *42*(8), 1247–1262. https://doi.org/10.1080/02602938.2016.1260091

Yan, Z., Brown, G. T. L., Lee, J. C. K., & Qiu, X. L. (2020). Student self-assessment: Why do they do it? *Educational Psychology*, *40*(4), 509–532. https://doi.org/10.1080/01443410.2019.1672038

Yan, Z., Lao, H., Panadero, E., Fernández-Castilla, B., Yang, L., & Yang, M. (2022). Effects of self-assessment and peer-assessment interventions on academic performance: A pairwise and network meta-analysis. *Educational Research Review*, *37*, 100484. https://doi.org/10.1016/j.edurev.2022.100484

Chapter 5: General Discussion

The Interaction Theory of Student Self-Assessment: The Roles of Cognition and Affect, Metacognition, and Emotions

Facilitating student self-assessment across education systems is imperative for enhancing learning for all students, yet there is currently no theoretical framework that articulates how students' actions, cognition and affect, metacognition, and emotions interact within the self-assessment process—a critical barrier to understanding the phenomenon of self-assessment and how it can be supported effectively in classrooms. Research has documented critical cognitive learning benefits of engaging in student self-assessment, including increased academic achievement (Lee et al., 2020; Yan et al., 2021; Yan, Lao, et al., 2022), creativity (Bolden & DeLuca, 2022; Yan, Lee, et al., 2022), self-regulated learning (Andrade, 2019; Panadero et al., 2017), and affective benefits including enhanced self-efficacy, positive emotions about learning, and motivation to learn (Panadero et al., 2023; Ross, 2006; Sitzmann et al., 2010). These findings signal the central role that students' capacity to self-assess plays in their learning and success in school and beyond (Hattie, 2023; Tan, 2007). Moreover, researchers have identified critical socio-cultural and societal benefits which heighten the need to foster student self-assessment in the face of unprecedented global challenges (e.g., climate change, social inequity). Student self-assessment can foster more inclusive classrooms by integrating students' values and voices in assessment and learning, reducing colonial power imbalances between learners and education systems (Ahmadi, 2022; Bourke & Mentis, 2013; Claypool & Preston, 2011; Nieminen, 2022) and is a requisite to developing students' global competencies—e.g., taking responsibility, self-regulated learning, creativity—which the Organization for Economic

Cooperation and Development (OECD) have identified as essential educational outcomes for students to address global crises (e.g., climate change, social inequity OECD, 2019, 2023). These socio-cultural and societal benefits have amplified researchers' and policymakers' calls to facilitate student self-assessment across education systems as a central goal of classroom assessment (e.g., Birenbaum et al., 2015; British Columbia Ministry of Education, 2023; Panadero et al., 2018).

Student self-assessment is theorized as a process comprised of actions, thoughts, and feelings that students mobilize to process information about, analyze, and generate self-feedback on their performance (Panadero et al., 2016, 2024). If researchers and educators are to develop pedagogical approaches to teaching student self-assessment, they must first understand the self-assessment process. Comprehensive descriptions of the phenomenon of student self-assessment can additionally help educators and researchers develop and implement tailored interventions and support within classrooms. Such descriptions are also needed to guide future research, as they inform researchers of all aspects of constructs and highlight key priorities for research (Greene, 2022). To date, researchers have empirically documented the actions and practices which comprise the self-assessment process (Panadero et al., 2024; Yan & Brown, 2017). These actions and practices, which include reading their work, comparing their work to indicators of quality (e.g., previous work, feedback, rubrics, exemplars), drawing on criteria, seeking external feedback via monitoring (e.g., completing a practice test, referring to a rubric), and seeking external feedback via inquiry (e.g., asking a peer for feedback), can be understood as behaviours that students engage in when self-assessing that have the student's work as the object of focus (i.e., that are directed at or provoked by the student's work). In addition to these behaviours, researchers have theorized that the self-assessment process also involves cognitive and affective

processes, metacognition, and emotions. However, to the best of my knowledge there is currently no theoretical framework which takes into consideration cognition and affect, metacognition, and emotions, and how they interact within the self-assessment process. Therefore, the purpose of this chapter is to advance the interaction theory of student self-assessment which integrates cognition and affect, metacognition, and emotions with self-assessment actions—and how these different dimensions of the self-assessment process interrelate—to inform an initial comprehensive theory of student self-assessment.

This interaction theory combining self-assessment actions, cognition and affect, metacognition, and emotions is essential for advancing self-assessment theory and supporting educational implications. While understanding the actions that comprise self-assessment is important, informing pedagogies and interventions that facilitate effective student self-assessment also requires understanding the deeper cognitive and affective processes that underpin and support these actions. Common and central self-assessment actions, such as comparing performance to rubrics and exemplars (Lin-Siegler et al., 2015; Nicol, 2021), have been shown to have differential effects on learning, with meta-analyses and reviews reporting of negative to large positive learning effects of the same self-assessment action (Andrade, 2019; Brown & Harris, 2013; To et al., 2021; Yan, Lao, et al., 2022). Researchers have theorized that much of this variance is related to how students engage in these actions and practices, including how they are supported in doing so. For example, the training and support students receive on how to use rubrics to self-assess their work may shape the effects of using rubrics on achievement (Brown & Harris, 2013). Moreover, students have reported different levels of cognitive load, which refers to the demand placed on cognitive and affective resources, associated with similar self-assessment actions, such as using rubrics and exemplars (Fastré et

al., 2012; Krebs et al., 2022; Lipnevich et al., 2023). These findings suggest that (a) there are underlying cognitive and affective processes associated with self-assessment actions and (b) these processes are central mechanisms shaping the learning outcomes of self-assessment actions (Andrade, 2019).

Additionally, researchers have theorized a metacognitive dimension of student self-assessment which encompasses students' monitoring and control of their own self-assessment processes. These researchers have emphasized the importance of students' capacity to articulate their self-assessment processes—i.e., their actions and thinking while self-assessing—as essential for making these processes visible to educators so that they can intervene and guide students' approaches (Nicol, 2021; Panadero, Lipnevich, et al., 2019). Metacognition theory holds that learning tasks activate a meta-level of cognition where learners maintain and continuously update a mental model of their cognition. This mental model allows learners to monitor and control their cognition, making decisions about thinking strategies based on their appraisals of their thinking (Flavell, 1979). When a student is engaged in self-assessment, metacognition theoretically involves students monitoring and controlling their own self-assessment actions. What distinguishes self-assessment metacognition from self-assessment actions is that actions have the student's work as the object of focus (e.g., read essay, revise the essay, compare the essay to a rubric) while metacognition has the self-assessment process itself as the object of focus (e.g., realizing something about the self-assessment process, intentionally selecting a self-assessment approach or strategy). Metacognition is therefore likely central to how students engage in self-assessment and the effects that self-assessment has on their learning. Understanding the metacognitive mechanisms whereby students monitor and control their self-assessment processes is a critical step in supporting students in making their self-assessment

processes explicit and identifying and selecting effective self-assessment actions in response to context.

Emotions that students experience during self-assessment also play a central role in the self-assessment process. Because it may relate to consequential learning contexts (e.g., graded assignments, exam preparation) and informs students' self-concepts, student self-assessment is theorized to be a highly emotional process (Baca & Smith, 2018; Nicol, 2021; Sitzmann et al., 2010). Like metacognition, emotions play an important role in learning. The emotions students experience during learning shape the strategies they employ (Boekaerts, 2010), influence how new information is encoded in memory (Genc Aksaray & Ozcelik, 2022; Pekrun & Perry, 2014; Winne, 2018), and direct students' attention (Belkaid et al., 2017) which shapes the cognitive processing of information (Stark et al., 2018). Importantly, self-assessment is a fundamental subprocess of self-regulated learning (Andrade, 2010; Boekaerts, 1991) and research has demonstrated that emotions shape the self-regulated learning strategies that students employ (Chevrier et al., 2019; Di Leo et al., 2019; D'Mello et al., 2014; Muis et al., 2015)—further indicating that the emotions students experience during self-assessment influence their self-assessment actions. These findings collectively signal that emotions play an important role in determining the effectiveness of students' approaches to learning and their perseverance with important learning tasks. Understanding the emotions that students experience as part of the self-assessment process and how these emotions may shape their processes can enable educators to support effective student self-assessment and engagement in the process.

Given the importance of these theorized dimensions for the efficacy of student self-assessment, an interaction theory of the student self-assessment process is needed to inform research and practice. The interaction theory below builds directly on the self-feedback model

developed by Panadero et al. (2024) by integrating empirical and theoretical evidence of how self-assessment actions, cognition and affect, metacognition, and emotions interact. Below, I summarize the theory and then describe and substantiate each of these dimensions with extant empirical evidence on self-assessment and explore theorized connections by drawing on other fields of scholarship—i.e., visual attention, metacognition, emotions, self-regulated learning—to elicit areas for future inquiry. Visual attention research offers an interpretive lens for understanding the dynamics of the cognitive and affective processes underpinning specific actions during self-assessment that involve the use of visual resources (e.g., students' work, rubrics, exemplars). Metacognition theory builds on empirical evidence of the ways in which students think about and control their own self-assessment processes. Research on self-regulated learning, a process whereby students manage their thinking, emotions, and behaviours in pursuit of goals, is predicated on self-assessment (Boekaerts, 1991; Panadero, 2017) and metacognition (Greene et al., 2024); therefore, self-regulated learning research and research on how emotions shape self-regulated learning extend empirical understandings of how emotions interact with other dimensions of student self-assessment.

The proposed theory adds substantively to self-assessment theory by illuminating core dimensions of the student self-assessment process and how these are interrelated. This interaction theory is partly a response to direct calls from researchers for exploratory research to more fully elucidate the self-assessment process, as many aspects are not yet well understood (Andrade, 2019; Brown & Harris, 2014; Nicol, 2021; Panadero et al., 2016; Yan et al., 2020). Additionally, it holds important educational implications. Decades of research from across education contexts has highlighted a persistent challenge that self-assessment is not yet widely practiced in schools. Educators and students have reported that student self-assessment is among

the least used pedagogical practices in their classrooms, citing a lack of guidance for facilitating student self-assessment as a primary barrier to its use (Brandmo et al., 2020; Kippers et al., 2018; Panadero, Fraile, et al., 2019; Veugen et al., 2021; Volante & Beckett, 2011). These findings mirror researchers' assertions that more empirical insights into the self-assessment process are needed to inform pedagogy (e.g., Brown & Harris, 2014). This interaction theory therefore is a response to a strong need to advance descriptive aspects of student self-assessment theory to more fully elucidate the self-assessment process so that practical guidance can be developed and enhanced.

The Interaction Theory of Student Self-Assessment

The interaction theory of student self-assessment is presented in Figure 1. The theory is depicted in horizontal layers, encompassed in dotted lines, with each layer articulating a dimension of the student self-assessment process. At the core of this theory is the "Self-Assessment Actions" layer, which is comprised of the three-phase self-feedback model developed by Panadero et al. (2024). The interaction theory thereby builds directly on Panadero et al.'s (2024) self-reflection model, providing an understanding of how self-assessment unfolds in time via interaction among the different layers. This self-feedback model articulates the actions, or behaviours, at the heart of the self-assessment process that have the student's work as the object of focus. In other words, it reflects what students do when they are engaged in self-assessment. The layer immediately below, "Cognitive and Affective Processes", depicts the deeper cognitive and affective processes underlying self-assessment actions. These cognitive and affective processes represent what students may be thinking about and feeling while using self-assessment resources, such as rubrics, exemplars, self-feedback, and their own essays, to facilitate self-assessment actions including reading, comparing, and assessing. The top layer of

the interaction theory, “Self-Assessment Metacognition”, represents the metacognitive dimension of self-assessment, encompassing the metacognitive experiences whereby students develop and maintain a mental model of their own self-assessment process, and the metacognitive skills they apply to control their self-assessment process. The “Emotions” layer illustrates the emotions students may experience during self-assessment which may directly shape their self-assessment actions or influence their self-assessment metacognition.

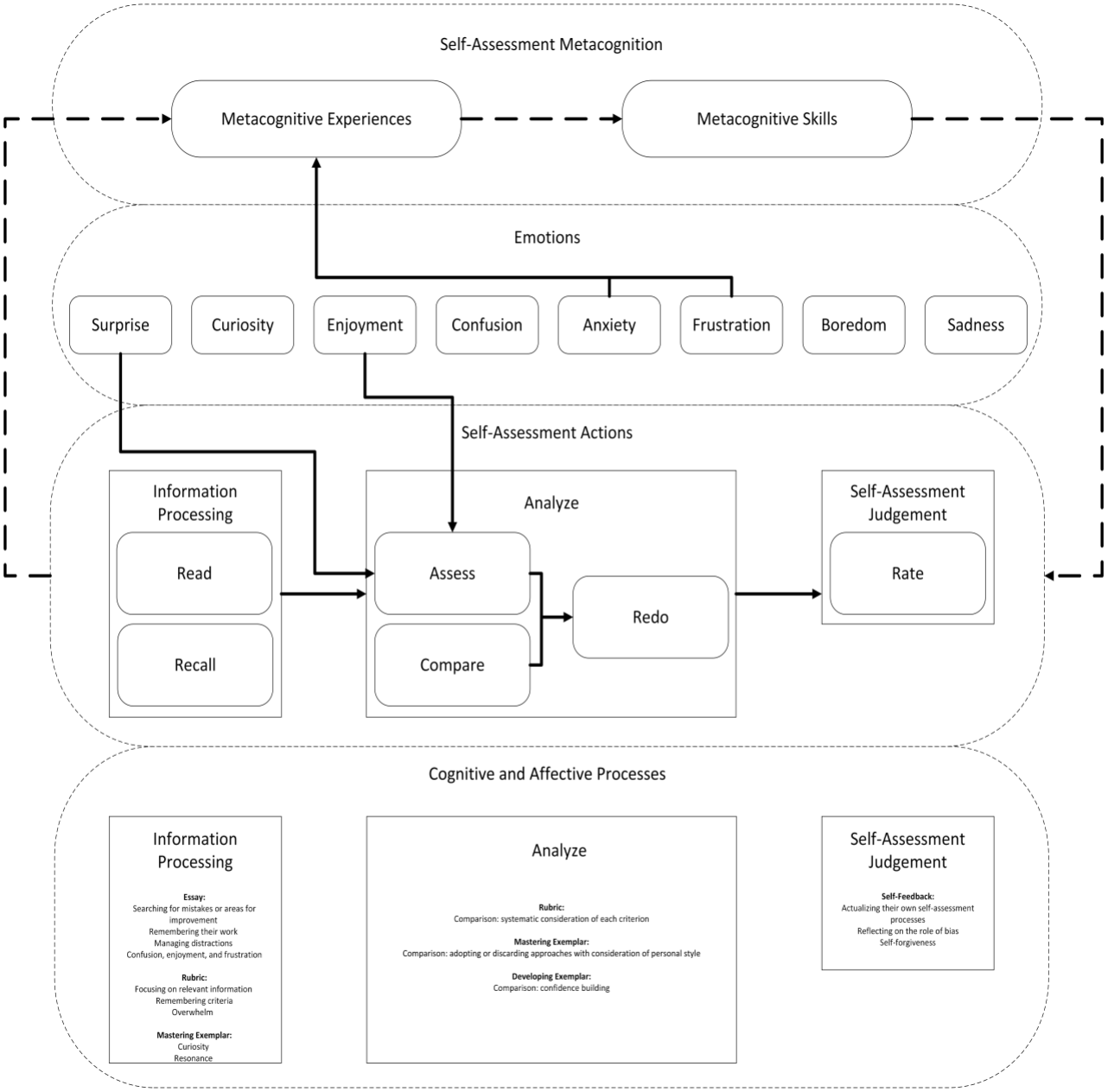
The interaction theory is presented in layers with broken borders, rather than a cycle or process, to indicate that interactions between the layers are highly recursive rather than reflecting a clear progression from one level of thinking or feeling to the next. What makes this an interaction theory is that its focus is on the interactions or connections between the dimensions of the self-assessment process, presented here as layers: i.e., the actions at the heart of the process; the cognition and affect underpinning those actions; the metacognition which monitors and controls the process; and emotions which arise during and shape the process. The solid arrows indicate directional relationships between actions, thoughts, and feelings that are supported by empirical evidence, while dotted arrows indicate theorized connections. In the below descriptions of each layer, I substantiate these connections by drawing on relevant scholarship from related domains, including self-regulated learning and the control value theory of achievement emotions (Pekrun, 2024).

Rather than describe the theory from the top layer to the bottom, I begin with the core of the self-assessment process—the actions that students employ with their work as the object of focus (Panadero et al., 2024)—i.e., the self-assessment actions layer. I then move to the cognitive and affective processes underpinning these actions, encompassed in the lowest layer titled cognitive and affective processes. Third, I describe the metacognitive layer as it is closely related

to self-assessment actions. Last, I describe the emotions layer which plays an important role in shaping both actions and metacognition.

Figure 1

The Interaction Theory of Student Self-Assessment



Student Self-Assessment Actions

The idealized student self-assessment process is comprised of six primary actions—i.e., distinct behaviours that have the student's work as the object of focus—which unfold across three overarching phases (Panadero et al., 2024). These actions are depicted in boxes within the self-assessment actions layer, contained within their respective phases (i.e., information processing, analysis, self-assessment judgement). Though these phases are understood to be recursive, ideally they progress from information processing to analysis to self-assessment judgement. The phases are derived from Panadero et al.'s (2024) self-feedback model and provide a way to understand how self-assessment actions relate to each other and progress towards a self-assessment judgement. The phases are not to be confused with layers in the interaction theory, which articulate different dimensions of the self-assessment process. While the phases are best described as unfolding chronologically, the layers can interact at any phase in the process.

First, students activate information processing where they engage with internal and external resources which can help them understand and situate the quality of their own work. The primary actions supporting information processing are Read (e.g., [re]read the essay, read the rubric, read the exemplar) and Recall (recall the essay). Next, analysis involves students applying their thinking about resources and criteria to their own work. Analysis is comprised of three actions: Compare, Assess, and Redo. Compare involves students explicitly comparing their work to a resource or indicator of quality, such as a rubric, feedback, or an exemplar. Assess involves students explicitly drawing on criteria to evaluate their work. This action is divided into two levels, depending on the nature of the criteria students use. Assess Level 1 refers to students drawing on internal and affect-based criteria, such as intuition or hindsight. Assess Level 1

describes students drawing on external criteria, such as those provided by a teacher or articulated in a rubric. Redo is an action that involves students considering or making revisions to their work which is also divided into two levels. Redo Level 1 refers to students considering how they would redo or revise all or parts of their work, while Redo Level 2 refers to students making actual revisions to all or parts of their work. The Redo action serves dual purposes of (a) improving the work and (b) helping the student better understand their performance by exploring alternative and ideally improved approaches, drawing on new information (e.g., from feedback or a rubric). The last phase, self-assessment judgement, is comprised of one action, Rate, which involves students making an overall rating of their work by synthesizing their information processing and analysis. This judgement may simply be a holistic rating of performance (e.g., good), a self-assigned grade or rubric score, or a decision about next steps in learning. Importantly, the objects of these self-assessment actions are the students' work (e.g., essay) or self-assessment resources—not the students' own self-assessment process (see "Self-Assessment Metacognition").

Panadero and colleagues developed the self-feedback model based on their analysis of think aloud and video recording data from hundreds of students in secondary and university education across three subject contexts: Spanish, math, and writing (2020, 2022, 2024). Through their analysis, they identified strategies (i.e., specific, content dependent actions; e.g., read the essay [writing], review the signs [math]) and criteria (e.g., intuition, rubric) that students drew on during self-assessment which they synthesized into the six core actions and three phases. Our recent research, which examined how emotions relate to self-assessment actions, supported the self-feedback model as an effective lens for interpreting students' self-assessment processes (see Chapter 4). In this study, we observed the same six core self-assessment actions—i.e., Read,

Recall, Compare, Assess (Levels 1 and 2), Redo (Levels 1 and 2), and Rate—in the self-assessment processes of a diverse sample of international students hailing from every continent. Notably, our study used gaze-cued retrospective think alouds, rather than the concurrent think alouds used to develop the self-feedback model, corroborating the model by using a different data collection method. Moreover, our design provided participants with the freedom to select and open or close a number of resources on the screen at any time during self-assessment, including an analytic writing rubric, a mastering (high quality) exemplar, a developing (low quality) exemplar, a blank document for self-feedback, the essay task instruction sheet, and their essays with tracked changes enabled. This design is significant because, in the studies that informed the self-feedback model, students followed standardized orders of stimuli (e.g., their work, feedback, rubric) to guide the self-assessment and did not have access to most of these resources. Our research, therefore, indicates that the self-assessment actions can be observed across different approaches to self-assessment, including more standardized ones and ones where students have more agency to select the resources they use.

Importantly, the self-assessment actions layer is an idealized process representing the full range of actions which comprise self-assessment. Students may not engage in all these actions. Panadero et al. (2020, 2022, 2024) developed self-assessment profiles to categorize self-assessment processes and found that most students tended to articulate no self-assessment or superficial self-assessment profiles. These profiles were characterized by primarily reading or describing work, with infrequent attempts to rate work holistically or draw on internal, gut-reaction criteria. This trend suggested that students may commonly engage in shallow or superficial self-assessment where they rush to rate their performance without deeply engaging with assessment criteria (Boud & Brew, 1995). Our findings corroborated this trend, as Read and

Rate were among the most frequently observed self-assessment actions (Chapter 4). Additionally, Assess Level 1 and Redo Level 1 (i.e., drawing on gut reactions and considering changes to work, respectively) were observed much more frequently than Assess Level 2 and Redo Level 2 (i.e., drawing on external criteria and making actual revisions). In fact, Redo Level 2, where students redo or revise parts or all their work, was the least observed self-assessment action.

Together, these findings suggest that students may use actions in the analysis phase of the self-assessment process—especially Assess Level 2 and Redo Level 2—less than actions in the information processing and self-assessment judgement phases. In other words, students may tend to review their work and rush to a self-assessment judgement without deeply analyzing the available information. This trend is concerning as researchers have argued that self-assessment may be most effective when it engages students deeply in self-regulated learning subprocesses and when students think critically about their own work (Boud & Brew, 1995; Brown & Harris, 2013). Additionally, Panadero et al. (2024) concluded that redo plays a particularly important role in the self-assessment process, allowing students to explore alternative approaches in order to (dis)confirm their original responses or better understand the quality of their performance in relation to success criteria. Further, making revisions based on self-assessment can help students see the value of self-assessment, enhance their self-efficacy for the tasks they self-assess (Ross, 2006), and learn to use self-feedback effectively to advance their learning (Andrade, 2019).

Cognitive and Affective Processes

The cognitive and affective processes layer describes the deeper cognitive and affective processing that students activate when engaging in self-assessment actions. Using joint display integration of eye tracking and think aloud data, we identified the cognitive and affective

processes that participants articulated when looking at five self-assessment resources: an analytic writing rubric, a mastering (high quality) exemplar, a developing (low quality) exemplar, their essays, and their self-feedback documents (Chapter 2). When integrated with Panadero et al.'s (2024) self-feedback model (i.e., self-assessment actions), these cognitive and affective processes can be aligned with specific self-assessment actions and distributed within the self-assessment actions layer. Some cognitive and affective processes are associated with reading resources and recalling information and therefore occur within the information processing phase. Cognitive and affective processes we found which were based on explicit comparisons (e.g., between resources and work, between multiple resources) occur within the analysis phase, when students are thinking more deeply about their work and available resources. Last, we found cognitive and affective processes that underpinned participants' self-generation of feedback, illuminating the thinking and feeling processes that support students' ratings and assessments at the first and second level. These findings elucidate the cognitive and affective processes students engage in when using self-assessment resources to enact self-assessment actions.

Information Processing

In the information processing phase, cognition and affect can be understood in terms of top-down and bottom-up attentional influences when students are using visual resources to self-assess (e.g., essays, rubrics, exemplars). In visual attention research, top-down influences refer to a person's higher-level cognitions or goals that drive their attention, whereas bottom-up salience refers to visual artifacts in the person's field of view that draw their attention. Navigating visual environments is an interplay of top-down and bottom-up influences on attention (Poltoratski et al., 2017). In our study (Chapter 2), participants' reading and recalling was driven by top-down influences where they set goals or intentions to guide these actions: *searching for mistakes or*

areas for improvement and *remembering their work* in their essays; and *focusing on relevant information* and *remembering criteria* in the rubric. These top-down influences guided their attention to resources and to specific parts of resources (e.g., information they deemed most relevant) and therefore guided their self-assessment actions. At the same time, participants reported bottom-up influences which guided their attention during reading and recalling: *managing distractions* and feelings of *confusion, enjoyment and frustration* when using their essays; feelings *overwhelm* when using the rubric; and feelings of *curiosity* and *resonance* when using the mastering (high-quality) exemplar. Self-assessment actions associated with information processing (i.e., read and recall) can then be understood holistically as a cognitive and affective balancing of tensions between top-down and bottom-up influence on attention.

This attention framing of information processing has important implications. Attention refers to a variety of mechanisms whereby the brain preferentially processes some environmental stimuli over others. The brain's finite resources mean it cannot, at any given time, fully process all the sensory information it receives (Fiebelkorn & Kastner, 2020). What visual resources a person attends to largely determines what information they process and internalize (Stark et al., 2018). For goal-directed top-down processes, working memory plays an important role as it refers to neural mechanisms which work in concert to hold information for readily available use in goal-directed processes (Thigpen et al., 2019), indicating the working memory plays an important role in student self-assessment when students are engaged in actions to process available information about their work (e.g., searching for mistakes in an essay, finding relevant information in a rubric).

Regarding bottom-up influences, research has indicated that visual cues can affect the information people attend to in working memory (Fang et al., 2019), suggesting that visual cues

in self-assessment resources (e.g., highlighting specific parts of an exemplar) could shape what a student attends to and therefore processes during self-assessment. Further, attention researchers have documented a phenomenon called predication error which learners experience when they detect a discrepancy between their expectations and observations in a visual environment which may focus the learner's attention on making sense of the discrepancy (Den Ouden et al., 2012). Prediction error could underpin the influence of bottom-up cognitive and affective processes, managing distractions and affective processes (e.g., overwhelm), that we observed when students are processing visual information to support self-assessment (Chapter 2). Emotions, another important bottom-up influence, have been shown to direct attention in a visual search task in the context of self-assessment (Belkaid et al., 2017), adding support to our findings that affect may shape attention during the information processing phase of self-assessment. These attentional dynamics, because they determine the information that students process, are important because they form the first phase of the self-assessment process and likely affect the following phases.

Analysis

In the analysis phase, students engage higher-level thinking about their performance. Drawing on the information they processed in the first phase, they may (a) Compare their work to other indicators of quality (e.g., rubric, exemplar), (b) Assess their work using internal (Level 1) and external (Level 2) criteria, and (c) Redo parts of or all their work, either hypothetically (Level 1) or concretely (Level 2; Panadero et al., 2024). In our research, participants articulated evidence of this higher-level thinking when they used the rubric and two contrasting exemplars (i.e., mastering and developing) to provoke comparison processes (Chapter 2). These findings substantiate theoretical perspectives and echoing previous studies which have implicated comparison as central to self-assessment and internal self-feedback processes (Butler & Winne,

1995; Nicol, 2021; Panadero et al., 2024; Yan, 2020). Our findings advanced these perspectives by illuminating different types of comparative processes participants engaged in while fixating on different self-assessment resources. When using rubrics, participants engaged in a systematic comparison process, comparing their work to one performance descriptor at a time. This finding builds on research which showed that students have reported using rubrics systematically, referring to a rubric before, during, and after assignment completion to ensure they addressed each criterion (Andrade & Du, 2007), adding that students may try to identify the most relevant criteria (i.e., in the information processing phase) and systematically consider how their work compares to each of these (i.e., in the analysis phase)—possibly as a way of managing overwhelm experienced in the information processing phase from the large amount of information contained in an analytic rubric (Chapter 2).

When using high-quality exemplars, students may employ comparison less systematically, instead making more holistic and intuitive comparisons with a focus on approximating the level of performance in the exemplar in ways that were authentic to their own writing style (Chapter 2). Researchers have noted that exemplars, especially when used in combination with rubrics, can provide students with concrete examples of success criteria which are often abstract (Lin-Siegler et al., 2015; Lipnevich et al., 2023; To et al., 2021). In our study, we found no evidence that students used exemplars systematically to identify examples of discrete success criteria, with students making few comparisons between the rubric and exemplars (Chapter 2). However, theoretically it is logical that students may search through high-quality exemplars to identify concrete examples of success criteria, so this could be an additional cognitive and affective process in the analysis phase.

Low-quality exemplars may provoke positive appraisals via comparison. We found that a

developing exemplar, in contrast to a high-quality mastering one, provoked students to make positive comparisons to their own performance, supporting feelings of confidence (Chapter 2). Boosts to confidence could enhance students' self-efficacy for the task they are self-assessing (Ross, 2006) and fortify their motivation to persist in self-assessment (Chevrier et al., 2019).

Self-Assessment Judgement

In the self-assessment judgement phase, students form a judgement about their work or performance which may be informed by the information they processed and analyzed (Panadero et al., 2024). In our research (Chapter 2), we found that when fixating on their self-feedback, participants engaged in cognitive processes to make sense of their feedback, including considering the role of bias in their thinking and considering the wider context of their work to support self-forgiveness. Participants also drew on their knowledge of their own personal approaches to self-assessment to determine how they structured their self-feedback. Self-judgements may also be supported by recalled relevant events or experiences (e.g., similar tasks in the past) and by setting reference points for evaluation (e.g., using a rubric, comparing to a peer's performance; Butler, 2018). Together, these findings demonstrate that students activate cognition, metacognition (e.g., reflecting on how bias may influence their self-judgements), and affect (i.e., self-forgiveness) to rationalize self-judgements. The quality and utility of these self-judgements may depend on the depth to which learners engaged in the analysis phase, as rushing to a self-judgement without thoughtful consideration of success criteria and alternate approaches could lead to inaccurate or superficial self-assessment (Boud & Brew, 1995; Panadero et al., 2024) or lack of development of self-regulated learning processes (Brown & Harris, 2013).

Self-Assessment Metacognition

Next, I move to the top of the interaction theory and describe the metacognition layer of

the self-assessment process which governs students' self-assessment actions. During self-assessment, in addition to thinking about their own work, students may be aware of and thinking about their self-assessment process. This awareness allows them to evaluate the effectiveness of their self-assessment actions and approaches and ultimately control the process in pursuit of their goals (e.g., improving their work, determining their current achievement). Metacognition theory, which describes the phenomenon of a learner thinking about features of their own cognition (Winne & Azevedo, 2014), provides a theoretical basis for this notion. Metacognition theory is founded on the idea that cognition involves both an object-level and a meta-level which maintains a perception of the object-level. Monitoring involves the flow of information from the object-level to the meta-level and enables learners to maintain and update their perceptions of their object-level thinking. Control refers to the opposite flow of information, where learners modify their object-level thinking based on monitoring (Nelson & Narens, 1990). In the context of student self-assessment, metacognition theory implies an object-level of cognition where the student's thinking is focused on the object of self-assessment, i.e., the work, performance, or skill that the student is self-assessing (e.g., essay), and a meta-level where the student's thinking is focused on the self-assessment process itself. The object-level is encompassed in the self-assessment actions layer of the interaction theory, as these information processing, analysis, and self-assessment judgement actions are focused on the student's work. Drawing on metacognition theory and our own empirical evidence (Chapter 3), we propose a self-assessment metacognition layer to describe the meta-layer and its role in the self-assessment process.

Assessment researchers have theorized a metacognitive dimension of student self-assessment. They have advocated for educators to support students in articulating their self-assessment processes, making these internal processes explicit so that educators and researchers

can develop interventions and pedagogical approaches to support students' self-assessment processes (Nicol, 2021; Panadero, Lipnevich, et al., 2019). This reasoning implies an ability to be aware of and control the self-assessment process—i.e., self-assessment metacognition. However, there have been few empirical investigations of self-assessment metacognition to describe how students monitor and control their self-assessment processes.

While metacognition can occur before, during, and after learning (Winne & Azevedo, 2014), how metacognition occurs during a learning process—like self-assessment—can be understood as unfolding via metacognitive experiences and metacognitive skills (Efklides, 2008; Flavell, 1979). Metacognitive experiences are cognitive and affective experiences that a learner is aware of during a learning task as they process information about the task, including metacognitive feelings (e.g., feelings of knowing, feelings of difficulty), judgements (e.g., judgements of learning, estimates of effort), and online task-specific knowledge (e.g., metacognitive knowledge about the learning task). Metacognitive skills are deliberate strategies learners use to control cognition during a learning task, driven by metacognitive experiences and metacognitive knowledge. In the context of self-assessment, metacognitive knowledge includes knowledge about self-assessment; metacognitive experiences are the feelings, judgements, and knowledge that students are aware of during self-assessment; and metacognitive skills are strategies students consciously deploy during self-assessment to control the process.

Metacognitive Experiences

During self-assessment, students may have a range of metacognitive experiences as they process information about the self-assessment process. Our empirical research has suggested that students may have nine distinct metacognitive experiences during self-assessment: *abduction*, *efficacy judgements*, *feelings of difficulty*, *feelings of (dis)satisfaction*, *judgements of learning*,

monitoring focus or concentration, *noting the time or time management*, *realizing their self-assessment process*, and *reflecting on task value* (Chapter 3). In this study, the most frequent metacognitive experience was feelings of difficulty which were provoked by participants' doubts about the accuracy of their self-assessments and struggles to think of revisions. Feelings of difficulty may serve an important monitoring function for participants. Interpretated as an instantiation of processing disfluency, feelings of difficulty may offer students a heuristic for making quick decisions about the self-assessment process (Reber & Greifeneder, 2017). Given how confusion can benefit learning when a learner is able to resolve the negative emotion by addressing a gap in understanding (D'Mello et al., 2014), feelings of difficulty could ultimately be beneficial for students' self-assessment process if they draw students' attention to perceived issues or gaps in understanding in relation to the self-assessment process which they can then resolve. Metacognitive experiences—particularly feelings of difficulty—may occur relatively frequently during self-assessment and may help students efficiently identify when their approaches to self-assessment are not working as intended.

Theoretically, these metacognitive experiences are the primary in-the-moment mechanisms whereby students perceive their self-assessment actions. Metacognitive experiences are cognitive and affective experiences that occur during task completion that a learner is aware of and that enable them to maintain a perceived model of their cognition (Efklides, 2008). This theoretical connection—i.e., monitoring—is represented by a dotted arrow in the interaction theory from the self-assessment actions layer to metacognitive experiences, reflecting the flow of information from the object-level (self-assessment actions) to the meta-level (the student's perceptions of their self-assessment actions). However, this connection has yet to be examined empirically.

Metacognitive Skills

Based on their metacognitive experiences, students may use metacognitive skills to modify their self-assessment process during self-assessment. In our research (Chapter 3), we identified seven metacognitive skills participants articulated during self-assessment: *goal setting/reasoning*; *considering a strategy or approach*; *considering if they are finished self-assessing*; *selecting a strategy*; *reflecting on initial states or personal predispositions*; *reflecting on their personal self-assessment process*; and *reflecting on the writing process*. The most frequent metacognitive skill was selecting a strategy whereby participants verbalized a specific strategy they were using to manage the self-assessment process. Participants articulated in-the-moment skills, such as making evaluative notes about the exemplars, and explicitly summarized their approaches. Selecting a strategy could be a particularly important metacognitive skill during self-assessment that students engage to control their self-assessment process. Supporting this notion, previous findings based in digital trace data collected during self-assessment have indicated that students pause frequently during self-assessment to consider their next steps (Rickey et al., 2023). Additionally, research in self-regulated learning has implicated metacognitive reflection (i.e., inviting students to articulate their learning approaches) as an important strategy for enhancing self-regulated learning (Dignath & Büttner, 2008). The participants in our study (Chapter 3) were university students studying abroad in an international program, suggesting that they were likely highly self-regulated learners and therefore may more habitually engage in selecting a strategy—a foundational strategy for self-regulated learning—than less self-regulated learners. These findings collectively point to the importance of selecting a strategy as a metacognitive skill for enabling students to control the self-assessment process.

Moreover, students may engage in goal setting/reasoning to control their self-assessment

process. In our research, goal setting/reasoning was the second most frequent metacognitive skill participants articulated during self-assessment (Chapter 3). Participants expressed various goals or reasons for engaging in certain actions, including generating fair feedback, confirming if they were satisfied with a judgement or revision, and orienting themselves to the spectrum of performance. Self-assessment has been posited to supporting goal setting by helping students understand success criteria and calibrate their goals in theories of assessment as the co-regulation of learning (i.e., regulation with the support of teachers, peers, and resources; Andrade & Brookhart, 2020), supporting that engaging in student self-assessment may involve students setting and monitoring goals. Moreover, metacognitive goal setting (i.e., setting goals for the self-assessment processes) has been traced in secondary students' self-assessment processes as they self-assessed their writing in an authentic English literature lesson (Rickey et al., 2023).

Theoretically, metacognitive skills are the main mechanisms whereby students control their self-assessment actions while engaged in self-assessment, drawing on metacognitive experiences, which provide a perception of their self-assessment actions, and metacognitive knowledge (e.g., prior knowledge about self-assessment). Metacognitive skills are strategies that learners employ to adjust their cognition using metacognitive information from metacognitive experiences and metacognitive knowledge (Efklides, 2008). This theoretical connection—i.e., control—is represented by a dotted arrow in the theory from metacognitive skills to the self-assessment actions layer, reflecting the flow of information from the meta-level (i.e., the student's perceptions of their self-assessment actions) to the object-level (i.e., self-assessment actions). The arrow is dotted because has it yet to be substantiated empirically.

Emotions

Students may experience emotions as a part of the self-assessment process. Based on

gaze-cued retrospective think aloud data, our research has suggested that students may experience eight different types of emotions during self-assessment: *surprise*, *curiosity*, *enjoyment*, *confusion*, *anxiety*, *frustration*, *boredom*, and *sadness* (Chapter 4). These emotions align with those reported by learners in complex learning situations, which involve higher order thinking skills (e.g., evaluation, application and transfer of knowledge, questioning, problem diagnosis and solving; e.g., D'Mello & Graesser, 2012; Muis et al., 2015), such as student self-assessment. The types and intensities of emotions students experience during self-assessment may be influenced by their perceptions of the task value and academic control (Pekrun, 2006; Pekrun & Perry, 2014). If they value self-assessment and believe it can help them improve their work, they may be more likely to experience positive emotions (e.g., enjoyment, curiosity). However, if they see little value in self-assessment or believe they have little control of the self-assessment proves or their own achievement, they may be more likely to experience negative emotions, like frustration which, if unresolved, could lead to boredom and disengagement from self-assessment (Chevrier et al., 2019).

When the stakes of self-assessment are relatively low, self-assessment may be an emotionally positive experience, with students experiencing positive emotions (e.g., enjoyment, curiosity) frequently and more intensely than negative ones (e.g., anxiety, frustration). In our study (Chapter 4), participants reported experiencing emotions at a low to moderate intensity. Enjoyment was the most frequent emotion participants expressed during self-assessment, while boredom was the least frequent. Participants reported that curiosity was the most intensely felt emotion and boredom was the least intensely felt emotion. Participants experienced negative emotions relatively more frequently than positive ones; however, they reported that negative emotions were less intense than positive ones. Overall, these findings depicted self-assessment to

be a generally positive emotional experience, a depiction supported by research showing that self-assessment enhanced enjoyment of learning in a language learning context (Aldosari et al., 2023). Additionally, providing students with rubrics to support self-assessment can increase positive emotions while decreasing negative emotions (Panadero et al., 2023), making the experience more enjoyable. Frequent feelings of enjoyment may help students persist with self-assessment (Chevrier et al., 2019) and may be related to enhanced self-efficacy or agency which self-assessment can provoke (Bourke, 2016; Ross, 2006).

However, if students are assessing coursework that may be graded or have material consequences for their lives (e.g., selection into programs of study), they may have a more negative emotional experience, with higher frequencies and intensities of negative emotions. Students have associated intense negative emotions, such as annoyance and frustration, with their assessed coursework in higher education (Wass et al., 2020). Frustration, if not resolved reasonably quickly, could lead to boredom and disengagement from self-assessment (Chevrier et al., 2019; Munzar et al., 2021). Whether the stakes are low or high, self-assessment appears to be an emotional activity (Baca & Smith, 2018), possibly because it has consequences for students' self-concepts and sense of identity (Nieminen, 2024).

The emotions that students experience during self-assessment may shape the self-assessment actions they employ. Specifically, surprise and enjoyment may provoke students to draw on internal criteria (e.g., intuition, hindsight) to assess their work, and anxiety and frustration may provoke students to have metacognitive experiences, especially feelings of difficulty. Using state-transition analysis to identify significant transitions between emotions and self-assessment actions and metacognition in think aloud data, we identified four statistically significant transitions between emotions and subsequent self-assessment actions: surprise to

Assess Level 1, enjoyment to Assess Level 1, anxiety to metacognitive experiences, and frustration to metacognitive experiences (Chapter 4). These transitions are depicted using solid arrows in the interaction theory (Figure 1).

Surprise – Assess Level 1

Surprise arises when learners detect a discrepancy between new information and prior expectations or beliefs (Chevrier et al., 2019). When students experience surprise during self-assessment, they may draw on internal criteria to make sense of their surprise in relation to their work. In our study, surprise – Assess Level 1 transitions were characterized by an expression of surprise followed immediately by an evaluative statement about their work which was guided by gut reactions (e.g., intuition). Importantly, surprise may be an immediate experience that quickly provokes follow up emotions, such as confusion or curiosity, as the nature of the surprising information comes to light for the student (see Muis et al., 2018). Moreover, the intensity of surprise or the emotion that immediately follows it may mediate the self-assessment actions that learners apply as a result (Chevrier et al., 2019).

Enjoyment – Assess Level 1

When students experience enjoyment, they may be more likely to transition to Assess Level 1 for two reasons. First, they may transition to Assess Level 1 as an instantiation of reflecting on an aspect of the quality of their work that sparked enjoyment (Chapter 4). The likelihood of enjoyment transitioning to Assess Level 1 is supported by research indicating that enjoyment can lead to deeper cognitive processing during learning (Obergriesser & Stoeger, 2020). Alternatively, transitions from enjoyment to Assess Level 1 may reflect feelings of enjoyment being interrupted by unrelated, negative evaluations of their work (Chapter 4). A similar phenomenon has been documented in the dynamics of emotions during complex learning,

where states enjoyment or flow are interrupted when learners detect impasses which they attempt to resolve to return to a state of enjoyment/flow (D'Mello & Graesser, 2012).

Anxiety– and Frustration–Metacognitive Experiences

During self-assessment, feelings of anxiety and frustration may be more likely to be followed by metacognitive experiences than other possible states or self-assessment actions (Chapter 4). In our research, anxiety– and frustration–metacognitive experiences transitions tended to be characterized by participants expressing anxiety or frustration and then immediately articulating a metacognitive feeling of difficulty where they attempted to make sense of or resolve an issue in their self-assessment process. While research has indicated that negative activating emotions, such as anxiety and frustration, tend to lead to shallow self-regulated learning strategies (e.g., rehearsal; Di Leo et al., 2019), studies have also demonstrated that frustration and other negative emotions (e.g., confusion) have led to metacognitive processes (Muis et al., 2018; Munzar et al., 2021). Muis et al. (2018) theorized that negative emotions may trigger metacognitive processes aimed at making sense of what caused the negative emotions, supporting the notion that anxiety and frustration may lead to metacognitive experiences during self-assessment. While the available empirical evidence suggests similar consequences of anxiety and frustration in self-assessment, an important distinction between these two emotions is that frustration, resulting from perceived persistent lack of progress, can overwhelm cognitive resources and, if unresolved, lead to boredom and disengagement from self-assessment (Chevrier et al., 2019; Park & Ramirez, 2022). Depending on the context, anxiety and frustration in self-assessment may lead to shallow self-assessment actions, where students rush from information processing to a self-assessment judgement (e.g., no self-assessment or superficial self-assessment profiles; see Panadero et al., 2024), or metacognitive experiences which may provoke deeper

thinking and analysis of the self-assessment process. One study involving grades 3-6 students completing math problems found that frustration predicted metacognitive strategies which directly predicted achievement (Munzar et al., 2021), suggesting that metacognitive experiences which follow negative emotions in self-assessment could lead to deeper analysis and have beneficial outcomes for learning.

Implications

The interaction theory of student self-assessment illuminates the complex and recursive nature of student self-assessment as it occurs in the moment. Self-assessment has been conceptualized on a macro level, where it is understood in the context of feedback literacy or as an going process throughout the span of a course or program of study (e.g., Jensen et al., 2022; Nicol, 2021; Sargeant et al., 2010; Yan & Carless, 2021). Such conceptualizations provide important contributions, particularly on the sources of information students draw upon to inform their self-assessments—e.g., their perception of feedback providers (Sargeant et al., 2010). However, the interaction theory of self-assessment advanced here articulates the theorized self-assessment actions, cognitive and affective processes, metacognition, and emotions that occur in the moment of self-assessment, drawing primarily on fine-grained data (e.g., think alouds, eye tracking) collected from students while students are self-assessing. This distinction is important because empirical evidence has suggested different cognitive mechanisms underpin retrospective and concurrent (i.e., in the moment) self-assessments (Eva & Regehr, 2011; Nowak et al., 2022; Zamora et al., 2018). For example, based on findings that learners' concurrent self-assessments are more accurate than global retrospective ones, Eva and Regehr (2011) proposed that global self-assessments may require learners to synthesize a variety of memories while concurrent self-assessments enable analysis of readily available evidence.

The primary theoretical implication of this theory is that it substantially expands previous understandings of what constitutes the phenomenon of self-assessment as a process which have primarily focused on the actions and practices students employ (Panadero et al., 2024; Yan & Brown, 2017). The interaction theory thereby expands the descriptive aspects of self-assessment theory, providing a more comprehensive description of what self-assessment encompasses which can guide future research and practice (Greene, 2022). Specifically, the interaction theory challenges researchers and educators to expand how they understand student self-assessment from a series of actions to a multidimensional process involving consequential interaction between actions, cognition and affect, metacognition, and emotions. These dimensions and interactions between them shape the process and therefore the outcomes of student self-assessment. Facilitating student self-assessment in education systems therefore requires researchers to incorporate emotions and metacognition when developing novel understandings of self-assessment and new interventions.

For educators, the interaction theory reinforces the importance of helping students make their self-assessment processes explicit (Nicol, 2021; Panadero, Lipnevich, et al., 2019) to understand how they are thinking and feeling and how those thoughts and feelings may be shaping their self-assessment processes. The interaction theory reveals that, during self-assessment, students' states shift rapidly between self-assessment actions, where they think about and act on their own work; metacognition, where they monitor and manage their self-assessment actions; and emotions, which arise from and shape self-assessment actions and metacognition. Using a digital trace data method to profile students' self-assessment processes in a secondary school English literature lesson, Rickey et al. (2023) similarly found that students' self-assessment processes were highly recursive, with rapid shifts between 22 cognitive and affective

operations. The recursive nature of student self-assessment indicates the importance of tailored in the moment interventions to support students' self-assessment processes.

Specifically, this interaction self-assessment theory holds three primary implications for teaching and learning. First, when students are engaged in self-assessment actions and practices (Panadero et al., 2024; Yan & Brown, 2017), they activate a range of cognitive and affective processes which may sometimes operate in tension. For example, they may set a goal to read a rubric to remind themselves of the success criteria, a top-down process, and then experience feelings of overwhelm due to the large amount of disaggregated information contained in the rubric, a bottom-up attentional influence, which they then use cognitive resources to address (Chapter 2). By being aware of these cognitive and affective processes and tensions, educators can support students during self-assessment by helping them become aware of their thinking and how to manage top-down and bottom-up influences to productively process available information for self-assessment. Further, as findings indicate that students tend to rush to judgements about their work without deeper analysis of their work using success criteria (i.e., analysis phase), educators can help students apply different approaches to comparison—such as systematic approaches using rubrics, or holistic approaches using exemplars—to foster deeper analysis. For example, when providing students with high-quality exemplars of writing, educators could help students explicitly consider ways that they could approximate the level of quality while maintaining their personal writing style (Chapter 2). Helping students develop and internalize productive cognitive and affective mechanisms during self-assessment could help students develop more sophisticated self-assessments that lead to improved performance.

Second, novel empirical evidence of the metacognitive dimension of self-assessment indicates that students actively monitor and control their self-assessment processes, in addition to

thinking about the object of self-assessment (i.e., their work; Chapter 3). Given the importance of metacognitive feelings of difficulty as a heuristic whereby students maintain and update their perceptions of their self-assessment processes, educators could support students' self-assessment processes by helping them identify and respond to feelings of difficulty, for example, by explicitly articulating the possible reasons for the feeling and seeking ways to address the issue (e.g., asking a peer for help, identifying a new self-assessment action or strategy). Additionally, as selecting a strategy and goal setting/reasoning seem to be important metacognitive skills for controlling the self-assessment process, educators may help students control their self-assessment processes by encouraging them to articulate their approaches and reasoning during self-assessment (see Panadero, Lipnevich, et al., 2019). Ultimately, helping students monitor and control their self-assessment processes can provide students with the tools to be agentic self-assessors.

Third, as students experience a range of emotions during self-assessment and because these may shape students' self-assessment actions and metacognition, educators can support students during self-assessment by helping them to recognize emotions that may arise and manage them to maintain engagement in self-assessment (Chapter 4). When students are self-assessing work that may have consequences for their lives (e.g., graded assignments), educators could recognize that students may experience intense negative emotions, such as frustration and anxiety, which may provoke shallow self-assessment processes or metacognitive feelings of difficulty. Educators can help students find strategies for managing negative emotions, such as using metacognitive processes to understand how they might adjust their self-assessment processes when they experience frustration—with the aim of cultivating feelings of enjoyment and curiosity, which can support persistence with self-assessment. Overall, knowing which

emotions might arise during self-assessment, and their potential consequences, can help educators and students understand responses to emotions effectively within the specific context of the self-assessment.

Areas for Future Research

The interaction theory of self-assessment provokes areas for future research. Overall, this theory advances the descriptive aspect of theory development, which describes phenomena to delineate their components. Studies focused on developing the descriptive aspects of theory are critical: comprehensive descriptions of phenomena are requisite to guiding explanatory aspects of theory, which examine how phenomena work and interact with other variables. Additionally, descriptive aspects of theory are often underdeveloped in education research, meaning that explanatory studies leave out key aspects of phenomena (Greene, 2022). Therefore, more work is needed to validate and extend this descriptive theory of student self-assessment in other education contexts, especially with different populations (e.g., secondary and elementary school students, different assessment and learning cultures).

In particular, further study of transitions between different layers and components of the theory would support deeper insights into how they are interrelated. We statistically examined transitions between emotions and self-assessment actions and metacognition (Chapter 4); however, larger sample sizes could permit sufficient observations to support state-transition analysis among more granular elements of the theory and validate our findings. For example, curiosity and confusion can be beneficial for learning by provoking sophisticated self-regulated learning strategies (e.g., elaboration, critical thinking; D'Mello et al., 2014; Muis et al., 2018; Munzar et al., 2021). Studies employing transition analysis could explore and verify the extent to which emotions, such as curiosity and confusion, provoke deeper processing in self-assessment

(e.g., redo, advanced self-assessment profiles; see Panadero et al., 2024). Additionally, while there is empirical evidence of transitions from emotions to self-assessment actions and metacognition, studies examining the transitions from self-assessment actions and metacognition to emotions could identify antecedents to emotions during the self-assessment process. It would be especially valuable to investigate transitions between self-assessment actions and specific metacognitive experiences (e.g., feelings of difficulty, judgements of learning) and metacognitive skills (e.g., selecting a strategy, goal setting/reasoning) and self-assessment actions to empirically verify the theorized connection presented in the interaction theory (see also Chapter 3). Additionally, significant contributions to the theory could be gained by examining transitions between metacognitive experiences and metacognitive skills to illuminate how monitoring informs control of the self-assessment process.

Further, studies could leverage sequential pattern mining approaches to identify significant patterns that are longer than dyads (e.g., Chen et al., 2017; Lyu et al., 2024; Mudrick et al., 2019), which could help illuminate how certain patterns develop and which ones are beneficial for certain learning outcomes. Such research would be particularly beneficial for informing pedagogical approaches to self-assessment, as pedagogies could be developed that teach and reinforce positive patterns. Specifically, as we found significant transitions between anxiety and frustration and metacognitive experiences, pattern mining could be used to identify the pathways and patterns that occur after the metacognitive experiences to determine if and how these metacognitive experiences support effective self-assessment actions or metacognitive control. Metacognition and emotions are interrelated (Efklides, 2017) and such studies could elucidate how metacognition and emotions shape each other during self-assessment—and how these relations shape students' self-assessment actions and the outcomes of self-assessment.

More evidence of the interrelations between self-assessment actions, cognition and affect, metacognition, and emotions would validate and extend this initial interaction theory of self-assessment and generate further educational implications.

Conclusion

To enhance the quality of learning in education systems and better equip students to address rapidly developing societal challenges, it is critical to facilitate student self-assessment in schools. Developing pedagogical approaches and tailored interventions for facilitating student self-assessment requires understanding the process of self-assessment. While the actions and practices that students employ during self-assessment have been empirically documented (Panadero et al., 2024; Yan & Brown, 2017), research has yet to map the cognitive and affective, metacognitive, and emotions dimensions of the self-assessment process. The purpose of this chapter was to advance an interaction theory of student self-assessment integrating self-assessment actions, cognition and affect, metacognition, and emotions. Building directly on Panadero et al.'s (2024) self-feedback model, the interaction theory of student self-assessment presented here draws on novel empirical multimodal data (Chapters 2-4) and previous empirical findings and theories to articulate the connections between self-assessment actions and cognitive, affective, metacognitive, and emotional dimensions of the student self-assessment process. This interaction theory advances descriptive aspects of self-assessment theory development (Greene, 2022) by providing an initial comprehensive model of the self-assessment process. Future research is needed to validate and expand the theory and form the foundations for explanatory aspects of theory—leading to better understandings of how and when self-assessment can support learning.

References

Ahmadi, R. (2022). Students' perceptions of student voice in assessment within the context of Iran: The dynamics of culture, power relations, and student knowledge. *Higher Education Research and Development*, *41*(2), 211–225. https://doi.org/10.1080/07294360.2021.1882401

Aldosari, M., Heydarnejad, T., Hashemifardnia, A., & Abdalgane, M. (2023). The interplay among self-assessment, using reflection for assessment, classroom enjoyment, and immunity: into prospects of effective language learning. *Language Testing in Asia*, *13*(1). https://doi.org/10.1186/s40468-022-00213-1

Andrade, H. L. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In H. L. Andrade & G. Cizek (Eds.), *Handbook of formative assessment* (pp. 90–105). Routledge.

Andrade, H. L. (2019). A critical review of research on student self-assessment. *Frontiers in Education*, *4*, 87. https://doi.org/10.3389/feduc.2019.00087

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

Andrade, H. L., & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment and Evaluation in Higher Education*, *32*(2), 159–181. https://doi.org/10.1080/02602930600801928

Baca, L., & Smith, S. R. (2018). The role of self-reflection and self-assessment in the psychological assessment process. In S. R. Smith & R. Krishnamurthy (Eds.), *Diversity-sensitive personality assessment* (pp. 3–24). Routledge.

Belkaid, M., Cuperlier, N., & Gaussier, P. (2017). Emotional metacontrol of attention: Top-down modulation of sensorimotor processes in a robotic visual search task. *PLoS ONE*, *12*(9). https://doi.org/10.1371/journal.pone.0184960

Birenbaum, M., DeLuca, C., Earl, L., Heritage, M., Klenowski, V., Looney, A., Smith, K., Timperley, H., Volante, L., & Wyatt-Smith, C. (2015). International trends in the implementation of assessment for learning: Implications for policy and practice. *Policy Futures in Education*, *13*(1), 117–140. https://doi.org/10.1177/1478210314566733

Boekaerts, M. (1991). Subjective competence, appraisals and self-assessment. *Learning and Instruction*, *1*, 1–17.

Boekaerts, M. (2010). The crucial role of motivation and emotion in classroom learning. In H. Dumont, D. Istance, & F. Benevides (Eds.), *The nature of learning: Using research to inspire practice* (pp. 91–111). The Center for Educational Research and Innovation, Organization for Economic Co-operation and Development.

Bolden, B., & DeLuca, C. (2022). Nurturing student creativity through assessment for learning in music classrooms. *Research Studies in Music Education*, *44*(1), 273–289. https://doi.org/10.1177/1321103X211054793

Boud, D., & Brew, A. (1995). Developing a typology for learner self-assessment practices. *Research and Development in Higher Education*, *18*, 130–135.

Bourke, R. (2016). Liberating the learner through self-assessment. *Cambridge Journal of Education*, *46*(1), 97–111. https://doi.org/10.1080/0305764X.2015.1015963

Bourke, R., & Mentis, M. (2013). Self-assessment as a process for inclusion. *International Journal of Inclusive Education*, *17*(8), 854–867. https://doi.org/10.1080/13603116.2011.602288

Brandmo, C., Panadero, E., & Hopfenbeck, T. N. (2020). Bridging classroom assessment and self-regulated learning. *Assessment in Education: Principles, Policy and Practice*, *27*(4), 319–331. https://doi.org/10.1080/0969594X.2020.1803589

British Columbia Ministry of Education. (2023). *A Framework for Classroom Assessment*.

Brown, G. T. L., & Harris, L. (2013). Student self-assessment. In J. H. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 367–393). Sage.

Brown, G. T. L., & Harris, L. (2014). The future of self-assessment in classroom practice: Reframing self- assessment as a core competency. *Frontline Learning Research*, *2*(1), 22–30. https://doi.org/10.14786/flr.v2i1.24

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

Butler, Y. G. (2018). Young learners' processes and rationales for responding to self-assessment items. In J. Davis, J. Norris, M. Malone, T. McKay, & Y.-A. Son (Eds.), *Useful assessment and evaluation in language education* (pp. 21–39). Georgetown University Press.

Chen, B., Resendes, M., Chai, C. S., & Hong, H. Y. (2017). Two tales of time: Uncovering the significance of sequential patterns among contribution types in knowledge-building discourse. *Interactive Learning Environments*, *25*(2), 162–175. https://doi.org/10.1080/10494820.2016.1276081

Chevrier, M., Muis, K. R., Trevors, G. J., Pekrun, R., & Sinatra, G. M. (2019). Exploring the antecedents and consequences of epistemic emotions. *Learning and Instruction*, *63*, 101209. https://doi.org/10.1016/j.learninstruc.2019.05.006

Claypool, T. R., & Preston, J. P. (2011). Redefining learning and assessment practices impacting

Aboriginal students: Considering Aboriginal priorities via Aboriginal and Western worldviews. *In Education*, *17*(3), 84–95.

Den Ouden, H. E. M., Kok, P., & de Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Frontiers in Psychology*, *3*, 548. https://doi.org/10.3389/fpsyg.2012.00548

Di Leo, I., Muis, K. R., Singh, C. A., & Psaradellis, C. (2019). Curiosity… Confusion? Frustration! The role and sequencing of emotions during mathematics problem solving. *Contemporary Educational Psychology*, *58*, 121–137. https://doi.org/10.1016/j.cedpsych.2019.03.001

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

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*(2), 145–157. https://doi.org/10.1016/j.learninstruc.2011.10.001

D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction*, *29*, 153–170. https://doi.org/10.1016/j.learninstruc.2012.05.003

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist*, *13*(4), 277–287. https://doi.org/10.1027/1016-9040.13.4.277

Efklides, A. (2017). Affect, epistemic emotions, metacognition, and self-regulated learning. *Teachers College Record*, *119*, 130305.

Eva, K. W., & Regehr, G. (2011). Exploring the divergence between self-assessment and self-monitoring. *Advances in Health Sciences Education*, *16*(3), 311–329. https://doi.org/10.1007/s10459-010-9263-2

Fang, M. W. H., Ravizza, S. M., & Liu, T. (2019). Attention induces surround suppression in visual working memory. *Psychonomic Bulletin and Review*, *26*(6), 1925–1932. https://doi.org/10.3758/s13423-019-01624-7

Fastré, G. M. J., van der Klink, M. R., Sluijsmans, D., & van Merriënboer, J. J. G. (2012). Drawing students' attention to relevant assessment criteria: Effects on self-assessment skills and performance. *Journal of Vocational Education and Training*, *64*(2), 185–198. https://doi.org/10.1080/13636820.2011.630537

Fiebelkorn, I. C., & Kastner, S. (2020). Functional specialization in the attention network. *Annual Review of Psychology*, *71*, 221–249. https://doi.org/10.1146/annurev-psych-010418-103429

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, *34*(10), 906–911.

Genc Aksaray, S., & Ozcelik, E. (2022). Attention mediates the effect of emotional arousal on learning outcomes in multimedia learning: an eye-tracking study. *Educational Psychology*. https://doi.org/10.1080/01443410.2022.2154317

Greene, J. A. (2022). What can educational psychology learn from, and contribute to, theory development scholarship? *Educational Psychology Review*, *34*(4), 3011–3035. https://doi.org/10.1007/s10648-022-09682-5

Greene, J. A., Bernacki, M. L., & Hadwin, A. F. (2024). Self-regulation. In P. A. Schultz & K. R. Muis (Eds.), *Handbook of educational psychology* (4th ed., pp. 314–334). Routledge.

https://doi.org/10.4324/9780429433726-17

Hattie, J. (2023). *Visible learning: The sequel: A synthesis of over 2,100 meta-analyses relating to achievement*. Routledge.

Jensen, L. X., Bearman, M., & Boud, D. (2022). Feedback encounters: Towards a framework for analysing and understanding feedback processes. *Assessment & Evaluation in Higher Education*, 1–14. https://doi.org/10.1080/02602938.2022.2059446

Kippers, W. B., Wolterinck, C. H. D., Schildkamp, K., Poortman, C. L., & Visscher, A. J. (2018). Teachers' views on the use of assessment for learning and data-based decision making in classroom practice. *Teaching and Teacher Education*, *75*, 199–213. https://doi.org/10.1016/j.tate.2018.06.015

Krebs, R., Rothstein, B., & Roelle, J. (2022). Rubrics enhance accuracy and reduce cognitive load in self-assessment. *Metacognition and Learning*, *17*, 627–650. https://doi.org/10.1007/s11409-022-09302-1

Lee, H., Chung, H. Q., Zhang, Y., Abedi, J., & Warschauer, M. (2020). The effectiveness and features of formative assessment in US K-12 education: A systematic review. *Applied Measurement in Education*, *33*(2), 124–140. https://doi.org/10.1080/08957347.2020.1732383

Lin-Siegler, X., Shaenfield, D., & Elder, A. D. (2015). Contrasting case instruction can improve self-assessment of writing. *Educational Technology Research and Development*, *63*(4), 517–537. https://doi.org/10.1007/s11423-015-9390-9

Lipnevich, A. A., Panadero, E., & Calistro, T. (2023). Unraveling the effects of rubrics and exemplars on student writing performance. *Journal of Experimental Psychology: Applied*, *29*(1), 136–148. https://doi.org/10.1037/xap0000434

Lyu, Q., Chen, W., Su, J., & Heng, K. H. J. G. (2024). Steps to implementation: The role of peer feedback inner structure on feedback implementation. *Assessment and Evaluation in Higher Education*, *49*(4), 572–585. https://doi.org/10.1080/02602938.2023.2291340

Mudrick, N. V., Azevedo, R., & Taub, M. (2019). Integrating metacognitive judgments and eye movements using sequential pattern mining to understand processes underlying multimedia learning. *Computers in Human Behavior*, *96*, 223–234. https://doi.org/10.1016/j.chb.2018.06.028

Muis, K. R., Chevrier, M., & Singh, C. A. (2018). The role of epistemic emotions in personal epistemology and self-regulated learning. *Educational Psychologist*, *53*(3), 165–184. https://doi.org/10.1080/00461520.2017.1421465

Muis, K. R., Pekrun, R., Sinatra, G. M., Azevedo, R., Trevors, G., Meier, E., & Heddy, B. C. (2015). The curious case of climate change: Testing a theoretical model of epistemic beliefs, epistemic emotions, and complex learning. *Learning and Instruction*, *39*, 168–183. https://doi.org/10.1016/j.learninstruc.2015.06.003

Munzar, B., Muis, K. R., Denton, C. A., & Losenno, K. (2021). Elementary students' cognitive and affective responses to impasses during mathematics problem solving. *Journal of Educational Psychology*, *113*(1), 104–124. https://doi.org/10.1037/edu0000460

Nelson, T., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125–173). Academic Press.

Nicol, D. (2021). The power of internal feedback: Exploiting natural comparison processes. *Assessment and Evaluation in Higher Education*, *46*(5), 756–778. https://doi.org/10.1080/02602938.2020.1823314

Nieminen, J. H. (2022). Disrupting the power relations of grading in higher education through summative self-assessment. *Teaching in Higher Education*, *27*(7), 892–907. https://doi.org/10.1080/13562517.2020.1753687

Nieminen, J. H. (2024). How does assessment shape student identities? An integrative review. *Studies in Higher Education*. https://doi.org/10.1080/03075079.2024.2334844

Nowak, A., Vallacher, R. R., Bartkowski, W., & Olson, L. (2022). Integration and expression: The complementary functions of self-reflection. *Journal of Personality*. https://doi.org/10.1111/jopy.12730

Obergriesser, S., & Stoeger, H. (2020). Students' emotions of enjoyment and boredom and their use of cognitive learning strategies – How do they affect one another? *Learning and Instruction*, *66*. https://doi.org/10.1016/j.learninstruc.2019.101285

OECD. (2019). *OECD future of education and skills 2030: OECD learning compass 2030: A series of concept notes*. https://www.oecd.org/education/2030-project/contact/OECD_Learning_Compass_2030_Concept_Note_Series.pdf

OECD. (2023). *Building the future of education 2023*.

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. L., & 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. *Australian Educational Researcher*, *45*(1), 13–31. https://doi.org/10.1007/s13384-018-0258-y

Panadero, E., Brown, G. T. L., & Strijbos, J. W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review*,

28(4), 803–830. https://doi.org/10.1007/s10648-015-9350-2

Panadero, E., Fernández, J., Pinedo, L., Sánchez, I., & García-Pérez, D. (2024). A self-feedback model (SEFEMO): secondary and higher education students' self-assessment profiles. *Assessment in Education: Principles, Policy and Practice*, *31*(3–4), 221–253. https://doi.org/10.1080/0969594X.2024.2367027

Panadero, E., Fernández-Ruiz, J., & Sánchez-Iglesias, I. (2020). Secondary education students' self-assessment: The effects of feedback, subject matter, year level, and gender. *Assessment in Education: Principles, Policy & Practice*, *27*(6), 607–634. https://doi.org/10.1080/0969594x.2020.1835823

Panadero, E., Fraile, J., Fernández Ruiz, J., Castilla-Estévez, D., & Ruiz, M. A. (2019). Spanish university assessment practices: Examination tradition with diversity by faculty. *Assessment and Evaluation in Higher Education*, *44*(3), 379–397. https://doi.org/10.1080/02602938.2018.1512553

Panadero, E., García-Pérez, D., Fernández Ruiz, J., Fraile, J., Sánchez-Iglesias, I., & Brown, G. T. L. (2023). Feedback and year level effects on university students' self-efficacy and emotions during self-assessment: Positive impact of rubrics vs. instructor feedback. *Educational Psychology*, 1–24. https://doi.org/10.1080/01443410.2023.2254015

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, *22*, 74–98. https://doi.org/10.1016/j.edurev.2017.08.004

Panadero, E., Lipnevich, A., & Broadbent, J. (2019). Turning self-assessment into self-feedback. In *The Impact of Feedback in Higher education: Improving assessment outcomes for learners* (pp. 147–163). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-25112-

3_9

Panadero, E., Pérez, D. G., Ruiz, J. F., Fraile, J., Sánchez-Iglesias, I., & Brown, G. T. L. (2022). University students' strategies and criteria during self-assessment: Instructor's feedback, rubrics, and year level effects. *European Journal of Psychology of Education*, *58*, 1031–1051. https://doi.org/10.1007/s10212-022-00639-4

Park, D., & Ramirez, G. (2022). Frustration in the classroom: Causes and strategies to help teachers cope productively. *Educational Psychology Review*, *34*(4), 1955–1983. https://doi.org/10.1007/s10648-022-09707-z

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Pekrun, R. (2024). Control-value theory: From achievement emotion to a general theory of human emotions. *Educational Psychology Review*, *36*(3). https://doi.org/10.1007/s10648-024-09909-7

Pekrun, R., & Perry, R. P. (2014). Control-value theory of achievement emotions. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 120–141). Routledge.

Poltoratski, S., Ling, S., McCormack, D., Tong, F., & Poltoratski, S. (2017). Characterizing the effects of feature salience and top-down attention in the early visual system. *J Neurophysiol*, *118*, 564–573. https://doi.org/10.1152/jn.00924.2016.-The

Reber, R., & Greifeneder, R. (2017). Processing fluency in education: How metacognitive feelings shape learning, belief formation, and affect. *Educational Psychologist*, *52*(2), 84–103. https://doi.org/10.1080/00461520.2016.1258173

Rickey, N., DeLuca, C., & Beach, P. (2023). Towards a new theory of student self-assessment: Tracing learners' cognitive and affective processes. *Metacognition and Learning*, *18*, 945–981. https://doi.org/10.1007/s11409-023-09359-6

Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research, and Evaluation*, *11*, 10. https://doi.org/10.7275/9wph-vv65

Sargeant, J., Armson, H., Chesluk, B., Dornan, T., Eva, K., Holmboe, E., Lockyer, J., Loney, E., Mann, K., & van der Vleuten, C. (2010). The Processes and dimensions of informed self-assessment: A conceptual model. *Academic Medicine*, *8*(7), 1212–1220.

Sitzmann, T., Ely, K., Brown, K. G., & Bauer, K. N. (2010). *Self-Assessment of knowledge: A cognitive learning or affective measure? 9*(2), 169–191.

Stark, L., Brünken, R., & Park, B. (2018). Emotional text design in multimedia learning: A mixed-methods study using eye tracking. *Computers and Education*, *120*, 185–196. https://doi.org/10.1016/j.compedu.2018.02.003

Tan, K. (2007). Conceptions of self-assessment: What is needed for long-term learning? In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education* (pp. 114–127). Taylor & Francis Group. http://ebookcentral.proquest.com/lib/queen-ebooks/detail.action?docID=292915

Thigpen, N., Petro, N. M., Oschwald, J., Oberauer, K., & Keil, A. (2019). Selection of visual objects in perception and working memory one at a time. *Psychological Science*, *30*(9), 1259–1272. https://doi.org/10.1177/0956797619854067

To, J., Panadero, E., & Carless, D. (2021). A systematic review of the educational uses and effects of exemplars. *Assessment and Evaluation in Higher Education*. https://doi.org/10.1080/02602938.2021.2011134

Veugen, M. J., Gulikers, J. T. M., & den Brok, P. (2021). We agree on what we see: Teacher and student perceptions of formative assessment practice. *Studies in Educational Evaluation*, *70*, 101027. https://doi.org/10.1016/j.stueduc.2021.101027

Volante, L., & Beckett, D. (2011). *Formative assessment and the contemporary classroom: Synergies and tensions between research and practice*. *34*(2), 239–255. https://doi.org/10.2307/canajeducrevucan.34.2.239

Wass, R., Timmermans, J., Harland, T., & McLean, A. (2020). Annoyance and frustration: Emotional responses to being assessed in higher education. *Active Learning in Higher Education*, *21*(3), 189–201. https://doi.org/10.1177/1469787418762462

Winne, P. H. (2018). Theorizing and researching levels of processing in self-regulated learning. *British Journal of Educational Psychology*, *88*(1), 9–20. https://doi.org/10.1111/bjep.12173

Winne, P. H., & Azevedo, R. (2014). Metacognition. In *The Cambridge Handbook of the Learning Sciences, Second Edition* (pp. 63–87). Cambridge University Press. https://doi.org/10.1017/CBO9781139519526.006

Yan, Z. (2020). Self-assessment in the process of self-regulated learning and its relationship with academic achievement. *Assessment and Evaluation in Higher Education*, *45*(2), 224–238. https://doi.org/10.1080/02602938.2019.1629390

Yan, Z., & Brown, G. T. L. (2017). A cyclical self-assessment process: Towards a model of how students engage in self-assessment. *Assessment and Evaluation in Higher Education*, *42*(8), 1247–1262. https://doi.org/10.1080/02602938.2016.1260091

Yan, Z., Brown, G. T. L., Lee, J. C. K., & Qiu, X. L. (2020). Student self-assessment: Why do they do it? *Educational Psychology*, *40*(4), 509–532.

https://doi.org/10.1080/01443410.2019.1672038

Yan, Z., & Carless, D. (2021). Self-assessment is about more than self: The enabling role of feedback literacy. *Assessment & Evaluation in Higher Education*, *47*(7), 1116–1128. https://doi.org/10.1080/02602938.2021.2001431

Yan, Z., Lao, H., Panadero, E., Fernández-Castilla, B., Yang, L., & Yang, M. (2022). Effects of self-assessment and peer-assessment interventions on academic performance: A pairwise and network meta-analysis. *Educational Research Review*, *37*, 100484. https://doi.org/10.1016/j.edurev.2022.100484

Yan, Z., Lee, J. C.-K., Hui, S. K. F., & Lao, H. (2022). Enhancing students' self-efficacy in creativity and learning performance in the context of English learning: The use of self-assessment mind maps. *Frontiers in Psychology*, *13*, 871781. https://doi.org/10.3389/fpsyg.2022.871781

Yan, Z., Wang, X., Boud, D., & Lao, H. (2021). The effect of self-assessment on academic performance and the role of explicitness: A meta-analysis. *Assessment & Evaluation in Higher Education*, *48*(1), 1–15. https://doi.org/10.1080/02602938.2021.2012644

Zamora, Á., Suárez, J. M., & Ardura, D. (2018). Error detection and self-assessment as mechanisms to promote self-regulation of learning among secondary education students. *Journal of Educational Research*, *111*(2), 175–185. https://doi.org/10.1080/00220671.2016.1225657

Appendix I

Questionnaires (Chapters 2-4)

Q1 Student Self-Assessment Processes: Unpacking the Critical Role of Affect - Participant Info

Start of Block: Participant Information

Q1 Here is your study code: **${e://Field/Random%20ID}** Please write this down as clearly as possible on your consent form.

Page Break

Q2 What is your home country?

▼ Afghanistan (1) ... Zimbabwe (195)

Q3 How old are you?

- ○ 17 (1)
- ○ 18 (2)
- ○ 19 (3)
- ○ 20 (4)
- ○ Other: (5) _____

Q4 What is your race?

_____

Q5 What is your gender?

_____

Q6 What is your first language?

_____

Q7 What is the highest level of education attained by your parent(s) or guardian(s)?

- ○ Primary school (1)
- ○ Secondary school (2)
- ○ Vocational training (3)
- ○ Apprenticeship (4)
- ○ Bachelor's degree (5)
- ○ Master's degree (6)
- ○ PhD (7)
- ○ Other: (8) _____

Q8 What was your grade in English last year?

- ○ 90-100% (1)
- ○ 80-89% (13)
- ○ 70-79% (14)
- ○ 60-69% (15)
- ○ 50-59% (16)
- ○ 0-49% (18)
- ○ A (19)
- ○ B (20)
- ○ C (21)
- ○ D (22)
- ○ F (23)
- ○ 10 (17)
- ○ 9 (2)
- ○ 8 (3)
- ○ 7 (4)
- ○ 6 (5)
- ○ 5 (6)
- ○ 4 (7)

- ○ 3 (8)
- ○ 2 (9)
- ○ 1 (10)
- ○ If your grade was reported using a different scale, indicate your grade here: (11) _____

Q9 What is your primary discipline of study?

- ○ Humanities (e.g., history, literature, philosophy) (1)
- ○ Arts (e.g., music, fine arts, visual arts) (2)
- ○ Sciences (e.g., biology, physics, chemistry) (3)
- ○ Engineering (4)
- ○ Social Sciences (e.g., education, psychology, sociology, geography) (5)
- ○ Languages (6)
- ○ Law (7)
- ○ Medicine (8)
- ○ Trades/technical (e.g., mechanics, cooking, hairstyling (9)
- ○ Other: (10) _____

End of Block: Participant Information

Start of Block: Eye Information

Q10 Are you extremely sensitive to light (e.g., from a computer screen, artificial lights)?

- ○ Yes (1)
- ○ No (2)

Q11 Do your eyes become watery after spending a lot of time in front of the screen?

- ○ Yes (1)
- ○ No (2)

Q12 Do you wear glasses?

- ○ Yes (1)
- ○ No (2)

Q13 Do you wear contact lenses?

- ○ Yes (1)
- ○ No (2)

Q14 Have you recently had eye surgery?

- ○ Yes (1)
- ○ No (2)

End of Block: Eye Information

Q2 Student Self-Assessment Processes: Unpacking the Critical Role of Affect - Lab

Start of Block: Self-Assessment Definition

Q15 Self-assessment happens whenever you reflect on your own work, learning process, or abilities. Sometimes, teachers or instructors may prompt you to self-assess, and they may provide you with resources (e.g., rubrics, exemplars) to help you self-assess. You may also self-assess without being asked to, either in your classrooms or at home (e.g., when working on an assignment). With that definition in mind, please answer the following questions.

*

Q16 Please enter your study code here:

_____

End of Block: Self-Assessment Definition

Start of Block: Academic Control

Q17 Please rate your level of agreement with each statement.

	Strongly disagree (1)	Somewhat disagree (2)	Neither agree nor disagree (3)	Somewhat agree (4)	Strongly agree (5)
I have a great deal of control over my academic performance in my courses. (1)	○	○	○	○	○
The more effort I put into my courses, the better I do in them. (2)	○	○	○	○	○
No matter what I do, I can't seem to do well in my courses. (3)	○	○	○	○	○
I see myself as largely responsible for my performance throughout my college career. (4)	○	○	○	○	○
How well I do in my courses is often the "luck of the draw." (5)	○	○	○	○	○
There is little I can do about my performance in university. (6)	○	○	○	○	○

When I do poorly in a course, it's usually because I haven't given it my best effort. (7)	○	○	○	○	○
My grades are basically determined by things beyond my control and there is little I can do to change that. (8)	○	○	○	○	○

End of Block: Academic Control

Start of Block: Task Value

Q18 Please rate how true each statement is for you.

	Not at all true of me (1)	Somewhat not true of me (2)	Neutral (3)	Somewhat true of me (4)	Very true of me (5)
In general, I find self-assessing my work very interesting. (1)	○	○	○	○	○
The amount of effort it takes to self-assess is worthwhile to me. (2)	○	○	○	○	○
In general, I think self-assessment is useful. (3)	○	○	○	○	○
I like self-assessing. (4)	○	○	○	○	○
Compared to my other activities, self-assessment is very useful for me. (5)	○	○	○	○	○
I feel that, to me, self-assessing is very important. (6)	○	○	○	○	○
Self-assessment is useful for my life. (7)	○	○	○	○	○

End of Block: Task Value

Start of Block: Self-Assessment Self-Efficacy

Q19 Please rate your level of agreement with the following statements.

	Strongly disagree (1)	Somewhat disagree (2)	Neither agree nor disagree (3)	Somewhat agree (4)	Strongly agree (5)
I know how to implement self-assessment. (1)	○	○	○	○	○
I can find reference materials (e.g., rubrics, examples) against which to assess myself. (2)	○	○	○	○	○
I have enough knowledge to implement self-assessment. (3)	○	○	○	○	○
I can design appropriate tasks to assess myself. (4)	○	○	○	○	○
I can tell when my work is actually good or bad. (5)	○	○	○	○	○

End of Block: Self-Assessment Self-Efficacy

Start of Block: Pause 1

Q20 Pause here! Please tell the researcher you have finished the first part of the questionnaire.

End of Block: Pause 1

Start of Block: Epistemic Emotions Scale

Q21 For each emotion, please indicate the strength of that emotion by selecting the option that best describes the intensity of your emotional response during the self-assessment.

	Not at all (0) (0)	Very little (1) (1)	Moderate (2) (2)	Strong (3) (3)	Very strong (4) (4)
Curious (1)	○	○	○	○	○
Bored (2)	○	○	○	○	○
Confused (3)	○	○	○	○	○
Surprised (4)	○	○	○	○	○
Interested (5)	○	○	○	○	○
Anxious (6)	○	○	○	○	○
Frustrated (7)	○	○	○	○	○
Inquisitive (8)	○	○	○	○	○
Dull (9)	○	○	○	○	○
Amazed (10)	○	○	○	○	○
Worried (11)	○	○	○	○	○
Happy (12)	○	○	○	○	○
Muddled (13)	○	○	○	○	○
Irritated (14)	○	○	○	○	○
Monotonous (15)	○	○	○	○	○

Excited (16)	○	○	○	○	○
Astonished (17)	○	○	○	○	○
Dissatisfied (18)	○	○	○	○	○
Nervous (19)	○	○	○	○	○
Joyful (20)	○	○	○	○	○
Puzzled (21)	○	○	○	○	○

End of Block: Epistemic Emotions Scale

Start of Block: Pause 2

Q22 Pause here! Please tell the researcher you have finished the second part of the questionnaire.

End of Block: Pause 2

Start of Block: Perceptions of Self-Assessment

Q23 Please rate your level of agreement with the following statements.

	Strongly disagree (1)	Somewhat disagree (2)	Neither agree nor disagree (3)	Somewhat agree (4)	Strongly agree (5)
This self-assessment activity helped me improve my work. (1)	○	○	○	○	○
This self-assessment activity helped me better understand the type of writing I should be aiming for. (2)	○	○	○	○	○
I will engage in this type of self-assessment activity again. (3)	○	○	○	○	○
The rubric was helpful in self-assessing my work. (4)	○	○	○	○	○
The good exemplar was helpful in self-assessing my work. (5)	○	○	○	○	○
The bad exemplar was helpful in self-assessing my work. (6)	○	○	○	○	○

Q24 Please rate how true each statement is for you.

	Not at all true of me (1)	Somewhat not true of me (2)	Neutral (3)	Somewhat true of me (4)	Very true of me (5)
In general, I find self-assessment very interesting. (1)	○	○	○	○	○
The amount of effort it takes to self-assess is worthwhile to me. (2)	○	○	○	○	○
In general, I think self-assessment is useful. (3)	○	○	○	○	○
I like self-assessing. (4)	○	○	○	○	○
Compared to my other activities, self-assessment is very useful for me. (5)	○	○	○	○	○
I feel that, to me, self-assessing is very important. (6)	○	○	○	○	○
Self-assessment is useful for my life. (7)	○	○	○	○	○

End of Block: Perceptions of Self-Assessment

Start of Block: Repeat Study

Q25 Would you be interested in completing the self-assessment activity again at a later date? If so, we will arrange with your teacher for you to complete the activity again.

- ○ Yes (1)
- ○ No (2)

End of Block: Repeat Study

Appendix II

General Research Ethics Board Approval

March 07, 2023

Dr. Christopher DeLuca
Queen's University

Title: "GEDCU-1151-23: Student Self-Assessment Processes: Unpacking the Critical Role of Affect;" TRAQ # 6038217

Dear Dr. DeLuca:

The General Research Ethics Board (GREB), by means of a delegated board review, has cleared your proposal entitled **"GEDCU-1151-23: Student Self-Assessment Processes: Unpacking the Critical Role of Affect"** for ethical compliance with the Tri-Council Guidelines (TCPS 2) and Queen's ethics policies. In accordance with the Tri-Council Guidelines (Article 6.14) and Standard Operating Procedures (405), your project has been cleared for one year.

You are reminded of your obligation to submit an annual renewal form prior to the annual renewal due date (access this form at http://www.queensu.ca/traq/signon.html/; click on "Events;" under "Create New Event" click on "General Research Ethics Board Annual Renewal/Closure Form for Cleared Studies"). Please note that when your research project is completed, you need to submit an Annual Renewal/Closure Form in Romeo/traq indicating that the project is 'completed' so that the file can be closed. This should be submitted at the time of completion; there is no need to wait until the annual renewal due date.

You are reminded of your obligation to advise the GREB of any adverse event(s) that occur during this one-year period (access this form at http://www.queensu.ca/traq/signon.html/; click on "Events;" under "Create New Event" click on "General Research Ethics Board Adverse Event Form"). An adverse event includes, but is not limited to, a complaint, a change or unexpected event that alters the level of risk for the researcher or participants or situation that requires a substantial change in approach to a participant(s). You are also advised that all adverse events must be reported to the GREB within 48 hours.

You are also reminded that all changes that might affect human participants must be cleared by the GREB. For example, you must report changes to the level of risk, applicant characteristics, and implementation of new procedures. To submit an amendment form, access the application by at http://www.queensu.ca/traq/signon.html; click on "Events;" under "Create New Event" click on "General Research Ethics Board Request for the Amendment of Approved Studies." Once submitted, these changes will automatically be sent to the Ethics Coordinator, GREB, at University Research Services for further review and clearance by GREB or the Chair, GREB.

On behalf of the General Research Ethics Board, I wish you continued success in your research.

Sincerely,

Chair, General Research Ethics Board (GREB)
Associate Professor and Distinguished Faculty Fellow of Marketing,
Academic Co-Director (Business), Master of Digital Product Management
Smith School of Business
Queen's University
chair.GREB@queensu.ca