Pedagogy and the Human Sciences

Volume 8
Issue 1 *2021*

Article 2

January 2022

How Can Assessment for Learning be Useful for Self-regulated Learning?: Four Approaches to Change of Assessment Conceptions from Individualistic to Contextualistic

Kohei Nishizuka
Tohoku University, Japan, kohei.nishizuka.s1@dc.tohoku.ac.jp

How Can Assessment for Learning be Useful for Self-regulated Learning?: Four Approaches to Change of Assessment Conceptions from Individualistic to Contextualistic

Abstract

Critics argue that theoretical frameworks for assessment for learning (AfL) and self-regulated learning (SRL) are sociocultural practices reliant on others/mediating artifacts than on individual psychological operations. However, if the broader context of learning is neglected, the developmental model for fostering evaluative judgment cannot cope well with the contextual complexity of the formal and informal aspects of learning. How students perceive assessment is strongly related to their learning outcomes as it represents how much they harbor a spirit of improvement, social stress, and cooperative efficacy. Focusing on the assessment conceptions can help us reconsider the purpose and function of AfL and SRL and their relationship. Therefore, this study clarifies how AfL can contribute to SRL by focusing on students' conceptions of assessment (SCoA) that are acquired and used in various learning contexts. The literature review of SCoA identified four paths for AfL to contribute to SRL: unpacking the ecological rationality of assessment, capturing assessment conceptions at the well-rounded development and community levels, clarifying challenges and differences in assessment conceptions, and strengthening AfL as a process theory. This paper argues for a focus on whether students are able to reconstruct assessment conceptions emergently in the community to exercise life-time evaluative judgment capability.

Keywords

assessment for learning, formative assessment, self-regulated learning, students' conceptions of assessment, sociocultural theory, pedagogy

Cover Page Footnote

*I have no conflicts to disclose. *This study was supported by Grant-in-Aid for JSPS Fellows (20J20092).

*Correspondence concerning this article should be addressed to Kohei Nishizuka, 980-8576, Kawauchi 27-1, Aoba, Sendai, Miyagi, Japan. Email: kohei.nishizuka.s1@dc.tohoku.ac.jp

Introduction

Pedagogical Approaches to Assessment for Learning and Self-regulated Learning

Learning and assessment had been separate research fields until recently, when a push to integrate learning and assessment theories appeared in the literature, aiming to increase assessment's contribution to educational goals (Baird et al., 2017). Assessment, the process of observing, collecting, and organizing information on the status and progress of learning, typically has two types: "assessment of learning (AoL)" and "assessment for learning (AfL)." Although AoL helps to summarize what students can do or not, and know or not mainly by using tests, it works to prevent students from learning. The key to the integration of learning and assessment is embedding assessment within the learning process via the AfL approach. AfL is the process of comparing a student's current level of development to the goal, and using the information that fills the gap to improve the quality of teaching and learning (Black & Wiliam 1998).

AfL is not a psychometric assessment, but a sociocultural and pedagogical approach. Pedagogy, in this paper, follows Daniels' (2001) definition and refers to the social practices that shape individuals' cognitive, emotional, and moral development. Teachers' roles involve not only providing one-sided feedback to students, but also adopting responsive teaching to monitor students' learning and adjusting teaching strategies (Black & Wiliam, 2006, 2009). Brown (2020) proposed that AfL should be viewed not as rigorous psychometric procedures, but as a teaching method. According to Brown, momentary interaction with students is an ad hoc and optimal instructional approach that does not minimize the impact of the interpretation of facts through observation. Moreover, for learner autonomy, the process of participating within a community is regarded as a learning experience, where teachers' role is to help students understand what is valuable for the community and encourage them to pursue it (Lave & Wenger, 1991; Willis, 2010). Willis (2010) posited that assessment is a culturally embedded pattern of interactive participation (i.e., participative pedagogy). Thus, the theoretical foundation of AfL is underpinned by the epistemological framework, which sees education as a mediated process that transcends schooling and is affected by social, cultural, and historical factors.

In recent years, the priorities of AfL research have shifted from test quality and knowledge acquisition to self-regulation, which refers to students' capability to efficiently develop competencies (i.e., learning-to-learn) (Brandmo et al., 2020; Brown et al., 2016; Clark, 2012; Greene, 2020; Nicol & Macfarlane-Dick, 2006). Self-regulated learning (SRL) is a theory of autonomous learning that emphasizes active involvement in metacognition, learning strategies and behaviors, and motivation to control the learning process (Zimmerman & Schunk, 2001). Zimmerman (2002) divided the phases of self-regulation into three stages: forethought (planning goals and learning strategies based on self-efficacy, expectations, and interests), performance (implementation, monitoring, and control of strategies for learning regulation), and self-reflection (individuals' appraisal of whether the learning strategies worked as expected, whether the goals were met, and why). Although some SRL logic models have been developed,

Panadero (2017) concluded that researchers and teachers should use different SRL models for students' different developmental stages and educational levels to understand the different effects of each model. Metacognitive feedback plays a vital role in SRL; it consists of metacognitive monitoring and control (Nelson & Narens, 1990). Metacognitive monitoring assesses current developmental achievement and previous learning progress, while metacognitive control modifies learning strategies or developmental goals to achieve them.

SRL's theoretical model has been shifting from mind and student-related issues to pedagogy in social interaction, implying that it should be incorporated into systematic teaching and curricula based on teachers' and students' interconnectedness (Andrade & Brookhart, 2020; Oates, 2019). Above all, the socially shared regulation of learning (SSRL) and co-regulated learning (CoRL) are sociocultural ideas that implement SRL theory (Hadwin et al., 2017 Panadero & Järvelä, 2015). SSRL emphasizes the development of learners' collective goals, and group's deliberate, strategic, and transactive planning, task enactment, reflection, and adaptation. In SSRL, subjects coordinate each other's activities collectively, including the interactive structuring of tasks and goals in teachers' and students' learning communities. Since what is desirable in regulating learning is continuously reassessed through a collaborative effort, teachers can prevent situations where they fail to facilitate the development of students' cognitive abilities and blame students. In contrast, CoRL, as a mediator between SRL and SSRL, emphasizes the development of learners' personal goals and the deployment of personnel proficient in encouraging SRL with others.

Is Assessment for Learning Really Effective in Self-regulated Learning?

AfL can enhance SRL because students learn to monitor the quality of work and use a repertoire of alternative strategies to improve said work (Brandmo et al., 2020; Sadler, 1989). Thus, as Panadero and Alonso-Tapia (2013) concluded, AfL and SRL influence the reinforcement of self-assessment, which students require to create their zone of proximal development (ZPD[1]). Self-assessment requires that students exhibit good evaluative judgment capability to make decisions about the quality of their work and that of their peers (Tai et al., 2018). AfL is an important condition for promoting SRL; conversely, SRL can make assessment practices more productive. Further, learning outcomes provide worth and meaning to AfL and SRL: AfL's logic model is effective in obtaining high test scores, while the applied mechanism of SRL may have to be adjusted to improve students' test performance.

AfL and SRL have been criticized because they are considered socially pedagogical

[1] ZPD is the difference between what one can do on one's own and what one can accomplish with someone else's help (Vygotsky, 1978). Specifically, the ZPD can only be accurately understood by considering the social foundations of problem-solving that are shared among those involved, such as the nature of the problem, the values involved in determining the appropriate ends and means, the intellectual tools available, and the institutional structure of the interaction (Rogoff, 1989).

practices that rely on others/mediating artifacts, rather than individuals' psychological operations alone. To explore how AfL may be useful in SRL, Panadero et al. (2018) stated that, although many studies on AfL and SRL tend to present their conclusions as though they were universally valid, more attention should be paid to different educational levels and contexts. Therefore, the axiom that AfL and SRL are universally valid must be dismantled through practical observation and rigorous research. If the broader context of learning is not considered, the developmental model for fostering evaluative judgment will not cope well with the contextual complexity of the formal and informal aspects of learning, learning beyond academic subjects, and learning within the family and the local community centered on school. Additionally, the marked differences between the historical background of AfL and SRL theories reveal that it is necessary to find a new way to link them (Brandmo et al., 2020).

While AfL focuses on pedagogical and instructional issues, SRL focuses on internal processes, such as cognition and emotion (Panadero & Alonso-Tapia, 2013). Simply combining the commonalities between both theories would be incompatible with real-world practice. Although Nicol and Macfarlane-Dick (2006) introduced seven AfL principles for teachers to foster students' self-regulation in the classroom (e.g., good feedback [goals, criteria, expected standards] helps students understand what good performance is), said principles were not theoretical but technical. These principles are not an exhaustive list, and they ignore teachers and schools' educational philosophy including educational purposes. Thus, how AfL links with SRL remains unclear.

Focus on the Brief's Underlying Assessment

Because both AfL and SRL consider students' ZPD, analyzing the beliefs underlying assessment can be beneficial. Beliefs determine whether assessment is viewed summatively or formatively as well as how students learn, their self-regulation goals, and where they must exercise self-regulation. The execution of AfL and SRL may also reinforce or reframe assessment beliefs by the impact of learning outcomes. Marshall and Drummond (2006) posit that the essence of AfL emerges in teachers' conceptualization and sequencing of students' tasks. Teachers' and students' formal adherence to procedures (e.g., questioning, feedback, disclosure of success criteria, and self-assessment) may overlook the underlying intent, which is deeply embedded in teaching. Similarly, teachers' knowledge, motivation, beliefs, and self-efficacy are involved in students' SRL (Karlen et al., 2020).

Generally, teachers' conceptions shape their purpose and beliefs, and are transmitted to the students (Thompson, 1992). Specifically, conceptions are usually tacit thought processes about the nature of phenomena (Brown & Gao, 2015). Views, values, and attitudes are established through their own or someone else's foreknowledge of what a given learning object really is and means. AfL should focus on teachers and students' conceptions and motivating students from different cultural backgrounds.

To clarify how AfL can contribute to SRL, it is necessary to examine the assessment conceptions that underlie evaluative judgment according to both theories. Accordingly, this

study's purpose is to find what the theoretical and empirical considerations in identifying how AfL can be useful for SRL are by focusing on students' assessment conceptions in various learning contexts. First, the author will examine research on students' conceptions of assessment (SCoA) to clarify how assessment conceptions can relate to AfL and SRL. Thereafter, challenges related to SCoA will be identified to determine how AfL can be useful in SRL. This analysis will reveal some issues that indicate the paths for future research and practice using a new theoretical framework that combines both AfL and SRL, beyond their epistemological differences.

Assessment Conceptions in AfL and SRL

Research on Assessment Conceptions

Several studies have examined what people feel about assessment. According to the Perceptions of Assessment Tasks Inventory (PATI; Dorman & Knightley, 2006), students' perceptions of assessment tasks are congruent with planned learning, authenticity, student consultation, transparency, and diversity. In short, it is desirable to learn where all students (diversity) are involved in the assessment process (student consultation) and examine the distance between their current development and their goals (congruence with planned learning) in a realistic context (authenticity). Vogl and Pekrun (2016) focused on students' emotions during peer assessment, noting that social and human factors strongly influence emotions. Specifically, reciprocal altruism influences students' mutual assessment; thus, if one student scores their peers highly, they will expect their peers to score them highly as well (Fujiwara et al., 2007).

The most widely used survey on students' perceptions of assessment is the SCoA; Brown, 2011). The latest version, SCoA-VI, is a 33-item self-report inventory using a six-point Likert scale. Sample items include "assessment is a good way to determine how much I have learned from teaching," "assessment results predict my future performance," "assessment is an engaging and enjoyable experience for me," "assessment measures the worth or quality of my school," and "our class becomes more supportive when we are assessed." The SCoA inventory is used to assess how students perceive the purpose and nature of assessment; it has been revised to its sixth edition. Although the majority of SCoA-related studies have been conducted with secondary school and college students, the measurement model's efficiency has been tested with upper-elementary school children (Brown, 2011; Brown & Harris, 2012). SCoA research began in New Zealand in 2003 and has continued in several other countries, including Hong Kong, Brazil, the United States, China, Tonga, Tanzania, Cyprus, Iran, Germany, the Netherlands, Portugal, and Japan. Compared with the PATI (which analyzes perceptions of assessment at the task level) and research focusing on the affective aspects of assessment, SCoA research considers the impact of various assessments and their relationship to the educational system as a whole, and can interpret assessment not only psychologically but also sociologically. Assessment conceptions discussed in SCoA research allow us to link AfL with SRL, which is becoming a

form of social action in the community. Hence, this study focuses on SCoA, thereby innovating in field-relevant literature[2].

The SCoA inventory is used to test whether students' assessment conceptions are compatible with AfL and SRL theories by assuming a theoretically feasible factor structure (second-order factor analysis model) as well as conducting exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). Additionally, Brown (2011) examined the correlations and causal relations between assessment conceptions and academic performance, by treating them as dependent variables.

Overview of SCoA Research

In the SCoA-I, Brown and Hirschfeld (2007) studied the association between the Assessment Tools for Teaching and Learning (asTTle[3]), mathematics achievement tests, and certain variables (e.g., gender and ethnicity) in New Zealand in 2003. Students with high test scores believed that assessment was beneficial and could make students accountable. Conversely, students who obtained lower scores believed that assessment interfered with learning and that the school or the teacher were accountable for the results.

Furthermore, Brown and Hirschfeld (2008) examined the relationship between test scores and asTTle reading comprehension. They analyzed the SCoA-I's 11 items and developed a shortened version (SCoA-II). Its structural model consisted of four factors: "student accountability," "school accountability," "fun," and "assessment is ignored," which were correlated; however, only "student accountability" showed a meaningful positive correlation with test scores. Students who found an assessment to be "fun" understood it to be "accountability on the part of the school." Thus, students' perception of assessments as enjoyable influenced their evaluation of schools' quality; however, this perception had no significant effect on test scores, suggesting that it was not so much related to the objective tests as it was to teachers and students' idea that objective tests must be used to assess achievement.

Subsequently, Brown and Hirschfeld (2005) created a third version (the SCoA-III) to reevaluate the first two versions. They conducted EFA to eliminate five items with low fit,

[2] Nevertheless, researchers sometimes modify their use of the SCoA. Flores et al. (2020) used a five-case method in the CFA of the SCoA-VI in a Portuguese university. Meanwhile, when Matos et al. (2013) administered the SCoA-VI to university students, they removed one item regarding the realities faced by Brazilian students who are the target of administration. When Wicking (2020) employed the Chinese version of the SCoA and translated it into Japanese (Brown & Wang, 2016), he used only two annotations: 1, "agree," and 2, "disagree." By doing this, Wicking's approach could lead to respondents misinterpreting the scale. Furthermore research is needed to determine the extent to which these degrees of methodological rigor distort the results of the SCoA.

[3] The asTTle is a free online test that is used for students in years 4–12 in New Zealand schools to track their academic progress.

resulting in three factors: “assessment is beneficial,” “assessment is bad,” and “assessment makes students accountable.” Additionally, a factor analysis was conducted, which found eight subfactors: “assessment makes teachers accountable,” “assessment is valid,” “assessment improves learning,” “assessment interferes with learning,” “assessment is ignored,” “assessment is worthless,” “assessment makes students accountable,” and “assessment is fun.” They then performed CFA to validate this hierarchical factor structure and identify its measurement model. Additionally, the structural equation modeling (SEM) output described the relationship between school level, gender, and asTTle achievement scores (reading comprehension and math). As in the SCoA-II, the SCoA-III suggested that students who view assessment as a resource for personal responsibility achieve higher grades, while those who downplay their value by blaming the school and teachers for assessment results obtain lower grades.

The fourth version (SCoA-IV; Brown, 2006) incorporated a question about assessment types and provided 12 options and open-ended fields: “When you hear the word ‘assessment,’ what kind of assessment practices do you think of?” Data were analyzed for ninth- and tenth-year students from four schools in Auckland. The SCoA-IV underwent CFA, and 21 items that were unsuitable for analysis were eliminated. For the remaining 39 items, a measurement model was obtained, consisting of six interrelated factors: “assessment makes students accountable,” “I use assessments,” “teachers use assessments,” “the public uses assessments,” “assessments are fun,” and “assessments are irrelevant.” These results are similar to those for the SCoA-III. Multidimensional scaling analysis (MDS) identified three assessment types: performance-oriented interactive assessment, teacher-controlled traditional assessment, and observation. No significant association was observed between assessment conceptions and assessment types, but a weak positive correlation emerged between the other factors and the interactive assessment types, except for “assessment is irrelevant.”

In the fifth version (SCoA-V), Brown et al. (2009a) conducted a more rigorous study in New Zealand using random sampling, considering schools’ size, region, socioeconomic status, and co-ed status. After reviewing the previous versions of the inventory and adjusting the 45 items, they aimed to identify and determine the relationship between assessment conceptions and assessment types used in the SCoA-IV. The factors expected to emerge from the 45 items were partly derived from the SCoA-IV and consisted of four factors: “accountable,” “affective/beneficial,” “improving,” and “irrelevant.” “Affective/beneficial” refers to the assessment’s emotional impact (e.g., fun) and usefulness to the student. Therefore, “affective/beneficial” and “irrelevant” were not related, but the other secondary factors were correlated. “Irrelevant” was also found to have a negative association with “personal enjoyment.” MDS was conducted to validate the assessment types and produced somewhat different results from those of the SCoA-IV study. Students interpreted the word “assessment” in two ways: interactive and informal assessment and test-like assessment. When examined with the findings of the SCoA-IV study (Brown, 2006), the former was defined as “performance-oriented interactive assessment,” while the latter was defined as “teacher-controlled traditional assessment.” Finally, they conducted a multiple regression analysis with these two assessment

types as the dependent variables and the eight factors of assessment conceptions as the independent variables. The teacher-controlled assessment was positively correlated with "teachers' improvement in facilitation of student learning (teacher improvement of learning)," and negatively correlated with "personal enjoyment." Meanwhile, interactive assessment had a weak positive correlation with "classroom environment (context)" and "should be ignored." A conclusion regarding students' thinking was reached on the basis of these results: if assessment were controlled by the teacher, students would not necessarily like it, but it could help improve teachers' instruction delivery. Additionally, while interactive assessments are beneficial for class dynamics, students do not pay much attention to them (Brown et al., 2009a).

The sixth version (SCoA-VI) had the same questions as the SCoA-V, but with a modified description of the "accountability" factor (Brown et al., 2009b). Specifically, "student accountability" included external attributions such as "assessments measure how much I have grown" and "assessments determine my future jobs," which are contrary to what students consider as the meaning of learning control. Therefore, the researchers decided to interpret "accountability," implying learning independence, as the "improvement" factor and renamed "accountability" in the SCoA-V as "external factors," while the subfactor "student accountability" was renamed to "students' future." Additionally, "school accountability" was renamed as "school quality." "External factors" are focused on school, parents, and the future, while excluding students' intrinsic attitudes toward taking ownership of their learning. Therefore, Brown et al.'s (2009b) hypothesis was that the correlation between the factors on the SCoA-VI list and achievement scores (AoL) would be positive only for "improvement" and negative for the other factors. Brown et al. (2009b) used data from the SCoA-V study in 2006 and from three new schools in New Zealand in 2007 to determine the fit of the newly established structure. They investigated the link between the model's degree of fit, the association between assessment types and conceptions, and the relationship between asTTle math scores. The results showed that the SEM, which explored the relationship between assessment types and conceptions, had the same structure as the SCoA-V. Meanwhile, a comparison of the 2006 and 2007 data showed that the relationship between assessment types and conceptions was not generalized and could be influenced by differences in learning environments. The association between math scores, assessment conceptions, and assessment types was also analyzed via SEM, although only for 2007. The results suggested that the stronger the tendency toward "external factors," the lower the score, while the stronger the tendency toward "improvement," the higher the score. In terms of assessment types, "test-like assessment" was positively correlated with scores, while "teacher-controlled assessment" was negatively correlated with "personal enjoyment" and positively correlated with "teacher improvement of learning."

The aforementioned students' assessment conceptions are shown in Table 1. These conceptions show that taking responsibility for one's own learning, recognizing assessment as a process for improvement, and believing feedback to be aligned with learning can enhance academic performance. Furthermore, some conceptions about assessment certainly have a positive impact on academic scores. Several studies have confirmed the inventory's measurement

invariance (Brown et al., 2014; Flores et al., 2020; Michaelides & Solomonidou, 2019) or reconsidered the SCoA items (Matos et al., 2019). Additionally, other psychometric enhancements to the SCoA-VI are underway. Because the widely used asTTle is a test that embodies public standards, it can provide valuable insights into the status and outcomes of national and local educational policies.

The following findings are noteworthy: specific assessment conceptions can be used to increase achievement scores; existing assessment conceptions can be culturally universal; SCoA research can help determine not only the effects of classroom practices but also policy and institutional effects; and SCoA surveys serve as a reference point for analyzing students' conceptions of assessment (power point) and what assessment provides (action point).

Table 1. *Rough Typology of Students' Assessment Conceptions in SCoA Research*

SCoA-I	SCoA-II	SCoA-III	SCoA-IV	SCoA-V	SCoA-VI
Improvement zone					
[Improvement] it is good for me		[Beneficial] improves	[I use it]	[Improvement] student improvement	[Improvement] student improvement
[Improvement] it improves teaching			[Teachers uses it]	[Improvement] teacher improves learning	[Improvement] teacher improves learning
[Useful] it is valid		[Beneficial] it is valid			
[Useful] it captures my thinking					
[Useful] reliable					
Irrelevant zone					
[Negative] it interferes	[Ignored]	[Bad] it interferes	[Irrelevant]	[Irrelevant] it is bad	[Irrelevant] it is bad
[Negative] it is ignored		[Bad] it is ignored		[Irrelevant] it is ignored	[Irrelevant] it is ignored
[Negative] it has errors		[Bad] it is worthless			
Accountability zone					
[Accountability] makes schools accountable	[School accountability]	[Beneficial] school accountability	[Public uses]	[Accountability] school accountability	[External factors] school quality
[Accountability] makes students accountable	[Student accountability]	[Makes student accountable] student accountability	[Students accountable]	[Accountability] student accountability	[External factors] students' future
Affect/Benefit zone					
[Improvement] it is fun	[Fun]	[Beneficial] it is fun	[Fun]	[Affect/Benefit] personal enjoyment	[Affect/Benefit] personal enjoyment
				[Affect/Benefit] class environment	[Affect/Benefit] class environment

AfL and SRL vis-à-vis Assessment Conceptions

Assessment conceptions dictate whether an assessment should be used in a formative or summative manner. Additionally, the assessment conceptions that drive AfL do not only involve improvement; affect/benefit is about what relationships, mainly social capital, lead to what kind of improvement content and methods, whereas irrelevant and external factors are about what political and social factors define the purpose of improvement. This means that if the content and weight of the affective/beneficial, irrelevant, and external factors change, the purpose, content, and methods of improvement will also change. For example, if students and schools' political accountability are high, AfL could lead to policy changes aiming to improve test scores, in line with the schools' standards. Irrelevant factors (which are negatively correlated with test scores) are influenced by students' negative emotions (e.g., fear and avoidance) and by parents, teachers, and classmates' conceptions of assessment and the external feedback received from them (Fujiwara et al., 2007; Nicol & Macfarlane-Dick, 2006; Vogl & Pekrun, 2016). Additionally, affect/benefit is also involved in determining learning styles and attitudes. If assessment can enrich classroom interactions, students should be able to develop in several ways that cannot be measured by tests because they will have opportunities to extend their learning beyond knowledge acquisition (Arimoto & Clark, 2018).

Assessment conceptions also serve to motivate each phase of the SRL cycle (Zimmerman's [2002] model is used as an example here). The belief that students will improve their learning is associated with internal feedback throughout the SRL cycle (Nicol & Macfarlane-Dick, 2006) and contributes to forming the foundation of the cycle. Particularly, the belief that teachers can improve learning means that students are aware of anomalies in the SRL cycle and request general support to ensure that the cycle runs smoothly. External factors (e.g., students' future and school quality) are related to SRL's goals and outcomes (i.e., the forethought and self-reflection phases), while affect/benefit is oriented toward forethought based on self-efficacy and interest. Among them, the class environment involving rich communication affects students' performance because it accumulates social capital and increases the number of options for learning strategies. Finally, students exhibiting high scores in irrelevant factors does not necessarily mean that assessment conceptions and SRL are unrelated. In short, students usually think that assessments and the pedagogies useful for learning are separate, and that the SRL cycle is designed to communicate to students the importance of understanding the purpose of assessment as a means for improvement or as irrelevant, depending on the outcome. The SRL cycle reinforces or restructures the framework of perceptions of assessment.

To engage in SRL, students must keep in mind their learning goals and compare their current performance to those goals. Thus, each assessment conception captured by the SCoA research shows the link between AfL and SRL. However, to clarify the transformative logic model that shows AfL's contribution to SRL, it is necessary to determine how AfL contributes to SRL, by modifying the perspective of conception of assessment. The following section will explore four challenges as the essence of self-regulation for AfL.

Four Paths for AfL's Contribution to SRL

Unpacking the Ecological Rationality of Assessment

Beliefs are rationally formed cognitive frameworks externalized in a context (Brown & Harris, 2012). Brown and Harris (2012) used the SCoA-VI to examine differences between elementary and high school students' assessment conceptions. They found that high school students were less likely to agree with SCoA-VI factors, such as "external factors," "improvement," and "affective/beneficial," and more likely to agree with SCoA-VI factors, such as "inappropriate." The students incorporated assessment methods and perceptions from the test-based learning environment into the SCoA-VI. A survey of university students in Brazil and New Zealand revealed that differences in belief systems within the ecosphere influenced the way conceptions were utilized (Matos & Brown, 2015). Assessment conceptions are formed and transformed in local contexts, and students' assessment conceptions are influenced by ecological rationality. Ecological rationality examines individuals' decision-making processes in response to their circumstances using information from the environment (Goldstein & Gigerenzer, 2002). Human thinking activities occur in specific ecosystems that link thinking, language, and activity. Thus, assessment conceptions are not used in isolation but are first externalized in the decision-making process according to ecological filters.

The objective of research in this field is to determine how to confirm assessment conceptions using ecological rationality. According to Mata et al. (2012), ecological rationality has three basic ideas regarding decision-making. First, the mind adapts its decision-making strategies to a particular environment. Whether a decision-making strategy is good is not determined by the content of the decision but by the environment in which such it is used. Second, a simple decision strategy may conflict with a complex strategy in certain environments (which will be discussed in the next section). Third, individuals often respond adaptively to tasks and environmental characteristics, frequently making decisions without adequate time for deliberation.

When these three ideas are applied to assessment, three points should be considered. First, the appropriateness of an assessment conception is determined by the environment (i.e., the purpose of the assessment [objects], the criteria within it, and the activities toward the object). Classroom assessment can be expressed in terms of specific skills and techniques (e.g., self-assessment and peer assessment); however, more appropriately, it is a process within a specific period aiming to transform an immediate problem or situation to realize a qualitatively better future. Teaching, learning, and educational goal theory should be analyzed from this perspective.

Second, humans have multiple (sometimes incompatible) assessment conceptions. Social structures (e.g., examination systems and work environments), educational needs from various perspectives (teachers' and students' views on teaching and learning), and other resources for mediation, place excessive stress on cognitive activities, causing a struggle among

assessment conceptions comprising various criteria of ecological rationality. Analyzing the biases of resources in local communities and school organizations and their decisional processes regarding what is most appropriate (considering multiple mediations), may be useful in identifying ways to select more appropriate assessment conceptions and modify existing beliefs.

Third, assessment conceptions can be demonstrated by identifying and deliberately confronting multiple assessment conceptions not only held by teachers and students but also at the level of a judgmental process that involves senses, connoisseurship, and practical wisdom. Below, the author will examine SCoA research vis-à-vis goal theory and tasks, tensions of assessment, and process theory, which originate from ecological rationality.

Capturing Assessment Conceptions from Well-rounded Development and Community Levels

The goal of assessment tasks should be to nurture agentic students. Regarding goal theory, AfL and SRL have recently been criticized for their tendency to address personal goals at the task level in terms of psychology and cognition.

In the SCoA-IV, a positive correlation was found between interactive assessment and the "fun" factor (Brown, 2006), which can also be identified in earlier versions of the inventory as interactive formative assessment is fun, and it lowers students' achievement scores. Brown explains that this may be because students do not respond positively or have access to the information gained from such interactions (Brown, 2006); that is, when students do not know how to use an assessment well or when they cannot use it formatively, assessment may be conceived as "fun" rather than as a means for improvement or learning accountability.

Assessment conceptions in SCoA research reveal a logic that is accepted only when linked to individual academic learning outcomes (e.g., test scores) and not interpreted in the dimension of community and well-rounded education based on environmental and mental interactions. In some situations, when students become aware of their collective ZPD and challenge that zone through communication, assessment can be conducive to development by enhancing the classroom atmosphere and immersing students in a sense of enjoyment (Arimoto & Clark, 2018). Elwood (2006) suggested that assessments tend to be conducted unilaterally by a teacher examining a passive student's mind, yet the nature of learning is collaborative (Elwood & Murphy, 2015). Pryor and Crossouard (2010) proposed an AfL process linked to tasks and identity, while Uebuchi (2007) noted that the developmental mechanism of metacognition can be better identified by exploring the relationship between SRL and identity. Meanwhile, Shinto (2017) indicated that SRL may affect students' interaction with society and attribute learning accountability to individual students.

Thus, alternatives for moving beyond the task level of assessment are likely to bring about new linkages between the components of assessment conceptions and learning activities. Although approaches in this perspective (e.g., SSRL and CoRL) are becoming more common, regulation may be required for students to become more flexible; while initially appearing to be optimally self-oriented and self-aware, students could actually be searching for the most skillful way to adapt to any context presented by authority figures (Engeström, 2015). Thus, assessments

could be used to maintain a community's discourse and reproduce its history.

When considering assessments that promote SRL, AfL encompasses self-efficacy, collective efficacy, feedback, and the normative context of the communication arena and the community. Clark (2012) encouraged teachers and students to become cultural change agents that create valid assessment conceptions tailored to fit the community. Accordingly, it is valuable to analyze how assessment conceptions are utilized in the broader social, cultural, and political context (Black & Wiliam, 2018). Assessment conceptions do not make AfL and AoL (at the end of a unit or grading period) mutually exclusive. Assessment conceptions should be tested over a long period in relation to factors that influence subjects' personality development (e.g., occupation, mental health, social capital, civic engagement, and willingness to learn and contribute to the company while working).

Identifying Tensions Between Assessment Conceptions

SCoA research has often problematized the tensions arising from teachers' assessment beliefs (Brown & Gao, 2015; Chen & Brown, 2016; Harris & Brown, 2009). Brown and Gao (2015) observed that Chinese teachers' freedom to use assessment is curtailed by "making exam marks the only standard for enrollment into higher levels of schooling; copying administration models and rules from factories and enterprises; ranking schools according to their resources and student performance in public examinations; and large class sizes" (p. 16). Additionally, Chen and Brown (2016) found a juxtaposition between cultural-level assessment (which emphasizes care and compassion for students) and institutional-level assessment (such as standards and compliance). Similarly, New Zealand teachers had conflicting assessment conceptions, such as student-centered vs. school standards-centered, improvement vs. accountability, and being happy with assessments vs. avoiding them (Harris & Brown, 2009).

While some teachers comply with the current education system, others have distanced themselves from it, developing a different idea of what education should be. Although Chen and Brown's (2016) study does not confirm whether teachers prefer teacher-, exam-, or standard-centered assessments, tensions clearly exist among their assessment conceptions, at least regarding policy issues. In China, AfL does not support development from a broader perspective unless the entire range of students' abilities is incorporated into how the testing system is measured (Brown & Gao, 2015). To alleviate the tension between different teacher assessment conceptions in China, all educational goals must be systematically controlled, while the scope of measurement in the testing system needs to be expanded. However, one can easily imagine that teachers' workload will be even greater because they are tasked with developing higher-order thinking skills in a context where both knowledge transfer and student-centered instruction exist, and where the time required to achieve proficiency and the pedagogy for efficient and reliable teaching are unclear. Providing opportunities for teacher agency and learning expansion in the process of confronting proprietary interest is better than avoiding tensions (Engeström, 2015). Thus, it is important to analyze which assessment challenges commonly frustrate teachers, researchers, and stakeholders, and how they can be overcome, by

developing solutions using a methodology of linear knowledge-creation intervention.

Strengthening AfL as a Process Theory

AfL is a process theory that explains a series of normative procedures. Therefore, satisfying the criteria for it to be considered a process theory is at the core of theory formation. Engeström (2016) identified five conditions in a process theory: (1) describing event successions and stages with a certain universality; (2) proposing evidence and principles to explain why said successions and stages occur in the order that they do; (3) explaining causal mechanisms that create transitions from one event to another; (4) eliminating the universalism of the theory and identifying its cultural and historical limits (the extent of its application); and (5) examining the gap between teaching and learning.

Andrade and Brookhart (2020) adopted Pintrich and Zusho's (2002) SRL model to examine the positive impact of classroom assessment on SRL. Pintrich and Zusho's (2002) model had four stages (forethought, planning, and activation; monitoring; control; and reaction and reflection) and four adjustment areas (cognition; motivation/affect; behavior; and context). They analyzed how assessment shapes each stage and adjustment area. They found that the first stage involves both teachers and students setting goals and students determining their self-efficacy for the assessment task, which involves teachers and students providing formative feedback on learning outcomes and processes to themselves and each other. In the last two stages, classroom assessment helps teachers adjust their instruction and assists students in adapting their learning strategies and activities in response to assessment information. Here, condition (1) is confirmed. Further, the theoretical basis for succession in condition (2) includes the "evolution from 'I cannot' to 'I can'" through feedback control. This is evident in Pintrich's (1995) analogy of a thermostat mechanism. Although it is not explicit in Andrade and Brookhart's work, Pintrich (2000) argued that "goal orientation" is the driving force that compares each stage using the criterion of the oriented goal and selects the appropriate stage. This may explain the causal mechanism of transition in (3).

Conditions (4) and (5) need to be addressed further. Both AfL and SRL theories are often regarded as good for us (Black & Wiliam, 2018; Panadero et al., 2018). However, regarding the aforementioned ecological rationality, there is no universal physical law that leads to success in AfL and SRL. Accordingly, Jakešová and Kalenda (2015) advocated for a break from universalism and a simple theoretical structure, from transcendental arguments for Bhaskar's critical realism to finding "an explanation of variability in the mechanisms of SRL through time and space in the character" (p. 188). Their preferred methodology for exploring SRL's causal mechanisms conforms to Bhaskar's new type of ontology. Meanwhile, others argue that theoretical inquiry must "explore and determine the relationship and irrelevance, respectively, between what we experience, what is actually happening, and the underlying mechanisms that produce various events in this world" (Danermark, 2016, p. 230).

Likewise, regarding condition (5), despite the fact that the basic premise of teaching-learning theory is the discrepancy between what a teacher teaches and what a student

learns (Engeström, 2015), and between what teachers want to teach and what they can teach, AfL entails learning improvement (Nishizuka, 2020), whereas SRL focuses only on students' self-regulation (Oates, 2019). According to Brandmo et al. (2020), one of the intentions of formative assessment "is to internalize external/environmental standards, so they become internal/personal and available for self-regulated learning" (p. 323). Accordingly, since AfL and SRL are built on the premise that learning is externally delivered to students, they focus on pedagogical interventions to make students reduce what they are incapable of accomplishing (incomplete) and improve what they can easily accomplish (complete). This way, the knowledge base of educational activities becomes less of an issue. Additionally, although the prerequisite of SCoA, which has an "improvement" factor for teaching, lies in task levels, the range of improvement does not include changing teachers' educational philosophy; SCoA determines what can be improved while preserving the current teaching system (single-loop learning), while the educational philosophy determines why the existing system should be revised (double- or triple-loop learning). Since SCoA research is insufficient to grasp the extent and target of improvement, it is necessary to capture the traces of actual educational and social activities.

Of importance is the need to deviate from a universalist view and thoroughly analyze the dialectical exchange between teaching and learning. Although SCoA is not developed in a decontextualized manner, it only considers assessment within a specific context, rather than as a cultural practice. Therefore, the factor structure of the SCoA (which has been revised throughout six editions) is one of all possible models. Because CFA increases the arbitrariness of the researcher, the results obtained when using the SCoA across cultures are the least common multiple of the assessment conceptions. In addition to cultural differences in existing scales, exploratory research in the opposite direction should be conducted to identify cultural uniqueness. The SCoA inventory, as an indirect assessment, questions the "normative representation (intention process)" rather than the "implementation process" of assessment, but the intention and implementation of human activities generally follow different trajectories (Engeström, 2016). Monteiro et al. (2021) suggest that teachers working within social and contextual constraints adopt assessment conceptions that are inconsistent with their practice. Hence, Panadero et al. (2018) questioned whether self-report questionnaires are a valid reflection of AfL and SRL theory. Researchers must welcome analytical methods that capture the environment in which learning occurs, such as qualitative analysis, which tracks and observes students' thought processes.

Preliminary Consideration

The four pathways consider the fundamental structures of schooling, such as large economic disparities, students from different social and family backgrounds, teachers with different perceptions of assessment, and large class sizes typical of East Asian countries. It also considers the global trend of the influx of market forces into schooling, such as test supremacy and fierce competition. Recently, the sophistication of students' learning goals and the accompanying

enlargement of the curriculum have also been increasing in OECD countries (OECD, 2021), and this is having a negative impact not only on the students but also on the teachers' physical and mental health. In Japan, for example, teachers are becoming increasingly busy in their daily work, and *karoshi* (death from overwork), which is caused by the inability to manage stress well, is also paid attention to (Kuwato & Hirano, 2020). It is impossible to separate the assessment conceptions from social structures and policies, so that the daily lives of teachers and students are always fraught with anxiety. Evidently, AfL is never easy, and there is a need for more cultural research in the future that will identify factors that influence teachers' and students' assessment conceptions (e.g., Arimoto & Clark, 2018).

One cue for students and teachers to find and use the appropriate assessment conceptions at any given time is to engage in learning that fundamentally changes social structures, such as the expansive learning proposed by Engeström (2015), which regards learning as concept formation while dialectically overcoming contradictions; that is, AfL that encourages SRL must be formed through expansive learning. Empirical research has provided insight into the process by which teachers acquire AfL through expansive learning (e.g., Nishizuka, in press). How these teachers' assessment conceptions are linked to students' assessment conceptions and SRL is still under investigation, but the process of forming and acquiring AfL will be able to reveal the unique culture of the school organization, classroom, or local community that supports the assessment conceptions.

As an example that covers all the four perspectives, the author outlines an empirical study which is currently being worked on. The study aims to understand how teachers can acquire AfL skills through their daily interactions. The author is focusing on three teachers (Akita, Kanazawa, Umeki, all pseudonym) of a first-year high school social studies course, "Contemporary Society," in a Japanese integrated junior high school. The year-long course incorporates inquiry-based learning, where students ask their own questions and solve problems, and teachers avoid a premade curriculum. The teachers hold subject committee meetings approximately once a week to exchange information and make decisions, such as reporting and reviewing the progress of instruction in all five classes and adjusting the curriculum. Since it is difficult for teachers to reconstruct the assessment conceptions on their own, the author, as a mediator, intervenes and gives teachers opportunities to reflect on the assessment conceptions.

In the following partial communication (June 26th, 2020), the teachers attempted to transfer the responsibility for assessment to the students. The author presented teachers with a contradiction that forced them to confront multiple conflicting judgments about whether the criteria in the rubric , a tool for assessing the quality of students' performance tasks, should be segmented or inclusive. Since the criteria were originally developed based on students' discussions, teachers thought that the more criteria they presented, the more students would believe that the assessment was consistent with their learning.

> 25 (Akita): Regarding "consistency of logic" (a higher-order criterion in the rubric), the "appropriate question" (a lower-order criterion) remains in the original proposal, but the

"argumentation process is appropriate" is below it.
26 (Kanazawa): Oh, it has increased.
27 (Akita): The mediator said "It is difficult to understand if it is 'appropriate or not,' so how about looking at the steps of claim, evidence, opposing view, rebuttal, and conclusion?" He said, "There is also a way to divide them like that," and, "It's not a matter of what to do, but which one is easier for the teachers to grade."
28 (Kanazawa): I see.
36 (Umeki): I wonder about scoring. I don't know if it's better to separate them.
53 (Akita): Yes. I think we can read more precisely if we separate them. In contrast, from the students' point of view, there are too many criteria.
56 (Umeki): I think the students would feel that they have understood various parts of the reports if there were more criteria.
57 (Kanazawa): Well, yes. We can say "I used the ones that everyone told us (in the class dialog before)."

The above research design is based on teachers' concept formation by expansive learning, which is a theoretical framework that views the formation and implementation of AfL as a process theory (Engeström, 2015). The author presented a contradiction regarding assessment conceptions, allowed teachers to discuss among themselves, and gave the final decision-making authority to them. It can be seen that the culturally optimized assessment, in which criteria should be presented more in line with students' suggestions, encouraged students to take more responsibility for the assessment. It is not about what is good from a scientific or academic perspective, but what is desirable for the students in front of them; the teachers were discussing the "ecological validity." The attempt to hold students accountable for assessment also encourages students to take control of their own learning, and works to let students know that assessment is not just concerned with testing. Thus, if researchers can intervene in a way that teachers are confronted with the assessment conceptions and are stimulated by their dialog, it may lead to students' acquisition of the assessment conceptions and the development of SRL.

Conclusion

The increasing proximity of AfL and SRL theories implies that the line between assessment and learning is beginning to blur; the two overlap significantly in both theory and practice. This paper seeks to clarify how AfL can contribute to SRL by focusing on SCoA based on cultural-historical activity theory. SCoA research provides many suggestions for identifying the underlying mechanisms of AfL and SRL. The four concerns confronting assessment conceptions research require a broader context—one that will force us to rethink the prevalent assessment conceptions. It is not realistic to hope for eclectic complementarity or a simple integration of research findings to make the best use of the theoretical strengths of AfL and SRL (Brandmo et al., 2020). Thus, integrating AfL and SRL implies a theoretical deconstruction and reconstruction.

Therefore, it is problematic that these challenges are not recognized in the existent literature.

For AfL to contribute to SRL, this paper presents three practical suggestions. First, it is necessary to promote actionable research in which the teachers involved, with the help of researchers and intermediaries, conceptualize assessment in a sociocultural way. Since there is a cultural gap between assessment conceptions at the theoretical and practical levels, teachers and researchers should clarify how to implement AfL for SRL and how to overcome its challenges. Second, it is necessary to reset the learning goals and assessment tasks. The collectivistic view of assessment is at odds with the individualistic view (which emphasizes examinations and test scores), thereby introducing additional complexity to this topic. Thus, the author suggests that AfL should be considered separate from summative function of assessment (AoL), and that teachers should set goals for collective (rather than individualistic) development. Third, the relationship between pedagogy and SRL must be examined. Specifically, CoRL and SSRL, which include the aforementioned suggestions, are clearly difficult issues. Teachers must reconsider what regulation is and its purpose, and how they can align it with their community's understanding of SRL, CoRL, and SSRL. When AfL is implemented to fulfill the purpose of regulation, and how it can be aligned with the community's understanding of SRL, CoRL, and SSRL, with respect to teaching and learning, teachers are able to improve the collective relevance of AfL.

Lastly, this study had two limitations. The first limitation is its lack of practical examples for future applications and research. This paper is mainly deepening theoretical claims, and the practical examples are from a Japanese high school. However, just as the SCoA study covers a wide range of schools, from elementary schools to universities worldwide, it is expected that the arguments here are equally applicable to any school level globally. This study's second limitation concerns its narrow range of considerations. As this study discusses how AfL is related to SRL, strictly from the standpoint of AfL, it is necessary to examine this issue from the standpoint of SRL.

Regarding the abovementioned practical suggestions, it is necessary to develop theoretical research based on an analysis of implicit assessment practices. The structural understanding of AfL, SRL, and assessment conceptions can be developed further by analyzing the assessment conceptions from an empirical perspective, as posited in the four approaches presented above. Hence, this paper contributed to a new potential theoretical framework for combining AfL and SRL.

References

Andrade, H. L. & Brookhart, S. M. (2020). Classroom assessment as the co-regulation learning. *Assessment in Education: Principles, Policy & Practice*, *27*(4), 350–372.

Arimoto, M. & Clark, I. (2018). Interactive assessment: Cultural perspectives and practices in the nexus of "heart or mind." In A. A. Lipnevich & J. K. Smith (Eds.), *The Cambridge*

handbook of instructional feedback (pp. 474–503). Cambridge University Press.

Baird, J., Andrich, D., Hopfenbeck, T. N., & Stobart, G. (2017). Assessment and learning: Fields apart? *Assessment in Education: Principles, Policy & Practice*, *24*(3), 317–350.

Black, P. & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74.

Black, P. & Wiliam, D. (2006). Developing a theory of formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 81–100). SAGE.

Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment Evaluation and Accountability*, *21*(1), 5–31.

Black, P. & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, *25*(6), 551–575.

Brandmo, C., Panadero, E., & Hopfenbeck, T. N. (2020). Bridging classroom assessment and self-regulated learning. *Assessment in Education, Principles, Policy & Practice*, *27*(4), 319–331.

Brown, G. T. L. (2006). *Secondary school students' conceptions of assessment: A survey of four schools.* Conceptions of assessment and feedback project report #5. University of Auckland. DOI: 10.13140/RG.2.2.24963.71202

Brown, G. T. L. (2011). Self-regulation of assessment beliefs and attitudes: A review of the Students' Conceptions of Assessment inventory. *Educational Psychology, 31*(6), 731–748.

Brown, G. T. L. (2020). Responding to assessment for learning: A pedagogical method, not assessment. *New Zealand Annual Review of Education*, *26*, 18–28.

Brown, G. T. L. & Gao, L. (2015). Chinese teachers' conceptions of assessment for and of learning: Six competing and complementary purposes. *Cogent Education*, *2*(1).

Brown, G. T. L. & Harris, L. (2012). Student conceptions of assessment by level of schooling: Further evidence for ecological rationality in belief systems. *Australian Journal of Educational and Developmental Psychology*, *12*, 46–59.

Brown, G. T. L. & Hirschfeld, G. H. F. (2005). *Secondary school students' conceptions of assessment*. Conceptions of assessment and feedback project report #4. University of Auckland. DOI: 10.13140/RG.2.2.11541.93921

Brown, G. T. L. & Hirschfeld, G. H. F. (2007). Students' conceptions of assessment and mathematics: Self-regulation raises achievement. *Australian Journal of Educational and Developmental Psychology*, *7*, 63–74.

Brown, G. T. L. & Hirschfeld, G. H. F. (2008). Students' conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy & Practice*, *15*(1), 3–17.

Brown, G. T. L., Irving, S. E., Peterson, E. R., & Hirschfeld, G. H. F. (2009a). Use of interactive—informal assessment practices: New Zealand secondary students' conceptions

of assessment. *Learning and Instruction*, *19*(2), 97–111.

Brown, G. T. L., Peterson, E. R., & Irving, S. E. (2009b). Beliefs that make a difference: Adaptive and maladaptive self-regulation in students' conceptions of assessment. In D. M. McInerney, G. T. L. Brown, & G. A. D. Liem (Eds.), *Student perspectives on assessment: What students can tell us about assessment for learning* (pp. 159–186). Information Age Publishing.

Brown, G. T. L., Peterson, E. R., & Yao, E. S. (2016). Student conceptions of feedback: Impact on self-regulation, self-efficacy, and academic achievement. *British Journal of Educational Psychology*, *86*(4), 606–629.

Brown, G. T. L., Pishghadam, R., & Sadafian, S. S. (2014). Iranian university students' conceptions of assessment: Using assessment to self-improve. *Assessment Matters*, *6*, 5–33.

Brown, G. T. L. & Wang, Z. (2016). Understanding Chinese university student conceptions of assessment: Cultural similarities and jurisdictional differences between Hong Kong and China. *Social Psychology of Education*, *19*(1), 151–173.

Chen, J. & Brown, G. T. L. (2016). Tensions between knowledge transmission and student-focused teaching approaches to assessment purposes: Helping students improve through transmission. *Teachers and Teaching*, *22*(3), 350–367.

Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review*, *24*, 205–249.

Danermark, B. (2016). Critical realism: An introduction. *Ritsumeikan Social Sciences Review*, *51*(4), 227–234. (In Japanese)

Daniels, H. (2001). *Vygotsky and pedagogy*. Routledge Falmer.

Dorman, J. P. & Knightley, W. M. (2006). Development and validation of an instrument to assess secondary school students' perceptions of assessment tasks. *Educational Studies, 32*(1), 47–58.

Elwood, J. (2006). Formative assessment: Possibilities, boundaries and limitations. *Assessment in Education: Principles, Policy & Practice*, *13*(2), 215–232.

Elwood, J. & Murphy, P. (2015). Assessment systems as cultural scripts: A sociocultural theoretical lens on assessment practice and products. *Assessment in Education: Principles, Policy & Practice*, *22*(2), 182–192.

Engeström, Y. (2015). *Learning by expanding: An activity-theoretical approach to developmental research* (2nd ed.). Cambridge University Press.

Engeström, Y. (2016). *Studies in expansive learning: Learning what is not yet there*. Cambridge University Press.

Flores, M. A., Brown, G. T. L., Pereira, D., Coutinho, C., Santos, P., & Pinheiro, C. (2020). Portuguese university students' conceptions of assessment: Taking responsibility for

achievement. *Higher Education*, *79*(3), 377–394.

Fujiwara, Y., Ohnishi, H., & Kato, H. (2007). An evaluation support system for impartial peer evaluation: "Reciprocal effect" occurred by selecting evaluators in case of peer evaluation of learning outcome. *Japan Journal of Educational Technology*, *31*(2), 125–134. (In Japanese)

Goldstein, D. G. & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, *109*(1), 75–90.

Greene, J. A. (2020). Building upon synergies among self-regulated learning and formative assessment research and practice. *Assessment in Education: Principles, Policy & Practice*, *27*(4), 463–476.

Hadwin, A. F., Järvelä, S., & Miller, M. (2017). Self-regulation, co-regulation and shared regulation in collaborative learning environments. In D. Schunk & J. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd Ed.). Routledge.

Harris, L. R. & Brown, G. T. L. (2009). The complexity of teachers' conceptions of assessment: tensions between the needs of schools and students. *Assessment in Education: Principles, Policy & Practice*, *16*(3), 365–381.

Jakešová, J. & Kalenda, J. (2015). Self-regulated learning: Critical-realistic conceptualization. *Procedia-Social and Behavioral Sciences*, *171*, 178–189.

Karlen, Y., Hertel, S., & Hirt, C. N. (2020). Teachers' professional competences in self-regulated learning: An approach to integrate teachers' competences as self-regulated learners and as agents of self-regulated learning in a holistic manner. *Frontiers in Education*, *5*(159).

Kuwato, M. & Hirano, Y. (2020). Sense of coherence, occupational stressors, and mental health among Japanese high school teachers in Nagasaki prefecture: A multiple regression analysis. *BMC Public Health*, *20*(1355).

Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge University Press.

Marshall, B. & Drummond, M. J. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education*, *21*(2), 133–149.

Mata, R., Pachur, T., von Helversen, B., Hertwig, R., Rieskamp, J., & Schooler, L. (2012). Ecological rationality: A framework for understanding and aiding the aging decision maker. *Frontiers in Neuroscience*, *6*(19).

Matos, D. A. S. & Brown, G. T. L. (2015). Comparing university student conceptions of assessment: Brazilian and New Zealand beliefs. In C. Carvalho & J. Conboy (Eds.), *Feedback, identidade, trajetórias escolares: Dinâmicas e consequências* (pp. 177–194). Universidade de Lisboa, Instituto de Educação.

Matos, D. A. S., Brown, G. T. L., & Gomes, C. M. A. (2019). Bifactor invariance analysis of student conceptions of assessment inventory. *Psico-USF*, *24*(4), 737–750.

Matos, D. A. S., Cirino, S. D., Brown, G. T. L., & Leite, W. L. (2013). A avaliação no ensino superior: Concepções múltiplas de estudantes Brasileiros. *Estudos em avaliação educacional*, *24*(54), 172–193.

Michaelides, M. P. & Solomonidou, G. (2019). Factorial structure, gender invariance, and predictive validity of the Students' Conceptions of Assessment-VI Inventory. *European Journal of Psychological Assessment*, *35*(2), 248–254.

Monteiro, V., Mata, L., & Santos, N. N. (2021). Assessment conceptions and practices: Perspectives of primary school teachers and students. *Frontiers in Education, 6*(631185).

Nelson, T. O. & Narens, L. (1990). Metamemory a theoretical framework and new findings. *The Psychology of Learning and Motivation, 26*, 125–174.

Nicol, D. J. & 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.

Nishizuka, K. (2020). A critical review of formative assessment research and practice in Japan. *International Journal of Curriculum Development and Practice*, *22*(1), 15–47.

Nishizuka, K. (in press). The Role of Asynchronous Formative Intervention in Formative Assessment Practice: Focusing on the "Contemporary Society" Subject Committee as a Third Space. *The Journal of Activity Theory Research*, *7*. (In Japanese)

Oates, S. (2019). The importance of autonomous, self-regulated learning in primary initial teacher training. *Frontiers in Education*, *4*(102).

OECD (2021). *Curriculum overload: A way forward*. OECD.

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

Panadero, E. & Alonso-Tapia, J. (2013). Self-assessment: Theoretical and practical connotations. When it happens, how is it acquired and what to do to develop it in our students. *Electronic Journal of Research in Educational Psychology*, *11*(2), 551–576.

Panadero, E., Andrade, H., & Brookhart, S. M. (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.

Panadero, E. & Järvelä, S. (2015). Socially shared regulation of learning: A review. *European Psychologist*, *20*(3), 190–203.

Pintrich, P. R. (1995). Understanding self-regulated learning. *New Directions for Teaching and Learning*, *63*, 3–12.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). Academic Press.

Pintrich, P. R. & Zusho, A. (2002). The development of academic self-regulation: The role of cognitive and motivational factors. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 249–284). Academic Press.

Pryor, J. & Crossouard, B. (2010). Challenging formative assessment: Disciplinary spaces and identities. *Assessment and Evaluation in Higher Education*, *35*(3), 265–276.

Rogoff, B., Mosier, C., Mistry J., & Göncü, A. (1989). Toddlers' guided participation in cultural activity. *Cultural Dynamics*, *2*(2), 209–237.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–144.

Shinto, T. (2017). The possibility of self-regulated learning theory: Focusing on motivation and individual differences. *Ritsumeikan Teacher Educational Studies*, *4*, 23–32. (In Japanese)

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.

Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (pp. 127–146). Macmillan Publishing Co, Inc.

Uebuchi, H. (2007). Self-regulated learning and metacognition: From the perspectives of intentionality, selves, and environments. *Japanese Psychological Review*, *50*(3), 227–242. (In Japanese)

Vogl, E. & Pekrun, R. (2016). Emotions that matter to achievement: Student feelings about assessment. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of human and social conditions in assessment* (pp. 111–128). Routledge.

Vygotsky, L. S. (1978). *Mind in society: Development of higher psychological processes.* Harvard University Press.

Wicking, P. (2020). Formative assessment of students from a Confucian heritage culture: Insights from Japan. *Assessment and Evaluation in Higher Education*, *45*(2), 180–192.

Willis, J. (2010). Assessment for learning as a participative pedagogy. *Assessment Matters*, *2*, 65–84.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, *41*(2), 64–70.

Zimmerman, B. J. & Schunk, D. H. (2001). *Self-regulated learning and academic achievement: Theoretical perspectives*. Lawrence Erlbaum Associates.