# Reconsidering the Assessment of Self-Regulated Learning in Foreign Language Courses

Kristen Sullivan, Shimonoseki City University, Japan

Corresponding author: kristenmareesullivan@gmail.com

**To cite this article**

**To link to this article**

Scroll down for article.

# Reconsidering the Assessment of Self-Regulated Learning in Foreign Language Courses

Kristen Sullivan, Shimonoseki City University, Japan

Abstract

This paper addresses the issue of how to assess learners' engagement with activities designed to develop self-regulatory learning strategies in the context of foreign language teaching and learning. The argument is that, if the aim of these activities is the *development* of learners' self-regulation, then the assessment practices used must also reflect this orientation. The problem herein is that traditional assessment practices are typically normative in nature, endorsing understandings of intelligence as fixed and failure as unacceptable. Using such approaches to assess learner engagement with self-regulated learning activities will undermine efforts to promote learner development, and may demotivate learners. This paper will discuss these issues through a critical reflection on assessment practices used to evaluate EFL learners' engagement with an assessable homework activity designed to develop their self-regulatory strategies. It is argued that learning-oriented assessment principles and practices are most suited to the evaluation of self-regulated learning in EFL. Potential issues related to the application of learning-oriented assessment in EFL contexts are also discussed.

*Keywords*: self-regulated learning, learner development, learning-oriented assessment, normative assessment practices, EFL

Today, there is a growing recognition among educators of all disciplines that one of our most important tasks as teachers is to help our learners learn how to learn. This special issue is testament to interest in this topic among foreign language educators of various backgrounds. Teachers who strongly believe in the importance of developing their learners' ability to self-regulate their learning typically act on these convictions by modifying their everyday teaching practices or introducing interventions that specifically target the development of these learning skills and strategies. The reader can find several examples of such practices in this special issue.

All in all, this is a positive development for the field of foreign language teaching and learning. However, in this paper I would like to draw attention to the little discussed issue of the assessment of learners' self-regulated learning (SRL). While a large amount of thought may go into the design of tasks and activities aiming to develop learners' ability to self-regulate their learning of foreign languages, the overall absence

of discussion on the matter suggests that less consideration is given to how to assess student performance and engagement with these activities. Helping our students develop their ability to learn may appear to be first-most in our minds. However, we must not forget that we are continuously assessing our learners' performance for various purposes: we assess to evaluate the success of our activities and levels of student understanding and interest; we assess to give students feedback on their performance and to give a score or grade to formally indicate level of achievement or engagement. Assessment is an issue for all educational contexts, but especially so in higher education, where formal assessment is omnipresent and grading, evaluation, and certification are inevitably foregrounded (Carless, 2007).

In this climate of ubiquitous assessment of learners and their learning, *how* we assess students is a question that cannot be overlooked. Here, the issue is not the method per se, although this too is important, but the philosophical approach to assessment that informs our practices. Especially in the case of activities or projects that higher education foreign language teachers set for their learners with the dual aims of facilitating learning of the target language and development of generic academic learning skills and strategies, it is crucial that a developmental or learning-oriented approach to assessment be taken. I argue that if a traditional assessment framework that is preoccupied with measurement against normative standards and certification is used, the evaluation of these activities may sabotage the very goals we are aiming to achieve through them (c.f. Benson, 2010; Dam & Legenhausen, 2010; Lamb, 2010).

My concern with the assessment of SRL has come directly from my colleague, Paul Collett, and my own experiences with introducing a self-regulated learning program into a series of English as a Foreign Language (EFL) oral communication courses at a university in Japan, and realizing, in hindsight, how our failure to address assessment principles head-on from the beginning led to the use of normative assessment approaches by classroom teachers in ways that did not necessarily support our objectives. Indeed, it is possible that the assessment approaches used had a negative impact on the students' learning, or, at the very least, their experience with the activity. In this paper, I will critically reflect upon our practices, and the normative approaches to assessment used by classroom teachers, through a discussion of theories of intelligence,

goal orientations, and learning-oriented assessment. The ultimate objective of this paper is to begin theorizing more sound ways to approach the assessment of students' engagement with tasks designed to develop their learning skills and strategies within a foreign language classroom context, and to encourage other classroom practitioners to think more deeply about their own assessment practices, especially in the case of institution-wide programs.

Although this article focuses on self-regulated learning strategy development, I believe the arguments made are also applicable to classroom interventions targeting learner autonomy, independent learning, and self-directed learning, as well as the evaluation of student use of self-access centers.

## Incorporating Self-Regulated Learning into EFL Classes: The Study Progress Guide

In 2009 we introduced a supplementary learning resource named the Study Progress Guide (SPG) into the first and second year oral communication courses on offer at our university. The SPG is linked to the course textbook through the inclusion of can do statements created specifically to outline the language learning goals of each unit. The overall aim of the SPG is to develop learners' ability to self-regulate their learning of English as a foreign language through having them experience a series of activities which require them to plan, monitor and evaluate their learning throughout the semester.

Specifically, the SPG asks learners to set their own goals and learning activities for each unit of work covered in the course textbook. This begins with a self-evaluation and analysis of strengths and weaknesses supported by a series of can do statements written to reflect key language learning points covered in the unit. After choosing the area they want to work on for the unit in question, and outlining a specific study plan or learning activity for this, students put this plan into action, and then reflect on the effectiveness of the activity or strategy used. This is repeated for each unit covered over the semester, and is accompanied by other activities designed to encourage learners to identify and reflect upon their personal goals for the course (see Sullivan & Collett, 2014, for a description of these activities). The majority of this work is conducted

outside of class as a homework activity. One page from the SPG is provided in Appendix A to give readers a general idea of the activities learners complete each unit, and the content of the can do statements. A more thorough description of the SPG can be found in Collett and Sullivan (2010).

For better or worse, the SPG is very much many things at once: a device to introduce self-regulated learning practices to learners, a chance for learners to engage with their language learning in a personalized and self-directed way, and an opportunity to revise class work. It is also important to note that while the SPG includes sections which explain what makes a good goal and directions on how to choose effective learning activities there is no specific instruction on this in class, and this very much compounds the issues discussed in this paper. We are working towards addressing this by incorporating the SPG more into classroom work, thereby creating opportunities for teachers to offer more guidance. Many of the ideas for improving our use of the SPG came from the presentations and subsequent discussions at the *Self-Regulation in Foreign Language Learning: Shared Perspectives* symposium, and we are indebted to all participants for inspiring these changes. See the papers by Hutchinson and Thornton in this issue for more about the role of the teacher and teacher guidance in self-directed learning.

The SPG is currently used in two courses consisting of 19 and 11 classes respectively, which are predominately taught by part-time teachers who were not necessarily involved in the creation of the SPG or the research that informed its development. The students' work on their SPGs accounts for 20 percent of their final grade for the course, and it is the predominately part-time course instructors who are required to evaluate the SPGs and assign a score out of 20, with each unit of work to be generally given a score out of 2. A very basic scoring scale (unattempted—0 points / unacceptable—0.5 points / acceptable—1 point / good—1.5 points / excellent—2 points) is included in table form within the SPG, and teachers are asked to use this to score student work unit by unit and provide feedback; note that only the descriptive evaluation and not the number of points appears in the SPG. Apart from this, other specific theoretical or practical guidelines for approaching the assessment of the SPG are not provided, and thus teachers have been required to draw on their own beliefs

about assessment when evaluating student work. (Obviously, all of this is inherently problematic and we have since introduced bilingual rubrics (see Appendix B) and pre-semester calibration workshops to begin to address these insufficiencies. There is also a concern that this approach to scoring is essentially normative and is thus contributing to the normative assessment practices problematized in this paper. This issue will be addressed in the following sections.)

### Teachers as Assessors: But What Kind of Assessors?

From our earlier interviews with students who had used the SPG, we realized that the classroom teacher can impact students' use and understanding of it (Collett & Sullivan, 2013a, 2013b). Thus, in order to learn more about this to improve the SPG and its use, at the end of the 2012 academic year we conducted an open-ended survey to investigate teacher perceptions of the SPG and student use of it, and to learn about how each individual teacher was actually using the SPG with their students. Responses given in this survey inadvertently revealed trends in teacher assessment practices and underlying assessment philosophies which directly motivated the current study.

A general theme arising from teacher responses was that many students had difficulty using the SPG, specifically with articulating strengths and weaknesses, choosing appropriate goals, and selecting learning activities related to their goals. Using the same learning activity each time was also raised as a common concern. (This is all to be expected, however, as these are precisely the skills that we are presuming most students have not yet fully developed, and this is what we are targeting through the SPG. The extent to which teachers realize and accept this will fundamentally influence their assessment practices, as we shall see in the discussion to come.) While some teachers explained how they tried to provide guidance in choosing more appropriate learning goals and activities, there seemed to be an overall belief, and a general resignation, that there was not much the teacher could do (or should do) for students not showing attempts to engage with the SPG homework. Teachers noted that students who engaged deeply with the SPG were typically students who liked English, and the effort these students made was praised. Teachers commented that it was those students who did not like English, and who would theoretically get the most out of using the SPG,

who used the SPG superficially or frequently forgot to do it. Threats, warnings and punishments were commonly used to coerce these students to engage with the SPG, although some teachers wondered if there was much point to this at all.

Teachers' feedback on the SPG, extracts of which are provided in Table 1, generally categorized learners into two groups: those who made efforts to use the SPG and those who did not or could not. Teachers typically praised the former group and penalized the latter. Threats about failing the course were often used to coerce students lumped into the latter group to use the SPG. In contrast, there were only limited accounts of teachers working with learners' SPG work to identify and discuss particular areas for improvement; it seems that there are very few opportunities for students in the latter group (those students labeled "bad learners") to get the feedback they needed to join the former. In other words, the idea of *developing* learners' skills was not being taken into account during the assessment process. The comments also suggest that teachers are tending to assess the learners as *people* rather than their *work* on the SPG tasks. My argument is that this approach to assessing students' SPG work is not conducive to achieving the development of SRL strategies.

Table 1. Extracts from Teachers' Feedback on the SPG

| Examples of Categorizing and Labeling Learners |
| --- |
| "The students who do use them [the SPG] well are the students who study hard, revise well, and score well on the tests anyway. The weaker students, who I assume are the main target, just never get a handle on how to use them effectively and, sorry to say, don't even desire to." – Teacher 6

"Most of the first year students did the [SPG] work in a timely manner. The second year students were far less timely in completing their work, although they did get it done in the end. First year students and serious students tended to do a much better job. Those who had a less positive attitude towards English did minimal work." – Teacher 4

"Some of the students really seemed to benefit from filling out their SPGs. Of course, there were also those who did very little in them, and did not put forth much effort when they did do something." – Teacher 7 |
| Examples of Teacher Feedback: Praise and Punishment |

| |
|---|
| "I gave bonus points if they [the SPGs] were well done, and tsk tsked or guilted them [the students] when it was late or not done." – Teacher 9 |
| "I gave praise to students who did [the SPG work], and penalized when they didn't." – Teacher 5 |
| Examples of Teacher Coercion |
| "I encouraged students to try different activities [for their SPG homework], even going so far as to warn them that the same activities would result in lower grades." – Teacher 4 |
| "I tried to remind them that failure to do the SPG-related work could actually lead to them failing the class." – Teacher 7 |
| "I just chanted the litany that 'it's part of your grade'." – Teacher 6 |

## Approaches to Assessment: Which Approaches Support SRL Development?

### *Traditional approaches to assessment*

I do not by any means believe that the approaches to assessment and evaluation demonstrated above are unique to these teachers. I think that these are the approaches that the vast majority of teachers take in the vast majority of cases. One could even go so far as to say that these are the approaches we are preconditioned to take within the educational culture that we belong. Moreover, it must be noted that the somewhat normative scoring system we asked teachers to use, combined with our failure to specifically encourage teachers to approach assessment in non-normative ways, no doubt reinforced this.

I believe that this dichotomized view of students as either good or bad, able or not, and the negative appraisal of the so-called "bad learners" is intrinsically linked to teachers' theories of intelligence and the goal orientations they bring into the classroom. There are generally two ways to view intelligence: as something that is fixed ("entity" theories of intelligence) or something that is malleable and can be changed ("incremental" theories of intelligence). Dweck and Master (2008) suggest that both theories are "equally popular" with "about 40% of adults and children endors[ing] an entity theory of intelligence, about 40% endors[ing] an incremental theory, and about 20% … undecided" (p. 32). It is important to recognize that these theories of intelligence "shape students' [and teachers'] goals and values, change the meaning of

failure, and guide responses to difficulty" (Dweck & Master, 2008, p. 32).

Achievement goal orientations are "cognitive representations of positive or negative competence-relevant possibilities that are used to guide behavior" (Fryer & Elliot, 2008, p. 55) and they are closely related to theories of intelligence. Performance-avoidance goals are based upon entity theories of intelligence and are characterized by a fear of failure and a host of other conditions and behaviors that can have a negative impact on students' academic performance and general well-being, such as superficial learning and self-handicapping (Fryer & Elliot, pp. 56-57). In contrast, mastery-approach goals are connected to incremental theories of intelligence and "give rise to positive processes and outcomes" such as intrinsic motivation, enjoyment of the learning process and increased self-regulation (Fryer & Elliot, 2008, p. 56).

Stobart (2014) argues that myths about fixed ability are still widely held in education, despite what we know about learning and the development of expertise. Indeed, the majority of the assessment that occurs in formal education is conducted against normative standards where correctness is praised and failure admonished; in other words traditional approaches to assessment take on a performance goal orientation (Ames, 1992; Tunstall & Gipps, 1996). This is in spite of the fact that there seems to be general agreement that performance-avoidance goal orientations are "highly problematic in achievement situations" (Fryer & Elliot, p. 56) and "should be discouraged at all costs" (p. 57).

### *Learning-oriented assessment*

There are ways to approach assessment that take on mastery goal orientations. This type of assessment is descriptive rather than evaluative. It provides feedback that specifies standards, areas of achievement, areas in need of improvement, and strategies to achieve this, while increasingly engaging the learner in a dialogue with the teacher about their learning, thus moving the responsibility for learning incrementally toward the learner (Tunstall & Gipps, 1996). As such, participation in this type of assessment helps learners to develop the skills necessary to become able to evaluate their own learning—i.e. it in itself contributes to the development of self-regulatory strategies.

Many terms are in use to refer to assessment that prioritizes and supports

learning over other functions, such as measurement and certification. I prefer the term learning-oriented assessment, which others have suggested helps to avoid conflicting definitions of formative assessment (Carless, 2007). Although learning-oriented assessment, formative assessment, dynamic assessment, and other variously termed non-normative approaches to evaluation have been applied in foreign language settings, studies into these applications tend to have a greater focus on the non-gradable evaluation of learner production of language during class activities, rather than on gradable assessment procedures (c.f. McNamara, 2014; Norris, 2014). This could perhaps be related to the fact that most of these studies have been carried out in primary and secondary school-based contexts, rather than within tertiary education settings where assessment takes on different purposes. This is not to say that these approaches to classroom-based assessment are not important; just that they do not offer much explicit guidance for those dealing with assessable tasks. (See the paper by Wilson in this issue for an example of good practice in classroom-based formative assessment.)

However, there is much discussion within the field of general higher education regarding sound assessment practices for supporting learning through actual assessment tasks. Moreover, there is a clear message from this growing body of work that these alternative approaches to assessment are very much geared towards the development of self-regulated learning strategies (c.f. Clark, 2012; Nicol & Macfarlane-Dick, 2006).

Here, I would like to introduce just two examples of work in this area to give the reader an indication of the suggestions being made regarding learning-oriented assessment in higher education. Various principles important for aligning assessment practices with learning have been suggested in the higher education literature. Carless (2007) argues that assessment that is learning-oriented needs to incorporate three interconnected strands or principles (pp. 59-60):

Principle 1. Assessment tasks should be designed to stimulate sound learning practices amongst students.

Principle 2. Assessment should involve students actively engaging with criteria, quality, and their own and/or peers' performance.

Principle 3. Feedback should be timely and forward-looking so as to support

current and future student learning.

Nicol and Macfarlane-Dick (2006) argue that good feedback practices support the development of learners' ability to self-regulate their learning. They offer seven principles of good feedback.

Good feedback practice:

1. helps clarify what good performance is (goals, criteria, expected standards);

2. facilitates the development of self-assessment (reflection) in learning;

3. delivers high quality information to students about their learning;

4. encourages teacher and peer dialogue around learning;

5. encourages positive motivational beliefs and self-esteem;

6. provides opportunities to close the gap between current and desired performance;

7. provides information that can be used to help shape teaching (p. 205).

I believe that incorporating these approaches into assessment will allow us to provide all learners, put simply in our case, those who are already self-regulating and those who are not yet self-regulating their learning of English as a foreign language, with the support they need to develop their abilities and move towards achieving their potential. (See the paper by O'Dwyer and Runnels in this issue for an example of learning-oriented assessment principles being applied in a process writing class. Also see Sullivan (forthcoming) for a description of learning-oriented assessment in a TOEFL preparation course.)

**Assessment of Self-Regulated Learning: Why Does it Need Special Consideration?**

Many would argue that all assessment should be conducted according to the principles of learning-oriented assessment. So, why is it of particular import when we talk about the assessment of student work on self-regulated learning tasks?

Firstly, underpinning self-regulated learning are theories of intelligence and ability as malleable and not fixed, so not engaging with learners who are not yet self-

regulating, and inadvertently labeling them as "bad learners", is antithetical to the aims of SRL and thus any classroom-based practices which attempt to develop SRL strategies. Secondly, when feedback is limited to "praise and punishment" it does not provide information to help students learn, which is a key aim of incorporating SRL practices into the language classroom. It also neglects the fact that the development of emerging ability requires mediation—through scaffolding, feedback, and modeling—from the teacher or more advanced peers (Lantolf & Poehner, 2011). This "co-regulation" of learning is something that our student interview participants specifically cited as being important for being able to effectively use the SPG (Collett & Sullivan, 2013a). One can easily envisage how the application of theories of intelligence as fixed and the limited provision of feedback feed into each other to negatively impact learners' motivation and self-esteem (Ames, 1992; Nicol & Macfarlane-Dick, 2006), which may cause learners to take on performance rather than mastery goals, setting off a vicious cycle.

***Bad assessment practices have the potential to harm: Examples***

Ames (1992, p. 264) argues that "the ways in which students are evaluated is one of the most salient classroom factors that can affect student motivation." Indeed, a major concern with the traditional or normative assessment of SRL tasks is that it will not only sabotage the development of SRL skills and strategies, but also demotivate learners.

One example that we encountered in student interviews was of Learner A, a first year male student who had just finished using the SPG for one year (see also Collett & Sullivan, 2013a). Learner A startled us with his eloquent theorization of the SPG as being purely a tool to assess students' participation and engagement with the course to help the class teacher compute a final grade. He was also adamant at the onset of the interview that we, the interviewers, must think of him as a "bad learner" as he was not using the SPG in the way that he presumed was being expected. We have no specific evidence to prove how Learner A formed this view of himself in relation to his use of the SPG, or how he came to see the SPG as an evaluation tool. However, it would not be too far a stretch to imagine this being at least partly related to feedback received from

a teacher during the course. Teachers must always remember that learners pick up and internalize messages (inadvertently) embedded in teacher evaluations in ways that we may never imagine (Farrell, 2014).

Teacher assessment of student work also has the potential to harm if it disparages examples of work that the student believes to be useful for them, without offering more appropriate ideas to guide student learning. Teacher assessment relies much on teacher perceptions of student effort and thinking, which are made based on subjective and simplistic criteria, without really taking into account the thought-processes and decision-making that occurs in the minds of learners and thus remains unseen. In the case of the SPG, sections that have not been filled in or completed, the choice of "simplistic" learning activities, and repeatedly using the same learning activity are some of the common teacher "warning signs" of superficial work. However, our students may not necessarily agree.

In our interviews, another first year student, Learner B, explained how she used different learning activities depending on whether she was working on a strong point or weak point. She explained how she would sometimes personalize the target structures by using them to write about her own experiences, but when the content became difficult she would write out the structures again and again in her SPG to try to remember them. When the content became even more difficult for her in the second semester, she said: "I felt that the only thing I could do was write out the bits that I couldn't understand, and try to remember them that way. Because my study approach was the same each time, it was annoying to have to write the same explanation for each unit, so I abbreviated that section [of the SPG]."

Just as teachers have their opinions about the value and effectiveness of learning activities, so do learners. The fact that learners may be choosing learning activities based on their perception of the difficulty of the learning point, and their past experiences of success with the activity, is often not registered by teachers. When judgments are made without dialogue, the teacher may only see the once diligent student who has started to cut corners, instead of the learner who is deeply thinking about and engaging with the activity in a way that she feels best reflects her current needs and goals—whether her choices are best for her or not is another question. If this

learner is clumped into the category of "bad learner" and punished for work that she is proud of, one can only imagine the confusion and demotivation that may follow.

As Ames (1992) has argued, the instructional practices informing task selection and evaluation practices "need to be coordinated, and … directed toward the same mastery goal" (p. 266). Incorporating SRL principles within learning tasks but not assessment tasks is not only ineffective for achieving educational objectives, but potentially detrimental in terms of learner motivation.

## Shifting from a Traditional to a Learning-Oriented Assessment Approach: In Practice

My aim through the discussion until now has been to show that applying traditional or normative assessment approaches literally sabotages teacher attempts to develop learners' self-regulated learning strategies. If teachers are incorporating SRL practices within their own courses, they must be aware that both the tasks they set and the methods they use to evaluate student performance need to be oriented towards learner *development*. Especially in situations like ours where there is a disconnection between curriculum development and instruction, and instruction is mainly undertaken by part-time teaching staff, building a shared assessment discourse is a crucial first step. It is unclear how widely the concept of learning-oriented assessment is known and understood, and whether it would be readily accepted by teachers so accustomed to working within a normative assessment framework. Influencing teacher beliefs about teaching (and assessment) practices is not easy (c.f. Borg, 2003) but it is crucial if SRL-based activities are going to work.

In addition to considering how to introduce ideas about alternative assessment practices to teaching staff, consequent practical issues related to such a change in assessment approaches will also need to be given consideration. Learning-oriented assessment calls for the provision of teacher feedback and chances for student-teacher dialogue. However, in cases such as ours where the majority of instructors are part-time staff who do not share the same L1 as their learners, the perpetual issues of time, language and space are obstacles to achieving this.

This could also generate inequality in assessment opportunities. Providing

opportunities for students to "complete the feedback loop" or "close the performance gap" through the resubmission of work based on teacher feedback is a typical learning-oriented assessment practice (Nicol & Macfarlane-Dick, 2006). However, given the obstacles introduced above, and other issues related to teacher-student relations, there is a concern that teachers may not be able to provide the same levels of quality feedback and the same chances for re-doing tasks equally to all students. These kinds of practical issues will also need to be given attention.

Finally, I believe there is a need to develop non-normative methods and tools of scoring student work. Even if we wish to conduct assessment using non-normative approaches, there is the possibility that we will slip into normative practices if we use traditional grading methods without some form of modification. One idea could be examining more specific methods of incorporating the appraisal of learner development over a course of work within grading systems and rubrics.

This paper by no means provides any answers. However, I hope it has drawn attention to the importance of the nature of assessment in self-regulated learning activities. Work conducted in general higher education is providing many practical ideas which should now be tried out in tertiary-level foreign language courses.

### Acknowledgments

### Notes on the contributor

Kristen Sullivan is a lecturer at Shimonoseki City University and co-writer of *Impact Conversation 1 & 2* (Pearson Longman Asia ELT). She is interested in the teaching, learning and assessment of speaking, as well as interactions between language learner identity and language use.

### References

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, *84*(3), 261-271. doi:10.1037/0022-0663.84.3.261

Benson, P. (2010). Measuring autonomy: Should we put our ability to the test? In A. Paran & L. Sercu (Eds.), *Testing the untestable in language education* (pp. 77-97). Bristol, UK: Multilingual Matters.

Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, *36*, 81-109. doi:10.1017/S0261444803001903

Carless, D. (2007). Learning-oriented assessment: Conceptual bases and practical implications. *Innovations in Education and Teaching International*, *44*(1), 57-66. doi:10.1080/14703290601081332

Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review*, *24*, 205-249. doi:10.1007/s10648-011-9191-6

Collett, P., & Sullivan, K. (2010). Considering the use of can do statements to develop learners' self-regulative and metacognitive strategies. In M. G. Schmidt, N. Naganuma, F. O'Dwyer, A. Imig & K. Sakai (Eds.) *Can do statements in language education in Japan and beyond: Applications of the CEFR* (pp. 167-183). Tokyo, Japan: Asahi Press.

Collett, P., & Sullivan, K. (2013a). Social discourses as moderators of self-regulation. In N. Sonda & A. Krause (Eds.), *JALT2012 Conference Proceedings* (pp. 255-265). Tokyo, Japan: JALT.

Collett, P., & Sullivan, K. (2013b). The social mediation of self-regulated learning. In M. Hobbs & K. Dofs (Eds.), *ILAC Selections 5th Independent Learning Association Conference 2012* (pp. 119-120). Christchurch, New Zealand: Independent Learning Association.

Dam, L., & Legenhausen, L. (2010). Learners reflecting on learning: Evaluation versus testing in autonomous language learning. In A. Paran & L. Sercu (Eds.), *Testing the untestable in language education* (pp. 120-139). Bristol, UK: Multilingual Matters.

Dweck, C. S., & Master, A. (2008). Self-theories motivate self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 31-51). Mahwah, NJ: Erlbaum.

Farrell, T. (2014, November). *Reflecting on practice*. Paper presented at the Japan Association for Language Teaching 40th Annual International Conference on Language Teaching and Learning, Kobe, Japan.

Fryer, J. W., & Elliot, A. J. (2008). Self-regulation of achievement goal pursuit. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 53-75). Mahwah, NJ: Erlbaum.

Lamb, T. (2010). Assessment of autonomy or assessment for autonomy? Evaluating learner autonomy for formative purposes. In A. Paran & L. Sercu (Eds.), *Testing the untestable in language education* (pp. 98-119). Bristol, UK:

Multilingual Matters.

Lantolf, J., & Poehner, M. (2011). Dynamic assessment in the classroom: Vygotskian praxis for second language development. *Language Teaching Research*, *15*(1), 11-33. doi:10.1177/1362168810383328

McNamara, T. (2014, October). *Discussant (2)*. Paper presented at the 3$^{rd}$ Teachers College Columbia University Roundtable in Second Language Studies, Roundtable on Learning-Oriented Assessment in Language Classrooms and Large-Scale Assessment Contexts, Teachers College, Columbia University, New York, NY.

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

Norris, J. (2014, October). *Discussant (1)*. Paper presented at the 3$^{rd}$ Teachers College Columbia University Roundtable in Second Language Studies, Roundtable on Learning-Oriented Assessment in Language Classrooms and Large-Scale Assessment Contexts, Teachers College, Columbia University, New York, NY.

Stobart, G. (2014). *The expert learner: Challenging the myth of ability*. Berkshire, UK: Open University Press.

Sullivan, K. (forthcoming). Test re-dos for supporting learner reflection and development. In G. Brooks & M. Grogan (Eds.), *The 2014 PanSIG Conference Proceedings*. Tokyo, Japan: JALT PanSIG.

Sullivan, K., & Collett, P. (2014). Exploiting memories to inspire learning. In N. Sonda & A. Krause (Eds.), *JALT2013 Conference Proceedings* (pp. 375-382). Tokyo, Japan: JALT.

Tunstall, P., & Gipps, C. (1996). Teacher feedback to young children in formative assessment: A typology. *British Educational Research Journal*, *22*(4), 389-404. doi: 10.1080/0141192960220402

**Appendices**

*Appendix A*
An Extract from the Study Progress Guide

*Appendix B*
Bilingual Grading Rubrics for the Study Progress Guide