Learner Autonomy, Self-Assessment, and Goal-Setting: The Accuracy of Learner Self-Assessments in Discussion Classes
Daniel Warchulski

ABSTRACT
The use of self-assessment and goal-setting activities is widely endorsed and used in language classrooms to promote autonomous learning. Despite the potential benefits of such activities, issues surrounding accuracy and reliability remain a concern. This study attempts to measure the accuracy of student performance with regards to function and communication skills use as compared to the instructor’s perception. The results indicate a strong correlation between student and instructor assessment scores. And although there was a difference between the participants’ and teacher’s overall perceptions, the students were relatively accurate in following the assessment patterns of the instructor. Overall, the participants’ view of the activity is positive and they regard self-assessments to be useful in helping them to set goals and to identify strengths and weaknesses.

INTRODUCTION
It is widely accepted that the successful implementation of communicative language teaching ought to include the promotion of learner autonomy since it is believed that this gives students more choices in the learning process and places an expectation on learners to take on a greater degree of responsibility for their own learning (Richards, 2006). This can include the routine use of self-assessment and goal-setting activities as part of an instructor’s teaching practices to encourage students to act more autonomously in the language learning process. Furthermore, the use of such activities can strategically be used to increase and sustain students’ motivation levels (Dörnyei, 2001). Given the potential role of self-assessment/goal-setting activities in supporting autonomous learning and enhancing motivation, this study is concerned with the ability of first year university students, enrolled in compulsory English discussion classes at a Japanese university, to assess their performance and to set goals using self-assessment/goal-setting worksheets as part of regular classroom activities underlying an ongoing feedback process (see Appendix).

The participants in this study were students enrolled in compulsory English discussion classes in the fall semester, consisting of fourteen 90-minute lessons. The English Discussion Course follows a unified curriculum that tries to ensure a consistent and equal learning experience for all students taking part in the classes. The aim of the curriculum and lessons is to improve learners’ communicative competence whereby lessons are structured around functional language. Specifically, students are required to use functions (i.e. changing topic, paraphrasing, balancing opinions, etc.) to express and share their ideas during 10 and 16 minute discussions, and at other stages of lessons, such as when doing various practice and preparatory activities. Additionally, students are required to use various communication skills. These include the ability to give reactions, agree or disagree with ideas, ask follow-up questions, and negotiate meaning. Both function and communication skills use provide structure during lessons and activities while enabling students to effectively participate in group discussions. Accordingly, instructors monitor, assess, and provide feedback to students regarding their ability to use functions and communication skills.

The notion that autonomy is an important component of successful language learning seems to be a widely agreed upon belief and a key feature of current pedagogical practices. The
introduction of autonomy as an indispensable feature of language education began with Holec (1981), who defined autonomy as “the ability to take charge of one’s own learning” (p. 3). Although numerous other definitions and theoretical frameworks have been proposed since Holec’s seminal definition, most of them are in some way grounded in or tied to Holec’s conception. An inherent feature of Holec’s theoretical framework entails giving more responsibility to learners in the management of various aspects of the learning process, including the monitoring and assessment of their progress and the setting of goals which require students to act in a self-deterministic manner.

Although there appears to be a general consensus among researchers that autonomous learners are those who understand the purposes of their learning, accept responsibility for it, share in or contribute to the setting of their learning objectives, implement appropriate learning strategies, and regularly evaluate their progress (Cotterall, 2000; Dickinson, 1995), some research suggests that there is a considerable degree of variation among educators and researchers in defining the precise nature of autonomy and difficulties in operationalizing the notion exist (Benson, 2006; Little, 2002). Accordingly, in implementing self-assessment and goal-setting activities in the classroom, it is important for educators to be aware of these findings, including some of the potential difficulties.

In recent years, new directions in the practice and growth of autonomy in language teaching and learning have led to the emergence of a variety of new views and theoretical frameworks in the area (Benson, 2006). These, in turn, have provided educators with a greater understanding and wider range of choices regarding the manner in which to promote and implement autonomy in their pedagogical practices. Additionally, since all institutional and learning environments, including the current one, are different and unique, promoting and viewing autonomy through an appropriate theoretical framework can have important implications.

Among some of the noteworthy concepts, theories, and models pertaining to autonomy that have emerged are notions of varying levels or degrees of autonomy, different versions of autonomy, and cultural variation in autonomous teaching and learning. For instance, there is a distinction in the literature between strong versions and weaker forms of autonomy. Here, Smith (2003) points out that strong pedagogies view autonomy as a trait that learners already possess with the focus being the co-creation of conditions that allow learners to exercise this autonomy, whereas weaker ones are premised on the idea that autonomy is something that students lack in some manner and need training in. Moreover, stronger versions tend to be associated with some of the earlier models, with the weaker forms having emerged more recently as a response and attempt to address some of the deficiencies of earlier conceptions by accounting for variation among learners and educational contexts (Benson, 2006).

With respect to varying levels of autonomy, Littlewood (1996) suggests that a student’s learning environment and cultural context can be key factors in the practice and promotion of autonomy. Within this framework, it is suggested that three main domains – autonomy as a learner, autonomy as a communicator, and autonomy as a person – are utilized as a basis for developing broad approaches which can be further broken down into specific practical strategies that allow educators to employ both stronger and weaker forms of autonomy.

Additionally, Littlewood (1999) proposes that educators in different contexts can decide what the optimal degree of learner freedom should be by referring to a distinction between “proactive” and “reactive” autonomy. The concept of proactive autonomy puts an expectation on learners to be actively engaged in all aspects of the learning process and often requires a radical change on the part of teachers and students. Conversely, reactive autonomy is seen as being more appropriate in many Asian educational settings, such as Japan, since it doesn’t necessarily create
its own direction and allows for a gradual and culturally-sensitive form of autonomy. Furthermore, reactive autonomy can be considered to be a preliminary step towards the proactive form or a goal in its own right since, “once a direction has been initiated, it enables learners to organize their resources autonomously in order to reach their goal” (p. 75). Accordingly, in designing and implementing the self-assessment/goal-setting activity, a weaker reactive model of autonomy was utilized since this was seen to be more institutionally and culturally appropriate. For instance, whereas a stronger proactive theoretical model may require students to make contributions in most aspects of their education, such as in the selection of course objectives and materials used, in the present study, a weaker reactive framework permitted course goals, particularly with respect to required language use, to be pre-selected by the instructor as per institutional requirements.

According to research on the topic, some of the potential benefits derived from self-assessment/goal-setting activities include increased productivity and autonomy, a reduction in frustration, the promotion of active learning, higher motivation levels, an awareness and perception of progress, and opportunities for the following: individualization, reflection, evaluation, and support (Saito, 2009; Harris, 1997; Rivers, 2001; Gardner, 2000). The use of self-assessments is also considered to be a vital learning strategy for autonomous language learning since it is widely accepted that it enables students to reflect on and monitor their progress, thereby playing a role in the process of identification of individual needs and setting of learning goals (Harris, 1997; Reinders, 2010). While the findings of existing research make the use of self-assessment activities sound attractive, issues pertaining to reliability and validity remain and need to be addressed since learners are not always capable of accurately assessing themselves (Saito, 2009). Of equal concern is that although self-assessment is widely discussed in the literature, research specifically comparing students’ self-assessment with that of an instructor is somewhat scarce (Stauffer, 2011). Furthermore, the results of the studies and literature on student and teacher comparisons of assessment are mixed with some research reporting moderate or high correlations and degrees of accuracy and usefulness, while others reporting low and insignificant levels of accuracy and utility (Gardner, 2000; Ross, 1998; Blanche & Merino, 1989; Matsuno, 2009).

Given the lack of research on the topic and the potential significance of accuracy in assessing one’s performance and abilities, this study seeks to address the following question: How accurate are students with respect to assessing their performance when compared to the instructor’s perception of their performance

This research is also concerned with students’ views of the self-assessment/goal-setting activity in enabling them to identify their strengths, weaknesses, and the setting of goals with respect to their language use. It is suggested that if students are able to assess themselves in an accurate and consistent manner, they should be able to identify their strengths and weaknesses and as such, they will likely perceive the self-assessment/goal-setting activity to be beneficial in assisting them to set goals. In other words, it is hoped that through the implementation of this activity, students will be more aware and better prepared to set meaningful goals.

It should also be noted that since self-assessment and goal-setting activities can be used for a wide range of purposes, the potential benefits from such activities depends on the reason for and the manner in which they are implemented. For instance, research has attempted to define different types of assessments in relation to their purpose. Generally, two purposes have been identified: 1) performance-oriented self-assessments, and 2) development-oriented self-assessments (Oscareson, 1989). A major distinction between the two is that whereas development-oriented assessments are concerned with the process of learning by incorporating self-managed activities and assess participants over an extended period of time to observe patterns of development and changes, performance-oriented assessments tend to sample performance at one
particular point in time and are concerned with measuring outcomes related to selection, placement, diagnosis, and so on (Saito, 2009).

In this study, a development-oriented self-assessment is utilized. And while it is hoped that the participants will benefit from many of the advantages of self-assessment/goal-setting activities, such as increased feelings of autonomy, higher levels of motivation, an understanding of lesson and course objectives, and the ability to perceive their progress, the primary purpose of this study is to measure the students’ accuracy of their perceived performance over a set period of time, as compared with the instructor’s. Accordingly, by being involved in this process, learners are required to engage in a form of self-monitoring and reflection, exposing them to a type of individualized student-generated feedback which will allow them to identify their strengths and weaknesses, thereby helping them to set achievable goals.

METHOD
The participants in the study were all first year students from various majors, including law, economics, tourism, sociology, community welfare, psychology, literature, and business. The classes were all mixed, comprising of at least one male and one female, and ranged in size from six to nine students. Language proficiencies of the participants also varied. The number of students that took part in the self-assessment/goal-setting activity and in the study totaled 97 (n = 97). The participants represented a fairly diverse group of learners with different language proficiencies and abilities, majoring in a wide range of subjects.

Although the activity was implemented beginning in Lesson 2 and ending in Lesson 12 (i.e. Lessons 2, 3, 4, 6, 7, 8, 10, 11, 12), the first three weeks were treated as a piloting period where the focus was on students’ understanding of and to do the activity. As such, the study commenced in Lesson 6 and data was collected from six lessons. Additionally, a short questionnaire was administered in the final lesson.

Harris (1997) suggests that to be effective, self-assessment must be practical in terms of time and should be integrated with everyday classroom activities. As such, the self-assessment/goal-setting activity was used as part of the regular procedure during the feedback stages of discussions one and two. In creating an assessment activity, to optimize its benefits Gardner and Miller (1999) suggest that they contain the following: the purpose of the assessment, benefits to the students, a procedure for conducting and marking it, a suggested marking scale, and a choice of follow up actions related to the score achieved. Accordingly, in both the creation and implementation of the self-assessment/goal-setting activity, the instructor ensured that these criteria were taken into consideration and included in the activity. Furthermore, the self-assessment/goal-setting worksheets and their implementation was based on a weak reactive model of autonomy whereby the educational context was taken into account, allowing for a lower degree of learner freedom than a strong proactive version would permit.

The instructor created weekly self-assessment/goal-setting worksheets that reflected the new language items for that particular week (as per curricular requirements). Before beginning the activity and during the assessment portion of the feedback process, the instructor provided guidance, clarification, and advice, as necessary, regarding evaluation. As part of a student-generated and learner-centered feedback process, participants were asked to reflect on and assess their performance after Discussions 1 and 2. At the same time, the instructor assessed the students’ performance. Specifically, participants were asked to score both individual items and their overall performance pertaining to function and communication skills use. Based on their assessment, participants then chose one or two goals to complete later in the lesson or in the following class. Examples of student goals included the following: Ask more follow-up questions, Ask others to
balance opinions more, Paraphrase more, and so on.

For the purposes of this study, assessment means either the students’ or instructor’s perception and evaluation of student performance with respect to function and/or communication skills use. Both students and the instructor used assessment criteria consisting of the following: 0 – did not use, 1 – rarely used, 2 – uneven use, 3 – good use, and 4 – superior use. Overall scores were used for comparative purposes in order to determine the relationship and accuracy of students’ assessments as compared with those of the instructor.

In the final lesson a short questionnaire was administered to provide information about students’ perceptions regarding the activity. As per practice in the field, the questionnaire was translated from English to Japanese and a back translation was performed to help ensure accuracy of the items (Dörnyei & Taguchi, 2010) and to make it easier for the students to complete. A 4-point Likert scale, ranging from 1 – strongly disagree to 4 – strongly agree, was utilized to measure student responses.

In addressing the primary focus of this research, a number of statistical measures – both descriptive and inferential - were performed to provide a variety of information about the data. During each lesson, students learned and were required to use either a new function or communication skill and were also expected to use previously learned ones. This was the focus of the assessment and student-centered feedback process. A Pearson product-moment correlation was performed to evaluate the relationship between students’ assessment of their scores and that of the instructor. As well, a paired-samples t-test was carried out to compare assessment scores for function and communications skills between students’ assessment scores and the instructor’s assessment scores. Finally, descriptive statistical measures – namely, the mean and range of responses – were used to summarize the student questionnaire, providing some valuable insight regarding students’ perception of the activity.

RESULTS
A Pearson product-moment correlation coefficient was computed to assess the relationship between the students’ and instructor’s overall assessment of student performance with respect to function use. There was a positive correlation between the two variables, $r = 0.87$, $p < 0.000$. Overall, there was a strong, positive correlation between the manner in which the students assessed themselves and the instructor’s assessment of student performance. Over the course of the study, the pattern of students’ self-assessment was similar to the instructor’s with increases and decreases in students’ scores being highly correlated with increases and decreases in the instructor’s scores. Whenever the instructor rated the learners’ performance well, the students did so too. Accordingly, when the teacher believed that the students used each lesson’s target language for function use well, the students also appeared to share a similar perception. Figure 1 summarizes this finding.

With respect to the evaluation between the participants’ and instructor’s assessment of target language use for communication skills, the relationship between the two variables was positive, $r = 0.61$, $p < 0.000$. On the whole, the relationship between the students’ assessment of their communication skills use and the instructor’s evaluation of their performance showed a strong, positive correlation. As such, whenever the participants viewed and positively evaluated their performance for target language use of communication skills, the instructor also perceived and evaluated the students in a similar manner. Figure 2 summarizes this result.

A paired-samples t-test was conducted to compare both function use and communication skills use scores in two conditions: 1) Students’ self-assessment, and 2) Instructor’s assessment. For function use, there was a significant difference in the scores for student assessments ($M = 2.75$, $SD = .63$) and the instructor’s assessments ($M = 3.20$, $SD = .45$); $t (94) = 13.16$, $p \leq 0.00$. 

261
Likewise, for communication skills use, there was also a significant difference in the scores for student evaluations ($M=2.78$, $SD=.53$) and those of the instructor ($M=3.35$, $SD=.43$); $t(94)=12.80$, $p \leq 0.00$. These results suggest that although the students’ and instructor’s evaluations were highly correlated for both function and communication skills use, there were differences in actual scores. Specifically, when evaluating language use, the students tended to consistently underestimate their performance and scores as compared with the instructor’s assessment. Moreover, when assessing function and communication skills use, the instructor perceived the students’ performance to be better than the students’ evaluation of themselves.

**Figure 1.** Comparison of average student/instructor function use assessment scores.

**Figure 2.** Comparison of average student/instructor communication skills use assessment scores.
Overall, questionnaire responses suggest that the students’ perception of the activity was very positive. It should be noted that in the final lesson the attendance was lower than usual. As such, only 73 students (n = 73) completed the survey. Figure 3 summarizes the results of participants’ views. There was relatively strong agreement among students that the self-assessment/goal-setting activity helped them to identify their strengths and weaknesses. Also, there was fairly strong agreement that the activity helped participants to set goals based on their assessment.

<table>
<thead>
<tr>
<th>Question</th>
<th>M</th>
<th>SD</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The activity helped me to identify my strengths.</td>
<td>3.30</td>
<td>.54</td>
<td>73</td>
</tr>
<tr>
<td>2. The activity helped me to identify my weaknesses.</td>
<td>3.58</td>
<td>.52</td>
<td>73</td>
</tr>
<tr>
<td>3. The activity helped me to set goals.</td>
<td>3.45</td>
<td>.53</td>
<td>73</td>
</tr>
</tbody>
</table>

*Figure 3. Descriptive statistics of questionnaire responses.*

**DISCUSSION**

With regards to the central focus of this research, the findings were quite positive. Overall, the data suggests that the participants were relatively accurate in their assessments, but tended to underestimate their scores as compared to the instructors’ evaluation. Pertaining to both function and communication skills use, a strong correlation was observed between the participants’ perceptions of performance and those of the instructor. So while the results also suggest that the students were not always necessarily accurate in evaluating themselves since they underrated their performance with regards to language use, they consistently followed the assessment patterns of the instructor.

Previous studies concerned with learners’ evaluative practices produced a wide range of results regarding students’ abilities to accurately assess their language abilities and performance. For instance, some research has shown moderate to strong correlations, concluding that learners are generally accurate in their evaluations and that self-assessment is beneficial and can be reliable (Tavakoli, 2010; Bachman & Palmer, 1989; Blanche, 1990). Conversely, other studies have found considerable divergence in accuracy ratings, a lack of reliability, and other issues (White, 2009; Matsuno, 2009; Thomson, 1996; Pierce, Swain, & Hart, 1993). The variability in findings can be difficult to account for since the wide range of variables, such as variation in sample size, age, cultural backgrounds, target language, language skills being tested, and environmental factors, may affect reliability (Gardner, 2000). In spite of this diversity in results, many of the studies, including those where accuracy and correlations of participants’ assessments were poor, point to positive aspects of using self-assessments for a variety of other beneficial purposes. In other words, the use of self-assessment activities is considered to hold many advantages despite students’ deficiencies in their ability to accurately assess their language skills.

Not surprisingly, like other research, this study produced mixed results. There was a strong correlation in terms of how the participants assessed themselves, yet there was a statistically significant difference in scores between the students and the instructor. Like other research on the topic, the instructor believed the self-assessment/goal-setting activity was positive and valuable in a number of ways. For instance, the students found the activity very useful in helping them to identify their strengths and weaknesses in function and communication skills use. Moreover, students felt that by assessing their performance, they were more prepared to and better able to set goals. As such, it is likely that this had the effect of promoting reflection and acting as an awareness or conscious-raising tool. The strong correlations in assessments suggest that the
students appeared to be self-aware, but perhaps a little modest in their scoring. This may possibly be the result of cultural or other factors. Future research on the topic could perhaps try to address the potential reasons for this.

Other ways the students may have benefitted from this activity include an increase in their motivation and participation, enhanced opportunities to act autonomously, and a greater understanding of course objectives. For instance, based on observations of the participants, it seemed that students were motivated and this motivation was maintained throughout the course. Additionally, it appeared as though students were able to set and achieve their goals, thereby improving their overall performance in subsequent discussions. It has been suggested that combining self-assessment with teacher-fronted feedback means that the latter is likely to be more effective (Harris, 1997). As such, instructor-fronted feedback seemed to complement and help reinforce students’ assessment and choice of goals.

Improving and ensuring that learners are accurate in their self-assessments can be difficult and challenging and in the present study, a desirable outcome would have been an increase in students’ perceptions of their performance, and thus scores. To improve reliability, instructors can help by making sure that students receive appropriate training and guidance from teachers in evaluating themselves, such as through awareness-raising activities (Harris, 1997; White, 2009). Additionally, it has been suggested that for students to be able to realistically self-assess their progress in language learning, teachers must carefully monitor both performance and progress, while providing constant feedback (Blue, 1994). Potentially, this has some important implications for classroom practices. For instance, to improve accuracy, instructors need to be very clear in terms of what exactly is being assessed and how. As well, teachers can ensure greater awareness of progress and continuity from lesson to lesson by making sure that students know and remember their scores and how these compare with the instructor’s evaluation. Improvements in learners’ accuracy ratings, while not necessarily providing an alternative to teacher grading, can be an extremely useful complement to it (Harris, 1997).

So while this study provided some valuable insights about students’ ability to self-assess their performance with respect to language use, there are a number of issues that can be addressed in future research. Perhaps most importantly, a deeper and more comprehensive analysis can be performed, such as through the use of a mixed methods research approach or by analyzing the data in greater detail. For example, an analysis and comparison of specific goals or individual item scores for each category can yield more insights about various relationships, such as whether students’ goals are accurate or the nature of the relationship between goals and scores for individual items.

**CONCLUSION**

The use of various activities to promote autonomous language learning in the language classroom is now commonplace. Autonomy in language teaching and learning is now endorsed by most researchers, teachers, and others. Implementing self-assessment and goal-setting activities hold great potential in enabling learners to act in a self-deterministic manner, thereby providing opportunities for autonomy. However, as has been found in this study, some potential issues exist. So while in this study the participants’ assessment patterns were strongly correlated with the instructor’s evaluations, the students tended to underestimate their performance. It would be interesting and useful to determine the cause of this and future research can address such issues.

Still, according to other research and studies, students are not always capable of accurately evaluating their performance. Despite this, given the wide range of benefits reported with respect to self-assessment and goal-setting activities, concerns over accuracy are unlikely to
negate their usefulness and utility. As such, educators would be well-advised to consider implementing self-assessment and goal-setting activities in their teaching practices.

REFERENCES


**APPENDIX**

**SELF-ASSESSMENT (Lesson 6)**

Please assess your performance in Discussions 1 and 2 and your overall performance according to the following grading criteria:

0 — Did not use.
1 — Rarely use.
2 — Unwilling use.
3 — Good use.
4 — Superior use.

**FUNCTIONS**

**Asking Others to Balance Opinions (this week)** What are the (dis)advantages of...? Are there any other (dis)advantages of...?

**Balancing Your Opinion (this week)** One (dis)advantage of... is...

Another (dis)advantage of... is...

**Paraphrasing Others/Yourself** Do you mean...? So, are you saying...? So, in other words...? I mean... What I'm saying... In other words...

**Changing Topic — Choosing a Topic/Closing a Topic** What shall we discuss first/next? Why don’t we discuss (TOPIC)? Is there anything to add? So, we agree/disagree about (TOPIC)

**Overall function use:** (Overall)

**COMMUNICATION SKILLS**

**Reactions** (Okay, I see. Yeah. Right. Uh-huh. Wow! That's great.)

**Follow-up Questions** (What...? What...? When...? Why...? Where...? Do...? Can...? Have you...?)

**Checking Understanding** (Do you understand? Sorry, I don’t understand...)

**Overall communication skills use:** (Overall)

I did the following well in Discussion 1: ____________________________

My goal(s) for Discussion 2: ____________________________

I did the following well in Discussion 2: ____________________________

My goal(s) for next class: ____________________________