Self-assessment and Learner Strategy Training in a Coordinated Program: Using Student and Teacher Feedback to Inform Curriculum Design

Heidi Evans Nachi

Kwansei Gakuin University,
School of Policy Studies

heidi@ksc.kwansei.ac.jp

This chapter investigates student and teacher attitudes regarding self-assessment instruments and procedures in one university English Language Program (ELP). Students in the final semester of a required four-semester ELP responded to a survey about self-assessment, and teachers answered a similar survey. Descriptive statistical analyses of responses and a qualitative analysis of open-ended question responses reveal some differences in attitudes among students; moreover, a comparison of student and teacher responses reveals further differences in relation to materials design and motivation. These results may have some relevance to future assessment instrument and curriculum design.

　本章は、ある大学の英語プログラム(ELP)における自己評価方法とその手順に関する学生と教師の考え方について論じる。4学期間に渡る必修のELPに、学生は自己評価についてのアンケートに回答し、教師も同様のアンケートに答えた。回答の記述統計分析および、自由回答式質問への回答の質的分析によって、学生の考え方におけるいくつかの相違点が明らかになった。さらに、学生と教師の回答を比較したところ、教材のデザインと動機づけとの関係において更なる相違が明らかになった。これらの結果は、今後の評価手段とカリキュラムデザインに対して何らかの示唆を与えるものかもしれない。

INTRODUCTION

Query students in your classes or teaching colleagues about alternative forms of assessment, such as self-assessment, and you're likely to hear a variety of responses:

- *Most of the students don't take it seriously and just write scores that they don't deserve to get.* Student

- *I think students and teachers often evaluate different aspects of tasks, so although their scores may be similar, I think 'intentions' are different, typically.* Teacher

- *I like self-assessment because we have time to review the class activities and personal work, so we can recognize what we need to do to improve for us.* Student

- *In general, self-assessment depends on the individual student. It's effective for some and a total waste of time for others.* Teacher

- *I think teacher and student scores will be different because they look at it (student learning) from different angle.* Student

- *I think that self-assessment is one of the most effective methods teachers can use to encourage thoughtfulness and autonomy in students.* Teacher

These responses, culled from student and teacher surveys I administered at the Japanese university where I teach, exemplify some of the complex range of views towards self-assessment that can be found. Although self-assessment is believed to promote language learning, learner awareness, and learner autonomy (Tudor, 1996), and self-assessment is widely used in current language education programs, inter-group (between students and teachers) opinions and intra-group (within a group of teachers or within a group of students) opinions both overlap and conflict. What does variation in student attitudes towards self-assessment mean? What does variation in teacher attitudes towards self-assessment mean? How can the answers to these questions impact curriculum development and materials design? Do these differences signal problems within the framework? My aim in this chapter is to explore these issues and offer some answers which can positively influence future materials development.

CONTEXT

The following study was conducted in the English Language Program (ELP) at Kwansei Gakuin University's School of Policy Studies. A brief explanation of the organization and aims of the ELP can help clarify my motivation in doing this research. The ELP is a rigorous, content-based English for Academic Purposes program enrolling approximately 1,000 students with TOEFL scores ranging from 350 to 600. Throughout the required four semesters, students learn how to formulate, support, and express various perspectives, in written and oral modes, on socio-

political and environmental topics through individual and group research, discussion, writing, and presentation projects. Commercially produced textbooks are not used in the ELP—all course materials used by ELP teachers are created by eight full-time Associate Lecturers of English using authentic sources for each skills class. Skills and topics are horizontally integrated across classes, and vertically integrated from semester to semester. In this way, ELP courses recycle and build upon skills and content, preparing students for success in upper division content courses and in their postgraduate careers.

CHALLENGES

As a member of the full-time staff, I am involved in the design and implementation of program-wide materials. I have witnessed both successes and challenges of using the same materials for all students, and I've become increasingly interested in how self-assessment and learner training materials in particular are used by a diverse teaching staff for a diverse student cohort. Course materials, for example, can be used by more than a dozen teachers for over 400 students. I developed this research study to explore how student and teacher attitudes and feedback in general might inform such curriculum design.

I have been fortunate to work in a context where promoting student autonomy—the ability to work independently—is a central goal. By "independently" I mean that students can, on their own and collaboratively in groups, identify topics, carry out research, and discuss, present, and write effectively about various topics. Autonomy is encouraged by giving students choice in their topics and assignments within the parameters of an overall class framework. Autonomy is further developed by giving students responsibility for critiquing their own and their peers' work in various stages of completing a project. In the first year, self- and peer-assessment are closely guided by instructors, and, in the second year, students are given more freedom in the criteria used for evaluation. Self- and peer-assessment instruments are used as a pedagogical tool to remind students of previously taught skills and topics, to facilitate goal-setting and encourage formulating plans to achieve those goals, to increase awareness and independence as students learn to evaluate their and others' work, and to help students see the value of giving and receiving feedback (Oscarson, 1989). Through training, practice, and discussion, teachers help students understand these benefits to remind them of the value of alternative forms of assessment.

At first glance, the ELP is a progressive program, embracing current methodology and pedagogy informed by 'cutting-edge' research to deliver an education designed to foster learner autonomy. However, since I joined the full-time staff more than 3 years ago, full- and part-time teachers alike have voiced concerns during teachers' meetings about current self-assessment materials and procedures, stating that (a) there seems to be little effort to coordinate self-assessment instruments for different skills classes, and (b) there are no explicit guidelines (for students or teachers) outlining how self-assessment should be carried out. Moreover, teachers have reported a number of informal observations regarding the use of self-assessment:

- Since self-assessment training is left up to individual teachers, students may receive very different experiences.

- Instruments which work well with higher proficiency students do not seem to work as well with lower proficiency students.

- Higher proficiency students and female students tend to score themselves lower.

- Students don't seem to see the value of self-assessment.

Hearing these claims both echoed my own classroom experiences and fueled my interest to investigate them more deeply in an effort to improve existing materials. At this point, responding to each claim in detail is beyond the scope of this chapter. Rather, my purpose is to begin exploring student and teacher attitudes toward self-assessment in an attempt to better understand how the stakeholders—teachers, students, and curriculum developers—perceive what has come to be a critical component of the curriculum. Ultimately, I'm arguing that the opinions of all stakeholders need to be considered to promote thoughtful, meaningful self-assessment and learner strategy materials design. To this end, I have focused on the following questions:

1. What attitudes do students and teacher have towards self-assessment in the ELP?

2. Do students and teachers share similar attitudes towards self-assessment in the ELP?

3. What impact, if any, does gender have on student beliefs?

4. Do students perceive some value in completing self-assessment?

5. Do students assess themselves honestly?

FOSTERING AUTONOMY—LEARNER AND CURRICULUM ISSUES

Thoughtful implementation of self-assessment is critical to effectively promote autonomous learning. To begin with, teachers must be aware of themselves as learners, believe and trust their students, and possess a sincere desire to foster learner independence (Breen & Mann, 1997). Teachers must also acknowledge the impact that learner diversity may have on classroom behaviors and procedures. Studies show teacher-learner interaction and learner-learner interaction can be affected by many factors, such as gender. More specifically, it has been shown that male students have more talk-time in the language classroom, and some evidence suggests that teachers expend more energy on male students (Shehadeh, 1999; Sunderland, 1992).

If male students talk more or if male students receive more teacher feedback, how might this affect self-assessment? Is this an issue in Japan? Thomson's research on learners studying Japanese takes into account gender and self-assessment, showing that Asian females tended to rate themselves lower, and ascribes this tendency to cultural background (Thomson, 1996). Yet to better understand the interplay of gender (as well as culture) with self-assessment, more research to investigate learner differences is needed (Vandrick, 2000). Ultimately, being mindful of these issues means additional research, combined with carefully planned self-assessment materials design and classroom procedures, is paramount: If students and teachers are dissatisfied with current self-assessment procedures, then current practices will be limited in making positive contributions towards learner autonomy.

METHODS

PARTICIPANTS

Both students and teachers in the ELP were surveyed for this study. Sixty-five students, nearly all of who were second-year students, took part in the study. They were all enrolled in Animal Issues, a Special Topics elective class taken during the last (fourth) semester of required English courses. Twenty-eight students were female, and 36 male. The students were spread across three, non-streamed classes, and represented a range of proficiency levels. Eighteen out of 22 ELP part-time and full-time teachers completed surveys.

COURSE DESCRIPTION AND SELF-ASSESSMENT TRAINING

The Animal Issues course was a reading-based discussion course focusing on endangered animals, wildlife trafficking, abandoned pets, and animal research among other topics. Students completed readings and internet research to support in-class discussions, as well as individual and small group project-based presentations. Assessment was based on completion of homework, vocabulary and reading comprehension quizzes, project work, presentations, and self- and peer-assessment scores. Students self-assessed discussion effort and participation as well as presentation preparation, and they also peer-assessed group members' presentation performance. The self-assessment component accounted for 5% of the final grade.

Students completed a self-assessment form three times during the semester immediately following in-class discussions (see Appendix A on the *Autonomy You Ask!* website). Students were asked to focus on their preparation for discussions, as well as their effort and participation during discussions, by responding to the following three prompts: (a) choose a numerical score reflecting your holistic effort and performance, (b) write a justification of the chosen score, and (c) write a plan outlining what and how you will improve for the next discussion.

To encourage active participation, some simple classroom practices were implemented. For example, at the beginning of the course, students reviewed features of effective group work and active classroom participation. Then, before all group work activities, they reviewed these features. The purpose was twofold: first, to enable students to co-construct, with teacher feedback, some standards for expected classroom behaviors, and second, to help students apply these standards when choosing and justifying their own scores. In addition, each member selected (or was assigned) a group role (leader, note keeper, reporter, English Police). Establishing these basic policies was instrumental in reminding students of their roles in their learning.

Activities designed to promote more thoughtful completion of self-assessment were conducted throughout the semester (see Figure 1 over). Before completing self-assessment for the first time, students brainstormed the rationale and potential benefits of self-assessment. Student answers were then discussed by the whole class with teacher feedback and support. Examples of good self-assessment forms were distributed and analyzed, and students brainstormed additional information, creating a menu of feedback phrases and content to draw on in future classes. These examples, in addition to classroom discussion, were aimed at giving students the tools they needed to choose and justify their scores. To encourage students to take charge of their learning, they formulated a plan to improve their effort and participation for the next class. Before beginning the next class discussion, students reviewed these plans and were asked to keep them in mind during the day's activities. The self-assessment forms were completed once again, collected at the end of class, and returned during the next class with feedback from the teacher. Throughout the semester, we revisited the evaluation criteria, reviewed good models, and discussed features of effective and less effective self-assessments.

I wanted to give useful feedback to students to further promote reflection and honesty. If students expressed themselves well with concrete examples and details, I praised their efforts. If student responses were vague and lacked detail, I wrote questions eliciting additional information, and made brief suggestions on how they might expand their answers. If I thought a student had given her/himself an overly high or low score, I wrote a question about their score. I never changed their scores, but I wrote a quick note based on their in-class performance, within our class-established evaluation framework, that might suggest their score could be higher or lower. I felt I could only give this kind of feedback in extreme situations: for example, for a student who did not interact and did homework in-class, yet scored themselves
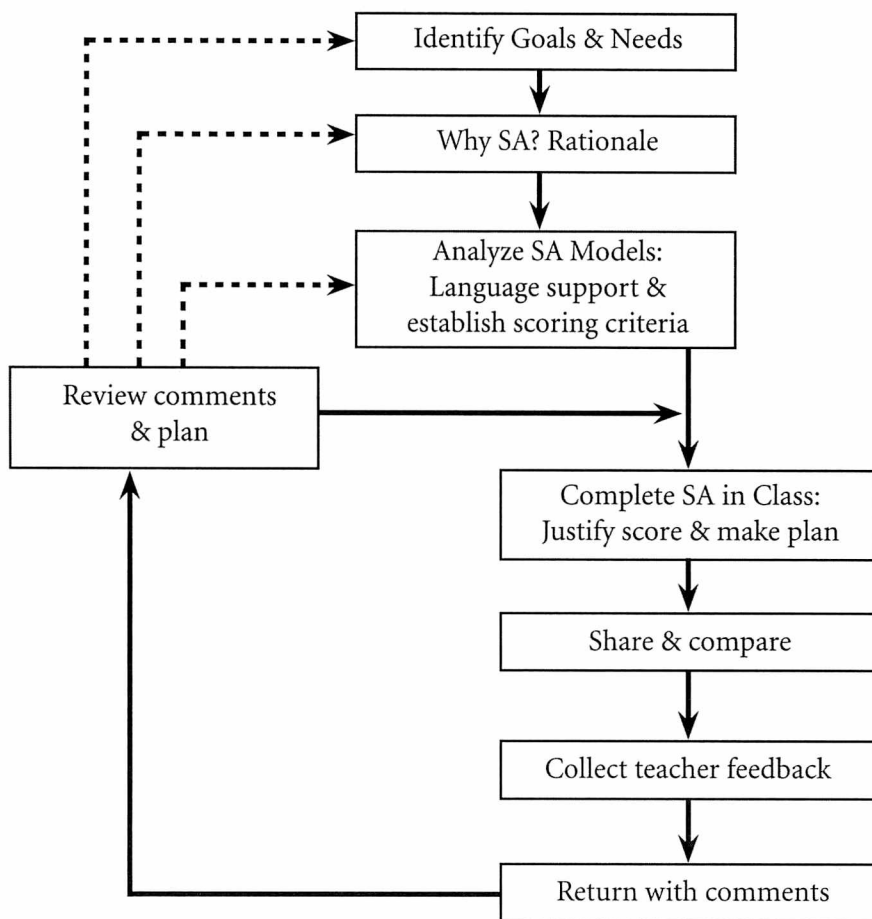
```
┌─────────────────────────┐
│   Identify Goals & Needs │
└─────────────────────────┘
           │
           ▼
┌─────────────────────────┐
│   Why SA? Rationale      │
└─────────────────────────┘
           │
           ▼
┌─────────────────────────┐
│   Analyze SA Models:     │
│   Language support &     │
│   establish scoring      │
│   criteria               │
└─────────────────────────┘

┌─────────────────┐
│ Review comments │
│    & plan       │
└─────────────────┘

┌─────────────────────────┐
│   Complete SA in Class:  │
│   Justify score & make   │
│   plan                   │
└─────────────────────────┘
           │
           ▼
┌─────────────────────────┐
│   Share & compare        │
└─────────────────────────┘
           │
           ▼
┌─────────────────────────┐
│   Collect teacher        │
│   feedback               │
└─────────────────────────┘
           │
           ▼
┌─────────────────────────┐
│   Return with comments   │
└─────────────────────────┘
```

Figure 1 Self-assessment Process Model

10; or for a very active student who scored themselves 6 or less. When self-assessments were returned at the beginning of the following class, students were encouraged to share their forms with each other and ask me questions regarding my feedback.

SURVEYS

Surveys were administered to students and teachers (see Appendices B and C on the *Autonomy You Ask!* website). The student survey consisted of 15 statements and 6 open-ended questions, in English and Japanese. Using a 4-point Likert rating scale, students were asked to agree or disagree with the 15 statements which were designed to investigate motivation, honesty, and attitudes towards self-assessment classroom procedures and materials, as well as student classroom behaviors. I decided to focus on these areas to explore whether or not students perceived value in self-assessment and believed they were honest. Ultimately, I expected that students (and teachers) who possessed positive attitudes towards self-assessment would agree with most statements. The student survey was given to students on the last day of class. A similar survey, made up of 12 questions and a comments section, was distributed to ELP teachers during the last week of class. Sharing 8 questions in common, the surveys aimed to

uncover stakeholders' attitudes towards self-assessment instruments and procedures currently in use in the ELP, with the ultimate goal of improving such practices and materials.

RESULTS

Using descriptive statistical analysis, I examined student and teacher responses and attempted to identify trends and differences between students and teachers. I was most interested in understanding, in more concrete terms, how students and teachers viewed self-assessment, namely: (a) whether students perceived some value in self-assessment; (b) if students completed self-assessment thoughtfully and honestly; and (c) on what points students and teachers broadly agreed or strongly disagreed.

QUANTITATIVE ANALYSES

Before reviewing the survey results, looking at student class scores and self-assessment scores may help situate my research focus. Since the research serves to explore teacher observations regarding self-assessment, I wanted to see how closely these claims matched the reality of my own classroom. Thus, I calculated average scores for each of the three times self-assessment was administered in class to see if any trends emerged.

Table 1 Average Student Self-assessment Scores

	Female students	Male students
SA#1	7.8	7.7
SA#2	8.4	8.2
SA#3	8.3	8.7
Overall SA average	8.1	8.2
Overall course grade (out of 100) average	84.0	82.0

Table 1 shows the average self-assessment scores for female and male students across the course. The average scores for all three self-assessments, as well as the overall average score, are very similar. Females scored themselves slightly higher than males for self-assessments #1 and #2, and slightly lower than males for self-assessment #3. In addition, female and male students received similar overall course grades, with female students earning slightly higher grades.

Bearing in mind student self-assessment scores and overall course performance, survey responses are outlined below. Tables 2 and 3 below summarize student and teacher responses for agree-disagree statements. For this survey, the four point Likert scale choices included: 1 = strongly disagree, 2 = disagree, 3 = agree, or 4 = strongly agree.

Table 2 Survey Responses: Student Attitudes (N=65)

Question	Mean		SD
1. I understand why self-assessment is used.	2.91		.66
	F = 3.11 (n = 28)	M = 2.83 (n = 36)	
2. Self-assessment encourages me to try harder.	2.73		.67
	F = 2.96	M = 2.61	
3. Students should have the opportunity to do self-assessment.	3.08		.65
	F = 3.25	M = 3.03	
4. Students choose their scores honestly.	2.63		.78
	F = 2.71	M = 2.64	
5. Student scores are similar to teacher scores.	2.48		.62
	F = 2.68	M = 2.39	
6. I would like more self-assessment in my English classes.	2.19		.61
	F = 2.25	M = 2.39	
7. I would like less self-assessment in my English classes.	2.19		.66
	F = 2.25	M = 2.19	
8. Writing a justification of my score was useful and valuable.	2.80		.76
	F = 3.04	M = 2.67	
9. Writing a plan for the next discussion was useful and valuable.	2.79		.76
	F = 2.93	M = 2.75	
10. If I write my self-assessment at home instead of in class, I will spend more time on it and write a more thoughtful self-assessment.	2.10		.79
	F = 2.32	M = 1.97	
11. I prefer to write my self-assessment in class.	2.77		.75
	F = 2.64	M = 2.94	
12. If I don't finish my self-assessment in class and take it home, I will probably forget to do it.	3.16		.83
	F = 2.93	M = 3.42	
13. It's OK to show my score, justification, and plan to other students.	2.54		.89
	F = 2.39	M = 2.53	
14. If someone gave themselves a score that I thought was too low, I could tell them to change it.	2.48		.89
	F = 2.50	M = 2.53	

Table 2 shows that students moderately agree with all statements, reflecting some satisfaction with the current system. The first three statements focus on how students view self-assessment, and whether or not they see its value and are motivated by self-assessment. Statements 1 ("I understand why SA is used"), 2 ("SA encourages me to try harder"), and 3 ("Students should have the opportunity to do SA") have means of 2.91, 2.73, and 3.08, respectively. In other words, means which approach 3 (equivalent to "agree" on the Likert scale) indicate positive attitudes towards self-assessment. Within these first three statements, however, is an interesting trend: Female students tended to have slightly higher means than male students, which may indicate that females agree more strongly with these statements than males. Statements 6 ("I would like more SA") and 7 ("I would like less SA"), which concern the amount of self-assessment used in classes, both have means of 2.19. Students moderately disagree with these

two statements, possibly indicating they regard the current amount of self-assessment as adequate.

Statements 8 to 10 center on classroom procedures. Statements 8 ("Writing a justification was useful") and 9 ("Writing a plan was useful") have means of 2.80 and 2.79, with female scores being slightly higher than male scores. Females' scores were, in fact, 3.04 and 2.93. Such agreement represents some degree of satisfaction of current self-assessment instruments. Statements 10 and 11 concern completing self-assessment at home or in class. Students tended to disagree with 10 ("I will complete my SA more thoughtfully if I can do it at home") with a mean of 2.10, representing moderate disagreement. In response to Statement 11 ("If I take my SA home I will probably forget to do it."), the mean was 3.16, suggesting agreement, with 2.64 for females and 3.42 for males. This statement has the highest mean of all survey statements, showing that students feel more strongly about this issue.

Several of the survey statements deal with issues of honesty. Statements 4 ("Students choose their scores honestly") and 5 ("Student scores are similar to teacher scores.") are slightly lower at 2.63 and 2.48, indicating less agreement with these statements. Females again tended to have slightly higher means. These figures become more meaningful when student responses to open-ended questions are considered. Likewise, the last three statements focus on honesty operationalized in classroom procedures. After students complete their self-assessments, they are asked to share their scores and explain their justifications to group members, as well as encouraged to feedback on fellow group members' scores. For Statement 13 ("It's OK to show my score to other students"), at 2.54, the mean rests in the middle. For 14 ("I could tell someone to change a score that is too low") the mean is 2.48 and for 15 ("I could tell someone to change a score that is too high") the mean is 2.06. As the means approach 2.0 (disagree), it could be inferred that these students may not feel comfortable with these classroom procedures.

Looking at the overall student means, as well as the means by gender, sheds some light on student attitudes towards self-assessment in general, as well as materials, classroom procedures, and classroom behaviors. Comparing teacher responses to student answers shows students and teachers broadly agree on some points and disagree on other points (see Table 3 below).

Table 3 Survey Responses: Teacher Attitudes (N=18)

Question	Mean	SD
1. Self-assessment increases motivation.	2.67	.75
2. Self-assessment promotes reflection.	3.28	.45
3. Students should have the opportunity to do self-assessment.	3.44	.69
4. Students choose their scores honestly.	3.00	.47
5. Student scores are similar to teacher scores.	2.72	.56
6. Writing the justification is useful and valuable.	3.33	.58
7. Writing a plan for the next discussion is useful and valuable.	3.17	.60
8. If students write self-assessment at home instead of class, they will spend more time on it and write a more thoughtful self-assessment.	2.50	.50

9. Having student show their self-assessment to other students promotes honesty and can be a valuable learning experience.	2.89	.66
10. Female students generally score themselves lower than male students.	2.78	.79
11. Lower level students tend to inflate their scores.	2.56	.69
12. Higher level students tend to score themselves lower than they should.	2.67	.58

Table 3 shows that teachers tended to agree more with statements than students on some points and less on other points, yet responses are generally more favorable, indicating teachers agree more broadly with the statements than students. Statement 1 ("SA increases motivation") with a mean of 2.67 is similar to Statement 2 ("SA encourages me to try harder") on the student survey. The student mean at 2.91 is slightly higher than the teacher mean, showing a slight difference in attitudes.

Teachers moderately to strongly agreed with Statements 2 ("SA promotes reflection"), 3 ("Students should have the opportunity to do SA"), and 4 ("Students choose their scores honestly"), showing positive attitudes towards self-assessment use in general. The teachers' mean for Statement 4 at 3.00 is higher than the students' mean of 2.63, suggesting that teachers do trust students and value self-assessment as a means of course evaluation. Teachers also agreed more strongly than students with statements connected to the self-assessment instrument itself (Statements 6 & 7), showing that teachers do feel asking students to justify their scores and formulate plans is a critical part of self-assessment process.

The last four statements, which center on honesty and classroom procedures, show moderate agreement. In particular, teachers seem to subscribe to the practice of having students share their self-assessments with group members (mean = 2.89). As expected, teachers also moderately agreed that students do not always score themselves fairly.

QUALITATIVE ANALYSES

The 6 open-ended questions on student surveys were coded to reveal patterns among responses. I organized responses from all students question by question, and looked for patterns in theme and vocabulary. I grouped responses together according to similar content and totaled the responses, as shown in Table 4 below. Student responses are organized here according to question. For Question 4, I divided these responses by gender because I wanted to see if gender had any impact on student perceptions of honesty. Since teachers claimed that female students' scores were "lower than they should be" and lower than male students' scores, I thought separating these responses by gender might shed some light on these claims.

Table 4 Survey Responses: Summary of Open-ended Student Responses

1. What do you like about self-assessment?	
I don't like it.	1
I can find strong point.	6
I can evaluate myself without someone's evaluation.	9
I can find weak point.	12
I can improve for next class.	26
I can look back at myself / review myself.	27

2. What do you not like about self-assessment?	
It's troublesome and takes time.	
It's too subjective; everyone has different standards.	8
Students lie and give themselves higher scores than they deserve.	35

3. What are the benefits of self-assessment?	
We can know our strong points.	3
We can know our weak points.	10
It can help us improve.	16
We can think about what we did and look back.	7

4. Do you think there are any differences between women and men's scores?	female	male
Yes.	1	2
No.	7	23
It depends on the person, not if they are woman or man.	6	5
Women are more honest / thoughtful / strict.		5
I have no idea.	2	2

5. Do you think there are any differences between teacher and student scores?	
No.	12
Yes. Students give themselves higher scores than they deserve.	12
Yes. Teachers and students are looking at different things; teachers and students have different viewpoints.	14

6. How can teachers improve self-assessment?	
Show more examples and give more support for what to write.	3
Give students more time to write.	4
Take more time to discuss self-assessment procedures.	6
Make time to compare with other students, and compare with teacher and peer-assessment scores.	7
I don't know.	8
Use "point" / "concrete" / "detailed" system.	9

From the responses in Table 4, we can see that self-assessment clearly helps students 'look back' and reflect on their performance and effort, even if they admit they do not always assign scores they deserve. Students view self-assessment as a method to inform them of their negative or weak points more than their strong points. Overall, students recognize that differences among scores exist not only among students, but also between students and teachers. Based on their experiences, students are in a position to deliver concrete feedback to impact instrument design.

All comments from teacher surveys are listed in Table 5 below. In contrast to the student survey, which included 6 open-ended questions, the teacher survey simply included a 'Comments' section. With very busy schedules, I didn't want to push teachers to write more than what they had time to thoughtfully complete.

Table 5 Survey Responses: Open-ended Teacher Comments

1.	I think it's quite valuable to ask students to assess themselves, though for Japanese students, the notion of "self-assessment" is really unfamiliar. So they may need time to feel comfortable with it.
2.	It's difficult to generalize! Generally, I think it's valuable, but not necessarily for the "score." I think it can help some, though not all, to promote responsibility in learning.
3.	Self-assessment can promote honesty but can also promote group dishonesty. Students are often reluctant to force another student to change a grade as it involves face loss and confrontation.
4.	If self-assessment is done too often, it loses value.
5.	I think self-assessment should be done immediately after the activity while it is still fresh. It may be useful to discuss a "group" self-assessment in order to hear other students' ideas about your performance.
6.	I think teacher feedback (such as "Your score is a little too high") on student score—especially with display of self-assessment of students' work and scores to classmates—also helps promote reflection, honesty, and bringing student scores in-line with teacher scores.
7.	I think self-assessment can be a motivator, but it depends less on the students and more on the teacher, e.g., how it has been introduced, what feedback is given.
8.	Women never grade themselves high enough. Self-assessment has value, but maybe only for highly-motivated students. Though on the other hand, maybe it's just something they need more training in…
9.	I have found reflection and planning the next discussion is very useful with high level students, but low level students don't reflect well. Perhaps there is a threshold level of English that you need!
10.	I think that self-assessment is one of the most effective methods teachers can use to encourage thoughtfulness and autonomy in students. Of course, the method and approach need to be actively considered and evaluated, so that the process retains credibility in the students' eyes.
11.	I think it's only necessary to write a justification or a plan—not both.
12.	In general, self-assessment depends on the individual student. It's effective for some and a total waste of time for others.
13.	I think students and teachers often evaluate different aspects of tasks, so although their scores may be similar, I think 'intentions' are different, typically.

The responses shown in Table 5 indicate a mixture of positive and negative opinions about self-assessment, as well as comments regarding procedures in the current self-assessment scheme. Broadly speaking, there is considerable overlap among these statements, showing that teachers generally view self-assessment in positive terms, as a useful tool to promote reflection and motivation. Teachers also recognize limitations and challenges of using self-assessment, and see the importance of careful, thoughtful implementation.

DISCUSSION

SELF-ASSESSMENT SCORES AND GENDER

Are there significant differences between female and male students' self-assessment scores? Remember that one area I wanted to investigate was whether or not teachers' claims that female students score themselves lower than they deserve and that their scores are generally lower than male students' self-assessment scores are valid. (For my class, I collected self-assessment scores only; I did not score students individually for each discussion. As a result, I cannot compare 'their' scores with 'my' scores.) What was striking to me was that, contrary to teachers' claims (and my own observations as well), female and male students' self-assessment scores were quite similar. The average male student's score of 8.2 is only slightly higher than the average female student's score of 8.1 (see Table 1 p. 163).

What can account for these results? It is possible that by the end of the last (fourth) semester of ELP classes, students' scores have leveled out, because they have had lots of experience and feedback using alternative forms of assessment. Moreover, the average score is slightly above the 'average' of 7. As 'veterans' of self-assessment, students may score themselves higher because they understand their own learning processes better and have become accustomed to completing self-assessment. Finally, since the course was an elective, and students were enrolled by choice, it's possible they were more motivated during discussions because of the content, so perhaps their scores reflect this active participation.

But what do students have to say about this issue? Do students believe gender has any influence on scores? According to Table 4 above, students do not presume that gender has any bearing on self-assessment scores. While three students do believe scores may be different, 40 students do not perceive differences stemming from gender. Rather, students ascribe score differences arising from students as individuals, not their gender. I still believe this issue needs further investigation, given that more female than male students (9 to 5) believe females are more honest, thoughtful, and strict (see Question 4, Table 4 above). Clearly, these accuracy and honesty issues deserve further consideration.

STUDENT AND TEACHER ATTITUDES

What differences in attitudes emerged between students and teachers? From students' perspectives, self-assessment is a tool that allows them to "look back" and "think about how to improve for the next class." I interpreted these responses to mean students see value in self-assessment as a method to reflect on their effort and participation and create a concrete plan to improve their performance. These responses also nicely mirror Dam and Legenhausen's claim (as cited in Cotterall, 2000, p. 112) that "Evaluation has a retrospective and prospective function, in which the learning experiences of the past are reflected upon and transformed into plans for future action." Essentially, students expressed the same idea, though not so eloquently. What is interesting, however, is how students view self-assessment as a way to understand their negative points more than recognize their successes. Identifying these weak points seems to be the primary benefit of self-assessment for students.

From the data, students aren't aware of all the benefits of using self-assessment, as outlined by Oscarson (1989). Oscarson believes that training students to evaluate themselves benefits their learning process and raises their awareness of not only their own performance, but also fosters goal setting and reminds them of previously learned content; that relying on students to provide insight into their own performance gives a more accurate picture of a learner's abilities and shares some of the assessment burden with the teacher; and ultimately that *self-assessment*

fosters independent learning beyond the classroom context. There is some indication that students understand that assessment is a shared process based on the claim that self-assessment allows students to "evaluate myself without someone's evaluation," which I interpret as their ability to provide another perspective, in addition to teacher and peer evaluation (see Questions 1 and 3, Table 4 above). Gaps in their overall understanding about why self-assessment is such a critical component of their learning can be overcome with teacher support and improved classroom procedures to make self-assessment more meaningful and valuable.

Based on the survey results and my experiences working closely with other ELP teachers, teachers typically support Oscarson's reasons for using self-assessment. I interpret teacher survey responses as a broad endorsement of self-assessment, when the materials are appropriate and procedures promote autonomy. Teachers value using self-assessment and are committed to its purposeful use as a tool to encourage reflection and foster autonomy. Clearly, teachers understand the reasons for using self-assessment, and, with improved materials, should be better equipped to help students more fully understand the rationale behind using self-assessment

Students and teachers hold slightly different views with regard to implementation of self-assessment. Teachers view writing a justification (mean = 3.33) and plan (3.17) to be useful and valuable while student means are slightly lower at 2.80 and 2.79, respectively. Moreover, whereas teachers somewhat agree with asking students to share self-assessments (mean = 2.89), students are less keen on it (mean = 2.54). These are hardly great differences, but I think it's important to note that perhaps some current classroom practices may need to be revised, as I will discuss more fully in the conclusion.

HONESTY

Are students honest? Responses from the open-ended questions certainly make the statistical data more meaningful. In response to "Students choose their scores honestly," the student mean was 2.6. By itself, this figure has little meaning. However, looking at the extracts from Table 4 below, we can see that clearly students believe that students are dishonest. It's not clear from the data who is lying. Are they lying? Or are other students lying? Regardless, student responses suggest that they believe scores are not (always) honest reflections of student effort and participation.

2. What do you not like about self-assessment?	
It's troublesome and takes time.	8
It's too subjective; everyone has different standards.	8
Students lie and give themselves higher scores than they deserve.	35

5. Do you think there are any differences between teacher and student scores?	
No.	12
Yes. Students give themselves higher scores than they deserve.	12
Yes. Teachers and students are looking at different things; teachers and students have different viewpoints.	14

Contrasting the results of the above two questions poses an interesting contradiction. To put it simply, 35 students claim that students lie and award themselves high scores. Yet with Question 5, only 12 students noted that students award themselves scores that are higher than they deserve. Shouldn't these numbers be similar? How can we account for this difference? Perhaps a hint of the answer can be found in the claim by 14 students that students and teachers look at different features, and if this is the case, then we can't expect scores to be equal because students and teachers have different perspectives. To truly understand these figures, however, follow-up investigation is needed.

LIMITATIONS

The results could be more powerful if data were subjected to more rigorous statistical analysis. For example, correlation between gender and achievement or proficiency with survey answers could be explored. Also, student and teacher responses could be compared quantitatively using discriminant analysis, a type of t-test which compares groups of different sizes to determine which features serve to distinguish group membership. These analyses might reveal trends not immediately evident in the present data. Finally, the number of cases is too small to generalize to a larger population. If this study were replicated using an improved survey with a 6-point Likert scale for a larger sample, perhaps the results would be more reliable, powerful, and generalizable to multiple contexts.

At this point, the results should be regarded as initial and applicable for one university context. Based on these preliminary findings, more questions have been raised than answered! Further investigation is needed to develop a deeper understanding and clearer picture of how these issues play out beyond this particular context. Conducting a longitudinal study, for example, that tracks changes in students' scores may reveal a deeper understanding of the relationship between self-assessment, gender, and proficiency, as well as show how experience impacts self-assessment procedures. Such a study would naturally include revised assessment tools and procedures, with follow-up surveys and interviews to investigate learner attitudes more fully.

CONCLUSIONS

The present study shows that attitudes towards self-assessment among teachers and students in the ELP are complex and wide-ranging. Teachers regard self-assessment slightly more positively than students. However, quantitative and qualitative results seem to confirm that students understand the aims and benefits of self-assessment in a broad sense. Evidence of this can be seen in the fact that students regard self-assessment as a tool to "look back on" their effort and a method to improve "for the next time," which is aligned with teacher comments. However, students also expressed concern over issues of fairness and accuracy. Students are well aware of their power to score themselves higher than they deserve, and they feel that some students take advantage of this power. Moreover, students (and teachers) believe that students and teachers are looking at learning from different angles, using different knowledge and criteria, so student self-assessment scores may very well be different from teacher-assessed scores. Ultimately, teachers can help students understand self-assessment and its benefits by providing opportunities for discussion and giving feedback.

RE-DESIGNING SELF-ASSESSMENT

One of my main purposes for investigating student and teacher attitudes towards self-assessment is to stimulate curriculum renewal. Based on the research results, I would now like to suggest a number of ways in which teachers in the ELP can improve self-assessment instruments *and*

procedures to help students have more meaningful experiences using self-assessment. These suggestions, based on my understanding of current pedagogical principles and informed by student and teacher feedback, represent a starting point for curriculum innovation for my teaching context. I hope that they may also be of some use to teachers in other contexts.

1. Teachers should spend more time facilitating discussions about self-assessment to help students discover the rational for self-assessment. This is an on-going process that needs to be re-visited as students become more proficient and experienced with self-assessment. This includes helping students understand the benefits of multiple assessment efforts, such as self, peer, and teacher assessment.

2. Self-assessment instruments should combine a global, holistic rating with specific performance-related ratings directly connected to current course skills and goals. Performance-related ratings give students and teachers a more concrete view of student performance and more fully inform holistic scores.

3. As students gain experience, they should be allowed to design or co-design self-assessment instruments.

4. Explicit evaluation standards and criteria for holistic and performance-related ratings should be co-constructed to help students choose their scores. These evaluation standards should be re-visited and revised to reflect changes in expectations as students' proficiency and experience increase.

5. Assessment instruments should aim to accommodate different perspectives among students, peers, and teachers, so students should only be responsible for assessing features which they can realistically assess. Teachers should take caution assessing features for which they lack insight (i.e., student effort) and stick with "external" features. However, it is likely that there will be overlap among various features assessed by students and their peers and teachers.

6. If a self-assessment instrument does not have a 'free' comments section, but students would like to write comments anyway, they should be encouraged and given time to write in English or Japanese if they feel it will benefit their learning.

7. Students should have the option to maintain a 'private' place to keep their comments, such as a personal notebook or journal which is not collected by the teacher or shown to other students.

8. Students need models and good examples especially in the early stages of self-assessment training. Analyzing models in class may benefit some students. Students of all levels may need additional language support to express their thoughts. Such language support requires lots of practice and should be revisited again and again.

9. Instruments and procedures should be realistic, achievable, and transparent. Students should have time to discuss materials with each other and with their teacher before writing.

10. Completing self-assessment can be an individual and a social effort. Students should have the opportunity to compare and discuss their self-assessment forms with other students (if they wish to do so). However, students shouldn't be forced to advise group members to change scores they perceive to be too high or too low. Grade deflation or inflation issues should be addressed by the teacher at first, and as students gain experience and get to know their peers better, it may be possible to ask students to feedback on each others' self-assessment scores. The key is to tread lightly here and provide options.

In the process of conducting this research, I have gained a number of insights into my own and my students' and colleagues' understanding of ELP self-assessment procedures. Materials developers can improve self-assessment instruments by reading books about curriculum design, subscribing to teaching and research publications, conducting action research, and discussing these issues with other teachers. But teachers should also survey students (and other teachers) and ask them directly what they think about materials and classroom procedures. Somewhere, between our experience using self-assessment in the classroom and our understanding of pedagogy and learning principles, we should consider our students' voices in self-assessment design and implementation. If we're using self-assessment to encourage students to become autonomous learners, then it simply makes sense that self-assessment materials and procedures should reflect their input.

CRITICAL READER RESPONSE 1

NOBUYUKI TAKAKI

What is most meaningful about this project is that it presents an 'institutionalized' self-assessment and learner training program at the university level where the faculty 'share' the same syllabus and the same self-created materials. This seems innovative for two reasons: First, assessment of students' abilities has been traditionally left up to individual teachers, who rarely communicate among themselves concerning their subject matter teaching (Takaki, 2002); second, learner training has been traditionally neglected in Japanese universities.

Unfortunately, however, it is not easy to get a good picture of the project's setting, such as the ratio of full-time vs. part-time and native vs. non-native speaker staff. This is an important issue because involving part-time staff or teachers with mixed nationalities in such an ELP can cause various conflicts. It would have been also useful if Heidi had specifically discussed the particular local constraints and conflicts that the faculty faced in implementing their program.

It seems that this research should be regarded as a pilot study by the faculty, for the data collected within such a limited time would be unsatisfactory to make suggestions about self-assessment. Why I say this is that self-assessment in Japan has started fairly recently in a restricted and superficial sense at best without contributing much to improving the classroom environment.

Heidi has not discussed teachers' reluctance in Japanese school culture to verbalize beliefs concerning classes, although she has cited some comments on the project voiced both by the teachers and the students. It seems to me that Heidi needs to consult more with both the students and the colleagues about their views and their changing understandings of 'self-assessment.' Moreover, since Heidi claims that the research is both quantitative and qualitative, she should give follow-up interviews to some of the examinees, citing some of their comments. Additionally, the gender issue in the survey seems out of focus and premature in her present paper, as it is briefly mentioned but not developed as a discussion.

One problem with Japanese students is that they have not been sufficiently educated to take responsibility in their own learning. A program like Heidi's should be designed to help them with their awareness-raising on this point. Thus, the 'Self-Assessment' form (see Appendix A on the *Autonomy You Ask!* website) should have been handed to the students on Day One of the course (if it was not), encouraging them to set their own goals in their learning in the course before discussing their self-assessment.

It is wonderful, on the other hand, that the students were given an opportunity to discuss the feedback from the teacher, but, to my disappointment, no example feedback discussions

are shown. And lastly, I would want Heidi to cite what the faculty has discussed in their collaborative reflection sessions, and I would also like to know how Heidi's project has influenced the staff's development as teachers, curriculum designers, and textbook writers. Despite my various criticisms, Heidi's project is very welcome as a possible contribution for finding innovative ways to evaluate students and teachers in Japan.

CRITICAL READER RESPONSE 2

STEPHEN H. BROWN

> Evaluation plays a pivotal role in the development of learner autonomy. The function of evaluation is on the one hand to ensure that work undertaken is discussed and revised, and on the other to establish a basis of experience and awareness that can be used in planning further learning. (Dam, 1995, p. 49)

Heidi Evans Nachi's focus on self-assessment touches on a key area of autonomous learning in practice, asking some important questions about the attitudes of students and teachers.

One of the things that struck me throughout Heidi's chapter was the barely-concealed tension stemming from the dual function of self-assessment in the ELP (the programme she describes), shared with many educational institutions worldwide. On the one hand, there is self-assessment as an instrument of *Assessment* (with a capital A, i.e., grade generation for official purposes); on the other, self-assessment is an important pedagogical tool—what I prefer to call, in common with Dam above, *(self-)evaluation*, (i.e., a tool for *reflection* on what has passed and for *future planning*). This second function is an essential element in fostering autonomy, evaluating what has been done in order to plan what comes next in the learning cycle. The challenge for a programme like the ELP is to find a way of reconciling this with the administrative, grade-generating function.

The results of Heidi's survey seem to suggest that her students understand the intended pedagogical function of self-assessment in their programme and value it, judging by the numbers responding with comments such as "I can look back at myself" and "I can improve for next class" to the question *What do you like about self-assessment?* The concerns they have seem to be related to their awareness of the other, administrative function (e.g., 35 responses of "Students lie and give themselves higher scores than they deserve"). The danger is that these concerns might impinge upon their view of the whole evaluation process.

One of the potential effects of any grading system is that learners are inclined to think *horizontally* ("How am I doing compared to the others?"), rather than *vertically* ("Where am I/we now compared to where I/we were one month ago, or where I/we want to be at the end of the semester?")—which is really how we want our learners to be thinking. In the real world, students are concerned with how others sees them, with their *performance* in relation to other learners—a natural human concern, and one which is also compounded by their previous educational experience.

Perhaps it's possible to separate the two functions and have the reflection / planning evaluation in the form of, say, worksheets, diaries or group reviews (something I've tried in my own classes, encouraging students to focus on their progress, rather than on the final grade). The self-assessment to generate grades would then be a separate process—still subject to the same inclination to compare with others, perhaps, but the process of reflective evaluation would be better protected. And that might offer a better environment to "include our students' voices" in the process.