

Student self-assessment sheets as feedback for teaching

WICKING, Paul, *Faculty of Foreign Studies, Meijo University*

GALLAGHER, Anthony Brian, *Faculty of Foreign Studies, Meijo University*

MEJIA, Justin, *Faculty of Foreign Studies, Meijo University*

MORRISH, Jaime, *Faculty of Foreign Studies, Meijo University*

Abstract

This study analyzes student self-evaluation sheets for the purpose of informing pedagogy and teaching practice in a discussion course. Self-evaluation has been the focus of many studies to date and has been shown to have various benefits for learning, although much less is known about the benefits for teaching. Potentially, there is a wealth of information that teachers can access through students' self-evaluation that can support teaching practice. This study posits that student self-assessments can serve as a pivotal resource for educators, offering nuanced insights into teaching efficacy and areas for pedagogical enhancement. In order to explore the benefits for teaching, data were collected from 125 first-year students enrolled in a discussion course at a university in Japan. After undertaking an assessment task, students completed a self-evaluation sheet that consisted of self-scoring rubrics and reflective comments. The data revealed how students viewed their strengths and weaknesses regarding discussion skills, and what future strategies they planned to adopt to develop those skills. This feedback was analyzed in order to evaluate its utility for teachers in determining lesson content and supporting curriculum development. It was found that the self-scoring rubrics provided feedback that was clear, informative, and could potentially expedite lesson planning. The reflective comments also gave an insight into learner perceptions and thought processes. This study underscores the untapped potential of student self-assessment in fostering a more responsive and student-centered educational environment, advocating for its broader integration into pedagogical practice.

Keywords: self-assessment, feedback, English as a foreign language

1. Introduction

Student self-assessment sheets are used in many university courses in Japan for the purposes of deepening knowledge and supporting the self-regulation of learning. Many studies have extolled the benefits accrued by learners when they engage in self-assessment (e.g., Benson, 2001; Andrade & Du, 2005; McMillan & Hearn, 2009; Hale, 2015; Murakoshi & Sato, 2015). However, there are very few studies concerning the benefits that self-evaluation sheets can provide for the teacher. Through the process of self-assessment, students give feedback to themselves concerning their own progress toward learning goals, and this feedback can be equally valuable for the teacher—perhaps even more so. Hattie and Clarke (2019) go so far as to argue that the feedback that a student gives to the teacher is more important than what the teacher gives to the student. This is because the fundamental role of a teacher is to gather and interpret evidence of students' current abilities, to discern where the gaps are between what they can do now and the goals of the course, and then to design instruction so that those gaps can be filled and student learning ameliorated. The gathering of feedback from students is integral to this process.

The feedback that is often prioritized in practice, however, is the feedback from the teacher to the student. As an example of this, recently many universities in Japan have

mandated that the methods used by teachers to give feedback be written on the official syllabus of each course. Perhaps this is in recognition of the important role that feedback from the teacher plays in self-regulated learning. When students reflect on the quality of their learning according to explicit goals, and then use that reflection to revise their learning behavior, students are taking control of and regulating their learning (Andrade & Heritage, 2018). Without students receiving evidence of their progress in the form of feedback from the teacher, classroom instruction risks being ineffective.

Likewise, teachers need feedback on assessment from students. This can potentially come from two sources: the assessment task itself and students' self-reflection on the performance of that task. Information that can be gleaned from the assessment task itself includes students' knowledge, capabilities, strengths, and weaknesses. Whether that assessment comes in the form of a short quiz, a written report, a presentation, or an examination, it provides insight into learner achievement. Timperley (2014) argued that this assessment information should also be viewed as an evaluation of teaching: how well knowledge and skills have been taught and what the teacher needs to learn to support the students in achieving more.

The second piece of legitimate evidence is student self-assessment. The information gathered from self-assessment can potentially be viewed as an episode of learning for the teacher. This is because it provides a number of important insights, such as what students think they know or understand, what kinds of errors they are making, and the cognitive framework they are employing in their learning activities. This information is crucial for teachers to know so that they can decide how the next stage of instruction should proceed (Ruiz-Primo & Brookhart, 2018). This action in response to feedback is an indispensable component of formative assessment. Merely giving and receiving feedback is not enough. The evidence of learning from the assessment task must inform or transform teaching and learning in some way, otherwise it cannot be called formative assessment.

However, studies have shown that while teachers often use formative assessment to clarify expectations and elicit student responses, they seldom use data from assessment to guide adjustments to teaching and learning. Veugen, Gulikers, and Brok (2021), in a survey of 96 secondary school teachers, found that within a 5-stage formative assessment cycle, the stage at which teachers were least capable was making use of assessment to adapt teaching. As they noted, this is perhaps the most important stage of formative assessment, as this is what distinguishes it from summative assessment. Teachers may have difficulties with this final phase of formative assessment due to feeling incapable and constrained by prescribed curriculum plans (Forbes et al. 2015), or unable to decide the next instructional steps (Heritage et al, 2009). Clearly, the development of an assessment tool that can aid teachers in understanding learner needs and also guiding instructional adaptation is warranted.

The present study examines evidence of learning provided by students to the teacher by means of a self-assessment sheet. It seeks to answer two questions. Firstly, how did students evaluate their own performance and that of their partners? Secondly, how useful is this information for the purposes of future lesson planning and course development?

2. Literature Review

Self-assessment in language learning has been a topic of interest for several years, particularly in the English language context. A substantial body of research exists in self-assessment, including studies done in Japanese higher education. Self-assessment has been praised for its potential to increase learners' awareness of their abilities and areas in need of improvement while promoting learner autonomy (Benson, 2001). Self-scoring rubrics, in particular, have been found to provide a structured way for learners to evaluate their language skills and set learning goals (Andrade & Du, 2005). This literature review will critically

explore the existing research on using self-assessment and self-scoring rubrics among Japanese university students studying English as a second language (ESL). A comprehensive assessment of various studies will focus on the efficacy, benefits, and potential drawbacks of self-assessment and self-scoring rubrics.

2.1 Self-Assessment in Language Learning

Self-assessment has been increasingly recognized as a crucial component of language learning, encouraging learners' autonomy and self-efficacy (Andrade & Du, 2005). Specifically, research has shown that self-assessment promotes reflection on personal performance and strategies, leading to more effective learning (McMillan & Hearn, 2009).

Scholars, including Gardner (1999), Matsuno (2009), Hale (2015), and Rivers and Ross (2020), have extensively explored the efficacy and advantages of self-assessment practices within education, particularly highlighting its implications for both students and educators. Among these, Hale (2015) provides a specific focus on the context of Japanese tertiary education, emphatically noting the positive reception of self-assessment among his students. He articulates 'that students appear to find value in the exercise' (Hale, 2015, p.1), thereby underscoring the perceived benefits and intrinsic value that learners attribute to self-assessment practice. This assertion not only underscores the relevance of self-assessment in fostering a reflective and autonomous learning environment but also suggests its potential to enhance the educational experience by aligning teaching strategies with students' self-perceived needs and competencies.

Other scholars have likewise observed the benefits of self-assessment in Japan. Murakoshi and Sato (2015) found that self-assessment helped students become more aware of their language competence and identify gaps in their knowledge. Moreover, Nakamura (2010) argued that student self-assessment, commonplace in many classrooms worldwide, not just in Japan, enables students to be more accountable for their learning.

2.2 The Role of Self-Scoring Rubrics

Self-scoring rubrics are crucial in self-assessment, providing learners with criteria to gauge their performance. Andrade (2005) defines rubrics as scoring guides to evaluate students' constructed responses. It is more likely that students will understand the expectations when they use self-scoring rubrics, allowing them to identify areas that require improvement (Panadero, 2017).

Spijkerbosch (2009) conducted a study on the influence of self-scoring rubrics on the English oral proficiency of students at a Japanese university, identifying two principal outcomes. Firstly, the adoption of self-scoring rubrics transitions the learner's role from being the examinee to assuming the evaluator's position, echoing the paradigm shift highlighted by Stobart (2006) and Black et al. (2003). This transformation not only empowers students by involving them in the assessment process but also encourages a deeper engagement with the learning material. Secondly, the nature of the assessment evolves from a low-stakes to a high-stakes context, thereby heightening the perceived importance and impact of the assessment on both students and educators.

The variability in scoring by different teachers introduces questions regarding the consistency and reliability of traditional assessment methods. In contrast, Spijkerbosch (2009) posits that a meticulously designed rubric can offer a more stable and dependable means of evaluation. The structured criteria within rubrics provide clear benchmarks for performance, thereby reducing subjective variations in scoring and enhancing the assessment's reliability. This assertion underlines the significance of well-constructed rubrics in achieving a more objective and fair assessment framework, which is crucial for the accurate evaluation of student competencies, particularly in the domain of oral proficiency.

Hawe et al. (2021) conducted an investigation that illuminated a positive correlation between the utilization of rubrics and students' proficiency in monitoring, evaluating, and regulating their learning processes. This correlation is not confined to the educational realm but extends to higher levels of academic achievement and success, a notion previously identified by Zimmerman (2000). The significance of rubrics transcends the immediate educational context, equipping individuals with the skills necessary to navigate their social interactions and laying the groundwork for a lifelong learning trajectory, as posited by Bandura (2006) and further corroborated by Kitsantas, Winsler, and Huie (2008).

The research by Hawe et al. (2021) underscores the multifaceted benefits of rubric use, highlighting its role not only as a tool for academic assessment but also as a mechanism for fostering essential life skills. The ability to self-assess and self-regulate, facilitated by rubric use, is instrumental in nurturing an individual's capacity to adapt and thrive in various contexts. This aligns with Zimmerman's (2000) assertion that self-regulatory skills acquired in educational settings are pivotal for achieving higher educational outcomes and success.

2.3 Effectiveness of Self-Assessment and Self-Scoring Rubrics

A notable factor contributing to the effectiveness of self-assessment and self-scoring rubrics is the metacognitive awareness they instil in learners. As learners actively evaluate their work, they develop metacognitive skills, enabling them to control, regulate, and assess their learning (Zimmerman, 2000). However, the accuracy of self-assessment remains a concern. Studies such as Topping (2003) suggested a discrepancy between students' self-assessment and teachers' assessment, implying the potential for over- or underestimating performance. Despite these challenges, studies conducted in Japan have shown a significant positive impact. For instance, Hirasawa (2013) found a positive correlation between self-assessment, using self-scoring rubrics, and English reading comprehension among Japanese university students.

Overall, the literature presents compelling evidence supporting the benefits of self-assessment sheets and self-scoring rubrics in ESL learning, particularly among Japanese university students. Despite potential limitations in self-assessment accuracy, the overarching impact on learner autonomy, metacognition, and motivation is significant. This paper aims to explore whether these benefits of self-assessment for learning can also be extended to benefits for instruction.

3. Methodology

3.1 Data collection

Upon approval from Meijo University's ethics committee, data were collected from 125 first-year students divided into eight classes, who were completing an English discussion course at a large university in central Japan. The data were initially gathered as part of an action research project into the application of the Common European Framework of Reference for Languages (CEFR) descriptors to enhance learning-oriented assessment (see Wicking, 2024). This paper examines a smaller subset of that data for the purpose of focusing on student self-evaluation. Namely, we examine the self-evaluation sheets completed by students after they had undergone eight weeks of instruction covering discussion skills and had also completed a transcription task.

The transcription task was conducted in two stages. The first stage entailed students having a group discussion and recording the audio. The second stage involved transcribing the audio of the discussion and doing an elementary discourse analysis of the content by applying annotations to the script. Once the task was completed, students filled in a self-assessment sheet with a rubric and three open-ended question prompts (see Appendix A).

These sheets were collected by each of the eight teachers in the discussion course and digitized. Descriptive statistics were used to analyze the data from the rubric, while the comments underwent thematic analysis.

3.2 Thematic Analysis

One approach that has been effective for analysing students' self-assessments is thematic analysis. Braun and Clarke (2006, 2014) proposed a flexible method for identifying, analysing, and interpreting patterns (themes) across qualitative data. It involves six stages: familiarising oneself with the data, generating initial codes, searching for themes, reviewing themes, defining and naming themes, and producing the report. Its flexibility and accessibility make it a valuable tool for analysing self-assessment data in various contexts, including Japanese university English students.

Several studies have applied Braun and Clarke's (2006, 2014) thematic analysis to explore students' perceptions and experiences in language learning self-assessment. For instance, Watanabe and Swain (2007) applied this method to investigate Japanese students' attitudes toward self-assessment. Their research found themes related to students' anxieties, lack of confidence, and positive attitudes towards learning autonomy and self-improvement. Hirasawa (2013) extended this work, focusing on using self-scoring rubrics among Japanese University English students, in a study which revealed themes, including the perceived helpfulness of rubrics in identifying strengths and weaknesses and the challenges in accurately assessing one's abilities. Smith (2015), meanwhile, highlighted the role of Japanese education's social and cultural context in shaping these self-assessment experiences. Themes of modesty, conformity, and hierarchical relationships were identified as influential factors in the self-assessment process. On a broader scale, Panadero, Brown, and Strijbos (2016) conducted a meta-analysis, concluding that thematic analysis can highlight patterns and inconsistencies in self-assessment, providing educators with insights to inform instruction and assessment design.

In conclusion, the existing literature indicates that Braun and Clarke's (2006, 2014) thematic analysis offers a robust and flexible framework for exploring self-assessment practices, particularly in the context of Japanese university English learners. However, as with all research methods, it is not without limitations; the subjectivity involved in interpreting themes may lead to researcher bias, and its inherent flexibility may make it challenging to replicate studies. Nonetheless, its ability to offer rich, detailed, and complex accounts of data makes it a valuable tool in the realm of self-assessment research.

4. Analysis of results

Questions on the self-assessment sheet were divided into two categories: those that allowed students to answer with checkbox-style responses and those that allowed students a chance to offer free responses.

4.1 Checkbox responses

The first checkbox response was a rubric for self-assessment. All ten skills covered in the course were listed, along with a description written in English and Japanese. Students were instructed to assess their ability in each skill by checking a box that stated, "I can't do this yet", "I can do this with difficulty", or "I can do this easily." The results can be seen in Table 1.

Table 1
Students' Self-assessment of Discussion Skills

Skill	Description	I can't do this yet	I can do this with difficulty	I can do this easily
Giving an opinion	Give a personal opinion and ask for others' opinions.	7	42	69
Follow-up questions	Ask follow-up questions to check that I have understood what someone intended to say, and to get further information.	17	70	31
Agreeing	Express my level of agreement politely and appropriately.	13	35	70
Disagreeing	Express my level of disagreement politely and appropriately.	29	55	32
Giving reasons and examples	Make my opinions and reactions understood, giving brief reasons and examples.	11	61	45
Asking for clarification	Ask someone to clarify or explain further what they have just said.	31	57	30
Active listening	Use active listening expressions to help keep a discussion going.	14	39	65
Repeating key information	Repeat back part of what someone has said to confirm mutual understanding and help keep the development of ideas on course.	19	45	54
Hesitation devices	Use hesitation devices to avoid awkward pauses and help keep a discussion going.	45	54	19
Circumlocution	Use a simple word or phrase to convey a concept when the most accurate word is unknown and invite "correction".	39	58	21

Notes.

(1) N=125.

(2) Some boxes were left blank, so not all totals add up to 125.

In the second checkbox response, students were asked, "Which skills do you most want to improve?" They were permitted to check as many boxes as they wished. The ten skills from which they could choose were the following: expressing an opinion (OP); asking follow-up questions (FUQ); expressing agreement (AGR); expressing disagreement (DIS); providing reasons and examples (R/EX); asking for clarification (CLAR); active listening (ACT); repetition of key information (REP); using hesitation devices (HES); and using circumlocution techniques (CIR). The results can be seen in Table 2.

Table 2*Skills That Students Reported They Most Wanted to Improve*

Skill	OP	FUQ	AGR	DIS	R/EX	CLAR	ACT	REP	HES	CIR
No. of students	30	64	15	54	39	51	31	37	40	43

Note. The “No. of students” represents the number of students who checked the box to indicate that this is a skill they wanted to improve.

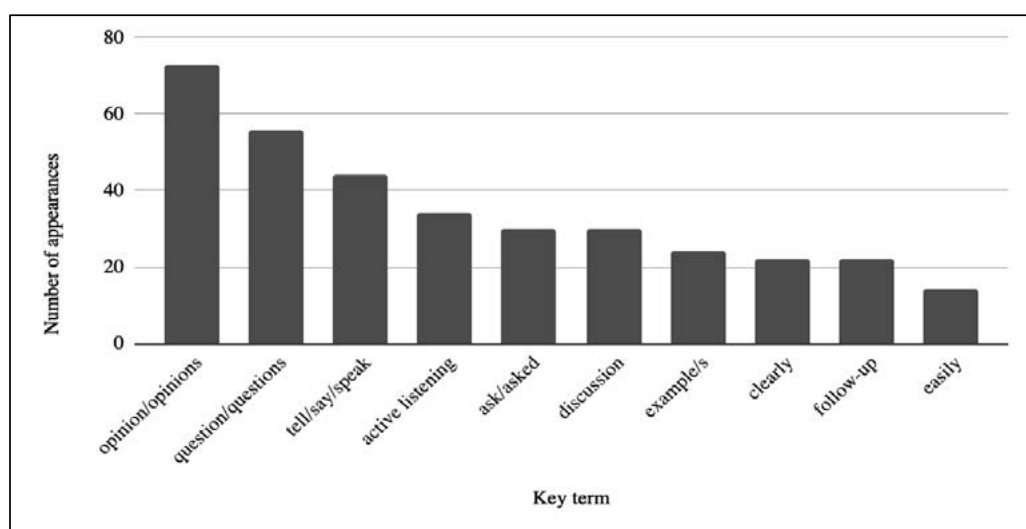
4.2 Free responses

In addition to the two checkbox response questions, students were also asked three free response questions: (1) What did your partners do well in this discussion?, (2) What did you do well?, and (3) What will you do to improve these skills (referring to the aforementioned ten skills) before your next speaking discussion test?

For these questions, an approach of identifying key terms that appeared repeatedly throughout the responses was used in the hopes of identifying major themes that might be present in the sentiments and attitudes of the participants. The most common key terms used in the responses to each question and the frequency with which they appeared can be seen in Figures 1-3. It should be noted that for the purposes of this analysis, articles, pronouns, prepositions, conjunctions, and words present in the questions themselves which may be deemed necessary by respondents to provide a complete answer (e.g. partners, well, could, improve, etc.) were discounted.

Figure 1

Most Common Key Terms from Responses to Question 1: What Did Your Partners Do Well in This Discussion?

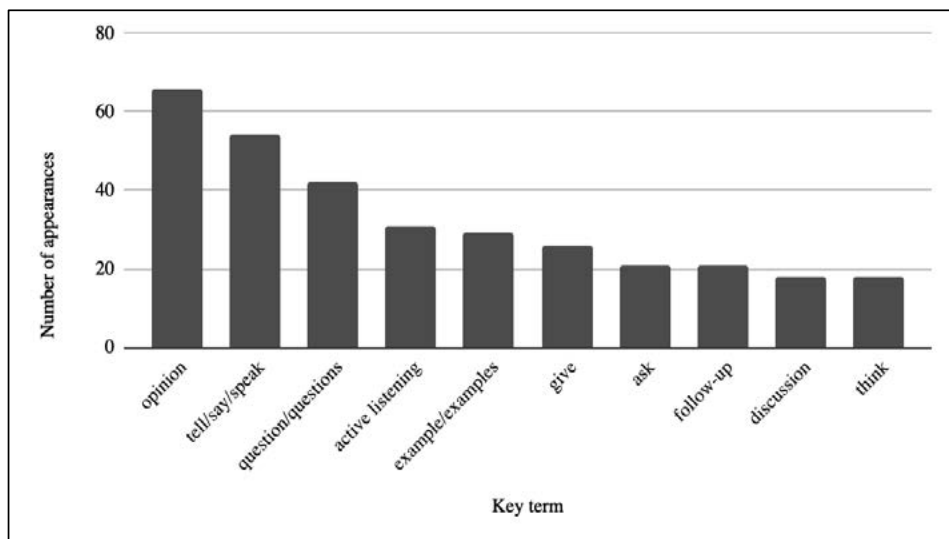


From the textual analysis, it's clear that when students evaluate the positive aspects of their partners' skills, the common themes revolve around communication skills, active listening, expressing opinions, using questions, creating a positive and supportive discussion atmosphere, and putting in effort. Some specific themes and examples extracted from the text can be seen in Table 3.

Table 3*Examples of Comments Evaluating What Partners Did Well.*

Theme	Example comment
Effective Communication Skills	My partners did well to do gesture, so I can understand the content easily.
Active Listening	They did do active listening in this discussion and they answered my question in detail. And while listening they didn't stop their reaction.
Expressing Opinions Clearly	My partners said their opinion and example clearly. With a lot of example, I was able to get it easily.
Use of Questions	They asked a lot of question to excite our discussion. They gave me some follow-up questions. Therefore, we could discuss more deeply.
Positive and Supportive Atmosphere	I'm comfortable discuss with the partners because everyone reacts to my opinion with laughing and speaking eye to eye. My partners could make good atmosphere while discussion.
Putting in Effort	They tried telling me their opinions even they didn't know how to tell in English.

These themes and examples illustrate how the participants in the discussions noticed the various communication and language skills done well by their partners. The key terms of students' self-evaluation can be seen in Figure 2.

Figure 2*Most Common Key Terms from Responses to Question 2: What Did You Do Well?*

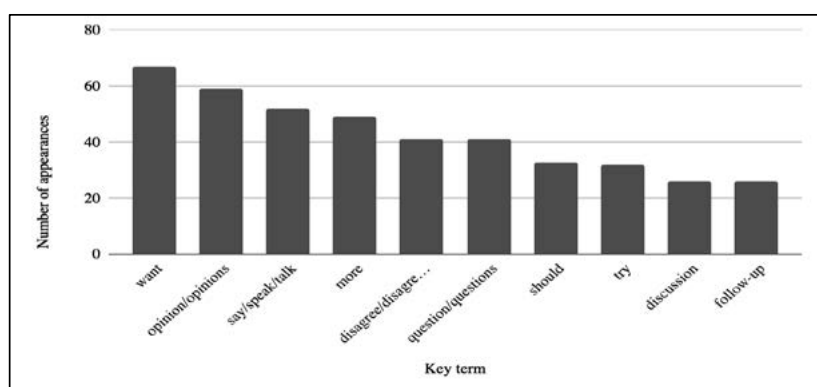
The common themes in the responses encompass various aspects of communication skills, active listening, expressing opinions, giving examples, and self-assessment. Table 4 shows the recurring themes and examples.

Table 4
Examples of Comments Made in Self-evaluation

Theme	Example comment
Effective Communication Skills	I could sum up the group's opinion and raise some examples.
Active Listening	I asked a question about the topic at the first and kept active listening. I can talk clearly because I use active listening expressions.
Expressing Opinions Clearly	I could tell my opinion and reasons easily. It is so good. I could say my opinion freely and give an example. I could disagree with my partner's opinion.
Use of Questions	I could ask them follow-up questions. For example, 'So, ~?' and 'What kinds of ~?'
Positive and Supportive Atmosphere	I tried to elicit opinions of the members.
Putting in Effort	I couldn't conversation with completely English sentences, but I could speak something. I asked for clarification to words which were unfamiliar to me.

These themes reflect the participants' efforts to engage in productive discussions, improve their language skills, and provide constructive input to their group discussions. In response to the question concerning strategies for improving discussion skills, the most common terms can be seen in Figure 3.

Figure 3
Most Common Key Terms from Responses to Question 3: What Will You Do to Improve These Skills Before Your Next Speaking Discussion Test?



The common themes in the textual data revolve around self-improvement in communication skills, language proficiency, and active participation in discussions. Some examples of these themes are shown in Table 5.

Table 5

Examples of How Students Said They Plan to Improve Their Skills

Theme	Example comment
Identifying Weaknesses and Areas for Improvement	I couldn't use several skills such as hesitation devices and asking for clarification. I will focus on being positive. I will give many opinions. I can't tell follow-up questions, because I can't understand other people's opinion clearly soon.
Vocabulary Building	I don't have enough vocabularies, so I feel difficulty to continue to conversation. I always say 'me too'. So I'd like to say other word like 'me too'.
Active Listening	I will make sure to pay close attention when my partner is speaking.
Expressing Opinions Clearly	I want to speak more clearly.
Disagreement	I can agree other's opinion easily, but I can't disagree, so I try disagreeing positively in this class. I'd like to be able to disagree every topic. I'm not good at disagreeing someone's opinion so I have to practice disagreeing politely!
Participation and Engagement	I will speak English a lot. I have to be more active.
Preparation and Planning for Discussions	I should come up with more idea before discussion. Also, I should prepare questions, as many as possible.

These themes reflect the participants' determination to enhance their communication skills, language proficiency, and active engagement in discussions. They acknowledge areas that need improvement and express their intentions to work on them in various ways, such as by using follow-up questions, improving language proficiency, and expressing opinions more clearly. Thematic analysis revealed their focus on self-improvement and their commitment to participating actively in discussions. This suggests that students would benefit from being provided with increased opportunities to use the vocabulary and expressions that they've been taught.

5. Discussion

As noted in the previous section, the self-evaluation sheet consisted of an assessment rubric to grade each skill, a set of checkboxes to identify which skills students most wanted to improve, and some free response questions to comment on their partners' performances, their

own performance, and their strategies for improvement before the next speaking test. The discussion will address the two research questions of this study: (1) How did students evaluate their own performance and that of their partners, and (2) How useful is this information for the purposes of lesson planning and course design?

5.1 Self-evaluation of discussion skills rubric

The first rubric prompted students to assess their use of ten discussion skills as “I can’t do this yet,” “I can do this with difficulty,” or “I can do this easily.” The data are very clear that there is a large gap between the top three skills which students felt were easy and the remaining skills. Those top three skills were expressing agreement (with 70 of the 125 respondents saying they could do it with ease), expressing opinions (69 respondents), and active listening (65 respondents). On the other hand, the skills that were most commonly checked as “I can’t do this yet” were using hesitation devices (45 respondents), using circumlocution techniques (39 respondents), asking for clarification (31 respondents), and expressing disagreement (29 respondents).

When looking for themes in these responses, a few key strands stand out. Firstly, students seem to evaluate themselves quite well on skills that maintain a supportive atmosphere in the discussion. This is evidenced by the extremely high number of students who chose “I can do this easily” for expressing agreement and active listening, which both display engagement, interest, and of course agreement. This could suggest that for these students perhaps the social aspects of the discussion are more important (or come easier to them) than producing the actual language content.

In terms of lesson planning and course design, the feedback provided in these responses suggests that there is a constructive alignment between course goals and students’ attitudes toward discussion. Teachers in the discussion course expend a lot of effort in encouraging students to create a safe and supportive atmosphere when speaking together. They are told not to worry about accurate grammar or expressions and are not assessed on grammatical form. As many students evaluated themselves positively on the social aspects of discussion skills, it’s an indication that teaching practice in this area has been effective.

Conversely, students rated themselves relatively low on their ability to use skills that could be viewed as admitting and drawing attention to their lack of understanding or general speaking ability. This doesn’t necessarily entail that students are actually poor at these skills, due to Japanese students’ tendency to underrate their own performance (Matsuno, 2009), but just that they perceived these skills were poor in relation to the other skills. Using hesitation devices and circumlocution techniques both essentially admit to the listeners that the speaker is unsure of what to say next or how to say it. Similarly, asking for clarification signals to the speaker that the listener does not understand what has been said. The inability or resistance to using these specific skills could be connected to ingrained cultural practices: it is common for Japanese speakers, even when speaking Japanese, to avoid revealing a lack of understanding (Miller, 1995).

This information is useful for course teachers in lesson planning. Admitting a lack of understanding and asking for clarification is an essential skill in discussion. What remains unclear are the reasons why students believe they “can’t do this yet”. It could be that they have not mastered the language forms needed, or that they lack the psychological disposition which supplies them the courage to admit their failure to understand. Moreover, as this discussion was an assessment task that contributed to the students’ final grades, students may have believed that drawing attention to their lack of comprehension would harm their scores. Either way, the fact that many students feel they are not able to do this provides impetus for teachers to spend more time developing this skill in class. This is also particularly useful information because it is not something that is easily observed by a teacher during the task. If

a student smiles and nods in a discussion task, it is unclear whether they have actually understood or not, a cognitive dissonance that we should seek to correct (Li & Zhang, 2023). The self-assessment on these sheets thus provided an insight into skill attainment which would not likely be evident through teacher observation alone.

Another major theme that can be drawn out of these responses is what might be called a contribution-interference divide. The skills that students felt most comfortable using all lend themselves to keeping a conversation flowing smoothly in one direction and even building on it. The top three most highly evaluated skills—expressing agreement, active listening, and expressing an opinion (especially when it is directly asked of a student, which was often observed in the transcript data)—are all examples of this.

In contrast, the lowest-rated skills, in a practical sense, all interfere with the natural flow of the conversation. Some, for example using hesitation devices and circumlocution techniques, may slow down or impede the progress of the discussion. Others, such as asking for clarification or expressing disagreement, interrupt and may even redirect it. This divide again shows that students may prioritize or perhaps simply feel more comfortable using certain skills due to the social aspects of maintaining the flow of the discussion. This too, may also be attributed to cultural tendencies of Japanese speakers, who tend to view discussion as a collaborative building project and therefore often prefer not to raise counterpoints or express contrary views (Miller, 1995).

One major consideration to be made here, however, is the design of the course itself. With the exception of expressing disagreement, all of the lowest-rated skills are taught in the second half of the course, with hesitation devices and circumlocution techniques taught in the final two lessons before the discussion test on which these sheets are based. The exact opposite is true of the highest-rated skills, with expressing opinion and agreement both being taught within the first three lessons of the course and only active listening being taught in the second half of the course. Taken in this light, student ease or difficulty with these skills could simply be due to the amount of time and practice they've had with the skills leading up to the test.

This insight from the student feedback provides a number of options for how teachers could adapt their lessons. One option could be to rearrange the order in which skills are taught. Front-loading the more difficult skills (such as hesitation devices and circumlocution) at the beginning of the course would allow increased time for practice and development. Another option is to create course materials that direct students' attention to the advantages of deploying these skills in a discussion. The provision of exemplars for students to imitate may also help improve these skills. Teachers can be confident that any efforts to develop these bottom four skills will meet the felt needs of students, as they all appear in the top five skills that students expressed a desire to improve.

5.2 Free response questions

The second section of the evaluation sheet prompted students to give free response answers to the following questions: (1) What did your partners do well in this discussion?, (2) What did you do well?, and (3) What will you do to improve these skills before your next speaking discussion test?

When analyzing the responses, this study took the approach of finding the most commonly used keywords in order to identify recurring themes. The first and most obvious thing to point out is that the responses to all three questions contained many instances of the discussion skills from the evaluation rubric. When the most common keywords were separated into nouns, verbs, adjectives, and adverbs, roughly 35% of the top ten key nouns were simply these skills. Specifically, opinions, (follow-up) questions, examples, and reasons appeared commonly in responses to each of the questions. For example, one comment states,

“My partners were able to speak opinion clearly. They were able to do active listening.”, while another is “I asked for clarification to words which were unfamiliar to me.” Some responses even used the exact abbreviations shown in the checkbox section, such as, “She always do ACT when I was saying the opinion, so I could talk actively”, and “I should do more KEY and FUQ.”

There could be two potential reasons for this. The first is simply the design of the survey itself. By having the checkbox section come before the free-response section, students may have been primed to frame their free responses in terms of the discussion skills. Another explanation could be that answering using the discussion skills terminology was the clearest and most concrete way for students to evaluate their own and their partners’ ability in English, especially given the fact that during the academic year leading up to this survey they had been taught to think of discussion using this exact language.

At first glance, this may seem like a defect in the design of the survey. However, Fujii and Kojima (2016) clearly observed that one of the biggest benefits of self-scoring rubrics was that students had precise criteria on which to evaluate themselves. With this in mind, the fact that responses to these questions tended to use the terminology of the rubrics can be taken in a positive light. It demonstrates that students understood the criteria of the assessment and the goals of the course in general, and were able to intelligently evaluate themselves and their peers accordingly, articulating their strengths and weaknesses.

5.2.1 Peer Evaluation

In addition to commenting on the use of the specific discussion skills taught in the course, an important theme that ran throughout the responses to Question 1 was an acknowledgement of and praise for their partners’ ability to create, add to, and maintain a positive and supportive atmosphere, echoing the importance of the social aspects of discussion as mentioned in section 5.1. This praise can roughly be divided into two main categories: increasing clarity/comprehensibility and encouraging engagement.

Students generally noticed and appreciated their partners’ efforts to communicate with clarity. One student wrote, “They agreed and disagreed clearly”, while another stated, “My partners said their opinion and example clearly. With a lot of example, I was able to get it easily.” This appreciation of clarity extended beyond verbal language to paralinguistic support for understanding, such as gesture: “My partners did well to do gesture, so I can understand the content easily.”

Students reported that their partners made an effort to encourage engagement in the discussion, with one commenting, “My partners could make good atmosphere while discussion.” This was done through the use of various strategies, such as asking questions: “They asked a lot of question to excite our discussion”, and “They gave me some follow-up questions. Therefore, we could discuss more deeply”. Another strategy was active listening: “She always do ACT when I was saying the opinion, so I could talk actively”. This includes the use of laughing and eye contact: “I’m comfortable discuss with the partners because everyone reacts to my opinion with laughing and speaking eye to eye.” Students valued their partners’ attempts to cultivate an environment conducive to discussion. This was seen as much more important than language form, such as accurate grammar or advanced vocabulary.

The data suggest that the peer evaluation section was useful for learning, as it prompted students to recognize those instances of practical skills that help to create a good quality discussion. However, it is not apparent whether this section provided any useful feedback for the teacher. By only asking about the positive aspects of their partners’ performance, no information was given about where weaknesses lie. It may be possible to perform a qualitative content analysis of the responses to determine what instructional steps

need to be taken, but this would probably be impractical for many teachers due to time and resource constraints.

5.2.2 Self-evaluation

In response to Question 2, which asked students to evaluate their own performance, students almost universally focused on their use of the specific discussion skills taught in the course. Five of the top ten most common nouns used in the responses to Question 2 are words taken directly from the labels given to these skills—opinion, question, example, reason, and expression. One student noted, “I asked a question about the topic at the first and kept active listening”, while another wrote, “I could tell my opinion and reasons easily. It is so good.” This may indicate that students were keenly aware of their own attempts to use specific skills. This can be seen in the following comment: “I could use follow-up phrases, follow-up questions, repeat back my partner’s opinion, agreeing, disagreeing, giving examples, active listening.” Students appeared to be focused on actively implementing the discrete skills that had been taught in the course, with one commenting, “I put more techniques that I learned at class into the test than usual test.”

In addition to this tendency to discuss their performance exclusively in terms of the discussion skills from the course, students also displayed a pattern of downplaying their performance in a way that was not seen in the responses to Question 1. They did this by stating that they attempted to use some skills rather than stating that they had successfully used them, as can be seen in these comments: “I tried to elicit opinions of the members”, “I tried to ask some questions to partners. Because I wanted to know more deeply. And it was connected discussion’s times”, and “I tried to explain that I don’t know words”. The word [try/tried] is one of the top ten most commonly used verbs in the responses to Question 2. Some comments also carried an implicit display of humility, such as this comment which admits a lack of vocabulary knowledge: “I asked for clarification to words which were unfamiliar to me.” The peer evaluations expressed more unreserved praise for their partners’ strong points, while the self-evaluations were more modest in scope.

Moreover, students were much more likely to be critical of themselves in their responses to Question 2, despite the fact that the question specifically asks them to answer with what they think they did well. This cognitive awareness, whether cognitive dissonance or cognitive consonance, shows that students were engaged in meta-cognition when responding to the prompt. One wrote, “I think I should practice more to have a good discussion”, while another stated, “I couldn’t conversation with completely English sentences, but I could speak something.” As with the tendency to value social aspects of discussion over the content of utterances, this inclination towards avoiding self-praise and even self-deprecation may be a result of certain Japanese social norms, as noted by Kobayashi and Brown (2003). In addition, the inclusion of many “should” statements in the responses to this question hints at the role of self-assessment surveys in cultivating personal accountability for learning, which was argued by Nakamura (2010).

As with the peer evaluation section, this self-evaluation section also has evident value for learning, in that students applied the concepts taught during the course as criteria to assess the quality of their performance. However, it is unclear how these comments can be used to inform teaching. One reason for this is that these surveys only provide a window into what students perceive. Although this is still very important data, the study lacks any way to confirm that the student self-assessment accurately reflects students’ strengths and weaknesses as determined by other assessed skills, like in Yamashita (2012), or by comparison with teacher evaluations of the same assessment. This leaves the researchers no way to determine if students have under- or over-evaluated themselves, a problem found by Topping (2003). Without a solid basis for comparison or confirmation of the self-assessments, it would be difficult for instructors to plan lesson content around these survey

results alone. Undoubtedly, the comments are useful for confirming students' understanding of discussion concepts and gaining insight into what they value as important. However, there is very little information here that can be used to focus future lesson content.

5.2.3 Learning Strategies

Similar to the responses to Question 2, some students displayed a disposition towards negatively evaluating their performance in their responses to Question 3 as well, despite the question not calling for this. Many responses focused on what the student could not do well or what they felt they were lacking, rather than stating what they planned to do to improve. One student, for example, commented, "I couldn't use several skills such as hesitation devices and asking for clarification." Some students attempted to identify the reasons for their poor performance, such as insufficient vocabulary: "I don't have enough vocabularies, so I feel difficulty to continue to conversation"; and difficulty understanding what the other students were saying: "I can't tell follow-up questions, because I can't understand other people's opinion clearly soon". Although these types of responses are outside of the scope of the survey, they are however in line with one of the main positive benefits of self-assessment surveys as found in both Murakoshi and Sato (2015) and Panadero (2017); namely, identifying areas of weakness and gaps in knowledge. Still, while it's undeniably important to diagnose the cause of poor performance, some students did not take the next step of planning how to improve that skill. Teachers would do well to notice this and guide students towards self-regulation and learning strategies that help students to take an active role in self-improvement.

Nonetheless, many students did express a desire to use specific skills that they either couldn't use or couldn't use well during the test. In particular, expressing an opinion, asking follow-up questions, providing reasons and examples, and expressing disagreement appeared in the most commonly repeated words throughout the responses (as seen in section 4). This is somewhat of an unexpected result as many students claimed to be able to do most of these skills (expressing disagreement being the exception) with ease or with some difficulty in their responses to the checkbox-style questions. However, asking follow-up questions, providing reasons and examples, and expressing disagreement are indeed three of the top five most common skills that students expressed a desire to improve on in the same section. Students wrote, "I want to improve giving reason and example skills", "I can agree other's opinion easily, but I can't disagree, so I try disagreeing positively in this class", and "I should do more KEY and FUQ."

Generally, more students showed a desire to improve their overall conversational abilities, again relating back to some of the social themes discussed in sections 5.1 and 5.2.1. Specifically, students expressed a desire to interact more deeply with their partners. Some positive actions that students said they would take to improve include paying more attention ("I will make sure to pay close attention when my partner is speaking"), speaking more ("I will try to discuss more, and take advantage of giving reasons and examples"), being positive ("I will focus on being positive. I will give many opinions") and speaking clearly ("I want to speak more clearly"). There were also occasions when students linked improved performance to the development of discrete skills covered in the course, with one student stating, "I want to increase follow-up questions. In order to increase them, I should understand my partners' opinion clearly with asking for clarification", and another, "I should remember how to continue discussion, such as follow-up question, hesitation devices."

Lastly, there were some students who did actually provide specific examples of how they might revise their learning behavior, an essential step in self-regulated learning (Andrade & Heritage, 2018). These tactics ranged in quality and method, however most focused on putting in extra study and practice outside of class time. For example, some students promised to make use of the university's self-access language center, called the

Global Plaza (GP), and viewing online videos. “I will go to GP in Meijo and listen to conversation on YouTube.” Another plan commonly expressed was better preparation before the test: “I should come up with more idea before discussion. Also, I should prepare questions, as many as possible.” Students are given the discussion topics before the test, so this is a realistic plan.

This section was rich in data useful for lesson planning. Many students commented on their weaknesses, which the teacher can use in addition to the self-assessment rubric to decide what to focus on more in class time. Additionally, students wrote down the skills they want to use more, and so opportunities for this can be provided in future lessons. Perhaps the most useful feedback, however, is the study strategies that students said they would employ before the next test. As found in McMillan and Hearn (2008), developing strategies for improvement is one of the key goals and benefits of self-assessment surveys. Although McMillan and Hearn focused on the benefit to students’ self-directed improvement, the responses are equally important to teachers since an essential role of the teacher is to help students close the gap between current levels of achievement and the goals of the course. By monitoring students’ study strategies, the teacher can dissuade poor practices and encourage those that are more likely to result in productive learning. Identifying poor practices can also help the teacher diagnose the cause of learning difficulties in some students, and then make moves to correct them.

6. Conclusion

This study examined data collected from students’ self-evaluation sheets, confirming their positive educational value for students. These sheets facilitated student engagement in learning activities likely to lead to positive outcomes. These include assessing their own progress toward the course goals, identifying their strengths and weaknesses, and also applying concepts taught in the course to assess their partners’ performance. Based on these assessments, they identified what skills they want to improve and considered future learning strategies to achieve that end which is clear evidence of cognitive consonance. These are all effective learning practices.

In addition, this study also revealed the potential benefits that self-evaluation sheets provide as feedback for the teacher. Teachers need evidence of students’ current learning achievement in order to plan instruction that guides and supports students in advancing to the next level of attainment. This was provided in the self-assessment rubric, which showed what students believed were their strengths and weaknesses. The self-evaluation sheets also indicated the skills that students most want to improve, together with their strategies for improvement, which is valuable information for focusing future lesson content and diagnosing sub-optimal study strategies. The free-response questions concerning self- and peer assessment, however, did not prove to contain feedback that was immediately useful for lesson planning, although they did identify elements of students’ performance ability and effort which could help focus future task design.

It goes without saying that the data from these self-evaluation sheets are essentially a gauge of student perceptions of their own learning. They are not in any way a statement about objective learning attainment. Students can often misjudge, downplay, exaggerate, or misunderstand their abilities, which is cognitive dissonance; but so too can the teacher. It is beholden upon educational practitioners to make use of a rich tapestry of feedback gained from multiple sources: classroom observation, test results, homework assignments, student conferences, teaching evaluation surveys, and others. In addition to these, this study has shown that self-assessment sheets can be a viable source of feedback. These sheets can provide insight into how students feel they are progressing towards course goals and thus help to guide the development of course content and inform other pedagogical decisions.

References

- Andrade, H. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching*, 53(1), 27–31.
- Andrade, H. L., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research, & Evaluation*, 10(3), 1–11.
- Andrade, H., & Heritage, M. (2018). *Using formative assessment to enhance learning, achievement, and academic self-regulation*. Routledge.
- Bandura, A. (2006). Towards a psychology of human agency. *Perspectives on Psychological Science*, 1(2), 164–180. <https://doi.org/10.1111/j.1745-6916.2006.00011.x>
- Benson, P. (2001). *Teaching and researching autonomy in language learning*. Longman.
- Black, P., Harrison, C., Lee, C., Marshall, B., & William, D. (2003). *Assessment for learning: Putting it into practice*. Open University Press.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Braun, V., & Clarke, V. (2014). What can “thematic analysis” offer health and wellbeing researchers? *International Journal of Qualitative Studies on Health and Well-being*, 9(1). <https://doi.org/10.3402/qhw.v9.26152>
- Forbes, C., Sabel, L., & Biggers, M. (2015). Elementary teachers’ use of formative assessment to support students’ learning about interactions between the hydrosphere and geosphere. *Journal of Geoscience Education*, 63(3), 210–221. <https://doi.org/10.5408/14-063.1>
- Gardner, D. (2000). Self-assessment for autonomous language learners. *Links & Letters*, 7, 49–60.
- Hale, C. C. (2015). Self-assessment as academic community building: A study from a Japanese liberal arts university. *Language Testing in Asia*, 5(1), 1–15. <https://doi.org/10.1186/s40468-014-0010-0>
- Hattie, J., & Clarke, S. (2019). *Visible learning: Feedback*. Routledge.
- Hawe, E., Dixon, H., Murray, J., & Chandler, S. (2021). Using rubrics and exemplars to develop students’ evaluative and productive knowledge and skill. *Journal of Further and Higher Education*, 45(8), 1033–1047. <https://doi.org/10.1080/0309877X.2020.1851358>
- Heritage, M., Kim, J., Vendlinski, T., & Herman, J. (2009). From evidence to action: A seamless process in formative assessment? *Educational Measurement: Issues and Practice*, 28(3), 24–31. <https://doi.org/10.1111/j.1745-3992.2009.00151.x>
- Hirasawa, A. (2013). Self-assessment and oral performance ability: The case of Japanese learners (ED572334). *Institute of Education Sciences*. <https://files.eric.ed.gov/fulltext/ED572334.pdf>
- Kitsantas, A., Winsler, A., & Huie, F. (2008). Self-regulation and ability predictors of academic success during college: A predictive validity study. *Journal of Advanced Academics*, 20(1), 42–67. <https://doi.org/10.4219/jaa-2008-867>
- Kobayashi, C., & Brown, J. D. (2003). Self-esteem and self-enhancement in Japan and America. *Journal of Cross-Cultural Psychology*, 34(5), 567–580. <https://doi.org/10.1177/0022022103256479>
- Li, D., & Zhang, L. (2023). Exploring EFL learners’ cognitive/emotional dissonance in content-based foreign language instruction: An ecological perspective. *System*, 114. <https://doi.org/10.1016/j.system.2023.103019>

- Matsuno, S. (2009). Self-, peer-, and teacher-assessments in Japanese university EFL writing classrooms. *Language Testing*, 26(1), 75–100.
<https://doi.org/10.1177/0265532208097337>
- McMillan, J. H., & Hearn, J. (2008). Student self-assessment: The key to stronger student motivation and higher achievement. *Educational Horizons*, 87(1), 40–55.
- Miller, T. (1995). Japanese learners' reactions to communicative English lessons. *JALT Journal*, 17(1), 31–50. <https://jalt-publications.org/files/pdf-article/jj-17.1-art2.pdf>
- Nakamura, Y. (2010). Self-assessment in autonomous language learning: A practical guide. *Studies in Self-Access Learning Journal*, 1(3), 238–253.
- Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, 8, Article 422.
- Panadero, E., Brown, G. T., & Strijbos, J. W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review*, 28(4), 803–830.
- Rivers, D. J., & Ross, A. S. (2020). L1/L2 communication self-efficacy beliefs and the contribution of personality. *The Language Learning Journal*, 48(6), 700–714.
<https://doi.org/10.1080/09571736.2018.1441895>
- Ruiz-Primo, M. A., & Brookhart, S. (2018). *Using feedback to improve learning*. Routledge.
- Smith, S. (2015). Accuracy in speaking self-assessment among Japanese-speaking English learners and its implications. *Polyglossia*, 27, 41–55.
- Spijkerbosch, P. (2009). Rubrics in the EFL classroom: A fresh look. *Studies in Language and Literature*, 1(29), 280–297.
- Stobart, G. (2006). The validity of formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 133–146). Sage Publications.
- Timperley, H. (2014). Using assessment information for professional learning. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing assessment for quality learning* (pp. 137–150). Springer.
- Topping, K. (2003). Self and peer assessment in school and university: Reliability, validity and utility. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimising new modes of assessment: In search of qualities and standards* (pp. 55–87). Springer.
- Veugen, M., Gulikers, J., & Brok, P. (2021). We agree on what we see: Teacher and student perceptions of formative assessment practice. *Studies in Educational Evaluation*, 70, Article 101027. <https://doi.org/10.1016/j.stueduc.2021.101027>
- Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research*, 11(2), 121–142.
- Wicking, P. (2024). Enhancing learning-oriented assessment through the practical application of CEFR descriptors. In G. Birch, N. Nagai, M. Schmidt, & J. Bower (Eds.), *Putting the CEFR into practice through action research* (pp. 187–217). Springer.
https://doi.org/10.1007/978-981-97-7545-3_6
- Yamashita, H. (2012). Japanese EFL students' perceptions of their English-speaking abilities: Implications for the use of self-assessment. *The Language Teacher*, 36(6), 11–16.
- Zimmerman, B. J. (2000). Attaining self-regulation: A social-cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–40). Academic Press.

Appendix A: Self-assessment sheet

Overall self-evaluation

Name: _____

Think about your transcript task and your overall discussion ability in English and complete the following check list.

Discussion Skill	Description	I can't do this yet	I can do this with difficulty	I can do this easily
GIVING AN OPINION	Give a personal opinion and ask for others' opinions. 自分の意見を言え、他人の意見を求めることができる。	☐	☐	☐
FOLLOW-UP QUESTIONS	Ask follow-up questions to check that I have understood what someone intended to say, and to get further information. 相手が伝えようとしていたことが理解できたか確かめたり、もっと情報を得るために補足的な質問をすることができ。	☐	☐	☐
AGREEING	Express my level of agreement politely and appropriately. 同意の度合いを失礼のないように伝えることができる。	☐	☐	☐
DISAGREEING	Express my level of disagreement politely and appropriately. 意見の不一致の度合いを失礼のないように伝えることができる。	☐	☐	☐
GIVING REASONS AND EXAMPLES	Make my opinions and reactions understood, giving brief reasons and examples. 自分の意見や応答を簡潔な理由や例を加えながら相手がわかるように伝えることができる。	☐	☐	☐
ASKING FOR CLARIFICATION	Ask someone to clarify or explain further what they have just said. 相手が言ったことについて明確化のための説明を求めることができる。	☐	☐	☐
ACTIVE LISTENING	Use active listening expressions to help keep a discussion going. ディスカッションが続くようにアクティブ・リスニングの表現を使うことができる。	☐	☐	☐
REPEATING KEY INFORMATION	Repeat back part of what someone has said to confirm mutual understanding and help keep the development of ideas on course. 共通の理解を確かめたりアイデアの広がりを助けるために相手が言ったことの一部を繰り返すことができる。	☐	☐	☐
HESTITATION DEVICES	Use hesitation devices to avoid awkward pauses and help keep a discussion going. 不自然なポーズを避けてディスカッションを続けるためにヘジテーション・デバイスを使うことができる。	☐	☐	☐
CIRCUMLOCUTION	Use a simple word or phrase to convey a concept when the most accurate word is unknown and invite "correction". 最適な単語が見つからないときにも概念を伝えるための単語やフレーズをつかうことができ、訂正を求めることができる。	☐	☐	☐

What did your partners do well in this discussion?

What did you do well?

Which skills do you most want to improve? *(More than one is OK.)*

- ☐ Giving an opinion
- ☐ Follow-up questions
- ☐ Agreeing
- ☐ Disagreeing
- ☐ Giving reasons and examples
- ☐ Asking for clarification
- ☐ Active listening
- ☐ Repeating key information
- ☐ Hesitation devices
- ☐ Circumlocution

What will you do to improve those skills before your next discussion speaking test?

Teacher Evaluation

		Did not do it	Did it	Did it well
MEDIATION	The student collaborated well with the group, showing interest and empathy, even if someone expressed a different viewpoint.	☐	☐	☐
ORAL PRODUCTION	The student made a significant contribution to the discussion, explaining their ideas reasonably fluently.	☐	☐	☐
ORAL INTERACTION	The student demonstrated confidence and ability in using various discussion skills to exchange information and opinions.	☐	☐	☐
ANNOTATIONS	The student demonstrated a clear understanding of key discussion concepts in their annotated transcription.	☐	☐	☐

Teacher Comments

Overall Grade

A+

This task was completed to an excellent standard.

A

This is very good work with no major problems.

B

There are some areas that need improvement.

C

There are many problems with this work.

F

This is well below a satisfactory standard.