

classroom presentation is that it removes the “assessor effect” that is present in interviews. It also allows students to demonstrate their ability to present spoken information logically in an extended turn.

One of the disadvantages of this task is that students may not have the need to give presentations (apart from undergoing assessments), and so the teacher should decide whether this kind of assessment is really an appropriate way to assess the student’s skills. Presentations are probably most relevant in academic or specific-purpose courses, where students are likely to be required to make individual presentations outside of the language class. One other disadvantage of presentations is the nervousness and anxiety many people experience in presenting publicly. In a classroom, students may be particularly anxious about addressing their peers; in which case, they are unlikely to give their most effective performances. One way a teacher can offset this potential problem is to ask students to record presentations, which can be evaluated after the time of the performance. Teachers then have the opportunity to involve other assessors and to ensure as much agreement as possible on the criteria used to judge the performance, and on the final grade.

Self- and peer assessment

In many language programs, students are increasingly being encouraged to take more responsibility for their learning and to develop autonomous learning strategies. Teacher assessment can be complemented well by self- and peer assessment. In situations where students are not used to taking responsibility for their learning, let alone assessing themselves, teachers may need to introduce them gradually to these concepts, and help them see the benefits for their learning of assessing themselves or each other. Students can be sensitized to the benefits through class-discussions, reflection sessions, and buddy systems, where they support each other in developing their speaking skills.

Self-assessment, as the term suggests, involves students in assessing their own performances. The most efficient way for students to do this in speaking classes is to have them record performances either inside the classroom, or preferably in a lab or self-access center, or elsewhere, such as at home, and then evaluate their performance according to a simple set of criteria. Figure 12.2 is one example of the kind of template that could be used. The students insert, in the box, the smiley face that best represents their response to the question, and can also add a comment.

Depending on the age of the learners and the speaking skills being taught, teachers can modify the assessment template to highlight different features of speaking. In addition, with or without the teacher’s assistance, students

How did I do on this task?

Name _____

Date _____




			
Did I speak fluently without too many pauses and hesitations?	Comment	Comment	Comment
Did I structure what I said accurately?	Comment	Comment	Comment
Did I use grammar appropriately?	Comment	Comment	Comment
Did I use vocabulary appropriately?	Comment	Comment	Comment
Was I able to use good strategies to keep the interaction going?	Comment	Comment	Comment

Figure 12.2: Example of self-assessment rating scale.

can be encouraged to develop their own self-assessment templates and to identify different spoken discourse features where they wish to improve. For example, they may want to improve their use of turn taking, giving feedback, or using communicative repair strategies. In this way, students can compare their own evaluations of their performances with their teacher's evaluations.

Peer assessment makes use of students as assessors of each other's performances and is increasingly being used, in speaking classrooms, as a supplement to teacher assessment, enabling students to get a more rounded picture of their achievements. Peer assessment allows students who are not involved in the speaking task to become more experienced in listening to spoken English, to engage more deeply with the learning goals and expected outcomes of the speaking course, and to learn from each other. Teachers also benefit from peer assessment, as they can share the task of assessment with their students and raise their own awareness of how assessment is carried out.

There are several challenges to introducing peer assessment in a speaking class. First, students may not be familiar with the concept of peer assessment

and may have doubts about its worth. Teachers may then need to spend time discussing with their students how peer assessment works, and its benefits for speaking development, and they may need to agree on some ground rules with the class for the role they will play during peer assessment. Second, students may not be used to working with explicit assessment criteria and will need to be given some initial training on what the criteria mean, how they apply to the speaking skills students are learning in class, and how to use the criteria during peer assessment. Luoma (2004) suggests that teachers avoid linguistic terms in the criteria and concentrate instead on criteria related to the task. She also suggests that developing the criteria jointly with the students will help them understand and use the criteria more effectively.

Rating and scoring assessment tasks

The final aspects of assessment we discuss in this chapter are the issues of rating and scoring. Rating has to do with assessing the student's performance against specified criteria, while scoring involves determining what mark or grade the student should get, based on the assessment against the criteria. Thus, rating and scoring are concerned with providing an evaluation of the student's speaking performance in the form of a grade or score.

Rating and scoring are the final stages in the ongoing cycle of assessment – from test need to test development and administration, and from the performance to the assessment of the performance. Although the focus in assessment often seems to be on the tasks themselves, as Luoma (2004: 171) argues, “The validity of the scores depends equally as much on the rating criteria and the relationship between the criteria and the tasks.” In other words, task criteria need to relate closely to the tasks themselves. Likewise, assessment tasks should be closely related to the goals and objectives of the course.

In order to put the criteria into practice, the process of rating needs to be carefully considered. How tasks are rated will depend on the structure of the assessment. Teachers will need to consider questions, such as: Does the assessment need to be scored task by task? Are short responses involved? Is there equal weighting between tasks? Points given to each task or task component can be in the form of scores (14 / 20, 80 / 100, and so on), which can then be categorized into achievement bands or grades. In longer performances, rating scales rather than scores are often used. Scales may be designed for different kinds of tasks, so that a narrative task might be rated for fluency, genre organization, use of past tense, and appropriate time

sequencing, and a discussion task might be rated according to comprehensibility, turn-taking ability, feedback strategies, and variety of vocabulary.

Figure 12.3 is an example of a numerical rating scale for a discussion task. This scale assesses students on categories of language, production, participation, expression, and coherence. In each of these categories, scores can be given for different features.

	1	2	3	4	5
Language: <ul style="list-style-type: none"> • Structure and organization. • Grammar and vocabulary. • Accuracy. 					
Production <ul style="list-style-type: none"> • Fluency. • Syllable / word pronunciation. • Intonation, stress, and rhythm. 					
Participation <ul style="list-style-type: none"> • Turn taking. • Maintenance of interaction. • Feedback. 					
Expression <ul style="list-style-type: none"> • Clarity of ideas. • Quality of ideas. 					
Coherence <ul style="list-style-type: none"> • Linking of ideas. • Justification of point of view. 					

Figure 12.3: Discussion task numerical-rating scale.

Very often rating scales are formulated in terms of “can-do” statements; for example, “Can organize the structure of the text appropriately,” or “Can produce fluent utterances.” It is useful for both teachers and students to use a rating form for assessment, where the criteria are listed, and numbers or letter grades can be indicated for each criterion. Forms can also include space for comments on the performance or for noting the reasons a particular

rating was allocated. Comments are a valuable way of providing feedback to students, because they provide more information about the quality of the performance and where improvements might be made.

Discuss it

With your colleagues, discuss the criteria you use to rate your students' speaking performance. Are you required to use generic ratings (a one-size-fits-all approach to assessing their performance), or do you use different ratings for different kinds of tasks?

The scores that result from the rating of a performance against the task criteria are usually provided in the form of a number or grade. Two types of scoring are generally used in assessment: *holistic* scoring and *analytic* scoring. Holistic scoring involves providing an overall score based on the performance as a whole, thus giving a holistic impression of the quality of the performance. The advantages of holistic scoring are that it does not take up as much time as analytical scoring, it provides an adequate account of the quality of the performance and the standards reached, and it gives students a single perspective on their achievement level. However, unlike analytic scoring, it doesn't pinpoint specific areas for development. Analytic scoring involves giving separate scores for different components of the task. The benefit of analytical scoring is that it breaks down areas of strength and weakness so that students have a better idea of what aspects of speaking skills they need to work on. Clearly, however, analytic scoring takes more time, something that for teachers is often in short supply. Thornbury (2006) warns also that although analytic scoring may be fairer and more reliable because several factors are taken into account, assessors may become distracted if there are too many criteria to consider and may lose sight of the overall performance picture. He recommends using approximately four or five categories for analytic assessment.

Try it

If possible, discuss with your students what kind of scores they prefer to get (you can do this in L1 if that makes the discussion more effective). Ask them to consider the advantages and disadvantages of receiving holistic or analytic scores. With your students, develop criteria for both types of scoring, and ask them to use these criteria for self-assessment.