ADAM J. SIMPSON

Sabancı University School of Languages, Turkey

Abstract

The aims of this chapter are to present the findings of research conducted on the subject of student perceptions as to what *is* and *isn't* effective pre-oral exam practice. The research indicates what students consider to be beneficial, what they find less useful and their overall feelings about oral assessment, while also highlighting the key differences in what we as teachers perceive as beneficial and what learners actually feel is helpful to them when preparing for a speaking test. Also discussed is the concept of awareness, in terms of the degree to which activities such as presenting grading criteria and grading each other's speaking practice actually help learners to enhance their performance.

Biodata

Adam has been living and teaching in Turkey for almost ten years, all of that time spent in the tertiary education sector in universities in Istanbul. His interests include corpus linguisitics, having been involved in the development of the Sabancı University School of Languages vocabulary syllabus, and in the assessment of oral skills.

Practice Makes…?

One vivid memory from my time as a student was a quirk of a fine lecturer of mine, who would let us know that our performance in presentations was inadequate by writing '5P' on a piece of paper and hand it to us. The term 5P referred to '*plenty of practice prevents poor performance.*' If you received a 5P, it meant that your performance wasn't up to scratch and that you would have benefited from having practiced more. I'm happy to report that I was never the recipient of such a note. The five Ps can of course be applied to any number of contexts throughout our lives, but what do they mean in language learning, and in the assessment of speaking in particular? In what situations do we want our students to practice speaking, and what are the implications on the testing of a learner's spoken skills. In this

chapter I will be exploring the notion of learners practicing the assessment format for speaking, what practice they are exposed to and their perceptions of how this did or did not benefit them in their exam performance. The findings will be drawn from the responses given by students currently attending the Sabancı University School of languages preparatory English program.

Difficulties in assessing speaking

When it comes to the assessment of speaking, a clear dichotomy exists. On the one hand, what assessors are hoping to see is a genuine, typical speaking performance from the learner. On the other lies the fact that any real expectation that the performance will be authentic is, quite frankly, ridiculous. As Underhill (1987:45) notes, '*our inherited attitudes to tests, and the way they are usually conducted, hold learners away from (assessors) at arms length.*' For this reason, familiarity with the test format is essential and should be one of the main caveats of the practice activities. Never the less, where do we draw the line between facilitating fair and reasonable test preparation and maintaining a premise of authenticity and unrehearsed performance?

The difficulties in assessing are myriad. Thornbury (2005: 124) notices the problems of assessing speaking from the perspective of the assessor; '*it considerably complicates the testing procedure, both in terms of its practicality and the way assessment criteria can be reliably applied… Moreover, different testers may have very different criteria for judging speaking, differences that are less acute when it comes to judging writing or grammar knowledge, for example.*' Knight (1992:294) reiterates, suggesting that, '*difficulties in testing oral skills frequently lead teachers into using inadequate oral tests or even not testing speaking skills at all.*' With issues such as these, what inevitably is required is a compromise that allows for a standardised approach, both in the administering of the test and in the assessment of those taking it. In order to achieve such standardization, at my university the tests the students are required to take are proficiency tests, defined according to Luoma (2004:3) as '*examinations that are not related to particular learning courses but, rather, they are based on an independent definition of language ability.*' The independent definition in question is that of the Common European Framework of Reference for Languages (CEFRL). So, having noted that assessment must negate the problems highlighted by Thornbury and

Knight by using a framework of reference as a basis for grading criteria, what other practicalities should be considered?

Practicalities of assessing spoken English

One of the key factors of a test is that it must be both practical *and* efficient. Luoma (2004:39) notes that group assessment, despite clear deficiencies, display these factors and that '*they can support learning quite well especially if learners also participate in the assessment process.*' This notion of learner involvement is a key element of the research described in this chapter and as such will be discussed in detail later. In declaring group or pair assessment to be suitable for these reasons, the drawbacks of such methods must also be considered if they are to be implemented with a sufficient degree of success. Norton (2005:287), Thornbury (2005:126) and Underhill (1987:46) are among those who note how the pairing or grouping of candidates may impact upon the language sample produced and could therefore affect the assessment process. In terms of the personalities of those being assessed, Underhill (1987:46) notes that there is, '*danger that a discussion/conversation technique will reward extrovert and talkative personalities rather than those who are less forthcoming.*' Thornbury (2005:126) also observes that the personality of one student can have an effect on others. The nature of the assessment task, the collaborative group discussion, leads to certain issues which must be considered when assessing individual performance. He notes that '*the performance of one candidate is likely to affect that of the others, but at least the learners' interactive skills can be assessed in circumstances that closely approximate real-life language use.*' Never the less, it is this notion of real-life use that Underhill (1987:45) is talking about when he notes that, although it is a perfectly natural everyday occurrence, discussion is '*one of the hardest (techniques) to make happen in the framework of a language test.*' Luoma (2004:39) expands further on the issue, stressing the importance of the length of the group assessment, noting '*all the participants must have a chance to talk for a sufficient length of time so their performance can be assessed.*' It is for reasons such as this that certain types of exam practice should occur, in order to ingratiate the participants into the assessment process.

The assessment process

Students are asked to respond to a set of discussion prompts on one theme that relates to the course material covered prior to the exam. For the oral assessment, the subject matter of the tasks relates

to the units covered in the course book. Students are informed that the subject matter of the exam can be taken from any of the appropriate units. For each oral discussion exam, between two and four separate themes with related discussion prompts are devised by the level instructors. During the exam, students may respond to all or only some of the prompts on the theme they are discussing. However, his or her total contribution to the discussion should be approximately equal in length to the contribution of the other participants. The assessment is administered by one teacher currently teaching the students and one not teaching the class. Prior to the exam, some class time is allocated to practicing, although it is not required for individual teachers to adhere to exactly the same practice techniques.

Students as participants in the assessment process

Thinking back to Luoma's assertion that learners benefit from being party to the assessment process, we must also consider the constraints that this participation must adhere to. As teachers, we are generally aware, through experience and training, what we are trying to achieve when we assess a learner. Learners, however, lack such understanding of the assessment process. Underhill (1987:22) notes that while all learners are able, to a certain extent, to evaluate their own oral proficiency, '*what they lack is the experience that enables the professional teacher or tester to compare that learner against an external standard.*' Involving the learner in the process, even if merely at the practice stage, therefore requires criteria that they are able to fully comprehend. Luoma (2004:39) stresses the importance of task clarity, noting, '*it is important that the task is sufficiently clear… for all participants.*' Therefore, knowledge of the criteria to be used in assessing the learner is vital at the practice stage. Explaining the criteria to students is consequently standard practice within my university. Luoma continues: '*If the participants also take part in the assessment, they need to know what criteria they should apply.*' Therefore, there needs to be certain similarities between the test situation and classroom activity. Thornbury (2005:125) notes how, ideally, '*the activities designed to test speaking are generally the same as the kinds of activities designed to practice speaking, there need be no disruption to classroom practice.*' Nevertheless, familiarity with the task through classroom practice needs to be separated from pure self assessment, for several good reasons. Often, pure self assessment is impracticable, not only due to the learners' lack of ability to measure to an external standard as mentioned by Underhill, but also because there is too much riding on the learner having to progress to a higher level to pass the course. Speaking assessment therefore has obvious implications for

the role of speaking in the classroom. Thornbury (2005:125) reflects on these repercussions: '*Where teachers or students are reluctant to engage in much classroom speaking, the effect of an oral component in the final examination can be a powerful incentive to do more speaking in class… the oral nature of the test 'washes back' into the coursework that precedes it.*' It therefore appears inevitable that classroom speaking activities will echo the test scenario to a great extent.

Underhill (1987:23) notes how various factors affect the ability of learners to effectively assess themselves, including *conscious* factors (an example being over-rating their ability to achieve a goal such as the aforementioned progression to a higher level) and *unconscious* factors (self-confidence and perceptiveness, for instance). Such factors make total self-assessment unworkable within many teaching contexts, mine included. Despite this, some degree of self-assessment, even if just at the practice stage, can play an important role, a role which perhaps cannot be replicated by other forms of testing. Underhill (1987:23) notes how self-assessment can be, '*introspective, where the learner is asked to reflect back on his foreign language experience and rate himself against some kind of scale; or it can be based on a specific speech sample.*' The role of effective practice should be to bring a degree of introspection into the evaluation of a learner's own proficiency, so that they may be able to better understand what steps they need to take to enhance their exam performance. This was a key factor in the undertaking of this research.
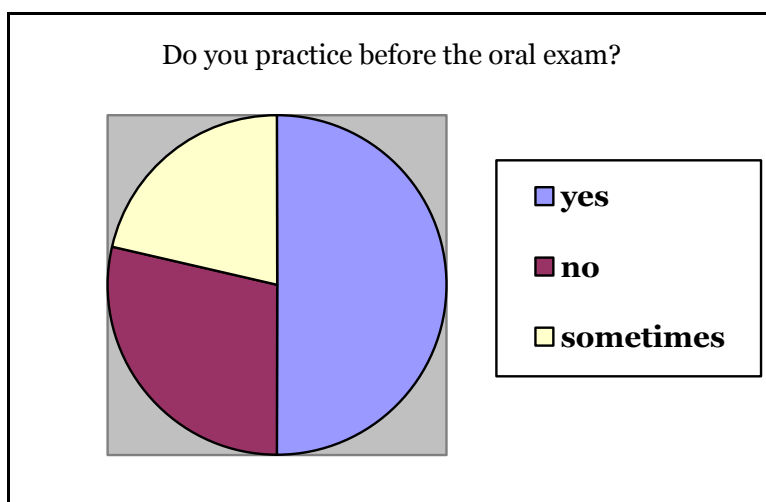
Previously mentioned was the process of presenting the assessment criteria to students prior to them taking the oral exam. This of course has the benefit of allowing participants to understand that there is a tangible scale against which they will be graded, but, in actuality, telling a group of students that they will be graded to a particular level if they are, for example, able to '*communicate in an appropriate style on a wide range of points, using complex language and correcting rare errors which may occur*' may not be as meaningful to them as it is to us as the assessor. This is especially the case when grading holistically, the previous quote being an example of our holistic grading. Thornbury (2005:127) describes holistic as being quicker and probably adequate for informal assessment, whereas analytic scoring, '*takes longer, but compels testers to take a variety of factors into account and… is probably both fairer and more reliable.*' Also noted by Thornbury on the downside is the '*wood for the trees*' phenomenon, wherein assessors get lost in the details and are unable to give appropriate feedback on the performance as a whole. Given that such phenomena could

affect the feedback given by a trained professional, what should we expect a learner to garner from being expected to assess themselves and their peers according to an analytical scoring system? Having said that, there is a degree of fairness and transparency in sharing the criteria with participants that should not be abandoned, regardless of the degree to which students are able to digest the system of grading. Having previously noted the importance of effective practice in bringing a degree of introspection into the evaluation of a learner's own proficiency, the aims of this research can be summarised as being an investigation into how the activities implemented by teachers achieve this.

Analysis of the Data

The data was collected between June and August, 2008, with fifty-six intermediate level students responding to the questions. The questionnaire can be accessed online (*http://www.surveymonkey.com/s.aspx?sm=t95VBsxXIRs8EpFN4zS4yA_3d_3d*).

The first question asked was whether or not the respondents were able to have any practice in class before an oral exam. Respondents were required to choose one answer.

Do you practice before the oral exam?

yes
no
sometimes

50.0% (28 respondents) stated yes, 28.6% (16) said no, while 21.4% (12) indicated that they sometimes practiced before an exam. The '*sometimes*' option was included as each respondent would have had at least two previous experiences of oral assessment at the time of answering, and so may well have had different experiences before each assessment. These results were intriguing as, while there is no standard practice before an oral assessment in

terms of specific classroom activities, there is always at least some form of practice. This issue of *awareness* is an important one that will be returned to later.

The second question asked the respondents which of the following occur before taking an oral exam. This list was compiled from discussions with colleagues about what activities they regularly employed when preparing students for the oral exam. Respondents were able to choose more than one option.

Action	#	%
A: The format of the questions is explained by the teacher.	20	35.7%
B: The criteria used by teachers to grade the exam are explained.	28	50.0%
C: Students can choose topics from the course book to practice.	48	85.7%
D: The range of topics that might be in the exam is explained.	24	42.9%
E: The teacher videos students and allows them to watch this video.	8	14.3%
F: Students can work in groups similar to the format of the exam.	48	85.7%
G: Students can watch other groups perform the task and comment.	8	14.3%

An '*other*' option was also offered, with one respondent offering another action, albeit one that actually appeared in the list of options above. One anomaly evident at this stage was that all of the students who had stated in question one that they received no practice opportunities proceeded to choose actions from the list above. This links to the issue of students' *awareness* of what we as classroom practitioners do with students. While we are perfectly aware that doing a particular activity in class is for the specific purpose of practicing for a test, are our students as aware of what we are trying to do? This was something that became more evident as the later questions are analysed.

The next question required the respondents to focus on one activity from the list in question two and give a reason why they regarded this particular activity as being beneficial. Respondents were able to choose more than one option.

Interestingly, the two activities that were considered the most beneficial correlated with the two from the list that occurred most often prior to exams, namely students choosing topics from the course book to practice (*C*) and students working in groups similar to the format of the exam (*F*). With regard to choosing topics from their books, comments such as '*because it*

helps to learn about topic[s] which may be in the exam' and *'because they are similar to exam format and they... prepare us'* were representative of the responses given. As for working in exam-type groups, remarks such as *'because it is [the] same style with [the] real exam'* and *'because if students can practice before exam they will feel more relax[ed] during the exam'* typified the reasons giving for preferring this task.

While not chosen as often as the aforementioned tasks, option *B*, an explanation of the criteria used during grading, was another significant choice, with half of the respondents indicating that this happened in class. *'Because it is giving you more information and you can speak longer'* and *'students can try [to] examine themselves before the real exam'* were representative of the responses given by those who chose this task. Thinking back to Underhill's and Luoma's assertions on informing those taking the test about what the test involves, it was interesting that this figure was not higher, given that it is standard practice to go through the criteria of any form of assessment with all students prior to any particular exam. Again, this raises the issue: are our students as aware as us of what we are trying to do?

The follow up question required the respondents to then do the opposite and focus on one activity from the list in question two and give a reason why they regarded this particular activity as less helpful. As with the responses given to the question asking about the most useful activities, the two activities that were considered less beneficial correlated with the two from the list that occurred least often prior to exams, namely the teacher videoing students and allowing them to watch this video (*E*) and students watching other groups perform the task and commenting (*G*). With regard to the videoing process, Luoma (2004: 39) highlights the advantages of recording the discussion, as they may be used in self reflection of speaking skills. However, any advantage overlooks the fact that students, teenage students in particular, may not like this method, as indicated by numerous responses. The idea that it heaps extra pressure on students is summarised thus, *'I think the other students' judgments about the others can make a pressure.'* Another issue pertaining to the videoing of practices is that this doesn't occur in the exam and therefore *'it is not helpful because of not [being] included in exam.'* When it comes to watching other groups, it seems that the respondents didn't always see the benefits of observing others completing the task. *'It doesn't develop our oral skills,'* noted several, while another popular response is typified thus: *'some students cannot be relax[ed] in front of other students when they are talking.'* Naturally, the age of the students, the majority being in their late teens and being extremely self-conscious in front of their peer

group, is also a factor in disliking these methods. Another, particularly thematic notion throughout the responses was exemplified in many answers to this question. '*We don't know what is good and bad*' was an answer given by many. Can we expect the students to be able to assess to any effective level using criteria, especially when they are undoubtedly very conscious of speaking out in front of their peers?

The next question asked the respondents to think of one thing that would benefit them if it were done before the oral exam and how would this help. Several obvious themes emerged from this question. The word '*practice'* appeared in almost every response, in some cases not defined any further than with this single word. However, this key concept of practicing was given greater explanation by many. The two main themes that appeared throughout the responses were those of 1) gaining experience of the exam situation by learning how to cope with group dynamics, and 2) gaining awareness of the possible exam subjects. Responses exemplifying the former were remarks such as '*practice which [is] like [the] oral exam can give some experiences before the exam*' and '*practicing for oral exam will be benefit for us because we can get some experience like the exam.*' For the latter, responses such as '*the teacher can help us about [the] topic, therefore students can learn and they can be successful in the oral exam*' was typical of numerous responses. These suggestions are rooted in the types of activities already being employed by teachers, while reinforcing the notion that pure self-assessment is impracticable: if the students had their way they would know exactly what they would be talking about and with whom. The consequences in terms of being able to assess a natural, true-to-life example of a student's oral ability in such a situation need no explanation.

The next issue the respondents were asked to consider was one thing about the exam format that doesn't help their performance. As with the previous question, two prominent themes emerged, the first relating to those doing the assessing and second to the topics that they would have to discuss in the exam. For many respondents, having teachers that they are not familiar with is a cause for concern: '*students can see the trainers during the exam - it makes them under stress.*' Furthermore, some felt that different assessors would bring about different grades: '*including different teachers in every class because their grade is very different for every student.*' This again relates to this issue of students simply not being as aware of everything that goes on in terms of assessment, i.e. that assessors will be working with a set of criteria. The second theme received many responses such as '*not knowing the topic before the*

exam makes [the exam] more difficult.' This issue of knowing the topic again reflects the impracticalities of self-assessment in this context, i.e. it falls on the wrong side of the line between facilitating fair and reasonable test preparation and maintaining the premise of authenticity and unrehearsed performance.

The next question asked the respondents how they feel about the oral exam, on a sliding scale from the lowest of not positive (1), to OK (3) through to very positive (5). The responses break down as follows:

not positive (1)		*OK (3)*		*very positive (5)*
30.8% (16)	7.7% (4)	53.8% (28)	7.7% (4)	0.0% (0)

The mean response to the question was 2.38, indicating an average of slightly less than Ok. Four respondents did not answer this question. The respondents were then asked to justify this in the following question, explaining why they feel like this about the oral exam. Again, some clear themes could be identified from the responses given, linking in this case to the aforementioned notion that the group discussion format favours particular personality types at the expense of others. The words *stress* and *anxiety* featured regularly among replies from the 'not positive' end of the scale, as well as the fear of making mistakes in front of their peers: '*When student[s] make a mistake in one thing, they can lose concentration very quickly and their grade can decrease because of that.*' At the other end of the scale were responses such as '*it helps us to see our oral ability*', '*I think it will be good because I trust myself and my friends*' and '*it's useful to speak fluently.*' Although the mean average veers towards the less positive end of the scale, there are significant numbers of responses at both ends of the scale, indicating that feelings about the oral exam are more based on the individual personality of the respondent.

Conclusions

Students tended to find the tasks that they had been exposed to most frequently to be the most useful when practicing for an oral exam. Conversely, those tasks to which they had received less exposure were deemed to be the least useful. So, do they benefit from activities that they are repeatedly exposed to or are there other reasons for these responses? Students are possibly

benefiting from the *washback* effect, whereby teachers' classroom practice is influenced by the means of assessment, and that the activities most often undertaken are by default, the most 'beneficial'. Furthermore, through trial and error teachers may use techniques such as videoing and having students watch and assess each other less frequently after receiving less than positive response to such tasks from students. Given the age group taking the exams, this is conceivable. Activities related to recreating the exam situation should, consequently, not harm the sensibilities of sensitive, teenage students.

Utilising a particular task or even explaining criteria to a group of students is no guarantee that the students will regard these actions as beneficial or even remember having done them in class. This perhaps reinforces the notion that they generally perceive as beneficial that which they do most often. Awareness of what we do, or rather lack of it, was a continuing theme throughout the responses, with answers to many of the questions asked indicating that the students do not always know what we are trying to achieve in class with any given activity.

The field of language assessment is a complex one, the complexity of which we as teachers don't fully realise ourselves until we get involved in this area of the field. With this in mind, we must appreciate the benefits of sharing assessment criteria and grading techniques with students while remembering that they may not be able to do very much with this information in terms of evaluating themselves or improving their classroom performance. Effective classroom practice when preparing students for an oral exam would, therefore, involve highlighting the fact that criteria will be used to assess the exam takers and that they will be assessed according to these descriptors without expecting them to use these to develop their performance to any great extent.

References

- Botsman, P.B. (1972) *Collective Speaking with Older Learners*, ELT Journal 26: 38-43, Oxford University Press.
- Cook. G. (1989) *Discourse*, Oxford University Press, Oxford, England.
- Council for Cultural Co-operation, Education Committee, Modern Languages Division, Strasbourg (2001) *Common European Framework of Reference for Languages: learning, teaching, assessment*, Cambridge University Press

- Farid, A. (1979) *Developing the Listening and Speaking Skills: A Suggested Procedure*, ELT Journal 33: 27-30, Oxford University Press.
- Knight, B. (1992) *Assessing Speaking Skills: a Workshop for Teacher Development*, ELT Journal 46: 294-302, Oxford University Press.
- Luoma. S, (2004) *Assessing Speaking*, Cambridge University Press, New York.
- Norton, J. (2005) *The Paired Format in the Cambridge Speaking Tests*, ELT Journal 59: 287-297, Oxford University Press.
- Thornbury. S, (2005) *How to Teach Speaking*, Longman, England.
- Underhill. N, (1987) *Testing Spoken Language*: a handbook of oral testing techniques, Cambridge University Press, New York.