

Fundamentals of ongoing assessment

Robert Croker

Discussion about Ongoing Assessment (OA) has become more prominent recently, for a number of reasons. As attention in language teaching continues to increasingly focus upon *process* as much as *product*, teachers are searching for ways to assess their students during the process of learning. As interest also grows in reflective teaching, ongoing assessment represents a means for the teacher to continually assess the efficacy of various instructional approaches and materials, and its impact upon the learner. OA also facilitates enquiry about the learners themselves, their motivation, learning style preferences and learner strategies. As teachers are expected to - and expect themselves to - give better targeted tuition, with more timely and appropriate feedback, ongoing assessment will increasingly represent a necessary classroom tool.

All teachers use OA in some form or another, to varying degrees and for various purposes. This short article hopes to provide a sound basis for understanding OA, and to help teachers become more aware of the alternatives they have regarding OA, and approach OA in a more consistent and structured manner.

After defining OA, the next section examines the relationship between course objectives and assessment criteria. The principles of OA are then considered, and this article concludes with some comments about reliability and validity.

Definitions

The more traditional form of language classroom assessment is *summative assessment* (SA), which is usually conducted at the end of the semester. The essential goal of summative assessment is administrative, to determine grades. This is usually accomplished by formally quantifying the product of learning, student knowledge, often through combinations of assignments, paper tests and perhaps formal oral interview. This form of assessment is commonly teacher-centred, the teacher deciding the form, timing, and criteria of assessment, and undertaking all the evaluation of learners. SA is principally for the teacher and administration.

In contrast, *formative* or *ongoing assessment* (OA) is carried out on a continual or periodic basis throughout the semester. Its essential goals are broader than SA, much more than the administrative task of determining grades. It has strong pedagogical goals also - diagnosing learner problems, assessing the efficacy of instruction, and to understand the learner herself.

Its emphasis is upon instructional process factors as much as instructional outcomes. It uses a range of collection procedures, both formal and informal, quantitative and qualitative. It is often learner-centred, having the potential for the learners and teacher together to determine the form, timing, and criteria of assessment, and often being undertaken by the learners themselves. OA has a broader purpose - and potential - than SA, as it goes beyond assessment to incorporate evaluation. OA is not only for the teacher but also principally for the learner. It is consistent with learner-centered instruction, and if structured appropriately, fosters learner autonomy.

Course Objectives & Assessment Criteria

Classroom instruction should be based upon course objectives. In the EFL classroom the most important objectives are language objectives, the language skills that students are expected to acquire in the classroom. Secondary objectives include strategic objectives, which refer to the development of strategies for communication and learning, and socioaffective objectives, which can refer to changes in learners' attitudes to the target language group. Whereas these first three sets of objectives are centred upon the learner, a further set of secondary objectives is focused upon instruction - process objectives, which consider the effectiveness of classroom instructional tasks, activities and materials (adapted from Genesee and Upshur, 1996).

Assessment should be tied to these course objectives. The first step is to determine course objectives in detail, and then to specify expected outcomes. The expected outcomes explicitly state for each objective what the learner should be able to know and do at each stage during the course. Some outcomes can be quite specifically stated, some generally. These outcomes form the basis of assessment criteria. Assessment criteria refer to the set of standards against which assessment judgements are made. Assessment criteria can also be either specific or general. Generally stated assessment criteria use holistic grading, or the assignment of a single overall score on the basis of an overall impression. Specific assessment criteria use analytical grading, which designate a separate score for each of a number of aspects of an expected outcome. Both holistic and analytical grading use a banding system, which sets criteria for each band (level) of competence of ability or skill.

There are advantages and disadvantages to using holistic and analytical grading. Holistic grading are rapid and easy to use. Yet they are often somewhat subjective, which reduces their reliability, and they also collapse together or conflate a number of different aspects of an outcome, which reduces validity. On the other hand, the main problem with analytical grading

is that it is time-consuming to consider each aspect. The advantage of using analytical grading is that the uneven development of skills can be addressed, essential for diagnostic purposes. Scoring tends to be more reliable and valid, as the teacher must address a number of different aspects of expected outcomes. So there is a paradox - analytical grading is more reliable and valid, but less practical; holistic grading is more practical but potentially less reliable or valid. Assessment criteria (AC) should be relevant, clear, and practical. That is, they should be relevant to the course objectives and expected outcomes on the one hand, and to the needs of the learners on the other. They should be clear and easy to understand both from the teacher's point of view and equally from the learners'. Learners that fully understand and internalized the AC have clear learning goals, which should closely reflect the course objectives. Lastly, AC must be practical. They must reflect the classroom context - the number of students in the class, the size and arrangement of the classroom, time constraints during the class, teacher preparation and evaluation time, materials constraints, and student patience and interest.

Principles

There are five sets of principles of OA - its purpose, its form, nature of data collection, degree of formality, and multi-dimensionality.

There are three purposes of OA - to focus upon learning and teaching outcomes, learning and teaching processes, and the learners themselves. That is, OA can be assessment-oriented, instruction-oriented, or learner-oriented. The focus of assessment-oriented OA is learning outcomes and teaching outcomes; that is, what the students have learnt, and what the students are having difficulty with. It is used to identify learner successes and problems, on the one hand as a basis for determining grades, and on the other to determine instructional priorities. This is the most commonly associated basic purpose of OA. The main audience for assessment-oriented OA is both the teacher and the learners.

The focus of instruction-oriented OA is the learning process and teaching process. It seeks to answer questions such as how effective instruction is, learner reaction to different teaching strategies, and learner preferences for different classroom activities. The purpose of this kind of OA is to inform instructional decisions, to facilitate more effective instruction. Most teachers are constantly doing this, consciously or subconsciously to varying degrees, and informally; the point is that OA represents a way to make this process more systematic and purposeful. Techniques include group discussions, feedback forms and questionnaires, and learner journals. The main audience for instruction-oriented OA is the teacher.

Learner-oriented OA seeks to better understand the learners in our classrooms - what motivates them, their learning style preferences, and the learner strategies that they employ. There are a number of techniques available - questionnaires, classroom observations of learners, reflective journals and learner journals. Again, teachers are constantly doing this informally, but there are techniques that allow this to be done more systematically. It can also facilitate increasing student knowledge and understanding of their own learning motivation, style preferences, and strategies, so again the main audience is both the teacher and the learners.

The second principle is the form of OA. A useful distinction can be made between response-based OA and observation-based OA. In response-based OA, the learners are assessed based upon responses that they make to certain questions or activities, set either by the teacher or by the learners in collaboration with the teacher. The learners are aware that they are being assessed, within a given framework and within a given time period. This is usually assessment-oriented OA. Response-based OA includes class tests, oral interviews, observed pair conversations, and portfolios.

By contrast, during observation-based OA, the teacher is observing the class during normal classroom activities, without setting specific tasks and activities. The learners may not be aware that they are being assessed. The instructor is observing either the whole class, sets of individuals, or certain individual only. This is usually instruction-oriented OA or learner-oriented OA, although assessment-oriented OA is also possible. Observation-based OA includes keeping classroom logs of observations of certain students, classroom observations of communicative activities, and teaching logs assessing the efficacy of various teaching strategies. Of course, there can be some degree of overlap between response-based OA and observation-based OA.

The third principle relates to the form of data collection - OA encompasses both qualitative and quantitative data collection. Quantitative data collection is data that can be easily counted - correct responses on a test, number of words spoken per minute, average length of phrase, length of pauses. Qualitative data is less easily quantified, but often more insightful - student description of learning strategies, student recollection or teacher observation of student reaction to different teaching strategies. Data is often both quantitative and qualitative, and this can be represented as a continuum. Whereas quantitative data is often considered to be more reliable, qualitative data may be more valid for the purposes at hand.

The fourth principle is the formality of data collection - OA includes both formal and informal data collection. Formal OA includes short class tests, reports and homework, assignments, and portfolios. Informal OA includes classroom observations, reflective journals,

and questionnaires. The distinction is that formal OA is designated as 'assessment' contributing in some defined way to grade determination. By contrast, informal OA is not designated directly as 'assessment' but rather as 'evaluation' - as an additional and supplementary means to further understand the instructional process and the learner.

Finally, OA can be multi-dimensional - it often relies upon more than one form of data collection. Ongoing formal paper tests may be augmented by portfolios, conferences, self- and peer-assessment, and classroom observations. OA may be multi-leveled, multi-dimensional, and longitudinal. This increases triangulation, which increases the reliability and validity of OA. It is to these issues that the next section turns.

Concerns

It is generally recognized that assessment can be reliable without being valid, but without reliability, there can be no validity. Traditional summative assessment tends to emphasize reliability over validity. Yet the varied and diverse nature of learning and instruction, and the diversity of teaching goals, means that for OA purposes, validity should be considered as important as reliability. The goal of OA should be to maximize both reliability and validity.

Reliability can be defined as 'freedom from nonsystematic error' (Genesee & Upshur 1996, p. 57). There are a number of practical ways to increase OA reliability. Assess on several different occasions, and use different methods of assessment. Make sure that the learner fully understands what is expected of them, and that they have had the opportunity to prepare. Try to make conditions as consistent as possible. Using a clear marking scheme helps the teacher to consistently compare different students. The marking scheme needs to be easy and quick to use, and easy to interpret. There is a trade-off between clarity, and the number of criteria that can be included. Therefore, the criteria being assessed at any one time needs to be limited. The advantage of OA is that assessment can occur regularly, over time allowing full coverage of a more complete set of criteria.

Validity, on the other hand, is a three-dimensional concept. Firstly, not all skills and abilities developed in the classroom can be assessed, so a representative sample should be taken. The more representative this sample, the stronger the content validity. It is necessary to carefully and systematically determine what skills and abilities are being developed in the classroom, then to design an OA system that assesses a range of them. The advantage of OA is that, as it assesses regularly, a wider range of skills and abilities can be sampled than under summative assessment. One problem of summative assessment is that it is usually limited to a short testing period, so relatively speaking only a very limited sample can be taken.

Criterion-related validity is the 'extent to which information about some attribute or quality assessed by one method correlates with or is related to information about the same or a related quality assessed by a 'different method' (Genesee and Upshur, 1996, p. 66). OA provides the opportunity to repeatedly assess, to use a number of different assessment procedures. A single language skill can be assessed by teacher interview, classroom observation, a short paper test, and a performance, and further checked using self-assessment and peer-assessment. Results should be reasonably consistent. Criterion-related validity can be increased by using a variety of assessment instruments, and checking the results across instruments systematically.

The third kind of validity is more difficult both to conceptualize and to operationalize. Construct validity means that the test measures exactly what it intends to measure; that is, that the items on a test operationalize an underlying theory of, for example, communicative competence, or lingual competence. This requires a clear theoretical and practical understanding of the underlying theory. To increase construct validity, the teacher must prepare thoroughly for OA, systematically linking an underlying theory of language to a clear set of course language objectives, testing purposes, and testing procedures. A test without construct validity is meaningless, because no meaningful conclusions can be drawn from the test results.

Conclusion

OA is a significant language classroom assessment tool. It focuses upon instructional process as much as product, teaching, learning, and the learner. It includes qualitative as well as quantitative, formal as well as informal data collection. Its multi-dimensionality bestows upon it a multi-functionality.

It is also rapidly becoming a necessary classroom assessment tool. In many learning contexts, the teacher is required to assess communicative competence as much as linguistic competence. This requires assessing not just accuracy not also fluency and complexity, not only discrete skills of grammar and vocabulary but also combined skills of speaking and listening. These discrete skills can be tested relatively passively, but the combined skills require active testing, as they emphasize performance rather than simply knowledge. OA is also suitable for criterion-referenced testing as well as norm-referenced testing. The variety of testing formats has evolved to match these broader requirements, and often these testing formats are more suited to OA than summative assessment. In short, OA is consistent with the broad trends of language testing in the classroom.

References

Allwright, D. & Bailey, K. M. (1990). *Focus on the language classroom: An introduction to classroom research for language teachers*. Cambridge: Cambridge University Press.

Genesee, F. & Upshur, J. A. (1996). *Classroom-based evaluation in second language education*. Cambridge: Cambridge University Press.

Hughes, A. (1990). *Testing for language teachers*. Cambridge: Cambridge University Press.

HTML: http://www.jalt.org/test/cro_1.htm **PDF:** <http://www.jalt.org/test/PDF/Croker1.pdf>