Dimensions of Oral Assessment

GORDON JOUGHIN, *Teaching and Learning Support Services, Queensland University of Technology, Brisbane, Australia*

ABSTRACT *An analysis of the literature on oral assessment in higher education has identified six dimensions of oral assessment: primary content type; interaction; authenticity; structure; examiners and orality. These dimensions lead to a clearer understanding of the nature of oral assessment, a clearer differentiation of the various forms within this type, a better capacity to describe and analyse these forms, and a better understanding of how the various dimensions of oral assessment may interact with other elements of teaching and learning.*

Introduction

While written forms of assessment now dominate summative assessment in higher education, oral assessment has had a long history and continues to form an important part of the assessment repertoire of universities (Brown & Knight, 1994, p. 80; Forrest, 1985, p. 3688; Hubbard, 1971, p. 93). Oral assessment is embedded in education for a number of professions, most notably medicine with its clinical assessment, law with its mooting or mock trials, and architecture with its 'design juries'. In other discipline areas, oral assessment is less common, often being considered as a form of 'alternative assessment'.

Oral assessment can be simply defined as assessment in which a student's response to the assessment task is verbal, in the sense of being "expressed or conveyed by speech instead of writing" (*Oxford English Dictionary*). As we shall see again later in this paper, the student's oral response may be combined with, or supplementary to, other forms of response such as a written paper or the demonstration of a physical skill. Assessment can be rightly considered as oral as long as a component of the student's response is verbal, and that component is being examined.

It is important to note a basic distinction between two different kinds of qualities that can be measured by oral assessment, namely:

- the student's command of the oral medium itself, i.e. the student's oral skills of communication in general or language skills in particular; and
- the student's command of content as demonstrated through the oral medium.

0260-2938/98/040367-12

Assessment in the first category is a long-established and well-accepted part of the assessment of language learning and oral communication skills. Oral assessment is unavoidable in these fields because that which is being examined is oral in nature. The literature on oral assessment in language teaching and learning is particularly extensive and well developed. This study does not seek to add to that literature but rather focuses on the second of the above categories and explores the nature of oral assessment in which the object of assessment is not the oral ability of the student but rather the student's cognitive knowledge, understanding, thinking processes, and capacity to communicate in relation to these.

Dimensions of Oral Assessment

If the distinguishing feature of oral assessment is the use of speech rather than writing as the exclusive, primary or at least a significant mode of student response, what are the principal attributes, or dimensions, of this form of assessment? The identification of such dimensions may lead to a clearer understanding of the nature of oral assessment as an assessment type, a clearer differentiation of the various forms within this type, a better capacity to describe and analyse these forms, and a better understanding of how the various dimensions of oral assessment interact with other elements of teaching and learning. The construct of 'dimensions of oral assessment' may therefore serve descriptive, analytical and research purposes.

The dimensions of oral assessment described in this paper were identified through an analysis of literature on oral assessment in higher education. This literature included specialist assessment texts (including Banta *et al.*, 1996; Brown & Knight, 1994; Nightingale *et al*, 1996; Rowntree, 1987), and 77 articles focusing on oral assessment and identified through standard educational literature searches. The attributes of oral assessment noted and/or discussed in each relevant study were listed individually. From the initial lengthy list of attributes which resulted, similar attributes were categorised according a common or underlying quality they were perceived to possess. Through this process, six dimensions of oral assessment were identified. The dimensions are listed in Table 1.

TABLE 1. Dimensions of oral assessment

Dimension	Range
1. Primary content type	Knowledge & understanding; applied problem solving ability; interpersonal competence; personal qualities.
2. Interaction	Presentation vs Dialogue
3. Authenticity	Contextualised vs Decontextualised
4. Structure	Closed structure vs Open structure
5. Examiners	Self-assessment; peer assessment; authority-based assessment
6. Orality	Purely oral vs Orality as secondary

Each dimension covers a range of practices. For four of these dimensions ('interaction', 'authenticity', 'structure' and 'orality'), the range has the quality of a continuum. The dimensions of 'primary content type' and 'examiner', however, are not continua but rather consist of more-or-less discrete categories.

The nature and range of each dimension is considered below by focusing on the variability towards each pole in those dimensions that are more-or-less continuous (i.e. 2, 3, 4, and 6) and by discussing the discrete entities in the non-continuous dimensions (1 and 5).

Primary Content Type

'Primary content type' is concerned with the object of assessment—what Rowntree (1987, p. 82) refers to as "what one is looking for, or remarking upon, in the people one is assessing". In the oral assessment literature, there is surprising agreement regarding what is tested by this form of assessment (Erhaut & Cole, 1993; Glowacki & Steele, 1992; Habeshaw *et al.*, 1993, p. 75; Kaplowitz *et al*, 1996; Levine & McGuire, 1970; Muzzin & Hart, 1985; Raymond & Viswesvaran, 1991). The learning which is assessed by oral assessment can be readily classified into four categories.

(i) Knowledge and Understanding

Here 'knowledge' is used in Bloom's sense to refer to the "recall of specifics and universals, the recall of methods and procedures, or the recall of a pattern, structure or setting" (Bloom, 1956, p. 201) while 'understanding' refers to the comprehension of the underlying meaning of what is known (OED). Knowledge and understanding are well exemplified by the first two components of the American Board of Orthopaedic Surgery's rating form, 'recall of factual information' and 'analysis and interpretation of clinical data' (Levine & McGuire, 1970, p. 65).

Oral assessment may be a preferred way of measuring a candidate's knowledge when there is a particular need to ensure that the responses are actually the candidate's, or where other forms of response are difficult, or where knowledge is being tested in association with other learning outcomes. Because oral assessment allows probing, it seems to be particularly suited to measuring candidates' understanding (Aiken, 1979, p. 3; Habeshaw *et al.*, 1993, p. 75; Lunz & Stahl, 1993, p. 174). In the context of measurements of student capabilities in professional fields, oral assessment is used to measure candidate's knowledge and understanding of facts, concepts, principles and procedures that underlie professional practice.

(ii) Applied Problem Solving Ability

The category of 'applied problem solving' in oral assessment has been variously described as "the ability to 'think on one's feet"' (Muzzin & Hart, 1985, p. 73), "the cognitive processes which constitute professional thinking" (Erhaut & Cole, 1993, p. 10), "the ability to think quickly and diagnose problems in novel situations" (Habeshaw *et al.*, 1993, p. 75), "clinical competence" (Solomon *et al.*, 1990), "problem-solving skills, application skills, interpretive skills" (Glowacki & Steele, 1992, p. 13), and "critical thinking" (Dressel, 1991). The applied nature of this category is explicitly noted by Hill, who, in the context of psychology, describes these abilities as "the capacity to apply the knowledge base" (Hill, 1984, p. 72), while Kaplowitz *et al.* (1996) in the context of medicine, claim that "the main advantage of the oral exam is that the examiner is able to ask students a series of related questions which can test not just their knowledge base, *but how well they can apply this knowledge to a clinical situation*" (author emphasis).

(iii) Interpersonal Competence

'Interpersonal competence' refers to a set of factors which Erhaut and Cole (1993, p. 10) describe as "the personal skills ... required for a professional approach to the conduct of one's work". These typically include communication or interview skills, though it must be emphasised that it is not skills *per se* but rather skills exhibited in relation to a clinical situation or problem solving exercise that are being considered in oral assessment (Gowaski & Steele, 1992: Habeshaw *et al.*, 1993; Levine & McGuire, 1970; Raymond & Viswesveran, 1991).

(iv) Intrapersonal Qualities

Oral assessment, particularly when designed for certification purposes to test candidates' suitability to enter a particular profession, is sometimes to used to measure a candidate's intrapersonal qualities. Muzzin and Hart (1985, p. 72) cite a number of personal attributes that oral assessment has been claimed to measure, including personality, alertness, reactions to stress, appearance, confidence and self-awareness. Others cite attitudes (Glowacki & Steele, 1992; Habeshaw *et al.*, 1993; Levine & McGuire, 1970) while Abrahamson (1983, p. 123) interestingly, and perhaps worryingly, includes what he refers to as "the subtle characteristic that I call his or her acceptability to the specialty for which the person is being examined".

Mixed Categories

Oral assessment may test only one of the above categories or it may encompass more than one category. The oral defence of the doctoral thesis focuses exclusively on knowledge and understanding, while clinical assessment in medicine typically encompasses at least the first three content categories (Levine & McGuire, 1970, p. 65).

Interaction

Interaction refers to reciprocity between examiner and candidate, with each acting on, responding to, and being influenced by the other. Most non-oral forms of assessment involve the student responding to a task which is presented at the beginning of the assessment process. The examiner sets the task, the student responds, and the response is then assessed by the examiner. In contrast to this, oral assessment creates the opportunity for a more complex set of interactions between examiner and student. It is this capacity that allows oral assessment to "probe a candidate's reasoning, ethics and knowledge" (Lunz & Stahl, 1993, p. 174), widely recognised as one of oral assessment's principal advantages (Brown & Knight, 1994, p. 75; Aiken, 1979, p. 3).

This interaction also gives rise to the possibility that the social interaction entailed in oral assessment may distort communication and affect both a candidate's performance and how that performance is perceived by the examiner/s (Moon, 1988, p. 2). Abrahamson (1983, p. 124) refers to this as "bias that is introduced in the interaction". This potential for distortion through interaction is the focus of numerous empirical studies (see, for example, Burchard *et al.*, 1995; Rowland-Morin *et al.*, 1991; Wigton, 1980).

The 'Presentation' Pole

At the 'presentation' pole of the interaction range, oral assessment may resemble written assessment in that a task is set, the student responds, and this response is then assessed. Oral presentations without subsequent questioning or discussion exemplify this pole (see, for example, Church & Bull, 1995; Hay, 1994; Juhl, 1996; Nightingale *et al.*, 1996, pp. 58–60). Another form is the oral assessment in which a series of pre-determined questions are posed, each followed by the student's response but with no other interaction occurring (e.g. Moon, 1988).

The 'Dialogue' Pole

The opposite end of this range is characterised by a high level of interaction between the examiner and student so that assessment takes the form of a conversation. Such interaction involves reciprocal statements by examiner and student in which each such statement includes a response to that made by the other participant. This interaction gives assessment an inherent unpredictability in which neither party knows in advance exactly what questions will be asked or what responses will be made.

Since the capacity for interaction appears to be one of the principal advantages of the oral assessment format, it is not surprising that most oral assessment include a significant interactive component. Thus interaction is a key dimension in (i) any assessment based on client interviews, whether authentic or simulated (e.g. Nightingale *et al.*, 1996, pp. 49–50 and 95–96), or (ii) in any assessment involving the probing of students' knowledge, understanding or problem-solving abilities by follow-up questions (e.g. Butler & Wiseman, 1993).

Intermediate points

An intermediate point on the interaction dimension is the formal presentation followed by questioning of the student on the content of the presentation (see, for example, Mandeville & Menchaca, 1994; Rogers & Stemkoski, 1995).

Authenticity

'Authenticity' refers to the extent to which assessment replicates the context of professional practice or 'real life'. Norman *et al.* (1985) nominate 'fidelity' or "the extent to which the simulation method resembles a real life experience" as the first dimension they use to discuss a variety of simulation methods, including oral assessment, in health sciences education. Authenticity is often discussed in the context of validity—a recurring theme in the oral assessment literature (Evans *et al.*, 1996; Hill, 1984; Holloway *et al.*, 1967; Levine & McGuire, 1970; Linn *et al.*, 1991).

The 'Contextualised' Pole

At the 'contextualised' pole, oral assessment is highly authentic, being conducted in contexts of genuine professional practice. A common example is the clinical examination in medicine which typically involves actual patients in hospital settings (see, for example, Raymond & Viswesvaran, 1991; Solomon *et al.*, 1990). Evans *et al.* (1966, p. 651) note that "oral examiners perceive the method as more nearly duplicating the usual

communication between colleagues, between doctors and patients, and between students and teachers".

The 'Decontextualised' Pole

At the opposite pole, an oral assessment may be 'decontextualised' or remote from the situation of professional practice. The oral defence of a doctoral thesis, conducted in a classroom or public auditorium and focusing on ideas abstracted from their context, would exemplify this form, as would any assessment that focuses on what Laurillard characterises as 'academic learning' (Laurillard, 1993).

Intermediate Points

Some oral assessment, while not able to be conducted in genuine settings, seeks to simulate significant aspects of authenticity. They are located close to the 'contextualised' pole and are often referred to as 'authentic' assessment since they contain many of the criteria for authenticity noted above. Simulations which culminate in oral presentations (and which themselves may be made to a simulated panel) and role plays represent common attempts to simulate 'real world' conditions in the context of assessment. Examples of such assessment abound, including simulated client interviews in law (Nightingale *et al.*, 1996, pp. 49–50); a 'real world' design project for accounting and engineering students who present their results to an industry-based panel (Rogers & Stemkoski, 1995), and mock administrative hearings in marine science (Evans *et al.*, 1992).

Structure

The dimension of structure refers to the extent to which oral assessment is based on a pre-determined, organised body of questions or sequence of events.

Numerous studies consider 'structure' as a fundamental dimension of oral assessment. Hill's report of oral methods for psychologists' licensing assessment in 39 American jurisdictions, for example, concluded that "the assessment strategies described varied on a continuum of less structure to more structure" (Hill, 1984, p. 69), while many studies have pointed out the need for structure as a means of increasing the reliability of oral assessment (Aiken, 1979; Muzzin & Hart, 1985, p. 89; Yang & Laub, 1983).

The 'Closed Structure' Pole

At the 'closed structure' pole, the assessment is tightly structured: the examiner asks a series of set questions in a given order, often following a carefully constructed set of protocols. The questions, the order in which they are asked, and the manner in which they are posed is not influenced by the student's behaviour. Moon's description of the use of oral assessment in General Certificate in Secondary Education (GCSE) Economics, in which set questions would be asked, followed by student replies, with no scope for examiners to deviate from their 'script', epitomises closed structure (Moon, 1988).

The 'Open Structure' Pole

At the 'open structure' pole, assessment follows a loosely structured agenda. The student may be given considerable freedom regarding how they present their ideas (for example, Dressel, 1991) or questions posed by the examiner may be dependent on the student's responses to previous questions, giving assessment the character of a free flowing dialogue.

Intermediate Points

An intermediate point on the closed-open structure continuum could take many forms. For example, students may be given firm criteria but have a choice of subject matter (Dressel, 1991); examiners may have a set of core questions but some latitude in posing follow-up questions (Kostick & Nixon, 1952–53); or students may have considerable freedom regarding their own response, while the examiner continues to follow a reasonably fixed pattern of response to the student's comments (Butler & Wiseman, 1993).

Examiners

The 'examiner' dimension concerns who judges the worth of the student's responses. Compared to written forms of assessment, the 'product' of an oral assessment is relatively 'public' and is accessible to whoever is present at the time of the assessment. Thus oral assessment lends itself to assessment by multiple examiners, including faculty-based panels, or, if the assessment is held in a class setting, by peers.

Self-assessment

The nature and role of self-assessment has received considerable attention in the literature of assessment in higher education (for example Boud, 1995). Self-assessment in the context of oral assessment involves procedures and issues similar to those that apply to other self-assessment formats.

Church and Bull (1995, p. 198) suggest that "there are a number of different methods for involving students in the marking process" and describe the use of self-assessment forms to identify how future oral presentations could be improved on the basis of feedback, while Joughin *et al.* (1995) describe an integrated self- and peer-assessment procedure for presentations. Trevitt, as reported in Nightingale (1996, p. 96) involves students in the assessment of their own interview performance.

Peer Assessment

Peer assessment occurs whenever other students are involved in responding to the candidate's work. This response can take a number of forms, ranging from peers being part of the audience to which a presentation is made to peers directly assessing the value of a presentation by giving written feedback based on set criteria (Joughin *et al.*, 1995). Students often contribute oral feedback during or following a presentation, which the student then incorporates in a final self-report (see, for example, Church & Bull, 1995; Nightingale, 1996, p. 22).

'Authority-based' Assessment

Most oral assessment, as in other assessment formats, does not take the form of self-assessment or peer-assessment but is exclusively conducted by teachers or others in a position of authority over the student. In oral assessment, however, such 'authority-based' assessment differs from most written assessment in (i) the frequent use of panels and (ii) the use of 'external' members of such panels.

In the discipline areas in which the use of oral assessment is most widely practiced, namely medicine, law and architecture, the use of panels is common practice. The 'design jury' in architecture, for instance, is by definition composed of several examiners (Anthony, 1991), while the literature on medical education suggests the use of pairs of examiners or larger panels is routine (see, for example, Muzzin & Hart, 1985; Raymond & Viswesvaran, 1991). Hill, in discussing standards for oral licensing assessment in psychology, simply assumes the use of panels (Hill, 1984, p. 75). In other discipline areas where oral assessment is less established, the use of more than one examiner is commonly reported (Church & Bull, 1995; Evans *et al.*, 1992; Holloway *et al.*, 1968; McDowell, 1995; Nightingale *et al.*, 1996, pp. 16–17, 49–50, 174, 184, 250; Platt, 1960; Rogers & Stemkoski, 1995).

It is quite feasible for oral assessment to utilise a mixture of examiner categories, with, for example, an element of self-assessment being conducted in conjunction with a final determination being made by a teacher, or self- , peer- , and teacher-assessment being used in combination (Joughin *et al., 1995).*

Orality

'Orality' as a dimension of oral assessment refers to the *extent* to which the assessment is conducted orally, ranging from the exclusively oral format of, for example, the clinical examination in medicine, to assessment in which the oral component is secondary to another component, for example, the oral presentation of a written paper.

The 'Purely Oral'

Many forms of assessment are purely oral, with no other medium being used. Any *viva voce*, in which the oral medium is deliberately substituted for the written, belongs in this category (e.g. Butler & Wiseman, 1993).

'Orality as Secondary'

In contrast to the purely oral examination, the oral component of assessment may be secondary to another component. The oral presentation of a written paper; the oral explanation of a physical work (e.g. an architectural design), or the oral defense of a written paper (e.g. a doctoral dissertation) are illustrative of this category.

Discussion

The dimensions of oral assessment identified in the study reported here constitute a comprehensive set of categories that probably can be applied to most instances of oral assessment to provide a thorough description of that instance. The adequacy of the dimensions for this purpose has been tested by retrospectively applying the dimensions

to some of the studies from which they were derived, as well as to studies that were not part of this initial set of cases. It is apparent that many of the instances of oral assessment reported in the literature would be considerably enhanced by addressing all six dimensions rather than the limited range of categories that are applied in most reported studies.

While the dimensions allow a comprehensive *description* of oral assessment, there are of course many important elements of assessment that are concerned with matters other than description. Thus, for example, of the five 'dimensions' which Rowntree uses to frame his exploration of assessment, only two are related to description-'what to assess' and 'how to assess' (Rowntree, 1987). His other categories, 'why assess, 'how to interpret' and 'how to respond' are concerned with non-descriptive matters. Such non-descriptive dimensions are of course vitally important, and the oral assessment literature is replete with studies addressing such issues as validity, reliability, rater bias, and student anxiety. The dimensions of oral assessment identified in the present study may help to focus studies which address such issues for while the dimensions of oral assessment are purely descriptive, the capacity of the dimensions to illuminate the oral assessment process may provide an important basis for research into these issues. For example, research into how oral assessment influences student approaches to learning suggests that the dimension of interaction may be quite significant: when students anticipate that their understanding will be probed through follow-up questions, they may tend to seek understanding of the material they are studying rather than relying on their ability to simply reproduce what they have read (Joughin, in press).

Validity and reliability are always important issues in considering any form of assessment. In the case of oral assessment, concerns about validity and reliability have been at the forefront of the discussion of this assessment format since at least 1929 when Barnes and Pressey published their frequently cited study on 'Reliability and validity of oral examinations' (Barnes & Pressey, 1929). The literature which forms the basis of the present study is replete with explicit considerations of validity and reliability issues (see, for example, Abrahamson, 1983; Aiken, 1979; Baker *et al.*, 1993; Carter, 1962; Evans *et al.*, 1966; Forrest, 1985; Holloway *et al.*, 1967; Hubbard, 1971; Kelly *et al.*, 1971; Levine & McGuire, 1970). This paper is not the place to consider these issues in detail. However, it is worth noting here how 'dimensions of oral assessment' may be useful in identifying and clarifying, for researchers and teachers alike, some of the issues relating to the validity and reliability of oral assessment.

The dimensions of 'primary content type' and 'authenticity' are particularly relevant to validity. The *content* validity of oral assessment is often considered to be one of the particular strengths of this format. As Evans and his colleagues note in the context of medical education, "Oral examiners perceive the method as more nearly duplicating the usual communication between colleagues, between doctors and patients, and between students and teachers" (Evans *et al.*, 1966, p. 651). Content validity would appear to be particularly supported when the categories of primary content type being examined are applied problem solving, interpersonal competence, or personal qualities, since instances of these can be directly observed during the examination. The dimension of 'authenticity' is also directly relevant to validity, since authentic assessment seeks to reduce, and in some instances eliminate, the discrepancy between the item of assessment and that which the item is intended to measure. When authenticity extends to performance-based assessment requiring the carrying out of genuine tasks in work contexts, the content validity of oral assessment would seem assured. The *predictive* validity of oral assessment is another matter. A student's capacity to 'think on her feet' in front of a panel of examiners may be quite different to that capacity in work settings, and even the most

authentic forms of assessment are distinguished from actual work performance by the very act of assessing that performance.

If the validity of oral assessment is considered to be one of the strengths of this format, the same cannot be said regarding reliability. If different assessors were to examine a candidate orally, are they likely to reach consistent conclusions about that candidate? Or if a single examiner were to examine a candidate at different times, will that examiner reach the same conclusion on both occasions? A consideration of the dimensions of 'interaction', 'structure', 'examiner' and 'orality' may help to clarify why the reliability of oral assessment is often called into question, while also suggesting how oral assessment can be carried out in ways that maximise reliability. Studies suggest that reliability is threatened when the 'interaction' dimension tends towards the 'dialogue' pole, when the 'structure' dimension tends towards the 'open' pole' so that there is inconsistency between the questions asked of different candidates, where 'examiners' are poorly prepared and their responses are not moderated, and where the examination is 'purely oral' rather than a combination of oral and written formats (Aiken, 1975; Baker *et al.*, 1993; Barnes & Pressey, 1929; Carter, 1962; Forrest, 1985; Hubbard, 1971, p. 97; Marshall & Ludbrook, 1972; Muzzin & Hart, 1985, p. 81). Each of these conclusions has its corollary—reliability can be improved by reducing dialogue and increasing presentation; by using a closed or predetermined structure; by using multiple and well-prepared examiners; and by using oral assessment to complement other formats.

One final comment should be made about these dimensions. They are based exclusively on teachers' or educational researchers' perspectives of oral assessment. Apart from a small number of studies on anxiety, the literature does not include studies of oral assessment from the perspective of students. A survey of students' descriptions of oral assessment may lead to a set of dimensions quite different to those generated from teachers' descriptions. Notwithstanding this important distinction, the dimensions identified from the teachers' point of view may usefully inform the process of listening to and describing the experience of oral assessment from the students' perspective.

Acknowledgements

The author is grateful to Professor Richard Bagnall, Faculty of Education, Griffith University for suggesting the usefulness of research into dimensions of oral assessment, and to Professor Bagnall and Professor Paul Ramsden, Director of the Griffith Institute for Higher Education, for their comments on draft versions of this paper.

Notes on Contributor

GORDON JOUGHIN is an education consultant to the Faculty of Law and the Department of Teaching and Learning Support Services at the Queensland University of Technology and the Griffith Institute for Higher Education in Brisbane, Australia. He is the co-author of two books on legal education, *A Framework for Teaching and Learning Law* and *Developing Print Materials for Flexible Teaching and Learning in Law*. His consultancy work focuses on flexible learning and the pedagogically effective use of technology to support learning, while his current research is on the influence of oral assessment on student approaches to learning. *Correspondence*: Gordon Joughin, Teaching and Learning Support Services, Queensland University of Technology, GPO Box 2434, Brisbane, Qld 4001, Australia.

REFERENCES

ABRAHAMSON, S. (1983) The oral assessment: the case for and the case against, in: J. S. LOYD & D. G. LANGSLEY (Eds) *Evaluating the Skills of Medical Specialists* (Chicago, American Board of Medical Specialists).

AIKEN, L. R. (1979) The case for oral achievement testing, ED 222 578.

ANTHONY, K. H. (1991) *Design Juries on Trial* (New York, Van Nostrand Reinhold).

BAKER, E. L., O'NEIL, H. F. & LINN, R. L. (1993) Policy and validity prospects for performance-based assessment, *American Psychologist*, 48(12), pp. 1210–1218.

BANTA, T. W., LUND, J. P., BLACK, K. E. & OBLANDER, F. W. (1996) *Assessment in Practice: putting principles to work on college campuses* (San Francisco, Jossey-Bass).

BARNES, E. J. & PRESSEY, S. C. (1929) Reliability and validity of oral examinations, *School and Society*, 30, pp. 719–722.

BLOOM, B. S. (Ed.) (1956) *Taxonomy of Educational Objectives: cognitive domain* (New York, McKay).

BOUD, D. (1995) *Enhancing Learning through Self Assessment* (London, Kogan Page).

BROWN, S. & KNIGHT, P. (1994) *Assessing Learners in Higher Education* (London, Kogan Page).

BURCHARD, K. W., ROWLAND-MORIN, P. A., COE, N. P. W. & GARB, J. L. (1995) A surgery oral examination: interrater agreement and the influence of rater characteristics, *Academic Medicine*, 70(11), pp. 1044–1046.

BUTLER, D. & WISEMAN, L. (1993) Viva the viva: oral assessment in Contract Law, *Legal Education Review*, 4(2), pp. 331–350.

CARTER, H. C. (1962) How reliable are oral examinations?, *California Journal of Educational Research*, XIII(4), pp. 147–153.

CHURCH, A. & BULL, P. (1995) Evaluating and assessing student oral presentations: a limited but effective role for employers in the geography curriculum, *Journal of Geography in Higher Education*, pp. 196–202.

DRESSEL, J. H. (1991) The formal oral group exam: challenges and possibilities—the oral exam and critical thinking, paper Presented at the Annual Meeting of the National Council of Teachers of English, Seattle, 22–27 November, (ED 347 527).

ERHAUT, M. & COLE, G. (1993) Assessment of competence in higher level occupations, *Competence and Assessment*, 21, pp. 10–14.

EVANS, J. C., DEAN, J. M. & CHAPAL, S. (1992) Expert witness or advocate: developing oral argument skills in the marine science student, *Journal of College Science Teaching*, December 1991/January 1992, pp. 149–153.

EVANS, L., INGERSOLL, R. W. & SMITH, E. J. (1966) The reliability, validity, and taxonomic structure of the oral assessment, *Journal of Medical Education*, 41, pp. 651–657.

FORREST, G. M. (1985) Oral assessment, in: T. HUSEN & T. N. POSTLETHWAITE (Eds) *The International Encyclopedia of Education*, pp. 3688–3690 (Oxford, Pergamon Press).

GLOWACKI, M. & STEELE, D. J. (1992) A synthesis of the research on alternative assessment methods in teacher education, paper presented at the Annual Meeting of the Mid-South Educational Research Association, Knoxville, Tennessee, 11–13 November (ED 355 257).

HABESHAW, S., GIBBS, G. & HABESHAW, T. (1993) *53 Interesting Ways to Assess Your Students* (Bristol, Technical and Educational Services).

HAY, I. (1994) Justifying and applying oral presentations in geographical education, *Journal of Geography in Higher Education*, 18(1), pp. 43–55.

HILL, D. S. (1984) Oral assessment: standards and strategies, *Professional Practice of Psychology*, 5(2), pp. 69–78.

HOLLOWAY, P. J., HARDWICK, J. L., MORRIS, J. & START, K.B. (1967) The validity of essay and viva-voce examining technique, *British Dental Journal*, 123, pp. 227–232.

HOLLOWAY, P. J., COLLINS, C. K. & START, K. B. (1968) Reliability of viva-voce assessment, *British Dental Journal*, 125(5), pp. 211–214.

HUBBARD, M. (1971) *Measuring Medical Education* (Philadelphia, Lea & Febiger).

JOUGHIN, G. (in press) Dimensions of oral assessment and student approaches to learning, in: S. BROWN & A. GLASNER (Eds) *Assessment Matters* (London, Kogan Page).

JOUGHIN, G., MCGRATH, F. & COLES, J. (1995) Cooperative peer learning in distance theological education, in: D. SEWART (Ed.) *One World Many Voices*, Vol. 2, pp. 258–261 (Milton Keynes, The Open University).

JUHL, L. (1996) General chemistry in technical education, *Journal of Chemical Education*, 73(1), pp. 72–77.

KAPLOWITZ, P. B., JENKINS, M. D. & NAIR, P. (1996) *The Oral Assessment* [http://vh.radiology.uiowa.edu/Providers/Societies/APA/GPCC/or al_Exa 12.96].

KELLY, P. R., MATTHEWS, J. H. & SCHUMACHER, C. F. (1971) Analysis of the oral examination of the American Board of Anesthesiology, *Journal of Medical Education*, 46, pp. 982–988.

KOSTICK, M. M. & NIXON, B. M. (1952–53) How to improve oral questioning, *Peabody Journal of Education*, 30, pp. 209–217.

LAURILLARD, D. (1993) *Rethinking University Teaching* (London, Routledge).

LEVINE, H. G. & MCGUIRE, C. H. (1970) The validity and reliability of oral assessment in assessing cognitive skills in medicine, *Journal of Educational Measurement*, 7(2), pp. 63–74.

LINN, R. L., BAKER, E. L. & DUNBAR, S. B. (1991) Complex, performance-based assessment: expectations and validation criteria, *Educational Researcher*, 20(8), pp. 15–21.

LUNZ, M. E. & STAHL, J. A. (1993) Impact of examiners on candidate score: an introduction to the use of multifacet Rasch model analysis for oral assessment, *Teaching and Learning in Medicine*, 5(3), pp. 174–181.

MCDOWELL, L. (1995) The impact of innovative assessment on student learning, *Innovations in Educational Training International*, 32(4), pp. 302–313.

MANDEVILLE, T. F. & MENCHACA, V. (1994) Group oral exams: exploring assessment techniques for new instructional paradigms, *Research and Instruction*, 33(4), pp. 319–325.

MARSHALL, V. R. & LUDBROOK, J. (1972) The relative importance of patient and examiner variability in a test of clinical skills, *British Journal of Medical Education*, 6, pp. 212–217.

MOON, R. (1988) Oral assessment in GSCE economics, *Research Papers in Economics, Number 14* (ED 307 199) (Institute of Education, London University).

MUZZIN, L. J. & HART, L. (1985) Oral assessment, in: V. R. NEUFELD & G. R. NORMAN (Eds) *Assessing Clinical Competence*, pp. 71–93 (New York, Springer).

NIGHTINGALE, P., WIATA, I. T., TOOHEY, S., RYAN, G., HUGHES, C. & MAGIN, D. (1996) *Assessing Learning in Universities* (Sydney, University of New South Wales Press).

NORMAN, G. R., MUZZIN, L. J., WILLIAMS, R. G. & SWANSON, D. B. (1985) Simulation in health sciences education, *Journal of Instructional Development*, 8(1), pp. 11–17.

PLATT, J. R. (1960) On maximizing the information obtained from science assessment, written and oral, *American Journal of Physics*, 29, pp. 111–122.

RAYMOND, M. R. & VISWESVARAN, C. (1991) Least-squares models to correct for rater effects in performance assessment, *ACT Research Report Series 91-8* (ED 344 947) (Iowa, The American College Testing Program).

ROGERS, R. L. & STEMKOSKI, M. J. (1995) Reality-based learning and interdisciplinary teams: an interactive approach integrating accounting and engineering technology (ED 392332).

ROWLAND-MORIN, P. A., BURCHARD, K. W., GARB, J. L. & COE, N. P. W. (1991) Influence of effective communication by surgery students on their oral examination scores, *Academic Medicine*, 66(3), pp. 169–171.

ROWNTREE, D. (1987) *Assessing Students. How shall we know them?* (London, Kogan Page).

SOLOMON, D. J., REINHART, R. C., BRIDGHAM, R. G., MUNGER, B. S. & STARNAMAN, S. (1990) Free-response formats for evaluating clinical judgment, *Academic Medicine*, 65(9), pp. 543–544.

WIGTON, R. C. (1980) Effects of student personal characteristics on the evaluation of clinical performance, *Journal of Medical Education*, 55, pp. 423–427.

YANG, J. C. & LAUBE, D. W. (1983) Improvement of reliability of an oral assessment by a structured evaluation instrument, *Journal of Medical Education*, 58(11), pp. 64–72.