

ASSESSMENT OF PUPIL ACHIEVEMENT

Distinguish between the terms 'criterion-referenced assessment' and 'norm-referenced assessment'. Describe and explain, from within your own educational practice, specific situations in which each type of assessment could be justified. Are these two approaches to assessment always mutually exclusive?

INDEX

Introduction	1
Education and Assessment.....	1
Frames of reference: Norm-referencing.....	2
Questioning past practice.....	4
Criterion-referencing.....	5
Norm & criterion-referencing: complementary or mutually exclusive?.....	7
Examples to illustrate positive & negative aspects of criterion-referenced assessment practices.....	7
Example to illustrate the need for, and potential benefits of, norm-referenced tests	8
Conclusion	9
References	11

Introduction

This assignment begins by acknowledging that assessment provides the information upon which many decisions are based, and so plays an important role in education, evoking strongly held views.

It then suggests that there is now a greater understanding of its role in, and effect on, the education of young people. This has led educators to question the widespread acceptance of norm-referenced approaches, leading many to believe that criterion-referencing may now be more appropriate.

In order to ascertain how far this is true, the distinguishing characteristics, the strengths and weaknesses, of each approach are considered. It is, however, recognised that in reality no assessment practice belongs entirely in a single referencing category.

This is followed by examples from the Argentine educational system, and my own teaching practice. These illustrate the fact that, though a criterion-referenced approach has immediate appeal in that it gives teachers information about actual learning outcomes, there are occasions when a mix, or a norm-referenced approach, would be appropriate. That it is not an 'either-or', but a 'fitness for purpose' issue.

Education and Assessment

If one accepts that the curriculum can at its best only be, to quote Lawton (1996), *a selection from the culture*, it would suggest that not only must choices be made in the process of selecting the ideas, concepts, values and skills a particular society considers valuable enough to transmit to, and expect to be further developed by, the next generation, but also that the effectiveness of the resulting educational programme must be studied - at the individual, institutional and national levels.

Assessment plays a crucial role in this process. However, as different practices usually reflect different ideological commitments, and a teacher's use of assessment will be heavily influenced by the expectations of the teaching system within which he is working, it is important to understand the relationship between changes in the socio-economic context and the ways in which these are reflected in changing assessment practices.

Having said this though, the prime purpose of assessment in education must be to support the teaching/learning process. It is the vital link between curriculum and teaching, providing feedback for pupils, teachers and others, *yielding a basis for planning the next educational steps in response to children's needs* (TGAT report in Black 1998:3) - as well as providing evidence that schools are indeed promoting the learning of their pupils. In this context, it can be defined as: *the process of gathering, interpreting, recording and using information about pupils' responses to an educational task*. (Harlen et al in Moon & Shelton Mayes 1994:273), functioning in a diagnostic or formative mode - both for individual pupils and the system as a whole - in what Stobart & Gipps (1997) call the *professional functions of assessment*.

Many writers do, however, recognise that assessment in schools has many other, sometimes conflicting, functions. Indeed, Broadfoot suggests that a basic definition of the term assessment is: *the deliberate and overt measurement of educational performance in order to provide information for purposes beyond the immediate interactive situation* (1996:6); and that the societal role of assessment can be summarised with three key words: certification, selection and control.

Whether one fully agrees with Broadfoot or not, assessment undoubtedly provides the information upon which many decisions are based. This added to the dramatic changes

witnessed in the late twentieth century, has prompted a great deal of talk about educational change in general, and the need to introduce new practices. In the field of assessment, this has led to a greater understanding of its role in, and effect on, the education of young people; and a growing recognition that assessment, as part of education, must be about promoting learning and opportunities, rather than about sorting people into social roles for society.

Frames of reference: norm-referencing

All types of assessment, regardless of what the eventual results will be used for, involve interpretation of a pupil's response against some standard of expectation. Even today, this standard of expectation is often set by the average performance of a particular section of the population or age group, because there is still a tendency to assume that comparisons must of necessity be made between individuals, and that this is a fair way of 'sorting' people into social roles for society.

Though acknowledging that distortions are inevitable when attempting to list salient points, if one were to highlight the main characteristics - directly or indirectly - attributed to the concept of norm-referencing, one could say that:

- The preoccupation of a norm-referenced system is with the public reporting of performance, measured against norms, in order that selection can take place, rather than formative feedback for pupils.
- It is based on a crucial assumption: attainment and mental characteristics are distributed across the population to be assessed.
- This second point tends to encourage students to attribute success to ability rather than effort. Those perceived as being less able will make less of an effort - their performance is, therefore, affected and reflected in the ranking. Thus, though one cannot suggest that this is necessarily a distinct characteristic of norm-referencing, it is indirectly perceived by many to be one.
- It is of little use on its own for diagnostic purposes - does not provide information about an individual's potential or his attitude towards a certain subject.
- Since ranking for selection is the major concern, norm-referenced tests or examinations must be designed to discriminate between pupils.
- Performance is measured against, and relative to, the average level of performance in the normative group.
- Improvement consists in moving up the rank order at the expense of other pupils.
- Competition rather than cooperation thus rewarded.
- Purpose: grading, ranking and comparing - it is essentially a screening device.
- Pupils who perform badly - and by implication at least half the pupils will be 'below average' - are failures.
- Pupils' life chances thus affected by results.
- Result: social reproduction and the perpetuation of social disparity. This may be a largely unavoidable outcome of assessment in general, whatever frame of reference is used. However, norm-referencing tends to 'legitimize' this outcome because it is perceived by many to be a scientifically neutral, fair, and reliable means of measuring learning. Thus, though once again not necessarily a distinct

characteristic of norm-referencing, it can be said to be indirectly accepted by many as such.

Filer (2000) stresses that public perceptions are crucial to the acceptance of systems of assessment, and that *the mass categorization and social differentiation of populations have needed to be accepted as broadly socially just, in particular by the losers in the assessment stakes* (Filer 2000:43). To understand why a norm-referenced approach is still socially accepted, even when, historically, the outcomes of mass assessment systems based on norm-referenced approaches have been economic and social rewards for some, but reduced access to educational and occupational opportunities for many, the approach must be seen as the logical extension of the theory behind intelligence testing. This suggested that the idea of norms, and the bell shaped curve, was scientific and objective and, thus, a fair system. Broadfoot, however, suggests that the theory underlying intelligence testing also served political needs, since *'intelligence testing, as a mechanism for social control, was unsurpassed in teaching the doomed majority that their failure was the result of their own inbuilt inadequacy'* (1996:35).

Given that, traditionally, assessment practices, particularly at secondary school, have been largely based on competition within a norm-referenced frame, the characteristics listed above do tend to support both Broadfoot's definition of assessment, and her suggestion that political needs are well served by the approach.

Picking out characteristics, though allowing one to see at a glance where the emphasis lies does, however, tend to distort the picture by highlighting the negative aspects. There are occasions when selection, or a degree of discrimination, are necessary and, thus, when a norm-referenced approach would not only be useful, but also fit for the purpose. Supporters of the approach, therefore, suggest that:

- Selection for various kinds of educational opportunity or career is necessary because no country believes it can afford to give every citizen all he might desire in the way of education.
- Bilingual schools may also have to 'select' and group pupils so that the teaching method – using first/second language approaches - matches both the level reached in the second language, and pupils' learning styles.
- It is cost effective and administratively convenient (though the counter-argument is that the emphasis is on the easily measured)
- It motivates (though it is also demotivating for those who 'fail')
- It prepares students for the pressures of 'real life' (even though it relies on competition and extrinsic rewards, rather than cooperation and inner satisfaction).

Despite the fact that globalization has tended to increase competition amongst nations, encouraging countries, particularly those with developing economies, to view selective, competitive assessment practices as a necessary pre-requisite for economic growth, the limitations and negative aspects embraced by a norm-referenced approach have led many to seriously question it.

Questioning past practice

Brown (in Moon & Shelton Mayes 1994), though supporting Broadfoot's argument that *selection* has been a major function of assessment, believes that in recent years the question has been raised as to whether there are not other, and more important, functions for it to fulfil. In particular, since it is part of the educational process, whether it should not have a more constructive role to play in teaching and learning.

Though writing over a decade ago, Lloyd Jones & Bray (1986) suggest that many of these questions have been raised as a result of a number of recognisable trends which have profoundly affected the way we view education in general, and the assessment process in particular. These include:

- An increasing recognition of individual differences.
- The emphasis on content giving way to an emphasis on learning process.
- Social, moral and aesthetic development now also considered important.
- Competition between pupils giving way to collaborative working.
- Major shift in priorities – focusing on the needs of the majority rather than selecting a minority.
- Pressure to raise standards of performance of all young people judged not through competitive examinations designed to select the ablest, but against descriptions of performance (criteria).

Not only is there an increasing concern about the proportion of young people leaving formal education without any record of what they have achieved, in an age when educational qualifications have become more and more important, but, if assessment is about producing useful information for teachers, students and others, norm-referencing has a further important limitation: while it enables comparisons to be made among the performances of individuals, it provides no evidence about *what* has been achieved. In addition, as deeper consideration is given to learning objectives, the negative 'backwash' effect on both curriculum and assessment procedures of high stakes examinations, has been more openly acknowledged.

Gipps (1994) believes that we need to develop a new way of thinking about assessment to deal with the issues that are emerging as a result of questioning past practice. She stresses that, as societies battle with the effects of a restructuring from an industrial to an information based society, assessment has taken on a high profile and is now required to achieve a much wider range of purposes.

Questions concerned with technical matters – **reliability and validity** – have also been raised. Messick, for example, has suggested that the traditional definition of validity: *whether an item measures or describes what it is supposed to measure or describe*, should be extended, to operate, as Gipps (1994) indicates, *with a notion of validity that relates to inferences drawn from test scores* (1994:59) - effectively extending the concept to include the potential consequences of test use. This recognises that it is not so much the test score itself that matters, but the inference that is made from it. The implication being that the responsibility for valid test use must shift to the user – suggesting, in particular, that the effects of the use of assessment outcomes on the subsequent behaviour of students, must also be considered part of the validation process.

This approach to validity, known as consequential validity, begins to address the issues of ethics in assessment by taking us *firmly into a scenario in which educational and social implications of assessment are addressed*. (Gipps 1994:66)

In an article intended to provide an overview of various technical issues in assessment, Wiliam appears to support Messick's views. He suggests that *the validity of an assessment depends on the use to which the results of the assessment are put*, and that, as a result, a test is never valid, only valid for a particular purpose. This leads him to conclude that: *validation is a never-ending process of marshalling the evidence to support the assertion that a particular assessment result has a particular meaning*. Consequently, *the onus is always on the users of the assessment information to establish that the inferences they make are warranted*. (1992:16) – making the measurable important is no longer a basis for an assessment to be considered 'valid'.

Much of this questioning has led to the acknowledgement that *any kind of function for assessment which aims to provide information which will help young people learn, or teachers to teach will require an evaluative description of what has been achieved*. (Brown in Moon & Shelton Mayes 1994:269). In other words, there has been a move towards criterion-referenced approaches replacing, or complementing, traditional norm-referenced systems.

Criterion-referencing

If the purpose of the assessment is to provide descriptions of what has, or has not, been achieved, rather than simply grading and ranking, criterion-referencing would appear to be of more use.

Though the notion of criterion-referencing can be traced back to the 1920s, Wood suggests that it was Glaser's paper published in 1963 which caught the imagination and marked *the point at which educational measurement began to detach itself from psychometrics*, (1991:83). Glaser defined criterion referenced testing in terms of its difference from norm-referenced testing. He hoped to *shift attention away from between-individual variation to within-individual variation* (Wood 1991:91) – also suggesting that measurement of achievement should be concerned with current levels of performance, rather than prediction. In the classroom context, as Marsh (1997) indicates, criterion-referencing can be seen as vital to any sequential teaching where it is essential to check that the student has achieved some prerequisite skill before going on to the next stage.

Gipps (1994:79) points out that, because criterion referencing *is concerned with an individual's growth (rather than variation between individuals)* ..., the concept of 'winners and losers' is not necessarily in evidence; as James (1998:146) says, *the assumption is that anyone who fulfils the requirements specified in the criteria, and meets the standards will have their achievement recognised*. James goes on to stress that, because everyone is regarded as having the potential to succeed, criterion-referencing is both more egalitarian and motivating than a norm-referenced system.

While once again acknowledging that distortions are inevitable when listing the salient points of any model or approach, for criterion referencing one could add to the above mentioned features that:

- It uses predetermined levels of performance
- If the criteria are made available to teachers (and students) assessment becomes an open rather than a hidden process.

- It focuses attention on performance resulting from instruction.
- Testing is linked to content matter taught
- Anyone who fulfils the requirements specified in the criteria, and meets the standards, will have their achievement recognized.
- As a result it has the potential to motivate, and cooperation rather than competition is encouraged
- It yields an explicit description of what is being measured and what has been learned
- Typically used for guidance and diagnosis; allows teachers to direct their teaching to the areas on which a pupil does not perform well
- Provides policy-makers and future employers with clear information about what students can actually do

Though the characteristics listed above tend to highlight the positive features, they should not blind us to the fact that, in practice, criterion-referenced systems of assessment are not problem free. Many writers on assessment suggest that difficulties commonly attributed to the approach include the following:

- Writing (criterion-referenced) tests may be extremely time-consuming, and the fundamental restructuring of teaching/learning makes considerable demands on teachers. (Lloyd Jones & Bray 1986:117)
- Defining learning activities in terms of tasks to be mastered can present difficulties. Certain subjects such as mathematics and topics such as motor skills and mapping are particularly amenable to this approach, but it is more difficult to establish criterion-referenced tasks for 'creative writing' or 'art'. (Marsh 1997:173)
- In actually writing a test, a high degree of precision is often required in deciding the level or criterion that is acceptable; this can lead to over-specification and a focus on narrow, tightly defined objectives. (Lloyd Jones & Bray, 1986:117 & Gipps 1994:93)
- Thus criterion-referenced tests may measure only short-term objectives; their use requires a level of pre-planning and inflexibility in the curriculum and assessment scheme that may be undesirable. (Lloyd Jones & Bray 1986:117)
- It is difficult to assess high-order objectives with them, and, as Ebel says – 'Knowledge does not come in discrete chunks that can be defined and identified separately.' (Lloyd Jones & Bray 1986:117)
- Even when the domain and performance criteria appear clearly specified, the tasks generated to test performance will inevitably differ somewhat from place to place and occasion to occasion, in terms of format, context and problem content. (James 1998:148)
- Despite the rhetoric about the robustness of criterion-referenced systems, they are operated by human beings who tend to 'interpret' using their own internalised, holistic conceptions about what constitutes competence; assessors tend to 'compensate' or 'make allowances' for performances in certain contexts. (James 1998:148)

- Aggregation, that is the collapsing of the detailed performance profile for each individual into a single reporting figure or grade, is a major problem for criterion-referenced assessment. (Gipps 1994:85)

Norm and criterion-referencing: complementary or mutually exclusive processes?

In reality, no assessment practice will belong entirely in a single referencing category, because assessment tends in practice to involve elements of both norm and criterion-referencing. The polarisation between norm-referencing and criterion referencing is, therefore, rather misleading – as Croll suggests, *the distinction between these two approaches is by no means as clear cut as is often supposed*. He goes on to explain that, *clearly norm referenced tests cannot be compiled without reference to educational criteria (and) conversely ... criterion referenced assessment must, in practice, take account of the performance of children*. (in Pollard 1996:269). He supports this by indicating that when we attempt to construct an educationally useful criterion relating to an element of skill and knowledge, we inevitably draw on at least an implicit model of what is appropriate for a child of a particular age, or with particular educational experiences. *Consequently, norm based criteria are almost certain to be part of any apparently criterion referenced assessments, even though this may not be made explicit. as William Angoff once remarked, if you scratch the surface of any criterion-referenced assessment you will find a norm-referenced set of assumptions lurking underneath ...* (in Dylan Wiliam 1992:17).

Black (1998), therefore, suggests that it is unhelpful to think of 'norm-referenced' and 'criterion-referenced' as phrases which describe two completely different kinds of test. Instead, norm and criterion should be seen as two different sets of properties possessed by every test, and/or, by the possible interpretations of its results. A point Lloyd Jones & Bray (1986) support, suggesting that it is more the way assessing is judged, than the instrument used, which indicates criterion or norm-referencing. This belief leads them to conclude that not only do the two approaches have more in common than is usually recognised, but that they can, and often do, co-exist. Thus, norm and criterion-referencing can be seen as complementary approaches, both of value at different points in a pupil's career.

Having said this, however, if the basic aim of assessment is to promote students' learning, one must agree with Loyd Jones & Bray who stress that *assessment can only be justified if it benefits, either directly or indirectly, our children's learning. It is not an end in itself ... assessment should always have a clear purpose*. (1986:12). This should be taken as a reminder that fitness for purpose is a key concept in assessment, and that no single mode is likely to fulfil all requirements.

Example to illustrate positive and negative aspects of criterion-referenced assessment practices

Many of the benefits, difficulties and problems encountered when assessment is criterion-referenced, are apparent in the Argentine national education system.

At primary level, for example, assessment is justifiably criterion-referenced because it is used in a diagnostic, formative mode to support the teaching learning situation. Testing is linked to content matter taught; all those who fulfil the requirements specified in the criteria, and meet the standards, have their achievement recognised. At the same time, the results can alert teachers, allowing them to direct their teaching to areas on which a pupil does not perform well. Criteria and level descriptors for the core content are suggested by the ministry – based on these minimum requirements, and the social context

within which a school finds itself placed, each school sends in its own curriculum proposal, including assessment criteria, for approval.

At an individual student, or school level, the benefits from such a system are self-evident. At a national level, however, it does mean that standards vary from district to district, and from school to school, effectively perpetuating social inequalities.

I felt this was clearly evident when talking to teachers at a small rural, primary school in Argentina earlier this year. Their curriculum proposal is less demanding than many an urban school's would be, reflecting the fact that some parents are illiterate and many are out of work. Further, teachers also tend to 'compensate' when interpreting criteria, and thus, even when the criteria are similar for both schools, achieving a 'very good' there will be different to a 'very good' at a school in the province of Buenos Aires. Despite this, the Year 4 teacher believed it was justifiable because she felt it was a way of motivating children and so keeping them in school - by staying in school the children are fed once a day and develop basic literacy skills.

The bilingual, fee paying school I work at in the province of Buenos Aires, provides a stark contrast to this. The curriculum proposal reflects the fact that most parents value an international education. Assessment in the primary school is also criterion-referenced, but the teachers development and interpretation of criteria is influenced by the fact that they expect, and demand, a higher standard from the pupils. A 'very good' in this context is not comparable to the rural school's 'very good'.

Assessment in the English department, also criterion-referenced, does, however include an element of differentiation, since it must take into account that children have not all had the same exposure to the second language. In Year 1, for example, the children are continuously assessed against criteria, developed by the school, that reflect the learning objectives for reading, writing, listening and speaking skills - using level descriptors (1 - 5), for each area. On the basis of the assessment results, children are grouped and re-grouped as different levels of fluency and different learning styles are observed. This approach links with psychological research on 'match', which suggests that for effective learning tasks must be pitched appropriately. If tasks are regularly too easy, the pupil will become bored; if too hard, the pupil may become demotivated.

Though the intention is not to 'stream' children into A and B groups, but group them according to whether they learn best through more or less structured approaches, parents do tend to view those being taught in a less structured way as the more advanced students and, as a result, undermine the potential benefits of this system.

Example to illustrate the need for, and potential benefits of, norm-referenced tests

Various assessment procedures are used by official bodies to monitor standards or evaluate the effectiveness of some particular method or policy - for, as Broadfoot (1996:7) indicates, *educational assessment is not just about judging individual potential and performance: it has always been just as much about judging institutional quality*. Particularly when changes have been introduced, the effectiveness of the curriculum as a whole needs to be constantly monitored in the interest of public accountability, and to provide information for further educational planning and policy-making. As Harlen et al (in Moon & Shelton Mayes 1995:281) suggest, *information about pupils' achievements is necessary in order to keep under review the performance of the system as a whole*. In this context, James' (1998) suggestion that judgements about an individual's progress in relationship to others is sometimes helpful in identifying whether there is an obvious

problem that needs to be tackled urgently, could also be applied to school and national levels.

The above points, added to the fact that it is becoming increasingly commonplace for governments the world over to link economic growth with educational performance, have been the driving force behind the introduction of norm-referenced tests, produced by the Argentine Ministry of Education, to be taken by all Year 12 students. The intention is that, by the year 2002, results of these tests will act as a selective device to determine university entrance. Though many problems still have to be overcome, the overall purposes are:

- To evaluate the existing educational system
- To compare results both between schools and across different areas of the country (urban/rural)
- To evaluate the implementation of the changes introduced by the Ley Federal de Educación 1993.
- To guide the allocation of resources and funds to those most in need.
- To raise standards and restore values: many young people do not take their secondary education seriously as entrance to university depends on internal school assessment systems. This would change with the introduction of national, competitive, selective tests.

Despite the fact that norm-referenced testing has been criticized for being *elitist and being obsessed with differentiating children and sorting them into a hierarchical ranking* (Croll in Pollard 1996:269), and that these tests would support Broadfoot's argument that assessment aiming at selection can lead to control of the actual content taught, the country cannot afford to offer a university education to all citizens. Thus, selection is seen by many as a justifiable way to improve the current situation faced by the national universities.

Conclusion

Assessment is inescapably central to any educational enterprise – indeed Rowntree (1987:118) believes that in education *assessment is intentional and of the essence*. It is, therefore, important for educators to understand the concepts that underpin it, so that they can explore the alternatives and dissipate confusion. This is particularly important if one acknowledges that there are many types of assessment, each suited to particular purposes, and that they cannot safely be used in other ways.

Perhaps, as Pollard (1996) suggests, the single most important source of confusion does, in fact, stem from misunderstandings about the purposes of different forms of assessment. Thus, reviewing the purposes of assessment is appropriate because all subsequent decisions about choice of method should be made on the basis of judgements about 'fitness for purpose' – that is, the methods of assessment chosen should be appropriate to the reason for the assessment and the type of information being sought. This is accentuated if one accepts that, as Lloyd Jones & Bray (1986) point out, the tools of assessment are crude and imperfect; that the various approaches to assessment are interlinked and overlaid in any practical assessment procedure.

Therefore, as far as norm or criterion referenced assessment is concerned, it is not an 'either-or' issue, but a 'fitness for purpose' issue. As Black (1998:77) stresses, *their*

theoretical value is as conceptual tools to identify and describe the ways in which different systems operate rather than as distinct categories within which any one system can be uniquely classified.

Finally, one must acknowledge that often changes in assessment are a response to a particular need in society, for educational change does not take place in a social or political vacuum. However, as Young (1998:5) indicates, when reviewing current practices it is *important not to fall into the trap that radicals can set for themselves of equating the past with what is bad and the future with what is good* (1998:5). We need to recognise the limitations of educational assessment, accepting it as a useful but, nevertheless, imperfect tool.

References

- Black, Paul (1998), *Testing: Friend or Foe? Theory and Practice of Assessment and Testing*, London, The Falmer Press.
- Broadfoot, Patricia (1996), *Education, Assessment and Society*, Buckingham, Open University Press.
- Brown, S. (1994), *Assessment: a changing practice*, in Moon, B. & Shelton Mayes, A. (Editors).
- Croll, P. (1996), *Norm and criterion referenced assessment*, in Pollard, A. (Ed.)
- Filer, Ann (2000), *Assessment: Social Practice and Social Product*, London, RoutledgeFalmer.
- Gipps, Caroline (1994), *Beyond Testing. Towards a theory of Educational Assessment*, London, The Falmer Press.
- Harlen, W., Gipps, C.; Broadfoot, P. & Nutall, D. (1994) *Assessment and the Improvement of Education*, in Moon, B. & Shelton Mayes, A. (Editors).
- James, Mary (1998), *Using Assessment for School Improvement*, Oxford, Heinemann Educational Publishers.
- Lawton, D., (1996), *Beyond the National Curriculum*, UK, Hodder & Stoughton.
- Lloyd Jones, R. & Bray, E. (1986), *Assessment: From Principles to Action*, London, Routledge.
- Marsh, C. J. (1997), *Planning, Management & Ideology, Key Concepts for Understanding Curriculum 2*, London, The Falmer Press.
- Moon, B. & Shelton Mayes, A. (Editors) (1995), *Teaching & Learning in the Secondary School*, London, Routledge.
- Pollard, Andrew (editor) (1996), *Readings for Reflective Teaching in the Primary School*, London, Cassell.
- Rowntree, Derek (1987), *Assessing Students: How shall we know them?* London, Kogan Page.
- Stobart, Gordon & Gipps, Caroline (1997), *Assessment. A Teacher's guide to the issues (3rd edition)*, London, Hodder & Stoughton.
- William, D. (1992) Some Technical Issues in Assessment: A user's guide, *British Journal of Curriculum and Assessment*, Vol. 2, No.3, 1992, pp11-20.

Wood, R. (1991), *Assessment and Testing: A survey of research*, UK, Cambridge University Press.

Young, Michael D. (1998), *The Curriculum of the Future*, London, The Falmer Press.