LITERACY AND TESTING: EXPOLORING ALTERNATE MEANS OF ASSESSMENT FOR ALL STUDENTS

Teresa Walsh   Therese Weir
University of Western Sydney Macarthur

Abstract
It has long been argued that formal or traditional assessment strategies often disadvantage many students due to the socio-political and cultural aspects of the procedures used. It has also been argued that traditional assessment has often been misused and its purpose not clearly related to the needs of the students. Tyler as early as 1942 (cited Horowitz 1994) argued about the inequities of assessment procedures and proposed a move away from the traditional forms towards a more qualitative assessment method There is now occurring within the education system, a shift away from the traditional assessment methods of standardised testing towards more processed focused, alternative and authentic assessment methods, which take into account the latest theories and research into how children learn and concentrate on encouraging literacy development and success for all learners. This qualitative research project is establishing a process where the divergent beliefs of teachers are mutually explored on the issues of assessment and reporting methods. It is a pilot study that is part of the requirements for degrees currently being undertaken by the researchers. The focus for the project is on the appropriateness of both traditional and authentic methods of assessment and reporting for literacy learning for a diversity of learners. The process exposes and clarifies the different beliefs of teachers and allows the building of an agenda for negotiation.

Traditional and Authentic Assessment - A Clarification of Terms

Throughout the discussion in this paper, assessment will be referred to in two ways, traditional assessment and authentic assessment. Traditional assessment refers to the use of testing as a form of assessment. It focuses on achievement at a particular point of time, known as the snap shot effect (Salinger 1990 as cited in Hancock 1994 p.11). This type of assessment tests knowledge only, and is not always linked to the day to day activities of the student in the classroom or to the processes of learning which are occurring. It has a sorting or ranking effect on learners and is used to identify weaknesses. It is summative in nature and includes such things as multiple choice, standardised and criterion referenced tests. Traditional assessment is quantitative based and is seen as a way of measuring what a student can do and is based on the assumption that -human nature can be quantified (Hancock 1994 p. 8)

Authentic assessment methods are qualitative in nature and refer to

those methods of assessment which focus on the processes of learning.
It assesses the knowledge, skills, attitudes and values of the learner
and is ongoing, relevant to classroom practices and is reported in a
descriptive way. This type of assessment includes the use of profiles,
portfolios, annotated work samples, conferencing, student self
assessment, interviews and observations. It allows for all those
involved in the learning process to be part of the procedure and
encourages the participation of the teachers, students and parents.

TRADITIONAL ASSESSMENT

A Historical Perspective
The focus of traditional forms of assessment has been mainly on the use
of tests, especially the use of Intelligence Quota (I.Q.), standardised
and norm referenced tests. It has been in use in the United States
since 1845 with the implementation of Binets intelligence test
(Glasser and Silver, 1993). The I.Q score achieved on that test often
determined what educational experiences, curriculum exposure and work
opportunities a student accessed throughout his/her educational career
(Darling-Hammond, 1994).

While the purpose of the widespread testing was reported to be to
improve social justice by providing all students with the curriculum
and learning opportunities best suited to their needs, the reality of
the situation was that testing was used as a tool of social control. It
allowed for the provision of alternative and less effective
educational opportunities to certain sections of the population based
on socio-political, racial and gender grounds (Darling-Hammond, 1994).
The tests being used were devised by members of the dominant cultural
and social group, those being middle to upper class, and the questions
and knowledge required to complete the test reflected the
understandings and culture of this group (Garcia and Pearson, 1994).
This type of testing created inequalities for the minority and
disadvantaged groups such as the Aboriginal, Native American, African
American, Latino and lower social classes. It resulted in the exclusion
and segregation of these students by placing them in the lower tracks
of the education system and in specialised educational and remedial
classes.

Terman (cited in Darling - Hammond, 1994, p10) used the scores on I Q
tests to prove that various cultural groups were of lesser intelligence
than the dominant or mainstream cultural group. This gave validity to
racial social and gender discrimination in the educational arena. His
results reinforced and extended the inequalities in education, allowing
them to become ingrained in the system. The bias and inequalities of
the test-like forms of assessment used in the past have persisted and,
as indicated by research carried out by Oakes (1990) and Garcia

(1991), are still apparent in the system today.

Research (Horwitz 1995, Darling-Hammond 1994, Garcia and Pearson 1994) has continued to support the concept that traditional assessment has often been misused in both the purpose for which it was employed and how results were used. Often the focus was not on benefits to the student. While the anomalies of traditional assessment were well known, traditional assessment continues to enjoy popularity, prestige and power (Bintz and Harste 1992) and still occurs on a grand scale in the schools today. In the United States in 1986-87 an average of one month in every nine was spent in public schools administering over 105 million standardised tests to 39.8 million students , an average of 2.5 standard tests per student per year (Harste 1990 as cited Bintz and Harste 1992). In Australia there has been a re-emergence of traditional forms of assessment as a method of evaluation. One example of this is the introduction of Basic Skills Testing in New South Wales in 1989 (Hancock, 1991, p2), it was seen as an attempt by the state to monitor the standards of the schools and a way of bringing about educational reform. In Victoria there has been the introduction of the Learning Assessment Project (LAP) which is a standardised form of assessment completed by State schools in years five and three and designed to ` evaluate levels of student achievement (Cook and Horwood 1996). The LAP has been criticised by teachers as being used as a tool of accountability with the possibility of adverse effects on students from diverse backgrounds.

Test Bias
Many of the traditional assessment tests that are still being used today are culturally, socially and to an extent gender biased and do not assess students equitably. Research (Monty, Neill and Medina 1989, Davidson 1994, Garcia and Pearson 1994) has indicated that in the development of new tests, a trialing procedure is implemented known as the norming process. This involves the use of a group of randomly selected participants as a test sample to view the appropriateness of the test strategies. The population selected tends to veer towards the mainstream with the sample population often consisting of 75 % from Anglo Saxon background (Monty Neill and Medina 1989). The sample is not necessarily a true representation of the population in general as it tends to include only a small proportion of participants from minority groups, and therefore gives a test a norm bias (Garcia and Pearson 1994). Nor are there any guarantees that testing is used on a group of participants who fit the characteristics of the original group used in the trialing process. Population compositions vary regionally and can change over a brief period of time. Davidson (1994), in his research demonstrated the problems that can arise with a change in population when he investigated the difficulties that Hispanic students had when completing a test that was normed on native English speaking children.

Another issue that has arisen in the use of the norming process, deals with the exclusion of certain test questions. One of the criteria of exclusion is that if the majority of the low scoring participants answer a question correctly, while at the same time the high scorers answer incorrectly then the question is discarded as the question does not fit the test norm. This will further reduce the chance of the minority group doing well as these are the questions most likely to be answered correctly by this group (Monty Neill and Medina 1989).

The issue of content bias has been raised by many researchers (Garcia 1994, Mpofu 1994) as a major area of disadvantage to minority groups. The content of many tests reflect the dominant cultures values, knowledge and understandings. Success on the test is determined according to the values and norms of that particular culture. Garcia (1991), found that for students to do well on tests they need to be culturally aware in terms of the dominant culture and have background knowledge of what the content of the test pertained to and have participated in the mainstream culture.

The analysis of data collected by Samuda (1975) and Williams (1975) (cited in Garcia and Pearson 1994 p 344) and Mpofu (1994) demonstrate the issue of content bias of tests for minority groups. The researchers found that the test questions are often biased towards the dominant social and cultural groups and disadvantage any who are not from this group. Many tests do not take into account the different ways of knowing or problem solving that are characteristic of different cultural groups. The test strategies do not reflect a variety of learning styles other than those of the dominant culture and therefore do not allow for answers that reflect a different way of thinking or knowing or answers that are derived from viewing a situation from a diverse cultural or social perspective. The QUASAR Project (cited Silver, Smith and Nelson 1993 p409) demonstrates the issue of diverse cultural thinking in problem solving, when reviewing the variety of answers to the bus ticket problem. While there was only one correct answer according to the test, many students gave the wrong answer but could justify their answer on cultural or economic grounds

Traditional Assessment and Students from Language Backgrounds other

than English
Research has shown (Garcia 1991, and Mestre 1984) that students who come form language backgrounds other than English (L.B.O.T.E.) are disadvantaged through the use of traditional assessment methods. Students need a firm understanding of the language used in the test in order to communicate their abilities. Students from L.B.O.T.E. backgrounds and those who speak non standard dialects of English such as Aboriginal (Shrubb 1989) and African American students do not always have a firm understanding of the formal use of the English

language as used in a test situation (Monty Neill and Medina 1989). These students face difficulties with interpretation of unfamiliar vocabulary and the use of culturally inappropriate synonyms (Schubb 1989). Problems also exist with translation of test questions into their first language and the time taken to translate questions and answers for encoding. The issue of misinterpretation of language and time restraints are demonstrated in Garcias (1991) comparison of test scores between bilingual Hispanic students and monolingual students. The Hispanic students were often too slow to complete the tests in the given time due to the extended thought processes they went through and subsequently scored poorly compared to the monolingual students. The mark given was not a true indication of what the student was able to do because of language and time constraints placed on them. Often a tests purpose ended up being to test the students level of English and not their learning (Monty, Neill and Medina, 1989).

Curriculum Development
The effect of using results of traditional assessment methods on the curriculum that is taught in schools has become an issue of concern (Hancock 1994). Research (Bracey 1996, Garcia and Pearson 1994) has indicated that the result achieved on a test score can be used to drive the curriculum being taught in a school and the way it is taught. This has led to further inequalities in education and resulted in those from non dominant backgrounds being disadvantaged. Schools have been known to align the curriculum to the content and structure of the test format (Bracey 1996), and the focus of teaching to be on students doing well in the tests. This has become known as teaching to the test and has resulted in a narrowing of the curriculum (Salinger, 1991). A survey carried out by The Centre for the Study of Testing, Evaluation and Educational Policy (1992, cited Garcia and Pearson, 1994, p355) found that teachers in low socio economic areas taught to the standardised, compulsory tests to ensure that the students did well. This resulted in the students in this area being exposed to a limited curriculum which was aimed at a student passing or doing well in a particular test. By the teachers teaching to the test, the students were taught lower order cognitive skills and little attention was paid to the development of conceptual knowledge or problem solving. This resulted in these students not developing higher order skills and limited their exposure to newer and more relevant teaching instruction. Madaus (1994) in a recent study, highlighted the practice of teachers working in low socio economic areas spending a large part of teaching time completing practice tests, working on test preparation and teaching about test strategies. By exposing students to frequent testing it has been shown that students scores will improve (Kika, McLaughlin and Dixon, 1991). However, while this may have improved a students score it also limited what was being taught to the student.

Colvin Murphy and Ross (1990) in their research into administrators attitudes towards testing, found that the majority of administrators

used the results of testing to identify weak curriculum areas, validate
the curriculum being taught and justify changes in the curriculum. The
decisions about what was to be taught and how it would be taught in

their schools were made based on the results of the tests and not on
the needs of the students in terms of their development. Colvin Murphy
and Rosss research further supported what Sproull and Zubrow(1981)
and Ruddell and Kinzer (1982) (cited Colvin-Murphy and Ross 1990, p3)
had found in the eighties and demonstrated that little change had
occurred since this time.

Specialised placement and tracking of students
Results of traditional assessment procedures such as tests, are used
to place students in both remedial education classes and in extension
classes. It has been shown that L.B.O.T.E. students and those from
diverse backgrounds tend to do poorly on tests, not due to an ability
or learning problems but rather due to the biases of the tests (Garcia
1991). It follows that the same students will do poorly on placement
tests and result in a disproportionately high percentage of these
students being wrongly placed in remedial classes and
disproportionately small numbers placed in extension classes (Glasser
and Silver 1993). Oakes (1990) in his research on the placement of
students in specialised classes found that students of a minority
background were seven times more likely to be identified as being of a
low ability and needing specialised placement as their white
counterparts. His research also concluded that placement in a special
needs class usually resulted in a decrease in performance on the part
of the students and a decrease in the level of educational instruction.
 Shepard (1983 cited Glasser and Silver, 1993, p499) found that the use
of testing to identify special placements had led to misidentification
of many students. The misdiagnosis was due to inadequacies in testing
and led to the wrongful labelling of many students from L.B.O.T.E. and
disadvantaged backgrounds. In Australia, one of the stated aims of the
LAP in Victoria is to pick up learning difficulties of a child in the
early stages so that the problems can be corrected (Cook and Horwood,
1996). There is a possibility that these children can end up with a
special placement in a support unit based on the use of the LAP.

Accountability
Traditional assessment methods have long been seen as a way of
monitoring teacher and program performance. The instrumental agenda
(Hancock 1991) sees assessment as a form of `accountability, selection
and standardisation where the level or standard of students is
measured as a reflection of how effective the teaching process is. The
concept of assessment in this situation is seen as ` an instrument for
broader contexts outside the school (Hancock 1991 p 3). This supports

the concept that the results of large scale testing achievement has
been used to compare the performance of individual teachers, classes
and schools. The higher echelons of the educational systems, both here
and overseas, have seen the results of testing as a means of holding
schools, principals and teachers accountable. Data collected by
Colvin-Murphy and Ross (1990) indicates that the purpose of
standardised testing was often perceived by principals as a process of
reviewing school programs and performance and was related more to
issues of accountability and politics than it was to improving
students learning processes. The purpose of wide spread testing was to
evaluate the schools and not necessarily to assess the students. Many
administrators and principals felt that the results of the tests were
seen as a measurement of their own performance. How well the school
competed in these tests was seen as a reflection of a principals or
administrators leadership and of the validity of the programs and
curriculum in place in the schools.

In turn, principals and administrators freely admitted to using the
breakdown of comparative test results to evaluate each teachers
individual performance in terms of how successful their students
performed on the tests. (Colvin-Murphy and Ross 1990). Results of the
standardised testing was considered to be a more accurate measure of
how well a teacher taught than as a measure of how well a student was
progressing.   At a regional level, results of testing were more likely
to be used to make decisions on educational policy and accountability
decisions with allocation of resources and the introduction of
specialised programs being based on overall test results. Within
schools results on testing have led to the allocation of rewards and
implementation of sanctions of teachers, and programs. (Darling Hammond
1994).

While this research was carried out in America, there is grave concern
that a similar situation could occur with the use of results from the
Basic Skills Test and the LAP. This is especially so since a change
occurred in the 1993 Basic Skills Manual which saw the omission of the
section on Confidentiality of Results. Previously it was clearly stated
that the results of a student may only be given to the parents and the
school were the child is enrolled. Now it is possible for the results
to be used to compare students and schools performance, as a way of
assessing teachers performance or as a way of distributing resources
While confidentiality has been assured, comparisons of various schools
achievement or of social or cultural groups within a school can occur
as can a comparison of teacher results.

The concept of the instrumental aspect of assessment can be seen to
have some merit. It has been promoted by education policy makers and
can be seen to be relevant in terms of ensuring that a ` broad and
balanced high quality education which is relevant(Excellence and
equity 1989,p 11) is being provided or there may be a legitimate

purpose for this type of assessment to be carried out by a researcher .

AUTHENTIC ASSESSMENT

Research has emphasised the inadequacies of traditional modes of
assessment for diverse learners and the need for the incorporation of
more equitable forms of assessment in terms of authentic assessment
methods. The need to cater for different cultural, social and language
differences amongst learners is paramount and has led to the
incorporation of a variety of authentic assessment methods into the
schools (Burke 1992). These methods include the use of checklists,
anecdotal records, observations, log books, product analysis
conferencing, portfolios and student self assessments.

The need for change has also been recognised by Government departments
and this has been demonstrated by the Eltis (1995) review into Outcomes
and Profiles which recommended that

..... in relation to assessment and reporting practices schools and
their communities be responsible for devising their own procedures,
with opportunities for community members to express views and indicate
their expectations . (Eltis 1995:v)

The focus of the NSW Department of School Education is seen to be on
schools developing assessment procedures which best suit the students
in their school, taking into account the needs of the students and the
community in general and by working collaboratively with the community.
This coincides with the shift in parameters of what the concept of
assessment is and the need to implement new and innovative modes.

Assessment is now about emphasising the achievement of students in
terms of learning outcomes and focusing in on the direction a students
learning now needs to take.

Authentic modes of assessment and the use of outcomes in assessment is
believed to be able to benefit diverse learners in a number of ways.
It has led to a rise in expectations of what the students can do, is
more focussed on educational purposes in terms of identifying student
strengths and weaknesses and planing learning experiences based on the
students needs. This approach identifies the progress of learning and
aims at fostering dignity and self esteem of the students through
acknowledging what each student has achieved in terms of their own
learning instead of in comparison to others.(NSW Department of School
Education 1994)

While the understanding exists that authentic assessment is beneficial
to diverse learners, concerns have been raised (Garcia and Pearson
1994) as to the lack of research that supports or that indicates the

extent of the effectiveness of authentic assessment. It is difficult to assess the impact it is having because authentic assessment is still in its infancy stage and more time is needed to monitor and gauge the effects. Authentic assessment is not without critics. Criticisms have related to the "ability of teachers to embrace the concept and change their instruction in the classroom to deal with authentic assessment" (Dewitz Carr, Palm and Spencer 1992 cited Garcia and Pearson 1994). Difficulties have also arisen with the understanding of what authentic assessment is and how these methods can be implemented in the classroom. Research into the implemtentaion of portfolios (Dewitz et al, Malone and Michelson 1993 cited Garcia and Pearson 1994) have indicated a varying response to the implementation process. Many teachers and administrators need to change their beliefs about assessment and some are finding it hard to leave behind the security they have found with the use of authentic assessment. Other arguments presented against authentic assessment have included that it is extremely time consuming on the part of the teachers and parents find it difficult to understand with many preferring the old ways. Results can be biased due to interpretation that teachers place on the students work. Qualitative assessment often relies on observations and the assessors intuitive insights, this can lead to problems with distortion, subjectivity and personal bias (Brookfield 1986 p.261) .

Where to now ?
The critical examination of the literature and research into the traditional forms of assessment such as testing indicates that there are many problems and difficulties associated with the fairness of tests to diverse populations and with the true purpose for which assessment takes place. It is obvious that traditional forms of assessment are not relevant to recent research into how children learn nor are they a useful tool in reviewing the processes that a child goes through to learn and are often misused. Many educationalists are supportive of the development of new assessment techniques and are exploring alternative, more authentic assessment. While this trend towards authentic assessment would appear to answer many of the problems associated with testing as a form of assessment there is a very real need to review what is available about authentic assessment and to research into how successful these methods are. Many articles have been written in support of authentic assessment techniques but actual research into the relevance and success of authentic assessment is still in its infancy stage and there is little research available at this stage which conclusively proves it to be successful. Does authentic assessment really counteract many of the problems of traditional assessment? Is it as effective for all students as many

educationalists claim ? What are the problems associated with this form of assessment or is it trouble free? How successfully can this form of assessment be implemented? Will the outcomes approach of authentic assessment result in improved teaching and learning? These are all

questions which need to be answered through further research and indicates that this is a relevant and frontier breaking area of research to be pursued.

Research plan and methodology

As the researchers want to "explore the construction ...., open each to critique in terms of other constructions and provide the opportunity for revised or entirely new constructions to emerge" (Guba & Lincoln 1989:89) a constructivist methodology is being used in this research. Maykut & Morehouse (1994:43-47) list eight characteristics of this approach. They are:
1) An exploratory and descriptive focus
2) Emergent design
3) Purposive sample
4) Data collection in the natural setting
5) Emphasis on human-as-instrument
6) Qualitative methods of data collection
7) Early and ongoing inductive data analysis
8) A case study approach to reporting research outcomes.

To ensure a trustworthy exploration of the focus of inquiry the data collection and data analysis procedures include several elements that increase the trustworthiness of the research findings. Multiple methods of data collection are being used to allow for emerging patterns to be further explored. An audit trail of all questionnaires, interviews, observations, field notes, tapes, photographs, and initial and ongoing analysis is maintained. The constructions of all participants and researchers are explored in a hermeneutic-dialectic process to "form a connection between them that allows their mutual exploration" (Guba & Lincoln 1989:149). This process ensures that the teachers' descriptions of their constructions have been accurately reported.

Methods
Sample
Using a purposive sampling approach the researchers chose a school in a socio-economically disadvantaged area in the south-western region of Sydney. The twenty five teachers have experience with both traditional and authentic assessment and reporting methods for literacy learning. A large proportion of the teachers are in the first five years of their teaching service and have been trialing a variety of assessment and reporting methods for the past four years.

Data Collection
As stated previously, the researchers employ a hermeneutic-dialectic process. This means that the participants are asked to present their constructions of the focus for inquiry and then analysed and validated.

Participants are introduced to the themes from others and invited to comment on those themes. This process has been repeated with new participants being added until a consensus is reached or an agenda for negotiation has been clarified. The data is being collected from three sources:
* Firstly, a questionnaire has been used with twenty-five consenting teachers. It consisted of both open-ended and a minority of closed questions.

* Secondly, informal interviews with four teachers are being conducted using the process described above. Permission to audio-tape has been sought prior to the commencement of the interviews. A typist transcribes the interviews for analysis, by the researchers, shortly after the interview.
* The interviewed teachers are also being observed in the classroom in the process of their day to day assessment and reporting procedures. The researchers take field notes during the observations with the permission of participants. Photographs have been taken by the researchers but these are for the purpose of report presentation only. Parents of children to be photographed have given their consent for this activity on the information sheet.

Data Analysis
The analysis of data will reflect the emergent design of the research in that it is ongoing and involves the constructions of all parties including the researchers. Patterns will be identified by the researchers, and constantly reviewed with further data from all sources with the aim of the process to reach a consensus or an agenda for negotiation. The computer software package Nud.ist is being used to assist with the organisation of the data.

Results to date - Consensus of themes

1) Teachers' Understanding Of Assessment And Reporting

Assessment is an integral part of the teaching / learning process. I know what my students can do and where they are going. I can gather and interpret information to aide this process.

2) Teachers' Understanding Of Why They Assess And Report

1) inform teachers about achieving outcomes
To gauge their progress - strengths and weaknesses in all K.L.As. It allows me to move a child along in their learning as well as finding the correct starting points.

2) inform teacher re future planning

I know what they can and can't do and enable me to program and provide
learning experiences which will be most beneficial to them, also it
helps me report to parents.

3) inform parents
.... so that I have documented evidence of my student's progress to
enable me to accurately report to parents both in interviews and for
formal reports.

3) Teachers' Construction Of The Ideal Methods Of Literacy Assessment
And Reporting

 In terms of process and product.
 *conferencing, both peer and teacher. A collection of work samples
presented with the progress and achievement comments. A 3 way pupil led
conference.
 * Task orientated, showing a student's progress along a learning
continuum. It should always be balanced and varied. An emphasis placed
on what a student can do
* continuous day to day observations and interactions.
* Those based on acknowledging student's individual progress and which

are part of a teaching and learning cycle. ie. student
portfolio/progress book with work samples or checklists that report
what a student knows and provides feedback about what type of learning
needs to take place.
* It should be an ongoing process documenting all progress made by the
chid. This should include work samples, anecdotal notes, checklists -
variety of assessment strategies to provide a balanced view to parents.
eg progress books or portfolios.

4) Effectiveness Of Authentic Assessment

* I am happy with them. They are outcome based and reflect students
individual progress.
The school policy and practices also allows for experimentation and
sharing of ideas so that reporting formats are continually being
refined and improved.
* I believe it provides a detailed, accurate representation of what
each child can achieve and where they are developmentally.
* Time consuming but effective if used in planning, teaching and
learning experiences.

5) Difficulties With Authentic Assessment

Time consuming and difficulties with parent understanding and

participation
* very time consuming, often ineffective as they depend on parent
input/participation.
* take too long and ineffective for parents
* pretty ineffective as it doesn't tell parents much

6) Traditional Assessment Appropriateness

a) Advantage for diverse learners
* as a one off, or a means- can be used to support idea of where child
is at. It is not the only means.

b) Disadvantage for diverse learners
* I would be against testing with my children as I have a clearer
understanding of their abilities through observations and work samples.
* performance is dependent on child's mood, health, emotional state,
attention
* span and time.
* traditional forms of assessment were not varied and unfair. They
compared children and did not show what the student could do.

7) The Relevance Of Outcomes On A Learning Continuum

a) tracking student learning
* I use the outcomes as a guide to where the children in my class are
at on the learning continuum.

b) assist in future planning
* asses in terms of where they are at in the outcomes and program
activities based on the information.
* .... they are the basis of my planning and programming.

c) assist in reporting
* .... all reporting is based on achievable outcomes with appropriate
samples of work

BIBLIOGRAPHY

Bintz, W. and Hartse, J. (1992) Whos Testing Who in Assessment and
Evaluation in Whole Language Program. Book of Readings, Integrated
Studies UWS Macarthur 1996.
Burke,K. (1992). Authentic Assessment: A Collection. Melbourne: Hawker
Brownlow.
Bracey , G.W. (1996). A Critical Look At Standards and Assessments.
Principal Magazine -Internet Explorer. May 1996
Brookfield, S. (1986).Understanding and Facilitating Adult Learning: a
Comprehensive Analysis of Principles and Effective Practices. San
Francisco: Jossey-Bass
Colvin-Murphy, C. and Ross, P. (1990). Attitudes of Administrators
Towards Standardised Tests: Perceptions of the Tests Impact on
Curriculum and Teacher Evaluation. Unpublished Paper
Connell, R.W. (1993) Schools and Social Justice. Pluto Press: Australia
Darling-Hammond,L. (1994). Performance -Based Assessment and
Educational Equity. Harvard Educational Review. Vol. 64 No.1. pp. 5-30
Davidson, F. (1994). Norms Appropriacy of Achievement Tests: Spanish-
speaking Children and English Childrens Norms. Language Testing Vol
11, No.1. pp 83-95
Eltis, K. (1995). Report on the Review of Outcomes and Profiles in NSW
Schooling. Sydney : Department of School Education
Garcia, G. and Pearson, D. (1994). Assessment and Diversity. Review of
Research in Education. Vol. 20 pp. 337-391
Garcia, E. (1991). Factors Influencing the English Reading Test
Performance of Spanish-Speaking Hispanic Children. Reading Research
Quarterly Vol.26 No. 4 pp. 371-390.
Glasser, R. and Silver, 1993. E. Assessment, Testing, and Instruction :
Retrospect and Prospect. Review of Research In Education. vol 20 pp.
393-419
Guba, E., Lincoln, Y. (1989). Fourth Generation Evaluation.   Newbury
California: Sage
Hancock, J. (1994). The Social-Political Nature of Assessment, Extract
from unpublished M.Ed (Hons) Thesis University of Wollongong.
Horowitz, R. (1994)A 75-Year Legacy on Assessment: Reflections From an
Interview With Ralph W. Tyler. The Journal Of Educational Research.
vol. 89 No. 2, pp. 68-75.
Kika, F., McLaughlin, T. and Dixon, J. (1991)Effects of Frequent
Testing of Secondary Algebra Students. Journal of Educational Research.

Vol.85, No.2 pp. 159-162

Madaus, G. (1991). The Effects of Important Tests on Students. Phi Delta Kappan. Nov. pp. 226-231

Madaus, F. A (1994). Technological and Historical Consideration of Equity Issues Associated with Proposals to Change the Nations Testing Policy. Harvard Educational Review Vol. 64 No.1 pp.76-92

Maykut, P., Morehouse, R. (1994). Beginning Qualitative Research: A Philosophic and Practical Guide. London: Falmer Press

Mestre,J. P. (1984). The Problem wit Problems: Hispanic Students and Maths. Bilingual Journal, 32, pp. 15-19

Monty Neill and Medina (1989). Standardised Testing: Harmful to Educational Health. Phi Delta Kappan May, Vol 70 pp. 688-695

Mpofu, E. (1994). Childrens interpretive strategies for class inclusion tasks. British Journal of Educational Psychology. Vol 64, pp 77-89

NSW Board of Studies, (1994). English K-6 Syllabus and Support Document. Board of Studies : Sydney.

NSW Department of School Education (1994) Introducing an Outcomes and Profile Approach. Sydney: Curriculum Directorate

Oakes, J. (1990). Multiplying Inequalities: The Effects of Race, Social Class and Tracking on Opportunities to Learn Mathematics and Science. Santa Monica, CA: Rand

Salinger, T. (1991). Educational Testing Service. Pre-Conference draft of Paper for Delaware Symposium.

Shrubb, G. (1989) Mass Testing: Toxic or Toxin. Australia : Australian Education Network.

Silver, E., Smith, M and Nelson, B. 1993. The QUASAR Project: Equity Concerns meet Mathematics Education Reform in the Middle School. In Fennema, E., Secada W. and Bird, L. New Directions For Equity in Mathematics Education. New York: Cambridge University Press

Woodward, H. (1993). Negotiated Evaluation. Newtown: Primary English Teaching Association.

14